Monash University

Faculty of Information Technology

Master of Data Science Minor Thesis

FIT5128 - Final Thesis Paper

A Design Exploration of Manual Segmentation Techniques for Immersive 3D Visualisation of Medical Images

---

Author:
Shuxian Qi

Student ID:
33930279

Supervisors:
Jiazhou 'Joe' Liu
Vahid Pooryousef
Yidan Zhang
Prof. Tim Dwyer

Total Words: 7765
June 6, 2025

MONASH UNIVERSITY

FACULTY OF INFORMATION TECHNOLOGY

MASTER OF DATA SCIENCE MINOR THESIS

Literature Review

3D Medical Image Visualisation and Segmentation in Immersive Environments

---

Author:
Shuxian QI

Student ID:
33930279

Supervisors:
Jiazhou 'Joe' LIU
Himashi PEIRIS
Vahid POORYOUSEF
Prof. Tim DWYER

Semester 2, 2024

Contents

1 Introduction

In the late 1800s, German physicist Wilhelm Röntgen discovered a mysterious form of radiation that could pass through human tissue but was obstructed by bone, leading to the creation of the X-ray. This groundbreaking discovery marked the beginning of medical imaging [1]. Since then, medical imaging has progressed from simple, blurry black-and-white images to today's high-resolution 3D scans that provide detailed insights into the human body's internal structures. These images allow physicians to not only identify abnormalities like lesions but also visualise the intricate three-dimensional anatomy of internal organs. With modern imaging techniques, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), healthcare professionals can use these detailed visuals for various medical applications, including diagnosis, treatment planning, and surgical guidance [2].

As these imaging technologies have advanced, detailed 3D images have become available, making effective visualisation and segmentation crucial for enhancing diagnostic accuracy and guiding precise treatments by clearly distinguishing healthy from diseased tissues [3]. Specifically, for medical experts or radiologists, such image visualisation provides a comprehensive view of anatomical structures, while segmentation on these images focuses on isolating specific areas of interest [4]. However, for data that is inherently 3D, such as brain MRI scans, displaying their visualisations on 2D screens is well known to suffer from issues of occlusion, perspective distortion and a loss of information [5]. Also, performing segmentation tasks of these 3D visualisations on 2D surfaces may lead to low efficiency because medical experts need to segment numerous 2D slices [6]. On the other hand, immersive technologies, such as Virtual Reality (VR) and Augmented Reality (AR), offer the opportunity to work with visualisation views via natural embodied interactions. Users can directly interact with or segment specific regions on visualisations using tracked devices or hand gestures. Recent advances in pass-through headsets, such as spatial computing on Apple Vision Pro (see Figure 1), may help to increase the adoption of immersive technologies across numerous domains (e.g., medical or health care). Yet, it remains unknown how to effectively and efficiently interact with 3D visualisations of medical images for segmentation tasks, which will eventually seamlessly integrate into existing medical workflows.

The aim of this literature review is to explore existing tools and software on how to visualise and segment 3D medical images, identify essential features

Figure 1: Apple Vision Pro. Source: `https://www.apple.com/au/apple-vision-pro/`

from the existing tools, and explore how immersive technologies (i.e., VR and AR) could benefit medical applications (e.g., diagnosis process).

Figure 2 shows an example of a scenario from ImmersiveTouch, where a 3D visualisation of brain structures is conducted using a VR interface. This allows for more intuitive interaction with anatomical models, providing a more immersive and precise environment for analysis and planning.

To achieve this aim and plan the research project, we will start with this comprehensive literature review to identify gaps in current segmentation techniques using immersive technologies like VR and AR. We will focus on existing interaction methods, segmentation accuracy, and challenges in medical integration. Based on these findings, we will design and develop prototypes to explore various interaction techniques to segment 3D MRI data of different body parts (e.g., brain and heart) in immersive environments. Then, we plan to evaluate these interaction techniques on accuracy, efficiency, and usability via a user study with experts and students from the medical field. Feedback from the study will guide further refinement of the techniques and their potential application in medical practice.

This literature review will begin with an overview of medical imaging modalities, including the capabilities and limitations of technologies such as CT and MRI (see Section 2.1.1). Next, it will focus on conventional medical image visualisation techniques for health applications, exploring their evo-

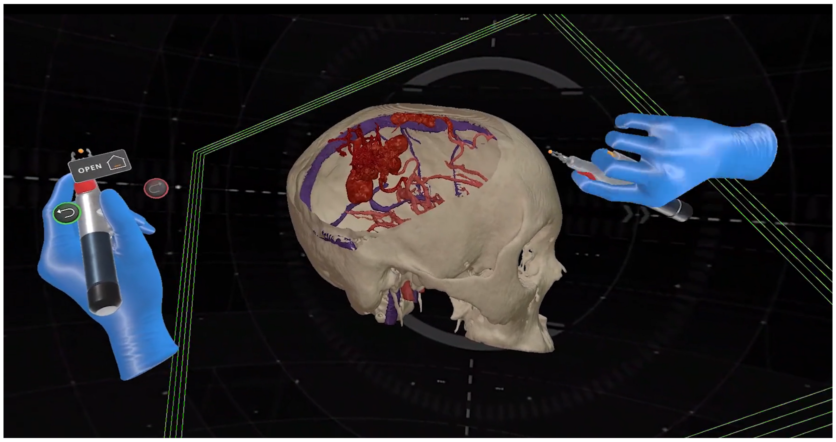


Figure 2: 3D Brain Visualisation Scenario from ImmersiveTouch [7].

lution and current use in medical practice (see Section 2.1.2). The review will then cover software tools for 3D medical image visualisation, analysing their functionalities and their role in enhancing medical workflows (see Section 2.1.3). Following this, the focus will shift to the application of medical image segmentation, highlighting its critical role in improving diagnostic accuracy and treatment planning (see Section 2.2.1), and the software tools specifically designed for segmentation tasks (see Section 2.2.2). Finally, the review will explore the use of immersive technologies, such as Virtual Reality (VR) and Augmented Reality (AR), in medical imaging, assessing their potential to transform visualisation and segmentation processes within medical settings (see Section 2.3.1 and Section 2.3.2). The review will conclude with a discussion of the current challenges faced in these applications and potential solutions.

2 Substantive Literature Review

2.1 3D Medical Image Visualisation

Medical imaging is a technology that provides detailed images of the body's internal organs and tissues to aid in the early detection of disease, accurately guide surgical procedures and monitor effective treatments by reducing the need for invasive procedures [8]. With the development of imaging technologies, 3D visualisation of medical images has become essential, as interpreting each individual slice manually would be extremely time-consuming for radiologists [9].

In this section, we'll first delve into the diverse imaging modalities that make 3D visualisation possible, then explore the traditional visualisation techniques used in healthcare, and introduce the software tools that bring these images to life, each with its own unique capabilities as well as limitations.

2.1.1 Medical Imaging Modality

Medical imaging technology has come a long way since the discovery of X-rays more than 120 years ago. Modern radiologists now use a variety of advanced modalities, including ultrasound, positron emission tomography (PET), computed tomography (CT), magnetic resonance imaging (MRI) and many others [10]. In this literature review, we focus on the most dominant modalities, CT and MRI, because of their ubiquity in clinical practice and their primary use in medical image segmentation tasks [11].

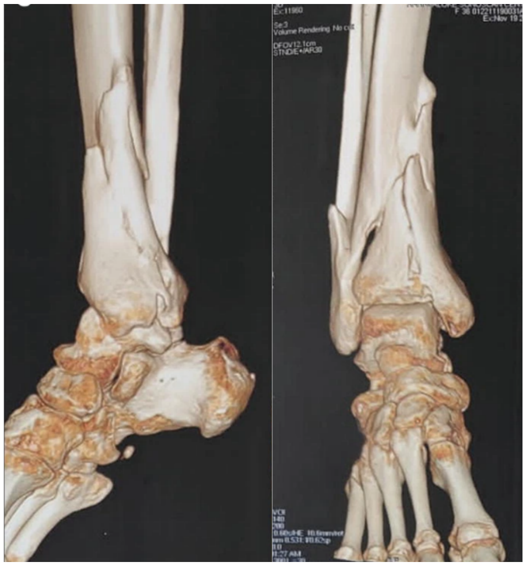


Figure 3: 3D CT Scan of Ankle Bones and Joint. From [12]

- **Computed Tomography**: CT scans, also known as "CAT" scans, utilise X-rays taken from multiple angles to generate detailed cross-sectional images of the body. This imaging technique constructs a three-dimensional representation by combining numerous X-ray images, offering a comprehensive view of internal structures. CT scans are particularly effective at visualising dense structures like bones (Figure 3) and organs due to their high contrast and resolution [13]. They are invaluable for diagnosing a wide range of conditions, including fractures, tumours, and internal injuries, and are essential in emergencies where rapid imaging is crucial for assessing injuries and guiding treatment decisions [14].

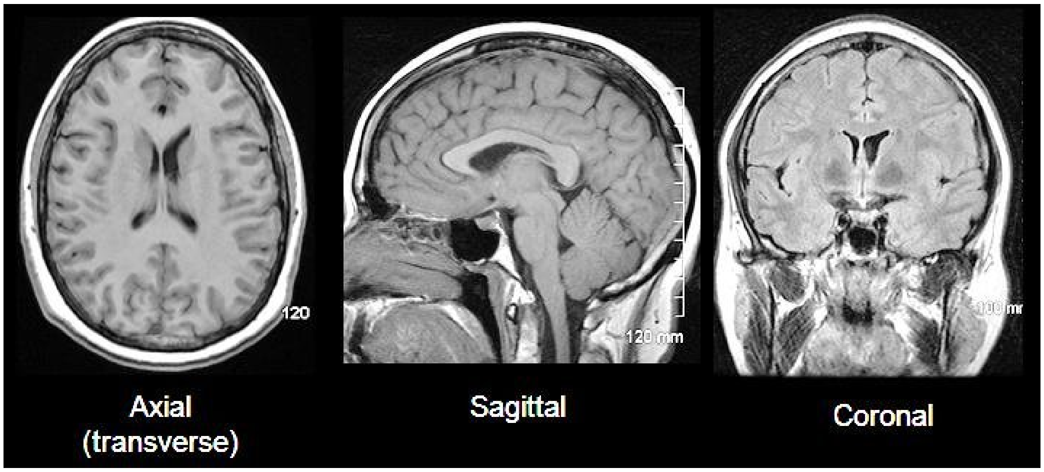


Figure 4: MRI Scans of the Brain in Axial, Sagittal, and Coronal Views. From [15]

- **Magnetic Resonance Imaging**: MRI uses powerful magnets and radio waves to create detailed images of the body's internal structures without involving ionising radiation, making it safer for repeated use, especially in sensitive populations like pregnant women and young children [16]. MRI can produce images in different planes, such as axial, sagittal, and coronal, providing a three-dimensional view of the body. It is particularly effective at imaging soft tissues, such as the brain (Figure 4), spinal cord, muscles, and joints, making it the preferred method for diagnosing neurological disorders, musculoskeletal injuries, and abdominal issues. Different MRI sequences, like T1-weighted images for anatomical detail and T2-weighted images for detecting fluid and inflammation, further enhance its versatility in diagnosing a wide range of conditions [17].

2.1.2 Conventional Medical Image Visualisation for Health Applications

With the development of medical imaging, healthcare applications have advanced significantly in areas such as diagnosis, surgery, and medical training [18]. In this section, we will discuss the applications and limitations of medical imaging in these fields.

- **Diagnosis**
  Medical imaging techniques have become indispensable for diagnosing almost all kinds of medical illnesses and abnormalities, offering detailed visualisations that enable healthcare professionals to accurately assess and understand their patients' conditions [19, 20]. Research has shown that CT significantly improves imaging of the brain and lungs, provides superior detection and monitoring of tumour mass changes as small as 1.5-2.0 cm in diameter during treatment [21, 22, 23]. Darty et al. has shown that MRI is primarily used for diagnosing skeletal metastases, cardiovascular diseases, and neurological conditions, and other soft tissues due to its detailed imaging capabilities and ability to differentiate between various tissue types without the risks of ionising radiation [24, 25].

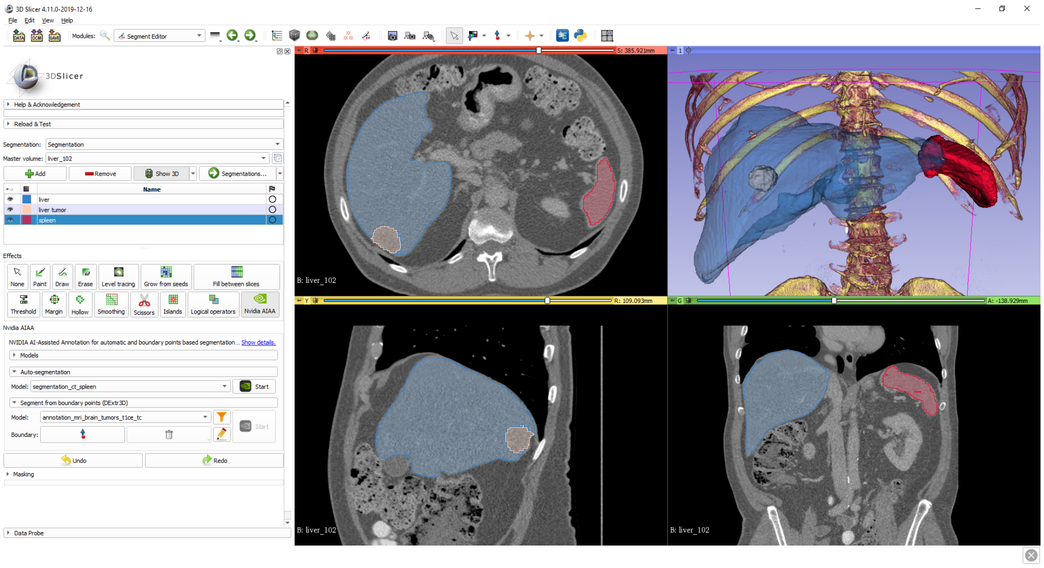
- **Surgery**
  Medical imaging is also vital in precision surgery, providing detailed visuals that guide surgeons in making precise incisions and enabling radiation oncologists to target tumours with high accuracy. G Gohla et al.'s study demonstrates that coronal T2-weighted images in intraoperative MRI have high diagnostic accuracy for detecting pituitary adenoma remnants during surgery [26]. However, limitations include the heterogeneity of intraoperative MRI image quality, affected by factors like head coil positioning and artifacts [27].

- **Medical training**
  Medical imaging techniques, such as CT, MRI, and X-rays, are widely used by medical specialists, from oncologists to internists [28]. Studies have shown that integrating radiological images into anatomy teaching significantly enhances student comprehension without adding to the subject's difficulty [29]. However, traditional anatomy teaching faces

limitations, such as ethical concerns and a shortage of cadaver donations [30]. This underscores the need for new technologies, like AR and VR, which offer interactive, vivid imagery that promotes more active and self-directed learning [31, 32]..

2.1.3 Software Tools for Medical Image Visualisation

Medical image visualisation tools enable users to easily view, analyse, and manipulate medical data from various imaging modalities. These tools increase accuracy, improve workflow efficiency, and support collaboration across healthcare. They can be categorised into 3D reconstruction, image fusion, and remote access platforms. Among these tools, 3D Slicer and PACS are the most representative, with 3D Slicer specialising in 3D visualisation and analysis, and PACS excelling in the management, storage, and remote access of medical images, particularly in teleradiology.



Segmentation: Segmentation
Master volume: liver_102

NVIDIA AI-Assisted Annotation for automatic and boundary points based segmentation... Show details.

Model: segmentation_ct_spleen
Model: annotation_mri_brain_tumors_t1ce_tc

Figure 5: User interface of 3D Slicer. From [6]

3D Slicer [1] is an open source platform for medical image visualisation and analysis with features such as multimodal imaging, image fusion, and cus-

---

[1] 3D Slicer: https://www.slicer.org/

tomisable interfaces for clinical and research use (see Figure 5) [6]. Its applications range from cancer research to real-time surgical guidance [33]. It also supports collaborative research with shared plug-ins that improve reproducibility [34]. Notably, Inoue et al. developed a neuronavigation system by integrating 3D Slicer with AR technology to assist surgeons in performing safer procedures [35].

PACS is a medical imaging technology designed for efficient storage and

Figure 6: Multi-Device Accessibility of PACS for Medical Imaging. From [36]

easy access to images from multiple modalities (see Figure 6) [37]. Its primary functions include replacing hard-copy storage, enabling remote access to images, integrating with other medical systems, and managing radiology workflows [38]. The advent of PACS has significantly advanced teleradiology, with studies demonstrating that remote readings via iPad can achieve similar diagnostic accuracy as traditional workstations [39]. However, image processing on mobile devices remains inefficient, highlighting the need for further improvements to fully realise PACS's potential for mobile and remote applications [40].

2.2 Medical Image Segmentation

Image segmentation is the process of dividing a digital image into multiple regions or segments to group similar pixels together [41], facilitating the identification of objects, boundaries, and features within the image [42]. Its main purpose is to simplify the image representation, making it easier to analyse and interpret [43], particularly in 3D visualisation systems [44]. The outcome of image segmentation is a set of distinct regions or contours that either encompass the entire image or highlight specific structures [45].

Research in image segmentation, particularly in medical image analysis, is a highly active and evolving field. Medical image segmentation is the process of dividing medical images from 2D/3D imaging modalities like MRI, CT, PET, and X-rays into distinct regions to isolate areas of interest, such as organs, tissues, or abnormalities (e.g., tumours).

Segmentation techniques can generally be divided into three main types based on the degree of required human interaction: manual, semi-automatic, and automatic methods [46, 47]. Manual segmentation involves experts manually outlining regions of interest (ROI) in medical images, either slice-by-slice or in 3D [48]. While advancements in semi-automatic and automatic segmentation using machine learning have reduced processing time [46], these methods still struggle with artifacts, low contrast, and irregular shapes, particularly in complex cases like tumour segmentation [49]. Therefore, manual segmentation remains essential. Research shows that anatomists prefer manual methods for their precision and control, making them the gold standard when automatic algorithms fall short [50].

Based on the importance of manual segmentation, this review focuses on traditional interaction methods and tools used in medical applications. Specifically, it explores how these interaction techniques enhance accuracy and user experience in manual segmentation tasks.

2.2.1 Application of Medical Image Segmentation

Medical image segmentation not only helps in diagnosing pathologies, segmenting organs such as the brain and liver, and assisting in treatment planning [51], but also improves the clarity of medical images. First, it allows for the precise delineation of regions of interest (ROIs) [52], which is crucial for tasks such as feature quantification, feature extraction, and statistical analysis, especially in radiomics [53]. Second, segmentation helps to visualise specific structures (e.g., tumours or organs) by highlighting them, while also removing unnecessary parts (e.g., bones in CT angiography) [49] that may hinder the observation of critical areas (e.g., vascular structures) [54].

Manual brain tumour segmentation involves experts manually outlining the tumour boundaries and labelling the regions of anatomical structures (see Figure 7) [55]. Traditional brain tumour segmentation techniques primarily utilise standard image processing methods, such as thresholding [56, 57]and region-based strategies [58], which are typically used for 2D image segmentation [59].

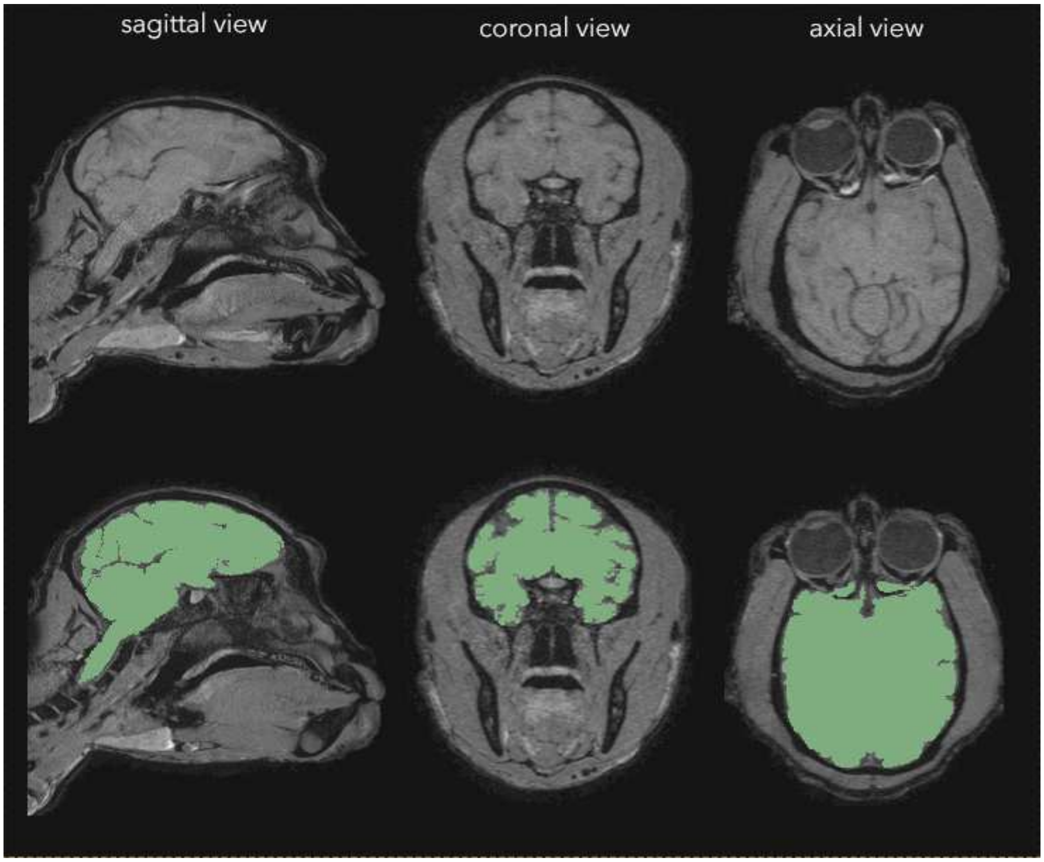


Figure 7: Manual segmentation of the brain (green regions) on MRI data using 3D Slicer. From [60]

Thresholding is often used as an initial step in brain tumour segmentation, which relies on intensity differences to isolate the tumour but struggle with low contrast or complex shapes [61], making them less effective in challenging cases. Region-based techniques like region growing and watershed segmentation focus on pixel similarity to define boundaries but are prone to issues like the partial volume effect and over-segmentation [62, 63].

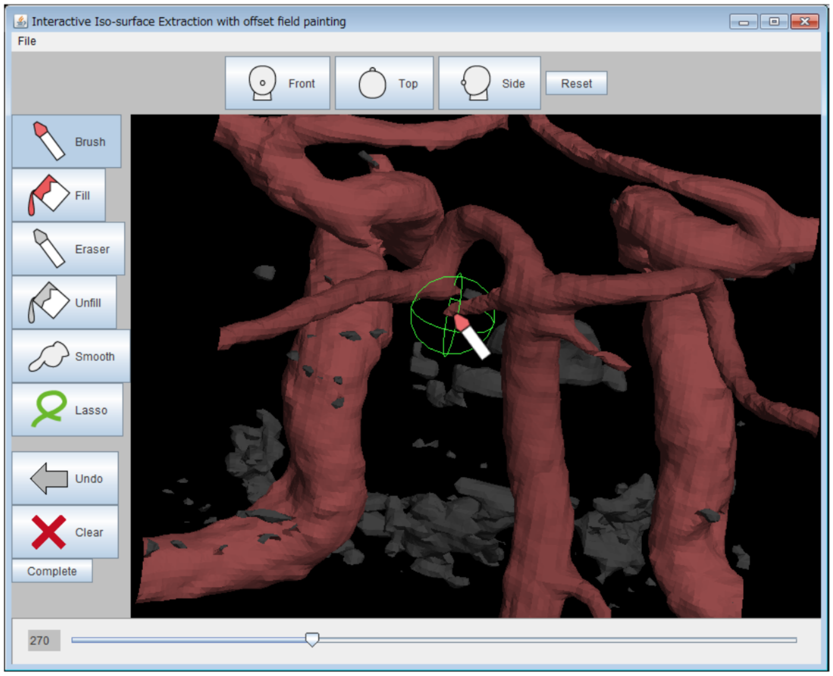


Figure 8: User Interface of the Threshold Field Painting system. From [64]

In 2016, Igarashi et al. developed Threshold Field Painting [64], a tool for manual segmentation and modelling of medical images (see Figure 8). The interface includes painting tools such as a brush for directly interacting with 3D volume data by assigning threshold values to specific regions. Users can adjust global thresholds and view segmentation results as isosurfaces, with colour coding to visualise different intensity levels. While the brush tool allows for precise control, the 2D interface is limited to single-user experiences and mouse-based input, restricting the user's ability to interact directly with 3D volumes in a more natural manner.

A key limitation of these traditional 2D methods is their inefficiency in handling 3D images [65], as slice-by-slice analysis often leads to inconsistencies and fails to capture the full spatial context of the tumour, underscoring the need for more advanced 3D techniques.

2.2.2 Software Tools for Medical Image Segmentation

Medical image segmentation tools are crucial for accurately identifying and isolating regions of interest, such as tumors or organs, in various imaging modalities like CT and MRI. These tools enable more precise diagnosis, treatment planning, and surgical navigation by allowing clinicians to interact with and manipulate complex medical data. In the following, we will focus on two widely used segmentation tools: 3D Slicer and ITK-SNAP.

3D Slicer is not only widely used for medical image visualisation but also plays a crucial role in medical image segmentation. This open-source platform supports a variety of segmentation tasks through its modular structure [66], allowing users to download extensions and perform both manual and semi-automatic segmentation. Modules such as Simple Region Growing Segmentation, EMSegment, and the Editor module [67] enable users to segment medical images based on intensity statistics or manual tools [6]. The platform accommodates multiple image formats and allows exporting segmentation outputs as NRRD [2] or NIFTI files, which is a very simple and minimalistic format [68]. 3D Slicer has been used in various studies, including the annotation of brain regions [69], segmentation of the hippocampal region [70], and extraction of abdominal organs [71], demonstrating its versatility in medical research. A key limitation of 3D Slicer is that it is not FDA approved for routine clinical use, restricting it to research purposes [72]. Additionally, its interaction capabilities rely mostly on 2D viewers, which can make handling complex 3D data less intuitive for users [6].

ITK-SNAP is another general-purpose tool for image visualisation and segmentation, primarily focused on offering intuitive, user-friendly semi-automatic and manual segmentation tools tailored for non-experts (see Figure 9) [66]. Its main features include tools like the polygon and paintbrush which provides a fast way to draw and refine edits using the mouse for manual segmentation [74], as well as active contour algorithms for semi-automatic segmentation [75]. A drawback of ITK-SNAP is that, as a general-purpose tool, it lacks optimisation for specific segmentation tasks, which leads to less precise results compared to tools that are specifically designed for particular segmentation problems [76].

---

[2] NRRD: Nearly Raw Raster Data, http://teem.sourceforge.net/nrrd

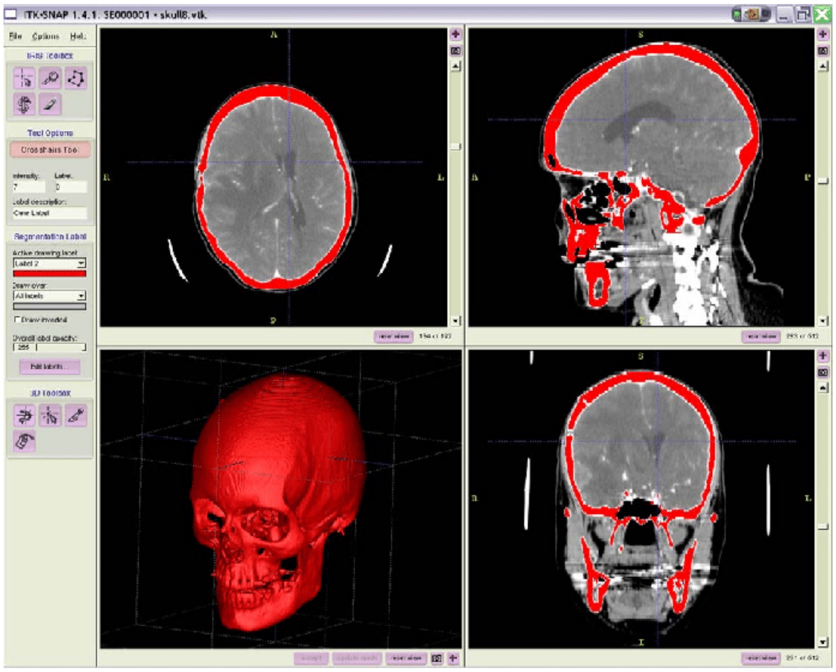


Figure 9: Segmentation of a skull in ITK-SNAP. From [73].

2.3 Medical Applications in Immersive Environments

2.3.1 Immersive Technologies

Immersive technology refers to virtual reality (VR), augmented reality (AR), and mixed reality (MR). These technologies are based on the reality-virtuality continuum (see Figure 10), which defines the degrees of immersion. According to Milgram et al. and Tang et al. [78, 79], there are four levels of immersion, determined by how much real and virtual elements blend through different display technologies.

VR is a technology that immerses users in a simulated digital environment by creating an engaging and interactive experience [80]. It is broadly categorised into two types: non-immersive and immersive. In non-immersive VR, users engage with the virtual world through a screen and standard devices like keyboards and mice [81]. Immersive VR uses head-mounted displays to fully envelop users in a virtual environment, enabling real-time interaction and thereby enhancing the sense of presence [82].

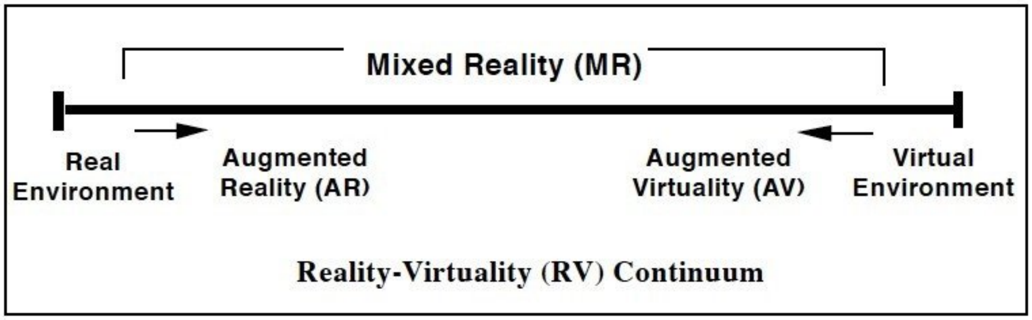


Figure 10: Reality-Virtuality Continuum. From [77]

AR is a technology that overlays virtual elements onto the real world, allowing users to interact with both environments simultaneously [83]. By incorporating digital content like visuals and sounds into the physical space, AR enhances the user's perception without replacing the actual surroundings [84]. This real-time interaction creates a richer, more immersive experience, blending the virtual and real worlds seamlessly to enrich the user's engagement with their environment [85].

There is ongoing debate in academia regarding the definition of **MR** [86]. Milgram et al.'s widely accepted Reality-Virtuality Continuum that we mentioned above describes it as a spectrum from fully real to entirely virtual environments. MR occupies the middle ground, blending real and virtual elements that coexist and interact. AR and Augmented Virtuality (AV) are part of this continuum, with AR closer to the real-world side and AV leaning toward the virtual [87].

In immersive visualisation, interactions often mimic how we handle everyday objects, hence the term "natural interaction techniques." These methods are designed to enhance immersion and ensure a smooth, uninterrupted flow during tasks [88]. According to Büschel et al., common approaches include touch-based interaction, sketching and pen input, tangible interaction, gestures, gaze-based control, and physical navigation [89]. Among these, **touch-based interaction** is intuitive for manipulating 2D and 3D data, such as rotating or zooming anatomical structures [90]. **Gestural interaction** uses hand gestures to control virtual environments, making it ideal for manipulating complex data in hands-free settings like surgery. Its main advantage is enabling remote manipulation without physical contact or handheld devices, which suits environments that require sterility or large public displays [91]. Additionally, **gaze interaction** frees the user's hands, making it useful for those with physical disabilities [92]. It can also indicate attention or enable multitasking by keeping hands available for other tasks [93, 94].

2.3.2 Medical Applications using Immersive Technology

With increased accessibility, immersive technologies like AR and VR are transforming how healthcare professionals interact with complex data [95]. They enhance the visualisation of surgical procedures, improve anatomy training for students, and aid in rehabilitation exercises, providing more effective and intuitive methods for both learning and patient care [96].

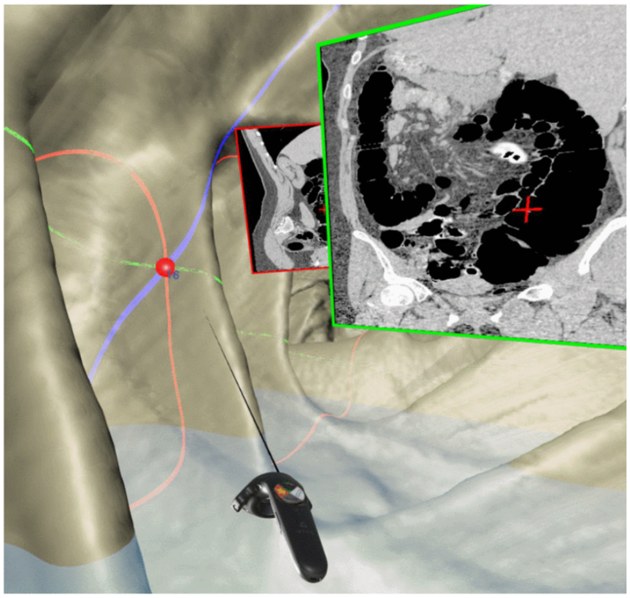


Figure 11: A view of IVC system. From [97].

Immersive technologies, especially virtual reality, are playing an increasingly important role in medical diagnostics. A notable example is immersive virtual colonoscopy (IVC), a noninvasive tool for screening colorectal polyps (Figure 11) [98]. IVC uses VR to create a 3D model of the colon from CT scans, providing an interactive, detailed exploration of inner surfaces. This combination of immersive visualisation and haptic feedback enhances diagnostic precision and aids in more effective preoperative planning [97]. In the field of radiology, VR provides an immersive environment for image interpretation. For example, Maurício et al. developed a VR-based reading room using head-mounted devices to assist radiologists in analysing medical images [99]. This system helps overcome environmental challenges, such as poor lighting or low screen brightness, which can hinder diagnostic accuracy. VR's immersive nature allows radiologists to view and manipulate 3D data, enabling faster and more accurate diagnoses [100].

Immersive technology gives surgeons the feel of performing open surgery [101]. In pediatric surgery, Souzaki et al.'s 2013 study demonstrated the effectiveness of AR navigation systems in identifying hard-to-detect tumours during endoscopic procedures, where limited visibility and lack of haptic feedback are common challenges [102, 103]. These systems, integrating preoperative CT and MRI scans, significantly improve detection accuracy and surgical outcomes [104]. Similarly, Kihara et al. developed a VR system for minimally invasive surgery, using head-mounted displays and 3D endoscopes to provide natural lines of sight, reduce distractions, and minimise surgeon fatigue during lengthy procedures [105].

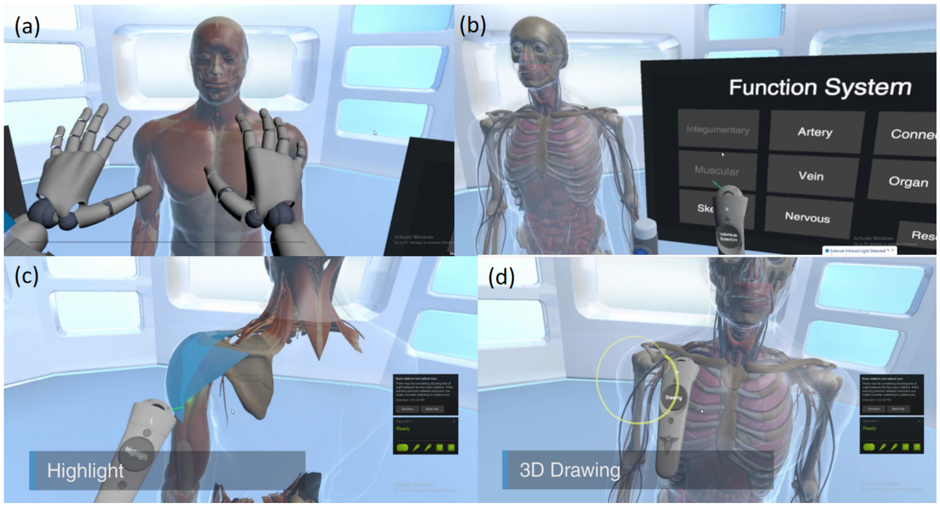


Figure 12: Interaction and Features of the Virtual Interactive Human Anatomy System. (a) The user's hands translated into the VR environment; (b) Interface showing different body systems; (c) Highlighting anatomical structures for collaborative learning; (d) Drawing in 3D to facilitate discussions with other learners. From [106]

VR plays a significant role in anatomy education. In 2017, Weiquan Lu et al. introduced Virtual Interactive Human Anatomy (see Figure 12), a system that allows users to interact with detailed anatomical structures using hand-based gestures through Leap Motion Controllers [106]. This method significantly improves the efficiency of anatomy learning, and follow-up studies have shown that virtual dissection can complement or even replace traditional cadaver-based learning [107, 108].

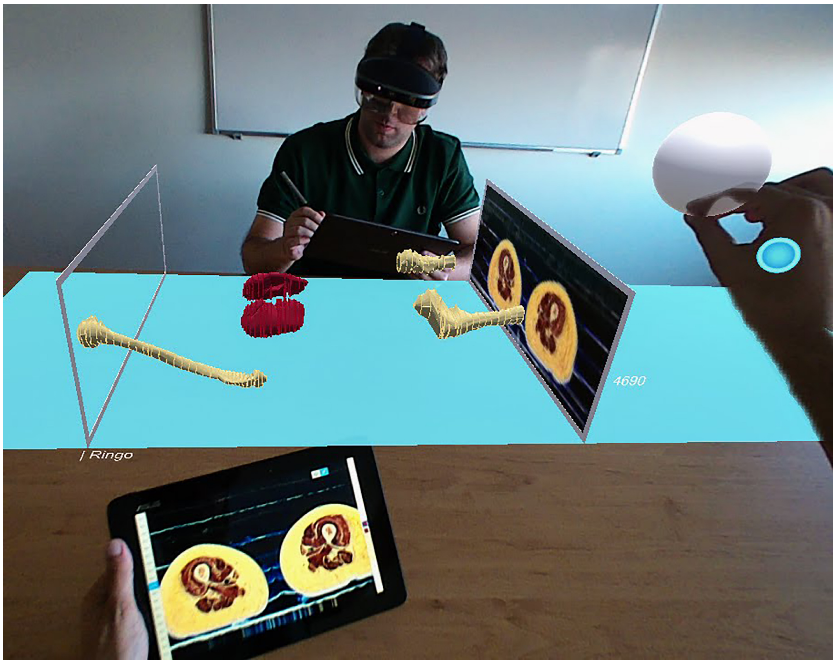


Figure 13: Overview of Anatomy Studio. From [109]

In 2019, Zorzal et al. introduced Anatomy Studio (Figure 13) [109], a collaborative mixed-reality (MR) dissection table designed to allow one or more anatomists to explore full anatomical datasets and perform manual 3D reconstructions. The system mirrors a drawing desk setup, where users sit equipped with head-mounted transparent displays, a tablet for 2D tasks, and a stylus (see Figure 14).

The familiar sketch-based interface of the tablet is used to outline anatomical structures, while simple gestures enable 3D navigation on the table. One of the key advantages of Anatomy Studio is the combination of 3D rendering with VR elements, allowing collaborative work and enhancing visualisation.

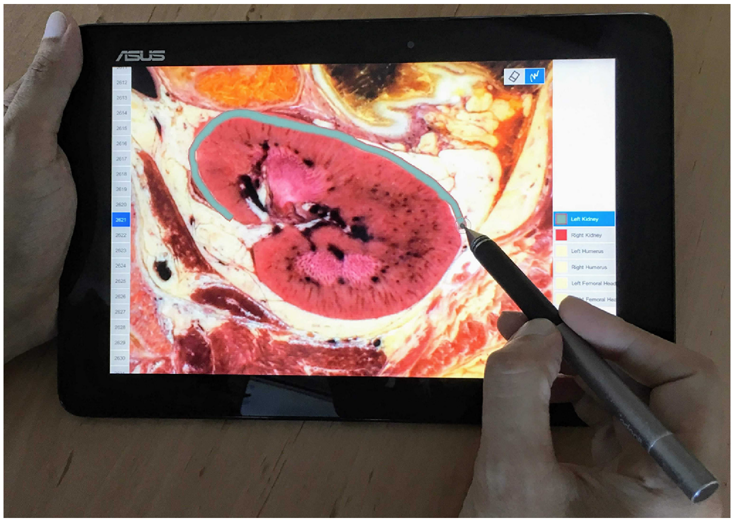


Figure 14: Tracing the contour of a kidney using a stylus on a tablet in anatomy studio. From [109]

Virtual Reality is also widely used in rehabilitation [110]; Peterd's study showed that VR significantly enhanced perception and participation in motor training for children with cerebral palsy [111], and Burke et al.'s study demonstrated the potential of VR to enhance the interactivity and enjoyment of rehabilitation therapy [112].

2.4 Discussion

In this section, we covered various medical imaging modalities, including CT and MRI, discussing their applications, for example, in diagnostics and treatment planning. We also explored the field of medical image segmentation, focusing primarily on manual segmentation techniques and tools such as 3D Slicer. Additionally, we examined the role of immersive technologies like AR, VR, and MR, as well as different interaction methods used within these environments, highlighting their applications in medical practice, such as surgical planning and anatomy education.

However, there are several limitations in each area. For instance, CT scans are associated with radiation risks and struggle to detect certain intraluminal abnormalities, which could be better visualised through more interactive methods like AR/VR. MRI, while offering high-resolution imaging, does not provide real-time feedback, which can limit its usefulness during surgeries. When it comes to manual image segmentation, one of the significant drawbacks is that most interfaces are 2D, which reduces efficiency in 3D segmentation tasks. This slice-by-slice approach is not only time-consuming but also introduces inconsistencies when attempting to reconstruct the full 3D structure of anatomical features.

Lastly, despite the potential of immersive technologies in improving user interaction, they are still underutilised in clinical practice, largely due to interface and interaction limitations, such as the difficulty of performing precise manual operations in fully immersive environments. Further research is needed to address these limitations and improve the efficiency of manual segmentation, as well as the adoption of immersive technologies in medical imaging.

3 Summary of the State of the Art

As previously discussed in the Section 2, 3D imaging modalities such as MRI and CT provide critical data for medical diagnosis and treatment planning. These modalities offer detailed insights into the body's internal structures, yet most 3D visualizations are still 2D projections displayed on conventional screens. This limits the ability to fully comprehend the spatial relationships within the data. By adding the ability to interactively change the point of view, selectively reveal or hide structures, and navigate around the 3D image, these visualizations become an invaluable tool for enhancing spatial understanding and precision in clinical tasks. Segmentation, which is essential for isolating anatomical structures like tumors or organs, relies on these inputs to provide accurate representations for diagnosis or treatment planning.

However, despite the potential of automated segmentation, achieving high accuracy remains a challenge, especially in complex or organ-specific cases [49]. As a result, manual and semi-automatic segmentation methods continue to be widely used. These methods allow experts to precisely define boundaries, but they are often constrained by outdated interaction methods. Tools like Threshold Field Painting [64] and Anatomy Studio [109] have introduced interactive elements to assist segmentation, yet they are primarily designed for 2D interfaces with mouse and keyboard, making slice-by-slice 3D image volume segmentation both labor-intensive and time-consuming.

Immersive environments present a significant opportunity to overcome these limitations. Unlike traditional 2D interfaces, immersive technologies allow users to engage directly with 3D data, providing a more natural and intuitive interaction. Anatomy Studio, for instance, enables 3D navigation using MR techniques, offering a step forward in interaction. However, the reliance on tablet-based input for manual segmentation still restricts its full potential. For example, stylus tools often generate overly thick contours and struggle with adapting to varying drawing orientations, leading to challenges in precision and control.

The key gap lies in the absence of tools specifically designed for fully immersive environments, which, we believe, would allow users to interact directly with 3D data in a spatial context, improving perception, accuracy, and overall user experience.

4 Research Project Plan

4.1 Research Question and Aims

The key focus of this research is to explore novel interaction techniques for medical imaging segmentation using immersive technologies. Current tools for segmentation mostly rely on 2D interfaces, which limit the efficiency and accuracy of the process. Immersive technologies, which offer greater spatial understanding and interaction, may have the potential to significantly improve segmentation tasks in medical imaging.

The primary research question of this project is: How can immersive technologies be used to design and evaluate novel interaction techniques for medical imaging segmentation to enhance diagnostic process and workflow efficiency?

To answer this research question, the following sub-questions will be investigated:

1. What segmentation tasks do medical experts need to perform when working with 3D medical images?
2. What segmentation techniques can be adopted from conventional 2D tools to immersive 3D environments?
3. What novel immersive segmentation techniques can enhance the efficiency of the segmentation tasks while maintaining their accuracy?

Segmentation is the process of delineating specific regions within medical images for defined purposes, such as identifying tumours or isolating organs. See Section 2.2 for more details.

Conventional 2D segmentation techniques, such as Threshold Field Painting, use tools like painting brushes to highlight regions based on threshold values, which might be helpful as well in immersive 3D environments.

4.2 Research Design and Method

4.2.1 Research Design

According to Van Aken and Joan Ernst's Design Science Methodology (DSM) [113], this research focuses on developing practical knowledge that addresses real-world challenges. In this study, the aim is to explore and evaluate novel interaction techniques for medical image segmentation in immersive environments. These techniques are intended to enhance the efficiency of segmentation tasks while maintaining their accuracy. Using a principled DSM approach, this research will develop a framework for immersive segmentation techniques that can be tested and evaluated.

Awareness of Problems: After reviewing the literature, the current state-of-the-art in medical image segmentation and immersive technologies will be identified. Limitations related to existing 2D interfaces and the lack of immersive interaction methods will also be discussed. A comprehensive literature review will be conducted to outline these challenges, and a research proposal will be formulated to address the identified gaps. In this research, the main problem is understanding how immersive technologies can improve segmentation efficiency and accuracy in medical imaging.

Suggestion: To address the problem, we will explore various interaction techniques and frameworks designed for immersive environments. This research will investigate different approaches to improving segmentation efficiency and accuracy, such as adapting existing 2D techniques to 3D immersive environments. The focus will be on evaluating which interaction methods, whether gesture-based, controller-driven, or others, offer the most intuitive and effective experience for medical image segmentation.

Design and Development: We will create prototypes based on the interaction techniques explored in the previous phases. These prototypes will be developed using immersive platforms, such as Unity, to allow users to interact with 3D medical images. Once the initial prototype is tested on a

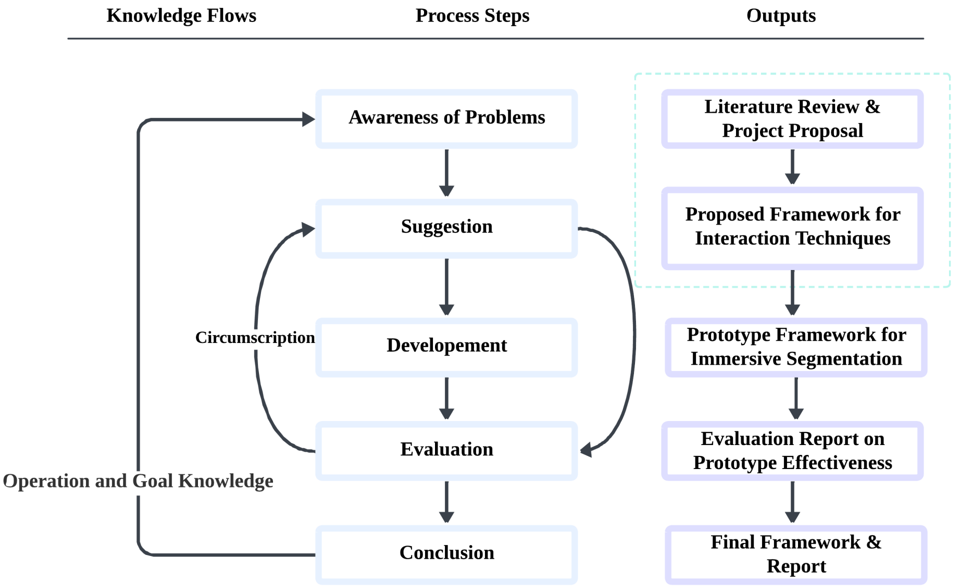


Figure 15: Design Science Methodology Process Flow.

desktop or immersive system, further development will focus on refining the interface and interaction methods for immersive devices, ensuring usability and efficiency in segmentation tasks.
Evaluation: Each prototype and design solution will be evaluated based on its performance in immersive environments. The evaluation will focus on assessing the efficiency and accuracy of segmentation tasks. A user study may be conducted with medical professionals or individuals familiar with medical image segmentation to gather feedback on usability and effectiveness in real-world contexts.
Conclusion: After several iterations of development and evaluation, we will conclude the findings of this research by preparing a final report or thesis. At this stage, a comprehensive exploratory framework for immersive medical image segmentation will be established, offering insights into how these technologies can enhance segmentation efficiency and accuracy.

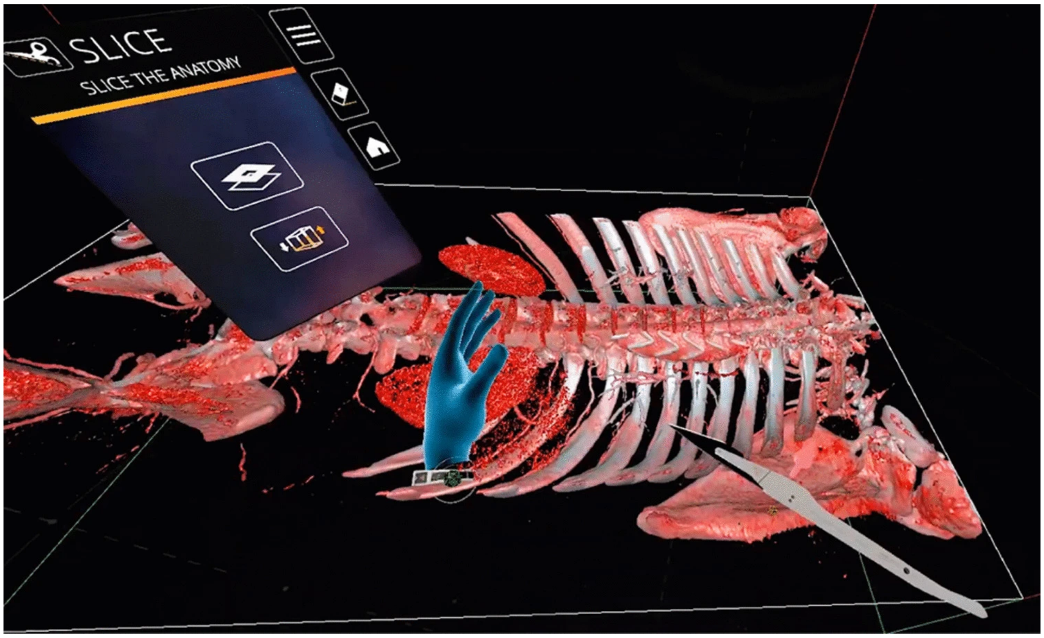


Figure 16: Medical image segmentation example. From [7]

4.2.2 Research Method

Medical image segmentation requires accurate identification of anatomical structures, which is typically done through various interaction techniques. In this research, we will explore the design space for interaction techniques in immersive environments, where users can intuitively engage with 3D medical data. By leveraging immersive technologies, we aim to enhance the efficiency and accuracy of segmentation tasks, allowing medical professionals to interact with 3D models of medical images in a more natural and spatially aware manner. This exploration will focus on developing and testing interaction frameworks that improve user experience and performance in medical segmentation.

4.3 Data Collection

Each prototype will be evaluated through user studies involving individuals with experience in medical image segmentation, such as medical professionals, students, or researchers. Different parameters such as segmentation task complexity, interaction techniques (e.g., gesture-based, controller-based), and the accuracy of the segmentation will be assessed. We will measure the performance of each prototype based on key criteria, such as task completion time, segmentation accuracy, and user satisfaction. The primary focus will be to identify the most effective interaction technique for improving segmentation efficiency while maintaining accuracy. Feedback from participants will also guide further refinement and development of the framework.

4.4 Experimental Goals

As medical segmentation tasks vary in complexity, we focus on improving the interaction methods used to perform these tasks in immersive environments. This experiment aims not only to find the most efficient technique for segmenting medical images but also to propose a promising approach for enhancing segmentation accuracy and efficiency in other medical fields. By leveraging immersive technologies, users can experience a more intuitive and spatially-aware way to interact with 3D medical data, offering potential improvements in both user experience and performance.

4.5 Ethics and Data Privacy

4.5.1 Medical Dataset

The data used in this research comes from the publicly available Medical Segmentation Decathlon dataset [114], which is fully anonymised and compliant with all relevant data protection regulations. Since the dataset is openly accessible and does not contain any identifiable personal information, no ethical approval or informed consent is required to use this dataset for developing our prototypes.

4.5.2 Ethics Application

In our evaluation phase, we plan to conduct a user study to test the usability of the implemented prototypes for segmenting 3D medical images. Because this study involves human participants, we will need ethical approval for our study. We will apply for the Monash Ethics Application via the Monash Ethics Review Manager. We will follow the Monash Ethics guidelines to provide a consent form for our participants to fill out before attending our user study. We will also provide the explanatory statement, which covers the study details, benefits and risks of this study, confidentiality, storage of data, and other information that participants need to know before attempting the user study. We will carefully follow the Ethics guidelines to design our study to minimise any potential risks during the study. Specifically for the experiment data, we will not use any personal information in the academic publications that will report the results of this study; the participants' responses (quantitative performance data and qualitative feedback) will be stored for the duration of our study; and the digital data will be store on a Monash Google Drive folder with access restricted to the investigators.

5 Conclusion

Medical image segmentation plays a vital role in fields such as radiology, surgery, and anatomical research. Radiologists and doctors require precise segmentation tools to accurately delineate regions of interest (ROIs), such as tumours, organs, or other anatomical structures, for diagnostic and treatment purposes. Therefore, the usability and accuracy of segmentation tools are crucial for ensuring efficient clinical workflows and accurate outcomes.

Current segmentation tools, which are predominantly based on 2D planes, have mainly focused on improving specific features, such as Threshold Field Painting, which introduced a brush function to outline ROIs on 2D slices. However, these tools overlook the significance of 3D visualisation and interaction, which are essential for improving both the accuracy of segmentation and the overall experience for medical professionals. For example, while Anatomy Studio integrates mixed reality (MR) for 3D data visualisation during the segmentation process, it remains constrained by the use of 2D contours. This reliance on 2D methods makes the segmentation process less efficient and limits its full potential.

Compared to existing tools, interaction techniques within immersive environments remain underdeveloped. For this research, we will explore various interaction techniques in a fully immersive environment. These techniques will be tested in an immersive setting, allowing users to directly engage with 3D medical images. An example of the proposed scenario prototype for improving segmentation efficiency in an immersive setting is illustrated in Figure 2.

The proposed framework, by enabling more intuitive and spatially aware interactions, will not only enhance the visualisation of 3D medical images but also improve segmentation accuracy, efficiency, and the overall user experience for medical professionals. This will ultimately lead to more effective medical diagnoses and better patient outcomes.

While the exploration of interaction techniques in immersive environments has demonstrated potential, further research and development are necessary. The next phase of this study will focus on conceptualising and testing various interaction techniques within immersive environments to better understand their applicability and limitations. This will allow for the gradual refinement of the proposed framework, ultimately leading to a more effective tool for immersive medical image segmentation, with the goal of enhancing both efficiency and accuracy for professionals in the field.

References

References

[1] William G Bradley. History of medical imaging. *Proceedings of the American Philosophical Society*, 152(3):349–361, 2008.

[2] Michael Correll and Laura Garrison. When the body became data: Historical data cultures and anatomical illustration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2024.

[3] Dzung L Pham, Chenyang Xu, and Jerry L Prince. Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1):315–337, 2000.

[4] Alireza Norouzi, Mohd Shafry Mohd Rahim, Ayman Altameem, Tanzila Saba, Abdolvahab Ehsani Rad, Amjad Rehman, and Mueen Uddin. Medical image segmentation methods, algorithms, and applications. *IETE Technical Review*, 31(3):199–213, 2014.

[5] Tamara Munzner. *Visualization analysis and design*. CRC press, New York, NY, USA, 2014.

[6] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, et al. 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic resonance imaging*, 30(9):1323–1341, 2012.

[7] ImmersiveTouch. Immersivetouch: Virtual reality for medical training and surgery. `https://www.immersivetouch.com/`, 2024. Accessed: 2024-09-14.

[8] James H Scatliff and Peter J Morris. From roentgen to magnetic resonance imaging: the history of medical imaging. *North Carolina medical journal*, 75(2):111–113, 2014.

[9] Kunio Doi. Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology. *Physics in Medicine & Biology*, 51(13):R5, 2006.

[10] Shah Hussain, Iqra Mubeen, Niamat Ullah, Syed Shahab Ud Din Shah, Bakhtawar Abduljalil Khan, Muhammad Zahoor, Riaz Ullah, Farhat Ali Khan, and Mujeeb A Sultan. Modern diagnostic imaging technique applications and risk factors in the medical field: a review. *BioMed research international*, 2022(1):5164970, 2022.

[11] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, 2024.

[12] Gaur Gautam Kar. How 3d ct scans are revolutionizing orthopaedic implantology. In *Handbook of Orthopaedic Trauma Implantology*, pages 1–21. Springer, 2023.

[13] Lee W Goldman. Principles of ct and ct technology. *Journal of nuclear medicine technology*, 35(3):115–128, 2007.

[14] Keith E Kocher, William J Meurer, Reza Fazel, Phillip A Scott, Harlan M Krumholz, and Brahmajee K Nallamothu. National trends in use of computed tomography in the emergency department. *Annals of emergency medicine*, 58(5):452–462, 2011.

[15] E Murali and K Meena. Brain tumor detection from mri using adaptive thresholding and histogram based techniques. *Scalable Computing: Practice and Experience*, 21(1):3–10, 2020.

[16] Ghulam Murtaza, Liyana Shuib, Ainuddin Wahid Abdul Wahab, Ghulam Mujtaba, Ghulam Mujtaba, Henry Friday Nweke, Mohammed Ali Al-garadi, Fariha Zulfiqar, Ghulam Raza, and Nor Aniza Azmi. Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. *Artificial Intelligence Review*, 53:1655–1720, 2020.

[17] Daisuke Kawahara and Yasushi Nagata. T1-weighted and t2-weighted mri image synthesis with convolutional generative adversarial networks. *reports of practical Oncology and radiotherapy*, 26(1):35–42, 2021.

[18] Qingyu Mao, Wenzhe Zhai, Xiang Lei, Zenghui Wang, and Yongsheng Liang. Ct and mri image fusion via coupled feature-learning gan. *Electronics*, 13(17):3491, 2024.

[19] National Research Council (US) and Institute of Medicine (US) Committee on the Mathematics and Physics of Emerging Dynamic Biomedical Imaging. *Mathematics and Physics of Emerging Biomedical Imaging*. National Academies Press, Washington, DC, 1996.

[20] Maggie A Flower. *Webb's physics of medical imaging*. CRC press, 2012.

[21] Muhm JR, LARRY R Brown, JOHN K Crowe, PF Sheedy 2nd, ROBERT R Hattery, and DAVID H Stephens. Comparison of whole lung tomography and computed tomography for detecting pulmonary nodules. *American Journal of Roentgenology*, 131(6):981–984, 1978.

[22] Muhm JR, LARRY R Brown, JOHN K Crowe, PF Sheedy 2nd, ROBERT R Hattery, and DAVID H Stephens. Comparison of whole lung tomography and computed tomography for detecting pulmonary nodules. *American Journal of Roentgenology*, 131(6):981–984, 1978.

[23] R Paxton and J Ambrose. The emi scanner. a brief review of the first 650 patients. *The British journal of radiology*, 47(561):530–565, 1974.

[24] Stephen N Darty, Mark S Thomas, Christina M Neagle, Kerry M Link, Deborah Wesley-Farrington, and W Gregory Hundley. Cardiovascular magnetic resonance imaging: As more patients undergo these examinations, will you know what to tell them about what to expect? *AJN The American Journal of Nursing*, 102(12):34–38, 2002.

[25] GD Honey, W Soni, ET Bullmore, M Varatheesan, SCR Williams, C Andrew, and T Sharma. Evidence of abnormal lateralisation of motor systems in schizophrenia using functional mri. *Schizophrenia Research*, 1(29):70–71, 1998.

[26] Georg Gohla, Benjamin Bender, Marcos Tatagiba, Jürgen Honegger, Ulrike Ernemann, and Constantin Roder. Identification of tumor residuals in pituitary adenoma surgery with intraoperative mri: do we need gadolinium? *Neurosurgical Review*, 43:1623–1629, 2020.

[27] Constantin Tuleasca, Henri-Arthur Leroy, Iulia Peciu-Florianu, Ondine Strachowski, Benoit Derre, Marc Levivier, Michael Schulder, and Nicolas Reyns. Impact of combined use of intraoperative mri and awake microsurgical resection on patients with gliomas: a systematic review and meta-analysis. *Neurosurgical Review*, pages 1–14, 2021.

[28] Marjan Laal. Innovation process in medical imaging. *Procedia-Social and Behavioral Sciences*, 81:60–64, 2013.

[29] Ramya Rathan, Hossam Hamdy, Salah Eldin Kassab, Miral Nagy F Salama, Anusha Sreejith, and Aji Gopakumar. Implications of introducing case based radiological images in anatomy on teaching, learning and assessment of medical students: a mixed-methods study. *BMC Medical Education*, 22(1):723, 2022.

[30] Shi Chen, Jiawei Zhu, Cheng Cheng, Zhouxian Pan, Lingshan Liu, Jianhua Du, Xinhua Shen, Zhen Shen, Huijuan Zhu, Jihai Liu, et al. Can virtual reality improve traditional anatomy education programmes? a mixed-methods study on the use of a 3d skull model. *BMC medical education*, 20:1–10, 2020.

[31] Leslie P Gartner. Anatomical sciences in the allopathic medical school curriculum in the united states between 1967–2001. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, 16(5):434–439, 2003.

[32] Horst-Werner Korf, Helmut Wicht, Robert L Snipes, Jean-Pierre Timmermans, Friedrich Paulsen, Gabriele Rune, and Eveline Baumgart-Vogt. The dissection course–necessary and indispensable for teaching anatomy to medical students. *Annals of Anatomy-Anatomischer Anzeiger*, 190(1):16–22, 2008.

[33] R Amaliya, St Aisyah, AP Hariyanto, F Jannah, A Sarasechan, A Rubiyanto, M Haekal, A Nainggolan, et al. Analysis study of doses distribution in lung cancer using 3d slicer. In *Journal of Physics: Conference Series*, volume 1943, page 012047. IOP Publishing, 2021.

[34] Csaba Pinter, Andras Lasso, An Wang, David Jaffray, and Gabor Fichtinger. Slicerrt: radiation therapy research toolkit for 3d slicer. *Medical physics*, 39(10):6332–6338, 2012.

[35] D Inoue, Byunghyun Cho, M Mori, Y Kikkawa, T Amano, A Nakamizo, K Yoshimoto, Masahiro Mizoguchi, M Tomikawa, J Hong, et al. Preliminary study on the clinical application of augmented reality neuronavigation. *Journal of Neurological Surgery Part A: Central European Neurosurgery*, 74(02):071–076, 2013.

[36] PaxeraHealth. Paxeraview pro. Accessed: 2024-09-14.

[37] Reinaldo Padilha França, Ana Carolina Borges Monteiro, Rangel Arthur, and Yuzo Iano. An overview of the impact of pacs as health informatics and technology e-health in healthcare management. *Cognitive Systems and Signal Processing in Image Processing*, pages 101–128, 2022.

[38] U Bick and H Lenzen. Pacs: the silent revolution. *European radiology*, 9:1152–1160, 1999.

[39] Sindhu John, Angeline CC Poh, Tchoyoson CC Lim, Elizabeth HY Chan, and Le Roy Chong. The ipad tablet computer for mobile on-call radiology diagnosis? auditing discrepancy in ct and mri reporting. *Journal of digital imaging*, 25:628–634, 2012.

[40] Rashid L Bashshur, Elizabeth A Krupinski, James H Thrall, and Noura Bashshur. The empirical foundations of teleradiology and related applications: a review of the evidence. *Telemedicine and e-Health*, 22(11):868–898, 2016.

[41] Krishna Kant Singh and Akansha Singh. A study of image segmentation algorithms for different types of images. *International Journal of Computer Science Issues (IJCSI)*, 7(5):414, 2010.

[42] DP Hudedagaddi and BK Tripathy. Quantum inspired computational intelligent techniques in image segmentation. In *Quantum Inspired Computational Intelligence*, pages 233–258. Elsevier, 2017.

[43] Ravi Anand, Ritesh K Mishra, and Rijwan Khan. Plant diseases detection using artificial intelligence. In *Application of Machine Learning in Agriculture*, pages 173–190. Elsevier, 2022.

[44] Detlev Stalling, Malte Westerhoff, Hans-Christian Hege, et al. Amira: A highly interactive system for visual data analysis. *The visualization handbook*, 38:749–767, 2005.

[45] Manasa Nadipally. Optimization of methods for image-texture segmentation using ant colony optimization. In *Intelligent data analysis for biomedical applications*, pages 21–47. Elsevier, 2019.

[46] Afshin Shoeibi, Marjane Khodatars, Mahboobeh Jafari, Navid Ghassemi, Delaram Sadeghi, Parisa Moridian, Ali Khadem, Roohallah Alizadehsani, Sadiq Hussain, Assef Zare, et al. Automated detection and forecasting of covid-19 using deep learning techniques: A review. *Neurocomputing*, page 127317, 2024.

[47] Stefan Bauer, Roland Wiest, Lutz-P Nolte, and Mauricio Reyes. A survey of mri-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology*, 58(13):R97, 2013.

[48] Brent Foster, Ulas Bagci, Awais Mansoor, Ziyue Xu, and Daniel J Mollura. A review on segmentation of positron emission tomography images. *Computers in biology and medicine*, 50:76–96, 2014.

[49] Bernhard Preim and Charl P Botha. *Visual computing for medicine: theory, algorithms, and applications*. Newnes, 2013.

[50] Anahita Sanandaji, Cindy Grimm, Ruth West, and Max Parola. Where do experts look while doing 3d image segmentation. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, pages 171–174. ACM, 2016.

[51] Dinesh D Patil and Sonal G Deore. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing*, 2(1):22–27, 2013.

[52] S Sudha, KB Jayanthi, C Rajasekaran, and T Sunder. Segmentation of roi in medical images using cnn-a comparative study. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, pages 767–771. IEEE, 2019.

[53] Jie Tian, Di Dong, Zhenyu Liu, and Jingwei Wei. *Radiomics and its clinical application: artificial intelligence and medical big data*. Academic Press, 2021.

[54] Martin Fiebich, Christopher M Straus, Vivek Sehgal, Bernhard C Renger, Kunio Doi, and Kenneth R Hoffmann. Automatic bone segmentation technique for ct angiographic studies. *Journal of computer assisted tomography*, 23(1):155–161, 1999.

[55] Dzung L Pham, Chenyang Xu, and Jerry L Prince. Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1):315–337, 2000.

[56] Hidetomo Suzuki and Jun-ichiro Toriwaki. Automatic segmentation of head mri images by knowledge guided thresholding. *Computerized medical imaging and graphics*, 15(4):233–240, 1991.

[57] Gordon J Harris, Patrick E Barta, Luon W Peng, Seong Lee, Paul D Brettschneider, Amish Shah, Jeffery D Henderer, Thomas E Schlaepfer, and Godfrey D Pearlson. Mr volume segmentation of gray matter and white matter using manual thresholding: dependence on image brightness. *American journal of neuroradiology*, 15(2):225–230, 1994.

[58] Rolf Adams and Leanne Bischof. Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, 16(6):641–647, 1994.

[59] C Vijayakumar and Damayanti Chandrashekhar Gharpure. Development of image-processing software for automatic segmentation of brain tumors in mr images. *Journal of Medical Physics*, 36(3):147–158, 2011.

[60] Xing Chen, Jessy K Possel, Catherine Wacongne, Anne F Van Ham, P Christiaan Klink, and Pieter R Roelfsema. 3d printing and modelling of customized implants and surgical guides for non-human primates. *Journal of neuroscience methods*, 286:38–55, 2017.

[61] Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu, and Yi Pan. A survey of mri-based brain tumor segmentation methods. *Tsinghua science and technology*, 19(6):578–595, 2014.

[62] Koon-Pong Wong. Medical image segmentation: methods and applications in functional imaging. In *Handbook of Biomedical Image Analysis: Volume II: Segmentation Models Part B*, pages 111–182. Springer, 2005.

[63] IEEE. *A gradient magnitude based region growing algorithm for accurate segmentation*, volume 3, 2000.

[64] Takeo Igarashi, Naoyuki Shono, Taichi Kin, and Toki Saito. Interactive volume segmentation with threshold field painting. In *Proceedings of*

the 29th Annual Symposium on User Interface Software and Technology, pages 403–413. ACM, 2016.

[65] Yichi Zhang, Qingcheng Liao, Le Ding, and Jicong Zhang. Bridging 2d and 3d segmentation networks for computation-efficient volumetric medical image segmentation: An empirical study of 2.5 d solutions. *Computerized Medical Imaging and Graphics*, 99:102088, 2022.

[66] Manar Aljabri, Manal AlAmir, Manal AlGhamdi, Mohamed Abdel-Mottaleb, and Fernando Collado-Mesa. Towards a better understanding of annotation tools for medical imaging: a survey. *Scientific Reports*, 81:25877–25911, 2022.

[67] Ron Kikinis and Steve Pieper. 3d slicer as a tool for interactive brain tumor segmentation. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6982–6984. IEEE, 2011.

[68] Xiangrui Li, Paul S Morgan, John Ashburner, Jolinda Smith, and Christopher Rorden. The first step for neuroimaging data analysis: Dicom to nifti conversion. *Journal of neuroscience methods*, 264:47–56, 2016.

[69] Sam Vickery, William D Hopkins, Chet C Sherwood, Steven J Schapiro, Robert D Latzman, Svenja Caspers, Christian Gaser, Simon B Eickhoff, Robert Dahnke, and Felix Hoffstaedter. Chimpanzee brain morphometry utilizing standardized mri preprocessing and macroanatomical annotations. *Elife*, 9:e60136, 2020.

[70] Boo-Kyeong Choi, Nuwan Madusanka, Heung-Kook Choi, Jae-Hong So, Cho-Hee Kim, Hyeon-Gyun Park, Subrata Bhattacharjee, and Deekshitha Prakash. Convolutional neural network-based mr image analysis for alzheimer's disease classification. *Current Medical Imaging*, 16(1):27–35, 2020.

[71] Jun Shen, Thomas Baum, Christian Cordes, Beate Ott, Thomas Skurk, Hendrik Kooijman, Ernst J Rummeny, Hans Hauner, Bjoern H Menze, and Dimitrios C Karampinos. Automatic segmentation of abdominal organs and adipose tissue compartments in water-fat mri: application

to weight-loss in obesity. *European journal of radiology*, 85(9):1613–1621, 2016.

[72] George Z Cheng, Raul San Jose Estepar, Erik Folch, Jorge Onieva, Sidhu Gangadharan, and Adnan Majid. Three-dimensional printing and 3d slicer: powerful tools in understanding and treating structural lung disease. *Chest*, 149(5):1136–1142, 2016.

[73] Silvio H Rizzi, P Pat Banerjee, and Cristian J Luciano. Automating the extraction of 3d models from medical images for virtual reality and haptic simulations. In *2007 IEEE International Conference on Automation Science and Engineering*, pages 152–157. IEEE, 2007.

[74] Paul A Yushkevich, Yang Gao, and Guido Gerig. Itk-snap: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 3342–3345. IEEE, 2016.

[75] Amanda Farias Gomes, Danieli Moura Brasil, Amaro Ilídio Vespasiano Silva, Deborah Queiroz Freitas, Francisco Haiter-Neto, and Francisco Carlos Groppo. Accuracy of itk-snap software for 3d analysis of a non-regular topography structure. *Oral radiology*, 36:183–189, 2020.

[76] Paul A Yushkevich, Artem Pashchinskiy, Ipek Oguz, Suyash Mohan, J Eric Schmitt, Joel M Stein, Dženan Zukić, Jared Vicory, Matthew McCormick, Natalie Yushkevich, et al. User-guided segmentation of multi-modality medical imaging datasets with itk-snap. *Neuroinformatics*, 17:83–102, 2019.

[77] Thammathip Piumsomboon, Arindam Day, Barrett Ens, Youngho Lee, Gun Lee, and Mark Billinghurst. Exploring enhancements for remote mixed reality collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*, pages 1–5. ACM, 2017.

[78] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[79] Yuk Ming Tang, Ka Yin Chau, Alex Pak Ki Kwok, Tongcun Zhu, and Xiangdong Ma. A systematic review of immersive technology applications for medical practice and education-trends, application areas, recipients, teaching contents, evaluation methods, and performance. *Educational Research Review*, 35:100429, 2022.

[80] Daniel A Guttentag. Virtual reality: Applications and implications for tourism. *Tourism management*, 31(5):637–651, 2010.

[81] Rafał Wojciechowski and Wojciech Cellary. Evaluation of learners' attitude toward learning in aries augmented reality environments. *Computers & education*, 68:570–585, 2013.

[82] Julia Beck, Mattia Rainoldi, and Roman Egger. Virtual reality in tourism: a state-of-the-art review. *Tourism Review*, 74(3):586–612, 2019.

[83] Pietro Cipresso, Irene Alice Chicchi Giglioli, Mariano Alcañiz Raya, and Giuseppe Riva. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. *Frontiers in psychology*, 9:2086, 2018.

[84] Ayoung Suh and Jane Prophet. The state of immersive technology research: A literature analysis. *Computers in Human behavior*, 86:77–90, 2018.

[85] Lauryn R Rochlen, Robert Levine, and Alan R Tait. First-person point-of-view–augmented reality for central line insertion training: a usability and feasibility study. *Simulation in Healthcare*, 12(1):57–62, 2017.

[86] Maximilian Speicher, Brian D Hall, and Michael Nebeling. What is mixed reality? In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–15, 2019.

[87] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. Spie, 1995.

[88] Bongshin Lee, Petra Isenberg, Nathalie Henry Riche, and Sheelagh Carpendale. Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2689–2698, 2012.

[89] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H Thomas. *Immersive analytics*, volume 11190. Springer, 2018.

[90] Martin Hachet, Jean-Baptiste de La Rivière, Jérémy Laviole, Aurélie Cohé, and Sébastien Cursan. Touch-based interfaces for interacting with 3d content in public exhibitions. *IEEE computer graphics and applications*, 33(2):80–85, 2013.

[91] Rose Johnson, Kenton O'Hara, Abigail Sellen, Claire Cousins, and Antonio Criminisi. Exploring the potential for touchless interaction in image-guided interventional radiology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3323–3332, 2011.

[92] Hiroshi Ishii, Carlo Ratti, Ben Piper, Yao Wang, Assaf Biderman, and Eran Ben-Joseph. Bringing clay and sand into digital design—continuous tangible user interfaces. *BT technology journal*, 22(4):287–299, 2004.

[93] I Scott MacKenzie. Evaluating eye tracking systems for computer input. In *Gaze interaction and applications of eye tracking: Advances in assistive technologies*, pages 205–225. IGI Global, 2012.

[94] Richard A Bolt. Gaze-orchestrated dynamic windows. *ACM SIGGRAPH Computer Graphics*, 15(3):109–119, 1981.

[95] Seth Johnson, Bret Jackson, Bethany Tourek, Marcos Molina, Arthur G Erdman, and Daniel F Keefe. Immersive analytics for medicine: hybrid 2d/3d sketch-based interfaces for annotating medical data and designing medical devices. In *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, pages 107–113. ACM, 2016.

[96] Richard M Satava. Virtual reality surgical simulator: the first steps. *Surgical endoscopy*, 7:203–205, 1993.

[97] Seyedkoosha Mirhosseini, Ievgeniia Gutenko, Sushant Ojal, Joseph Marino, and Arie Kaufman. Immersive virtual colonoscopy. *IEEE transactions on visualization and computer graphics*, 25(5):2011–2021, 2019.

[98] Lichan Hong, Shigeru Muraki, Arie Kaufman, Dirk Bartz, and Taosong He. Virtual voyage: Interactive navigation in the human colon. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 27–34. ACM, 1997.

[99] Daniel Simões Lopes, Daniel Medeiros, Soraia Figueiredo Paulo, Pedro Brasil Borges, Vitor Nunes, Vasco Mascarenhas, Marcos Veiga, and Joaquim Armando Jorge. Interaction techniques for immersive ct colonography: A professional assessment. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11*, pages 629–637. Springer, 2018.

[100] Maurício Sousa, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge, and Daniel Simões Lopes. Vrrrroom: Virtual reality for radiologists in the reading room. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 4057–4062. ACM, 2017.

[101] Ezequiel R Zorzal, Maurício Sousa, Pedro Belchior, João Madeiras Pereira, Nuno Figueiredo, and Joaquim A Jorge. Design requirements to improve laparoscopy via xr. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 425–429. IEEE, 2022.

[102] Satoshi Ieiri, Munenori Uemura, Kouzou Konishi, Ryota Souzaki, Yoshihiro Nagao, Norifumi Tsutsumi, Tomohiko Akahoshi, Kenoki Ohuchida, Takeshi Ohdaira, Morimasa Tomikawa, et al. Augmented reality navigation system for laparoscopic splenectomy in children based on preoperative ct image using optical tracking device. *Pediatric surgery international*, 28:341–346, 2012.

[103] T Iwanaka, M Arai, M Ito, Ho Kawashima, K Matoba, and S Imaizumi. Challenges of laparoscopic resection of abdominal neuroblastoma with lymphadenectomy: a preliminary report. *Surgical endoscopy*, 15:489–492, 2001.

[104] Ryota Souzaki, Satoshi Ieiri, Munenori Uemura, Kenoki Ohuchida, Morimasa Tomikawa, Yoshiaki Kinoshita, Yuhki Koga, Aiko Suminoe, Kenichi Kohashi, Yoshinao Oda, et al. An augmented reality navigation system for pediatric oncologic surgery based on preoperative ct and mri images. *Journal of pediatric surgery*, 48(12):2479–2483, 2013.

[105] Kazunori Kihara, Yasuhisa Fujii, Hitoshi Masuda, Kazutaka Saito, Fumitaka Koga, Yoh Matsuoka, Noboru Numao, and Kazuyuki Kojima. New three-dimensional head-mounted display system, tmdu-s-3d system, for minimally invasive surgery application: Procedures for gasless single-port radical nephrectomy. *International Journal of Urology*, 19(9):886–889, 2012.

[106] Weiquan Lu, Suresh Pillai, Kanagasuntheram Rajendran, Yoshifumi Kitamura, Ching-Chiuan Yen, and Ellen Yi-Luen Do. Virtual interactive human anatomy: dissecting the domain, navigating the politics, creating the impossible. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 429–432. ACM, 2017.

[107] Young Hyun Yun, Hyeok Yi Kwon, Su Kyoung Jeon, Yu Mi Jon, Min Jung Park, Dong Hoon Shin, and Hyung Jin Choi. Effectiveness and satisfaction with virtual and donor dissections: A randomized controlled trial. *Scientific Reports*, 14(1):16388, 2024.

[108] Haya Almaree, Roland Fischer, René Weller, Verena Uslar, Dirk Weyhe, and Gabriel Zachmann. Enhancing anatomy learning through collaborative vr? an advanced investigation. *Computers & Graphics*, 123:104019, 2024.

[109] Ezequiel R Zorzal, Maurício Sousa, Daniel Mendes, Rafael Kuffner dos Anjos, Daniel Medeiros, Soraia Figueiredo Paulo, Pedro Rodrigues, José João Mendes, Vincent Delmas, Jean-Francois Uhl, et al. Anatomy studio: a tool for virtual dissection through augmented 3d reconstruction. *Computers & Graphics*, 85:74–84, 2019.

[110] Mahasak Ketcham, Thittaporn Ganokratanaa, and Patiyuth Pramkeaw. Learning game with stroke paralysis rehabilitation for virtual reality. *International Journal of Emerging Technologies in Learning*, 17(21), 2022.

[111] Denise T Reid. Benefits of a virtual play rehabilitation environment for children with cerebral palsy on perceptions of self-efficacy: a pilot study. *Pediatric rehabilitation*, 5(3):141–148, 2002.

[112] James William Burke, MDJ McNeill, Darryl K Charles, Philip J Morrow, Jacqui H Crosbie, and Suzanne M McDonough. Optimising engagement for stroke rehabilitation using serious games. *The Visual Computer*, 25:1085–1099, 2009.

[113] Joan Ernst Van Aken. Management research as a design science: Articulating the research products of mode 2 knowledge production in management. *British journal of management*, 16(1):19–36, 2005.

[114] Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M Summers, et al. The medical segmentation decathlon. *Nature communications*, 13(1):4128, 2022.

A Design Exploration of Manual Segmentation Techniques for Immersive 3D Visualisation of Medical Images

SHUXIAN QI, Monash University, Australia
VAHID POORYOUSEF, Monash University, Australia
YIDAN ZHANG, Monash University, Australia
TIM DWYER, Monash University, Australia
JIAZHOU LIU, Monash University, Australia

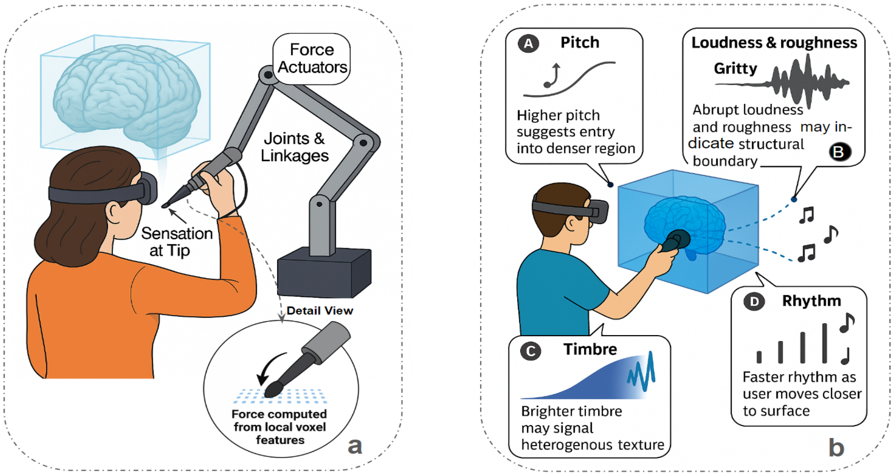


Fig. 1. Overview of the two designed systems for immersive 3D medical image segmentation. (a) Stylus-based haptic system: the user explores a brain MRI volume with a force-feedback stylus that conveys resistance based on local voxel features. (b) Auditory system: the user navigates using a VR controller and hears real-time sonification, where pitch, loudness, timbre, and rhythm map to intensity, gradient, texture, and depth to support spatial perception and segmentation.

Manual segmentation of 3D medical images remains a cognitively demanding process, particularly in immersive environments where visual-only workflows often fail to support accurate spatial perception. Although emerging tools offer 3D volumetric rendering and novel interaction methods, current systems still lack structured guidance for designing effective segmentation interfaces beyond 2D slice-based paradigms. To address this, we introduce MedSegWeave, a design space that maps key patterns and variations in manual segmentation across different systems and highlights design gaps based on a review of existing tools. We observe that most immersive systems still rely only on visual feedback. To address this, we propose two conceptual designs that explore haptic and auditory feedback to support more effective segmentation, offering new directions for building more intuitive and multimodal segmentation experiences.

CCS Concepts: • **Human-centered computing** → **Virtual reality**.

Additional Key Words and Phrases: Image Segmentation, 3D Medical Imaging, Virtual Reality, Interaction Design

---

Authors' Contact Information: Shuxian Qi, sqii0014@student.monash.edu, Monash University, Melbourne, Australia; Vahid Pooryousef, vahid.pooryousef@monash.edu, Monash University, Melbourne, Australia; Yidan Zhang, Monash University, Melbourne, Australia, Yidan.Zhang@monash.edu; Tim Dwyer, Monash University, Melbourne, Australia, tim.dwyer@monash.edu; Jiazhou Liu, Monash University, Melbourne, Australia, joe.liu@monash.edu.

1 Introduction

3D medical image segmentation plays a critical role in clinical tasks such as tumour localisation, anatomical quantification, and surgical planning [46, 67]. It allows clinicians and researchers to extract structural information from volumetric data acquired by computed tomography (CT), Magnetic Resonance Imaging (MRI), or Positron Emission Tomography (PET) scans. A range of segmentation techniques have been developed, including manual, semi-automatic and fully automated approaches [3, 57]. Although deep learning-based methods have gained attention for their scalability and efficiency [57], they remain challenged by image artifacts, low contrast, and irregular anatomical shapes - particularly in complex cases such as tumour segmentation [49]. Moreover, reliable model training requires large datasets of expertly annotated 3D images, which require significant manual effort [38]. As a result, manual segmentation remains essential in both clinical practice and the development of robust AI models. Manual segmentation is commonly performed using 2D slice-based tools like 3D Slicer[1], where users delineate anatomical boundaries plane by plane, typically axial, sagittal, or coronal, using a mouse or stylus [22]. Despite its widespread use, this process places a high cognitive burden on users, who must mentally reconstruct 3D structures from 2D slices [69]. The workflow can be time-consuming and susceptible to errors and inefficiencies [17].

With the advancement of immersive technologies such as virtual reality (VR) and augmented reality (AR) [39], medical experts are now able to visualise volumetric data directly in 3D space. This enhanced immersion and realism facilitate a better understanding of complex spatial relationships, which is particularly valuable in medical applications [24]. Building on these capabilities, recent systems such as VRContour [12] and Elucis [4] have introduced diverse interaction techniques, demonstrating that immersive 3D visualisations can improve segmentation accuracy and efficiency while alleviating cognitive load, frustration, and overall effort [4, 12]. Despite promising results, most immersive tools still rely solely on visual feedback, which can become overloaded when interpreting densely layered anatomy or resolving artifacts such as noise, occlusion, and intensity ambiguities [50]. The interaction is further challenged by a limited field of view, depth perception issues, and ambiguous spatial cues, which hinder precise boundary localisation and increase disorientation [34].

To address these challenges, researchers have explored incorporating multimodal feedback, such as haptic and auditory, into immersive segmentation systems to support navigation and enhance feature recognition [31, 56], as exemplified by Pooryousef et al. [48], who investigated the use of haptic gloves to assist with shape selection during immersive manual segmentation. However, a structured framework for understanding and designing such multimodal segmentation systems remains largely absent.

This paper presents MedSegWeave, a design space for manual 3D medical image segmentation, which organises existing systems across key dimensions of data representation and interaction techniques. This design space is derived from a structured review of both traditional desktop-based and immersive segmentation tools. Based on this framework, we identify underexplored opportunities for sensory feedback and propose two conceptual designs that demonstrate the integration of haptic and auditory feedback into immersive segmentation workflows.

Our contributions include: (1) a design space that organises manual medical image segmentation systems across key dimensions of data representation and interaction techniques; (2) an analysis of current tools identifying gaps, particularly the limited use of sensory feedback; and (3) two conceptual designs demonstrating how haptic and auditory feedback can support segmentation tasks in immersive environments. With this work, we aim to support future research on multimodal segmentation systems and provide design guidance for immersive medical applications.

---

[1] 3D Slicer: https://www.slicer.org/

2 Background

2.1 Segmentation of 3D Medical Images

The development of 3D medical imaging modalities such as MRI and CT has greatly improved clinical diagnostics and treatment planning by providing volumetric views of anatomical structures [16]. MRI offers strong soft tissue contrast, while CT is effective for visualising bone and dense regions [26, 35]. These modalities form the basis for segmentation tasks that isolate regions of interest (ROIs), including tumours, organs, and lesions. Medical image segmentation is the process of dividing an image into distinct regions to highlight relevant anatomical or pathological structures. This process plays a central role in various clinical tasks, including diagnosis, surgical planning, and quantitative extraction of characteristics for radiomics[46]. Depending on the level of human input, segmentation methods are classified as manual, semi-automatic, or automatic [3, 57].

While machine learning enables efficient segmentation with minimal input, it still struggles in complex cases—such as brain tumours—where human experts consistently outperform AI due to better handling of artifacts, anatomical complexity, and image quality variations [27, 49]. As a result, manual segmentation remains essential in domains like paediatric oncology [61], and continues to serve as both a clinical standard and a benchmark for evaluating new algorithms [1, 38].

Traditional manual segmentation is commonly performed on desktop platforms using picture archiving and communication system (PACS) or open-source tools such as 3D Slicer. Users navigate 2D slices and trace anatomical boundaries to isolate regions of interest (ROIs) with a mouse or stylus in a slice-by-slice workflow. During segmentation, experts focus primarily on ROIs and their boundaries, often shifting attention between the volume and interface, suggesting the need to reduce interface distractions and provide focused, context-aware feedback near the ROI to support accurate and efficient segmentation.

Segmentation Techniques such as thresholding, region growing, and edge detection are widely used but require experts to mentally reconstruct 3D structures, resulting in high cognitive load and increased error risk. To address these issues, Igarashi et al. [32] introduced Threshold Field Painting, which enables interactive thresholding in 3D volumes via a brush-based interface with colour-coded isosurfaces. Although this improves feedback, the interaction remains constrained to 2D screens, where environmental factors such as glare or low brightness may hinder visibility and precision.

2.2 Manual Segmentation in Immersive Environments

Immersive technologies such as Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) enable users to interact with 3D medical data via head-mounted displays, motion-tracked controllers, and spatial sensors [9, 23]. These systems support a range of spatial operations—such as navigating through volumes, manipulating viewpoints, selecting regions of interest, and outlining anatomical boundaries—which are particularly suited to segmentation tasks requiring depth perception and fine motor control[33].

In this context, interaction is typically mediated through controllers, styluses, or hand gestures, and can be broadly categorised into navigation, selection, annotation, and manipulation. The quality, granularity, and responsiveness of these techniques directly affect segmentation precision, cognitive load, and workflow efficiency. For instance, Anatomy Studio [70] combines 2D stylus tracing with real-time 3D exploration to support more collaborative and accurate reconstruction. More advanced platforms such as Elucis [4] offer full 3D segmentation using tracked controllers, achieving high accuracy and significantly reduced contouring time. Although immersive environments provide a more embodied and intuitive workspace compared to 2D desktop platforms, they also introduce new challenges. Users may experience mid-air fatigue from prolonged gestures, spatial disorientation due to limited field of view, and difficulty perceiving depth cues, especially when handling occluded or fine-grained structures [50].

Most existing immersive systems, such as Elucis [4], operate primarily within visual-interactive paradigms. While these systems support intuitive spatial interactions, they often suffer from visual limitations—especially when rendering multiple volumetric datasets—resulting in cluttered scenes and reduced perceptual clarity. Even when anatomical structures are visible, their depth and organisation may remain ambiguous due to overlapping content and limited lighting models [34]. Cognitive psychology research further suggests that such complexity increases search latency and attentional demands [43, 63], posing challenges for tasks like manual segmentation that rely on precise boundary detection in occluded or deeply nested regions.

These findings highlight the limitations of vision-only systems and point to the potential of multimodal interaction—integrating haptic, auditory, or other sensory feedback—to enhance spatial perception, reduce cognitive load, and support more accurate and ergonomic workflows [50]. However, despite the growing diversity of immersive segmentation tools, there remains no systematic framework to analyse their interaction strategies or compare how different visualisation and input modality combinations impact user experience and task effectiveness. This motivates the development of a structured design space to classify existing approaches and guide future innovations in immersive manual segmentation.

2.3 Multimodal Feedback in Immersive Segmentation

Multimodal feedback integrates multiple sensory modalities—typically visual, auditory, and haptic—to support interactive systems. In immersive environments, overreliance on visual input can increase cognitive strain and reduce efficiency [40]. By distributing perceptual demands, multimodal systems improve hand-eye coordination and interaction fluency [5], while enhancing signal clarity and robustness for more accurate task execution [10, 14, 66]. These benefits are critical in medical image segmentation, where haptic and auditory feedback can complement visual cues to reduce perceptual load and broaden interactive expression.

Haptic feedback enables interaction through touch by simulating tactile or kinaesthetic sensations via forces, vibrations, or motion [37, 51]. A key subtype is force feedback, delivered through devices like gloves or controllers, often using impedance-based control for simplicity and cost-efficiency. Studies show that haptics improve proprioception and visuomotor coordination in immersive tasks, particularly when visual cues are limited [41, 48]. In medical field, haptics have supported virtual dissection and anatomical exploration [31]. However, most systems use static data; few explore real-time force feedback for guiding segmentation. Our system addresses this gap by using force cues to support tactile boundary tracing in immersive environments.

Auditory feedback offers temporal precision and perceptual clarity that complement visual and haptic inputs. It supports attentional focus, memory, and spatial orientation [2, 8, 28, 62], and enhances motor coordination during spatially demanding tasks [53]. A key technique is sonification—the transformation of data into sound—commonly implemented via parameter mapping sonification (PMSon), which links data features to sound parameters [56, 64]. PMSon has been applied in PET/CT segmentation and anatomical prototyping [25, 56], improving spatial understanding, especially for non-experts. However, most use pre-segmented data; our system addresses this by exploring real-time auditory mapping to guide manual segmentation in immersive environments.

3 MedSegWeave: A Design Space for Manual Medical Image Visualisation and Segmentation

MedSegWeave presents a design space for visualisation and segmentation of manual medical images. Identifies key design aspects across both traditional desktop-based tools and emerging immersive systems, focusing on how medical images are visualised and segmented in 2D and 3D environments. MedSegWeave distills common practices and novel approaches into a unified framework that captures how segmentation tasks are carried out through combinations of visual representation, interaction techniques, input modalities, and sensory feedback. This section formalises the design space by surveying representative systems and organising their features into orthogonal dimensions. The resulting framework enables researchers, designers, and system developers to analyse current segmentation solutions, compare them across environments, and uncover underexplored opportunities, particularly in immersive settings. By examining patterns and gaps in this structured space, MedSegWeave supports the design of more effective and multimodal medical image segmentation tools.

3.1 Approach

We construct our design space using a structured methodology grounded in Zwicky's General Morphological Analysis [52], a method to systematically organise complex problem domains through orthogonal dimensions derived from taxonomical analysis. This approach has been applied to the design of 2D information spaces in immersive environments by Ens [20], and to the visualisation view management by Liu [36]. In our context, it provides a generative framework for identifying and organizing key dimensions of manual segmentation systems across both 2D and immersive environments, and points to combinations that remain theoretically possible yet underutilised in practice.

Our process follows three methodical steps:

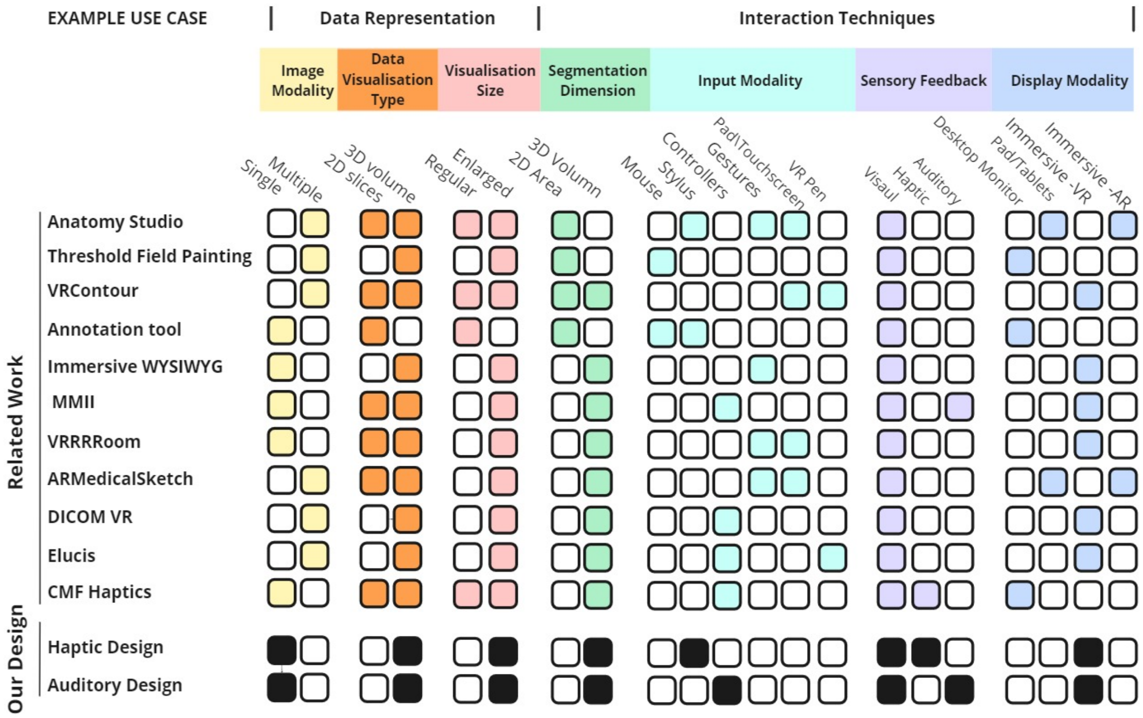
- Review of existing segmentation systems to identify recurring visualisation and interaction features;
- Categorization of these features into orthogonal dimensions to reveal patterns and gaps;
- Generation of new design opportunities through systematic recombination of dimension values.

3.2 Paper Selection

To construct our design space, we conducted a systematic review of representative systems for manual medical image segmentation. Our selection focused on systems that feature explicit interaction designs (e.g., segmentation, navigation, annotation) rather than purely algorithmic or backend processing. We included both desktop-based 2D tools and immersive 3D applications, but excluded systems that offered no direct user-facing segmentation interaction. We searched recent proceedings from CHI, UIST, ISMAR, MICCAI, and other VR-related venues, and supplemented this with citation tracking from key review papers. From this process, we identified and selected eleven representative systems that provided diverse perspectives on immersive medical image segmentation. The eleven systems were retained for analysis and are listed in Appendix. Following a bottom-up coding process, we extracted an initial set of candidate dimensions and iteratively refined them through merging, elimination, and abstraction. Dimensions that were overly narrow, nested within other concepts, or were not broadly applicable were excluded. This process resulted in a final set of seven core dimensions, grouped into two major categories: *Data Representation* and *Interaction Techniques*.

3.3 Data Representation

The data representation category refers to the way medical imaging data is structured and presented visually for segmentation tasks. It captures the fundamental characteristics of the input data and its display, which in turn influence



	EXAMPLE USE CASE	Data Representation								Interaction Techniques												
		Image Modality		Data Visualisation Type		Visualisation Size		Segmentation Dimension		Input Modality						Sensory Feedback			Display Modality			
		Single	Multiple	2D slices	3D volume	Regular	Enlarged	2D Area	3D Volumn	Mouse	Stylus	Controllers	Gestures	Pad\Touchscreen	VR Pen	Visaul	Haptic	Auditory	Desktop Monitor	Pad/Tablets	Immersive -VR	Immersive -AR
Related Work	Anatomy Studio	□	■	■	■	■	■	■	□	□	■	□	■	■	□	■	□	□	□	■	□	■
	Threshold Field Painting	□	■	□	■	□	■	■	□	■	□	□	□	□	□	■	□	□	■	□	□	□
	VRContour	□	■	■	■	■	■	■	■	□	□	□	□	■	■	■	□	□	□	□	■	□
	Annotation tool	■	□	■	□	■	□	■	□	■	■	□	□	□	□	■	□	□	■	□	□	□
	Immersive WYSIWYG	■	□	□	■	□	■	□	■	□	□	□	■	□	□	■	□	□	□	□	■	□
	MMII	■	□	■	■	□	■	□	■	□	□	■	□	□	□	■	□	■	□	□	■	□
	VRRRRoom	■	□	■	■	□	■	□	■	□	□	□	■	■	□	■	□	□	□	□	■	□
	ARMedicalSketch	□	■	■	■	□	■	□	■	□	□	□	■	■	□	■	□	□	□	■	□	■
	DICOM VR	□	■	□	■	□	■	□	■	□	□	■	□	□	□	■	□	□	□	□	■	□
	Elucis	□	■	□	■	□	■	□	■	□	□	■	□	□	■	■	□	□	□	□	■	□
	CMF Haptics	■	□	■	■	■	■	□	■	□	□	■	□	□	□	■	■	□	■	□	□	□
Our Design	Haptic Design	■	□	□	■	□	■	□	■	□	■	□	□	□	□	■	■	□	□	□	■	□
	Auditory Design	■	□	□	■	□	■	□	■	□	□	■	□	□	□	■	□	■	□	□	■	□

Fig. 2. Use cases from literature (top section) and our proposed designs (bottom two rows) mapped onto the proposed design space. A filled cell indicates that the corresponding system supports the given design dimension attribute. The dimensions are grouped into two categories: *Data Representation* and *Interaction Techniques*. The proposed Haptic and Auditory Designs—illustrated in Figure 1—represent our extensions to underexplored areas of sensory feedback.

the types of interaction and interpretation strategies required from users. This category includes three design dimensions: **Image Modality**, **Data Visualisation Type**, and **Visualisation Size**.

Image Modality – This refers to whether a system supports segmentation for a *single modality* or *multiple modalities*. Different image types, such as MRI, CT, or PET, vary in anatomical detail and clinical focus, requiring different segmentation strategies. *Single modality* systems, such as Anatomy Studio, support only one type of image and are often optimised for its specific characteristics [21, 45, 56, 59, 65]. In contrast, *multiple modality* systems, like Elucis, handle two or more types and enable comparative or flexible workflows [4, 12, 32, 60, 68, 70], although this can introduce added complexity in both visualisation and interaction.

Data Visualisation Type – This refers to the way medical image data is rendered for segmentation tasks, either as *2D slices* or a *3D volume*. *2D slices* present anatomical structures in cross-sectional views aligned along anatomical planes, supporting detailed slice-by-slice annotation [21]. *3D volumes*, in contrast, provide a continuous spatial representation, allowing for holistic exploration and spatial reasoning [4, 32, 60, 65]. Several systems support both views simultaneously, allowing users to cross-reference planar and volumetric information [12, 45, 56, 59, 68, 70].

Visualisation Size – This describes whether the image data is presented at a *regular* or *enlarged* scale within the system interface. *Regular* size displays preserve anatomical proportions and are commonly used for overview or multi-slice layouts [21]. *Enlarged* views magnify specific regions to support fine-grained inspection or precise boundary tracing [4, 32, 56, 59, 60, 65]. Some systems allow dynamic scaling between the two, balancing global context with local detail [12, 45, 70].

3.4 Interaction Techniques

The interaction techniques category concerns how users interact with and manipulate medical images during segmentation. This includes the dimensional context of the segmentation environment, the physical tools used for input, the types of sensory feedback provided, and the visual display through which the interaction occurs. This category covers four design dimensions: **Segmentation Dimension**, **Input Modality**, **Sensory Feedback**, and **Display Modality**.

Segmentation Dimension – This denotes whether segmentation tasks are carried out within a *2D area* or across a *3D volume*. *2D area* segmentation is typically conducted on flat displays such as desktop or tablet displays, using slice-by-slice annotation techniques [21, 32, 70]. *3D volume* segmentation occurs in immersive environments, where users interact directly with volumetric structures in spatial context [4, 45, 56, 59, 60, 65, 68]. Some systems, such as *VRContour* [12], support both segmentation dimensions, enabling flexible workflows.Studies showed that combining 2D and 3D segmentation inputs offered promising flexibility for immersive systems, although reducing mode-switching overhead remained a key design challenge [18].

Input Modality – This dimension concerns the physical channel through which users interact with the segmentation interface. *Mouse* and *Stylus* are widely used on the desktop, offering pixel-level control and familiarity for the annotation of 2D slices [21, 32]. *Controllers* are common in immersive environments and enable fast and gross movements in 3D space, while the *VR Pen* input offers greater precision through a pen-like grip, supporting fine-grained interaction with volumetric data [13]. The system proposed by Wang [65] employed *Gestures*, which allow hands-free interaction but generally suffers from lower accuracy due to limited tracking fidelity [13]. *Pad/Touchscreen* input enables direct 2D contouring and is often used in combination with other input modalities such as gestures or VR pens [12, 59, 68, 70].

Designing input modalities involves balancing precision and speed. Studies show that *VR Pen* outperforms both *Controllers* and *Gestures* in precision tasks, making it preferable for detailed boundary tracking [47]. However, pens may induce fatigue and reduce speed during broader movements [13]. *Controllers*, on the other hand, support faster interaction and are effective for coarse segmentation or spatial navigation. For instance, *Gestures* may suit tasks such as region selection or view adjustment, while a *Stylus* or *VR Pen* is better suited for illustrating complex anatomical boundaries. Task-adaptive strategies are therefore recommended, assigning high-precision tasks to pen or styluses, and assigning coarse or repetitive operations to controllers or gesture-based inputs [13].

Sensory Feedback – Sensory feedback refers to the information users receive through their senses as a result of their actions during segmentation. It plays a crucial role in guiding user actions, improving precision, and enhancing spatial awareness. Existing 3D medical image segmentation systems predominantly rely on *Visual* feedback, with some also incorporating *Haptic* or *Auditory* cues.

Visual feedback typically includes perceptual updates, such as highlighting segmented regions or displaying real-time contours, whether on a flat screen or within an immersive environment [4, 12, 21, 32, 59, 60, 65, 68, 70]. In some systems, visual feedback is complemented by *Haptic* or *Auditory* feedback to improve the richness of interactions. *Haptic feedback* provides physical sensations, such as resistance or vibration, to simulate the tactile qualities of anatomical boundaries [45]. *Auditory feedback* delivers information through sound cues, helping users track segmentation progress or identify anatomical features [56]. As most existing systems are based solely on visual feedback, the integration of haptic and auditory modalities presents an interesting and underexplored opportunity to improve segmentation experiences in immersive environments.

Display Modality – This dimension captures the type of visual environment through which users perceive and interact with medical images during segmentation. The display setup influences how spatial information is conveyed, how users perceive depth and orientation cues, and how naturally they can perform segmentation tasks in different contexts. We consider three common settings: *Desktop Monitors*, which are typically used for 2D interactions and represent a traditional mode of medical image segmentation [21, 32, 45]; *Immersive Environments*, including VR and AR, enable direct 3D spatial interaction and have been adopted in several systems such as VRContour, VRRRRoom, and Elucis [4, 12, 56, 60, 65]. *Pad/tablets* support touch-based manipulation and are occasionally integrated with AR displays, as demonstrated in Anatomy Studio and ARMedicalSketch [68, 70].

3.5 Summary and Design Opportunities

Our design space analysis reveals that even among systems designed for immersive or volumetric segmentation, many still retain interaction paradigms rooted in traditional desktop workflows. For example, the tools proposed by Igarashi [32] offer a novel threshold painting interface for volumetric data, but the interaction remains entirely bound to a 2D slice-based desktop environment, relying on mouse input and flat-screen displays. This reliance on planar segmentation and traditional input modalities persists in several systems, reflecting the inertia of conventional segmentation tools that prioritize familiarity and precision over spatial immersion. Although recent systems support 3D data and immersive displays, segmentation interactions are often confined to 2D spaces or static overlays, rather than leveraging the full potential of interactive 3D workflows.

In terms of interaction techniques, visual remains the primary, if not exclusive, feedback to convey the state of the system and guide segmentation actions. Despite the availability of VR hardware that supports rich haptic and auditory output, very few systems explore these modalities. Our analysis shows that haptic and auditory feedback is minimally integrated, leaving a significant opportunity space to improve spatial awareness, segmentation precision, and user engagement.

Furthermore, input modality patterns suggest a lack of adaptive strategies for the task. Although many systems provide controller-based navigation or stylus-based contouring, few explicitly tailor input types to segmentation subtasks. For example, high-precision interactions, such as boundary tracing, may benefit from stylus or pen input, while coarse navigation or volume filtering could be better served by gesture or controller-based input.

These gaps point to a broader challenge in immersive segmentation: bridge the cognitive load imposed by 3D interaction. As Norman describes in his model of interaction [44], users must continuously bridge the *gulf of evaluation*—understanding the current system state—and the *gulf of execution*—deciding and enacting the next action. Most current systems rely on visual feedback alone, which can overload users' perception, particularly when field-of-view constraints and ambiguous depth cues make it difficult to maintain spatial coherence.

To address these limitations, we propose two complementary designs that enrich sensory feedback using underexplored modalities: a *haptic design* that simulates tissue resistance during segmentation and an *auditory design* that sonifies voxel characteristics in real time. These designs aim to reduce ambiguity, support precise segmentation in 3D contexts, and ultimately help users navigate and manipulate volumetric medical data more effectively. The following sections describe each design in detail.

4 Exploring Haptic–Visual Feedback Design for 3D Medical Segmentation

4.1 Haptic–Visual Feedback Design

This design aims to help users perceive boundary resistance, tissue stiffness, and spatial transitions during manual segmentation of 3D MRI volumes. Although visual feedback is crucial for identifying anatomical structures on 2D interfaces, it can face challenges in 3D immersive environments , such as spatial distortions, occlusions, or insufficient depth information. To address this gap, we propose a conceptual design solution to integrate haptic feedback through a force-feedback stylus, enabling users to feel dynamic resistance based on real-time voxel-level information. This provides a complementary sensory approach that reinforces spatial awareness, guides boundary tracing, and reduces cognitive load when operating in regions of uncertainty or low contrast.

Feature Selection from Raw MRI Data. To generate meaningful haptic responses, the system computes voxel-level features directly from raw MRI volumes, without requiring pre-existing labels or segmentation. The following features are selected on the basis of their perceptual relevance to tissue density and structural transition, and their suitability for mapping to force dynamics.

- Voxel intensity: Serves as a proxy for tissue stiffness. Higher intensity regions are assumed to be denser and more resistant to penetration.
- Gradient magnitude: Indicates the sharpness of intensity changes, corresponding to boundaries or abrupt material transitions.
- Depth (z-axis position): Encodes spatial layering, allowing for depth-dependent force modulation.

These features are computed in real time within a small neighbourhood around the stylus tip (e.g., 5×5×5 voxel cube), enabling dynamic, context-sensitive force rendering as the user explores the volume.

Stylus-Based Haptic Device Design The haptic interaction is mediated by a stylus-based force-feedback device, exemplified by devices such as the Geomagic Touch [2]. This device features a desktop robotic arm with six degrees of freedom for tracking the position and orientation of a handheld stylus, and three degrees of freedom for force output. Component Breakdown:

- Stylus Tip: Tracks 3D position and orientation with submillimetre accuracy and serves as the point of interaction with the volume.
- Joints and Linkages: Contain encoders that capture spatial motion and orientation in real time.
- Force actuators: Generate reactive directional forces, applied through the arm to the tip of the stylus, simulating resistance.
- Grounded Armature: Ensures physical stability and isolates force transmission from the user's arm.

As shown in Figure 3, the stylus-based haptic device consists of four key components that enable precise spatial tracking and directional force rendering.

This configuration provides a stable point-based interaction metaphor analogous to a virtual scalpel or probe. The design supports precision and fine-grained manipulation and is in line with familiar clinical practices where radiologists use pen-like instruments to examine images.

Force Feedback Mechanism The generation of haptic force in our system is based on the principles of physical-based modelling [54]. Specifically, we simulate interaction between the user-controlled stylus and the volumetric dataset by treating virtual anatomical structures as deformable or resistive materials. When the tip of the stylus penetrates a

---

[2] Geomagic Touch Haptic Device: https://shop.gomeasure3d.com/products/geomagic-touch-haptic-device

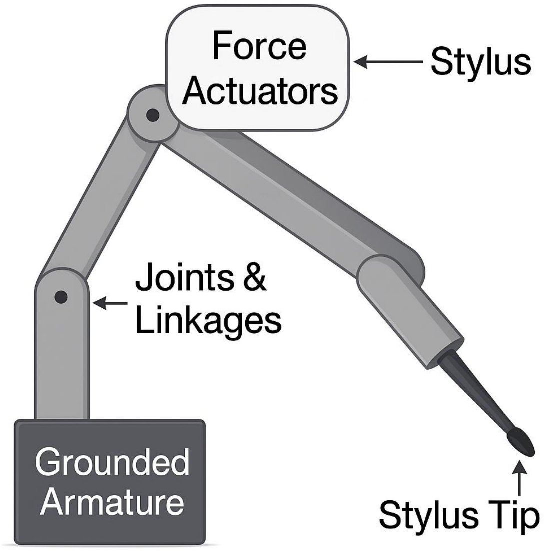


Fig. 3. Hardware structure of the stylus-based haptic device, illustrating its key components: grounded base, articulated joints, force actuators, and stylus tip.

region of interest, such as a tumour boundary or a high-density tissue zone, the system computes a reactive force that pushes back against the motion of the stylus, emulating the sensation of physical resistance. This is modelled using a simplified elastic contact model, where the magnitude of the force is proportional to the depth of penetration into the virtual structure [54]. The reactive force vector $\vec{F}$ is calculated as

$$\vec{F} = -k \cdot d \cdot \hat{n}$$

where:

- d represents the penetration depth of the stylus tip into the target voxel region,
- $\hat{n}$ is the estimated surface normal vector at the point of contact,
- k is a spatially-varying stiffness coefficient derived from local voxel-level features (e.g., intensity or gradient).

This formulation draws on the principles of elastic spring modelling, where the force increases linearly with displacement - an idea commonly used in haptic rendering [15, 29, 55]. In our system, voxel intensity or gradient magnitude is used to modulate k, enabling the simulation of softer or firmer regions depending on the characteristics of the local image. For example, homogeneous soft tissue produces low k values, producing gentle resistance or no force at all. In contrast, tumour cores or structural boundaries with steep gradients yield larger k, producing a stronger tactile response.

The directionality of the force (encoded by $\hat{n}$) ensures that the feedback is applied oppositely to the user's motion, effectively limiting further penetration into virtual structures. The magnitude and vector orientation jointly contribute to the perception of physical constraint, enabling the user to sense virtual material transitions as if interacting with tangible anatomy.

By translating scalar voxel features into continuous force vectors, this model allows users to physically perceive anatomical variation in real time. This improves spatial understanding during segmentation and reduces cognitive reliance on visual estimation alone.

The relationship between voxel-level features and corresponding haptic parameters is summarised in Table 1.

Table 1. Feature-to-force mappings used in this study. Haptic parameters were selected to provide intuitive tactile cues during segmentation.

Image Feature	**Mapped Haptic Parameter**
Voxel intensity	Force magnitude
Gradient magnitude	Force fluctuation
Depth	Progressive damping

To operationalise this physical model in real-time interaction, the system follows a structured pipeline from stylus tracking to force rendering, as outlined below.

1. The position of the stylus is tracked in the 3D space.
2. A local neighbourhood of voxel features is sampled.
3. Feature values are mapped to force vectors using pre-defined transfer functions (e.g., Hookean spring model for stiffness).
4. The resulting force is applied to the stylus actuator, allowing the user to feel the resistance as if touching a physical structure.

4.2 Use Case

Dr. Maya, a radiology resident in a teaching hospital, performs a manual segmentation of a brain MRI volume using our immersive VR system. Equipped with a stylus-based force-feedback device, she enters the virtual environment with the goal of delineating a suspected tumour and exploring the structural complexity of surrounding tissues. The haptic cues provided by the system are designed to assist in tracing boundaries, detecting transitions, and navigating volumetric depth.

Feeling the transition into denser regions. As Maya begins to probe the peritumoural region with the stylus, she encounters only minimal resistance, indicating low-density tissue. Gradually, as she moves toward the tumour core, the stylus begins to push back with increasing force. This growing resistance alerts her that she is entering a stiffer region. She slows her hand and starts to carefully trace the boundary where resistance increases most dramatically, relying on the haptic cue to guide her along the edge of the lesion.

Identifying abrupt structural boundaries. When sweeping the stylus across a suspected boundary zone, Maya suddenly feels a sharp fluctuation in force feedback. The stylus reacts momentarily in a different direction before returning to a smooth motion. This discontinuity suggests a high-gradient transition, probably a boundary between tissue types. The system reaction provides her with a tactile signal of significance of the anatomical segmentation, prompted her to mark this zone for closer inspection in the next segmentation pass.

Perceiving fine tissue heterogeneity. As Maya explores deeper into the parietal lobe, the stylus tip begins to subtly jitter and fluctuate in resistance, even though the overall force remains low. These microvariations reflect local irregularities in voxel intensity and gradient—potentially areas with mixed or disorganised tissue. Although not visually apparent, tactile feedback leads her to interpret this region as structurally heterogeneous. She switches to a more refined tracing mode and segments this area at a higher granularity.

Using depth resistance to sense volumetric layers. Maya adjusts her slicing depth and begins to segment deeper into the volume. As she presses the stylus downward, she notices that the resistance gradually increases, not sharply but with a damping-like effect. This feedback corresponds to her descent into deeper tissue layers, mimicking the sensation of pushing through anatomical density. She interprets this progressive force buildup as a cue to move cautiously, aware that deeper regions may contain compacted structures or lie close to critical anatomy.

Confirming segmentation with tactile consistency. After completing an initial boundary pass, Maya retraces her path in several areas. She probes the same locations again and finds that the force feedback remains consistent: edges resist, cores push back more, and soft zones remain unimpeded. This tactile stability gives her confidence in the segmentation accuracy and reassures her that her annotations align with the actual structural changes embedded in the volume data.

5 Exploring Auditory–Visual Feedback Design for 3D Medical Segmentation

5.1 Auditory–Visual Feedback Design

This design aims to support users in localising and interpreting anatomical structures during manual segmentation of raw 3D MRI volumes. Since no segmentation labels are available prior to interaction, auditory feedback is generated in real time based on directly extracted voxel-level features. The goal is to improve spatial awareness, facilitate boundary detection, and reduce reliance on overloaded visual channels, particularly in immersive environments where multiple layers of volume data may be present simultaneously. To provide intuitive and continuous auditory feedback during manual segmentation, we adopt the parameter mapping sonification approach (PMSon), in which individual image features are directly mapped to specific auditory dimensions. This method has been widely used in interactive systems due to its flexibility and perceptual transparency.

Feature Selection from Raw MRI Data. To generate informative and perceptually meaningful auditory cues from unsegmented MRI volumes, we extract a set of low-level features directly from the voxel grid. These features are selected based on three key criteria: (1) computational efficiency to ensure real-time responsiveness, (2) perceptual relevance for sonification, and (3) their ability to reflect meaningful anatomical variation in the absence of labelled data [19, 30].

- Voxel intensity: reflects signal strength in MRI, serves as a scalar descriptor of tissue composition and proton density [6].
- Gradient magnitude: quantifies the spatial rate of change in intensity values and is a well-established indicator of structural transitions or anatomical boundaries [58].
- Local variance: captures the degree of complexity or heterogeneity of texture within a localised 3D neighbourhood, making it a useful cue to identify irregular structures such as lesions or tumours.
- Sampling depth (e.g., ray penetration distance): defined as the sampling distance or ray intersection depth in the volume, provides spatial context, and supports egocentric navigation and orientation in immersive environments.

These features are computed in real time at the probe–volume intersection, typically aggregated within a localised 3D window (e.g., a 5×5×5 voxel cube), enabling dynamic estimation and continuous auditory feedback as the user explores the data.

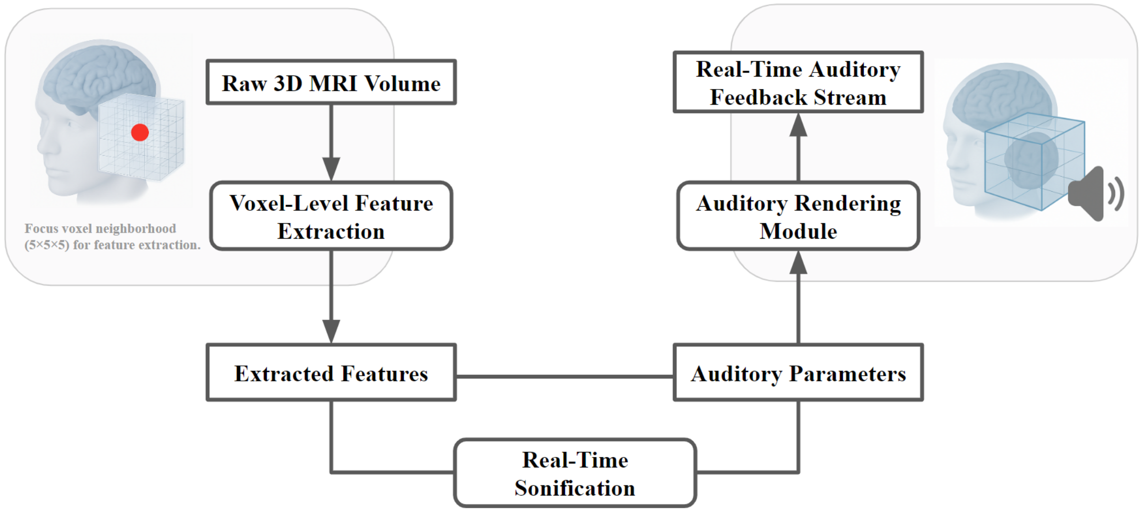


Fig. 4. Overview of the auditory feedback pipeline. A 3D MRI volume is sampled in real time to extract voxel-level features, which are mapped to auditory parameters via a real-time sonification engine. These parameters are rendered into spatialised sound to support immersive manual segmentation.

Feature-to-Sound Mapping Strategy

The selection of auditory parameters in our design is informed by the systematic review conducted by Dubus et al. [19], who analysed 179 published sonification projects to identify recurring trends in feature-to-sound mappings. In their review, the authors classified auditory dimensions into five high-level categories: pitch, loudness, timbre, spatialisation, and temporal — and reported that pitch and loudness were the parameters most frequently used to represent physical quantities. Timbre and spatialisation were also commonly employed, especially in contexts involving material quality, texture, or directionality.

Drawing from these findings, we selected a set of auditory parameters that are perceptually distinct, computationally lightweight, and suitable for real-time interaction. Our goal was to ensure that each type of image feature could be conveyed through sound in a way that is both meaningful and non-intrusive, enhancing spatial understanding without disrupting visual focus.

The specific feature-to-sound mappings used in our system are summarised in table 2.

Table 2. Feature-to-sound mappings used in this study. The choice of auditory parameters is guided by the taxonomy reviewed by Dubus and Bresin [19].

Image Feature	Mapped Sound Parameter
Voxel intensity	Pitch
Gradient magnitude	Loudness / Roughness
Local variance	Timbre
Depth	Rhythm / Spatialisation

In the following, we describe the motivation and perceptual reasoning for each mapping, along with its intended role in supporting manual segmentation.

Voxel intensity → Pitch. The intensity of the voxel is assigned to the pitch, as the pitch has been shown to be an effective auditory dimension to convey scalar variations in physical properties [30], and has frequently been used in sonification to represent continuous data such as pressure and density [11]. In our design, higher voxel intensity values are rendered with higher pitches, leveraging pitch's high perceptual salience to make local density differences more readily detectable in immersive environments.

Gradient magnitude → Loudness / Roughness. We map the magnitude of the intensity gradient to both loudness and timbral roughness, so that stronger gradients produce louder and more spectrally complex sounds. This design draws on natural perceptual associations: abrupt visual or physical transitions are often perceived as more forceful or energetic in auditory terms [19]. Loudness increases linearly with gradient strength, while roughness is introduced by modulating the harmonic structure of the tone, e.g., via noise overlay or spectral inharmonicity. Together, these mappings help users detect anatomical boundaries by making edge regions sound more intense and texturally distinct.

Local variance → Timbre. We encode local variance using timbre, specifically modulated through brightness and spectral centroid, which correspond to perceptual and acoustic dimensions of timbral quality. As discussed by Dubus and Bresin [19], auditory mappings can be described at varying levels of granularity, from general perceptual categories such as timbre to precise signal-level controls such as spectral shaping. Within this framework, brightness and spectral centroid have been frequently employed to convey differences in material composition and structural irregularity, making them appropriate parameters for representing local heterogeneity in medical image data. Regions exhibiting high local variance, potentially corresponding to tumours, lesions, or other irregular structures, are rendered with noisier or sharper timbral characteristics, supporting early recognition of structures even before visual confirmation.

Depth → Rhythm / Spatialisation. Depth is encoded using a combination of rhythm and spatial delay, allowing users to perceive proximity through temporal auditory variations. This strategy is inspired by Bologna et al. [7], who demonstrated that depth can be effectively sonified by using rhythm or sound duration in perceptual mobility aids, allowing blind users to detect nearby objects. Specifically, faster rhythms were intuitively associated with closer objects, while slower rhythms indicated greater distances. Our mapping is designed to support users' perception of depth and spatial relationships during immersive segmentation tasks.

These mappings are designed to be perceptually grounded, intuitive, and continuously responsive, allowing users to interpret structural features of the data through real-time auditory feedback. Each mapping reflects a deliberate choice of auditory dimension based on the nature of the underlying image feature and established sonification practices. In the following, we describe how these mappings are implemented through a real-time sound synthesis engine.

Sound Generation Mechanism To create interactive and perceptually rich sonification, we implement a hybrid synthesis approach using a real-time audio engine (for example, Unity's FMOD system [3]). All sounds are synthesised procedurally, rather than pre-recorded, ensuring immediate auditory feedback in response to user interaction.

Pitch and loudness are controlled through parameter-mapped sine-wave oscillators, where the voxel intensity and gradient magnitude dynamically modulate frequency and amplitude. The timbre is shaped by applying real-time filters, such as low-pass or band-pass filters, with cutoff frequencies modulated by local texture features (e.g., variance or spectral complexity). In regions of high variance, additional broadband noise layers are introduced over the base tone to simulate irregularity and structural roughness.

To encode depth-related cues, rhythmic patterns are generated using envelope modulation and trigger scheduling. Specifically, deeper sampling points reduce the trigger frequency or add a stereo delay, creating the perceptual illusion

---

[3] Unity's FMOD system:https://www.fmod.com/unity

of distance. This mapping is consistent with established auditory perception principles, where rhythm and delay are often associated with spatial separation or temporal pacing [7].

All auditory parameters are recalculated in real time within the local sampling neighbourhood (e.g., a 5×5×5 voxel cube). As the user moves the probe or navigates through the volume, the feature values are updated and mapped to the synthesis parameters without latency, providing continuous feedback that adapts dynamically to the data context.

5.2 Use Case

Leo, a medical imaging researcher, is exploring a 3D brain MRI volume using our immersive VR-based auditory feedback system. His objective is to identify structurally distinct regions and guide preliminary segmentation decisions based on real-time acoustic cues generated directly from voxel-level features. The VR environment is equipped with a standard controller that allows free navigation through the volumetric dataset.

Detect Uniform Structures Leo begins by scanning a region near the corpus callosum. As he moves the VR controller through space, he hears a stable midrange tone with minimal variation. The sound remains smooth and even, indicating a homogeneous tissue area. He continues to sweep the controller in a wide arc, confirming the region's consistency. The lack of tonal change ensures that this area does not contain significant structural variation and can probably be bypassed during detailed manual segmentation.

Sense Density Changes As Leo navigates toward the left temporal lobe, the pitch of the tone starts to rise gradually. At first, the change is subtle, but it soon becomes more pronounced as he nears a high-density cluster. The increasing pitch suggests a gradual intensification in voxel values, possibly indicating a tumour mass or another type of dense tissue. Leo slows down his exploration and mentally marks the region for closer investigation, using the pitch trajectory as an early warning cue.

Identify Structural Boundaries While sweeping across a sulcal boundary, Leo suddenly hears the tone grow louder and rougher in texture. A brief distortion effect cuts in, and the overall volume spikes slightly. This auditory feedback signals a sharp local gradient, prompting Leo to halt his movement. The gritty timbre and sudden volume shift tell him that he has likely crossed a significant anatomical edge. Adjusts the position of the controller, probing the area with short, careful strokes to map the direction of the boundary.

Highlight Heterogeneous Areas Leo then moves into a medial zone near the thalamus. The sound transforms from a clean sine tone into a brighter, more harmonically rich sound. Subtle vibrato and layered modulation become audible. This change in timbre alerts him to an area of internal complexity—perhaps mixed tissue types or partial volume effects. Although the visual contrast is low, the auditory richness suggests that this region warrants closer segmentation. He slows down and makes a mental note to revisit this section during boundary tracing.

Maintain Depth Awareness As he transitions deeper into the volume, Leo notices the auditory rhythm slows down: The pulses are farther apart and more deliberate. When he rises toward the top of the volume, the rhythm accelerates to a steady beat. This rhythmic modulation helps him to sense spatial depth, providing a proprioceptive anchor in the absence of strong visual landmarks. When the rhythm suddenly changes, Leo realises that he has crossed into a previously uninspected region. He pauses to reassess his orientation and consider backtracking.

Throughout the exploration, the layered auditory cues provide Leo with a constant stream of perceptual feedback, reinforcing his spatial awareness and supporting segmentation planning. Each shift in pitch, loudness, texture, or tempo serves a specific purpose: to alert him to a feature boundary, to guide the focus to an irregular zone, or to maintain vertical orientation within the volume.

6 Discussion and Design Implications for Haptic–Visual and Auditory–Visual Feedback

Comparing Haptic and Auditory Feedback Haptic and auditory feedback represent two distinct but complementary modalities to support manual 3D medical image segmentation. Although both aim to extend user perception beyond the visual channel, they differ in their underlying mechanisms and interaction styles. Haptic feedback delivers kinaesthetic cues tightly coupled with user actions, making it well suited for conveying shape and boundary information through direct contact. In contrast, auditory feedback offers continuous, omnidirectional cues, enabling users to passively monitor spatial variation without explicit movement.

These differences influence how users allocate attention and manage the cognitive load. Haptic feedback demands precise motor coordination and is typically limited to localised points of contact, providing high-fidelity control for tracking structures. On the other hand, auditory feedback distributes perceptual load over time, offering greater situational awareness and reducing visual or manual strain, especially in immersive environments where spatial resolution and field of view are limited. Together, these two modalities are not competing but synergistic: auditory alerts can preemptively signal areas of interest, while haptics can offer confirmatory force cues during boundary refinement.

Limitations of Each Approach Despite their advantages, both approaches face limitations. The haptic design depends on hardware constraints such as limited force resolution, workspace size, and the need for continuous physical contact. This makes it less effective for detecting distant or volumetric structures and may induce fatigue during prolonged exploration. Furthermore, because force feedback is inherently local, users must actively scan each region to construct a complete mental map, an effort-intensive and time-intensive process.

Auditory feedback, while more spatially diffuse, comes with its own challenges. Mapping multiple voxel-level features (e.g., intensity, gradient, density) to sound parameters such as pitch, timbre, and roughness can lead to perceptual overlap and cognitive confusion, particularly in noisy or highly heterogeneous datasets. The parameter mapping sonification approach (PMSon) used here is also sensitive to rapid voxel changes, resulting in abrupt audio transitions that may disorient users. Moreover, continuous auditory output—especially in dense volumetric scans, can become fatiguing or distracting over time, particularly for users who are more sensitive to sound. In addition, the interpretation of auditory cues often requires training, as users may struggle to form immediate associations between sonified patterns and anatomical structures without adequate training.

In addition, both systems currently lack global integration capabilities. Haptic interaction remains confined to surface-level contact, while auditory sonification offers limited structural context. Without higher-level abstractions or adaptive modulation, their usefulness diminishes in complex segmentation cases involving layered or ambiguous tissue boundaries.

Opportunities for Multimodal Integration These limitations also point to opportunities for future development. A key direction lies in the integration of haptic and auditory cues into a unified multimodal framework. By aligning feedback across modalities, such as synchronizing haptic confirmation with auditory alerts, designers can reinforce spatial information through redundant sensory channels, improving user confidence, and reducing ambiguity. For example, auditory cues could guide exploration in uncertain regions, while haptics could provide precise validation upon contact.

Another promising area is lightweight, context-sensitive feedback modulation. Rather than treating the feedback as static, the system could adapt the output based on the segmentation phase, anatomical characteristics, or user behaviour. For example, haptic resistance could increase when boundary tracing is detected, while auditory feedback could be minimised in uniform regions and selectively enriched in high-gradient zones. Although not relying on full AI segmentation, such minimally assisted adaptation - driven by context-awareness cues - can reduce cognitive burden. Allowing users to customise sensitivity thresholds or toggle modalities may further enhance usability and accommodate individual preferences.

In summary, this study highlights the value of haptic and auditory feedback as perceptual enhancements for immersive medical segmentation. Although each modality offers unique strengths, their combination—guided by perceptual design principles and adaptive mechanisms, can enable more intuitive, precise, and user-centred workflows in complex 3D environments.

7 Conclusion and Future Work

This paper introduces MedSegWeave, a design space framework for 3D medical image visualisation and segmentation. While the primary focus is on immersive systems, the design space framework also incorporates design patterns from traditional manual segmentation tools to enable comparative analysis across paradigms. It organises these patterns along critical dimensions such as data representation, interaction techniques, input modalities, and sensory feedback. Through this comparative mapping, MedSegWeave identifies a consistent reliance on visual-only workflows and highlights underexplored opportunities for multi-sensory support in immersive environments.

To explore these opportunities, we presented two conceptual designs that integrate haptic and auditory feedback into immersive segmentation workflows. The haptic system leverages force-feedback stylus input to deliver tactile cues based on voxel properties, enabling users to feel structural resistance during segmentation. In parallel, the auditory system employs parameter mapping sonification to translate voxel-level features into continuous acoustic cues, supporting spatial orientation and feature recognition through sound. Together, these designs demonstrate the potential of multimodal interaction to enhance user control, perceptual clarity, and task efficiency.

Through our conceptual exploration, we identified key characteristics and trade-offs in sensory integration that must be balanced according to task demands and user context. Haptic feedback offers precise boundary guidance, but may lead to physical fatigue or be constrained by hardware resolution. Auditory cues, on the other hand, provide ambient awareness and directional hints without requiring direct contact, yet they require careful tuning to avoid perceptual overload. These findings highlight the importance of aligning the choice of modality with task-specific requirements and developing adaptive systems capable of dynamically modulating feedback.

Future work may extend this framework by empirically evaluating the effectiveness of each feedback modality through systematic user studies. Such investigations should incorporate well-established evaluation metrics, including segmentation accuracy (e.g., Dice Similarity Coefficient), task efficiency (e.g., completion time), cognitive workload, and user learning curves across different levels of expertise, with particular emphasis on metrics such as Dice, Intersection-over-Union (IoU), Sensitivity, and Average Hausdorff Distance (AHD), which, as recommended by Müller et al. [42], provide robust means for assessing performance, particularly in the presence of class imbalance and variable contour complexity. In addition, more advanced multimodal coordination strategies—such as synchronised visual–auditory or haptic–auditory cues—could be explored to reduce perceptual ambiguity and enhance learnability. Beyond structural MRI, this framework may also inform the development of interactive segmentation tools for other volumetric imaging modalities, including CT and functional MRI, thus supporting a broader spectrum of diagnostic and interventional workflows. We envision MedSegWeave as a foundation for future multimodal segmentation systems that are both perceptually rich and clinically robust.

Acknowledgments

I would like to sincerely thank my supervisors, Jiazhou (Joe) Liu, Vahid Pooryousef, Yidan Zhang, and Tim Dwyer, for their continued support, guidance, and encouragement throughout this research. Their dedication and insightful advice have greatly strengthened the depth and clarity of this research.

References

[1] Khalid A Al-Kofahi, Ali Can, Sharie Lasek, Donald H Szarowski, Natalie Dowell-Mesfin, William Shain, James N Turner, and Badrinath Roysam. 2003. Median-based robust algorithms for tracing neurons from noisy confocal microscope images. *IEEE Transactions on information technology in biomedicine* 7, 4 (2003), 302–317.

[2] Lucas Battich, Merle Fairhurst, and Ophelia Deroy. 2020. Coordinating attention requires coordinated senses. *Psychonomic bulletin & review* 27 (2020), 1126–1138.

[3] Stefan Bauer, Roland Wiest, Lutz-P Nolte, and Mauricio Reyes. 2013. A survey of MRI-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology* 58, 13 (2013), R97.

[4] Jason Belec, Justin Sutherland, Matthew Volpini, Kawan Rakhra, Dal Granville, Dan La Russa, Teresa Flaxman, Eduardo Portela De Oliveira, Rafael Glikstein, Marlise P Dos Santos, et al. 2024. A Pilot Clinical and Technical Validation of an Immersive Virtual Reality Platform for 3D Anatomical Modeling and Contouring in Support of Surgical and Radiation Oncology Treatment Planning. *Journal of Imaging Informatics in Medicine* (2024), 1–16.

[5] David Black, Christian Hansen, Arya Nabavi, Ron Kikinis, and Horst Hahn. 2017. A survey of auditory display in image-guided interventions. *International journal of computer assisted radiology and surgery* 12 (2017), 1665–1676.

[6] Jerzy Bodurka, F Ye, Natalia Petridou, Kevin Murphy, and Peter A Bandettini. 2007. Mapping the MRI voxel volume in which thermal noise matches physiological noise—implications for fMRI. *Neuroimage* 34, 2 (2007), 542–549.

[7] Guido Bologna, Benoit Deville, and Thierry Pun. 2010. Sonification of color and depth in a mobility aid for blind people. In *Proceedings of the 16th International Conference on Auditory Display (ICAD2010), Washington, DC, USA*.

[8] Fabiano Botta, Valerio Santangelo, Antonino Raffone, Daniel Sanabria, Juan Lupiáñez, and Marta Olivetti Belardinelli. 2011. Multisensory integration affects visuo-spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance* 37, 4 (2011), 1099.

[9] Douglas A Bowman. 1999. *Interaction techniques for common tasks in immersive virtual environments: design, evaluation, and application*. Georgia Institute of Technology.

[10] Lizhou Cao, Huadong Zhang, Chao Peng, and Jeffrey T Hansberger. 2023. Real-time multimodal interaction in virtual reality-a case study with a large virtual interface. *Multimedia Tools and Applications* 82, 16 (2023), 25427–25448.

[11] Daniel Cesarini, Thomas Hermann, and Bodo Ungerechts. 2014. An interactive sonification system for swimming evaluated by users. In *Sonification of Health and Environmental Data-York 2014. Conference Proceedings*.

[12] Chen Chen, Matin Yarmand, Varun Singh, Michael V Sherer, James D Murphy, Yang Zhang, and Nadir Weibel. 2022. VRContour: bringing contour delineations of medical structures into virtual reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 64–73.

[13] Chen Chen, Matin Yarmand, Zhuoqun Xu, Varun Singh, Yang Zhang, and Nadir Weibel. 2022. Investigating input modality and task geometry on precision-first 3d drawing in virtual reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 384–393.

[14] Dapeng Chen, Yi Ding, Hao Wu, Qi Jia, Hong Zeng, Lina Wei, Chengcheng Hua, Jia Liu, and Aiguo Song. 2025. Visually Impaired People Learning Virtual Textures Through Multimodal Feedback Combining Vibrotactile and Voice. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2025).

[15] J Edward Colgate and Gerd G Schenkel. 1997. Passivity of a class of sampled-data systems: Application to haptic interfaces. *Journal of robotic systems* 14, 1 (1997), 37–47.

[16] Michael Correll and Laura Garrison. 2024. When the Body Became Data: Historical Data Cultures and Anatomical Illustration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.

[17] Kunio Doi. 2006. Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology. *Physics in Medicine & Biology* 51, 13 (2006), R5.

[18] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. Vrsketchin: Exploring the design space of pen and tablet interaction for 3d sketching in virtual reality. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–14.

[19] Gaël Dubus and Roberto Bresin. 2013. A systematic review of mapping strategies for the sonification of physical quantities. *PloS one* 8, 12 (2013), e82491.

[20] Barrett Ens, Juan David Hincapié-Ramos, and Pourang Irani. 2014. Ethereal planes: a design framework for 2D information space in 3D mixed reality environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. 2–12.

[21] Pedro M Ferreira, Teresa Mendonça, Jorge Rozeira, and Paula Rocha. 2012. An annotation tool for dermoscopic image segmentation. In *Proceedings of the 1st International Workshop on Visual Interfaces for Ground Truth Collection in Computer Vision Applications*. 1–6.

[22] Brent Foster, Ulas Bagci, Awais Mansoor, Ziyue Xu, and Daniel J Mollura. 2014. A review on segmentation of positron emission tomography images. *Computers in biology and medicine* 50 (2014), 76–96.

[23] Yi Gao, Cheng Chang, Xiaxia Yu, Pengjin Pang, Nian Xiong, and Chuan Huang. 2022. A VR-based volumetric medical image segmentation and visualization system with natural human interaction. *Virtual Reality* 26, 2 (2022), 415–424.

[24] Christina Gillmann, Noeska N Smit, Eduard Gröller, Bernhard Preim, Anna Vilanova, and Thomas Wischgoll. 2021. Ten open challenges in medical visualization. *IEEE Computer Graphics and Applications* 41, 5 (2021), 7–15.

[25] Letizia Gionfrida and Agnieszka Roginska. 2017. A novel sonification approach to support the diagnosis of Alzheimer's dementia. *Frontiers in neurology* 8 (2017), 647.

[26] Lee W Goldman. 2007. Principles of CT and CT technology. *Journal of nuclear medicine technology* 35, 3 (2007), 115–128.

[27] Lennart Johannes Gruber, Jan Egger, Andrea Bönsch, Joep Kraeima, Max Ulbrich, Vincent van den Bosch, Ila Motmaen, Caroline Wilpert, Mark Ooms, Peter Isfort, et al. 2024. Accuracy and precision of mandible segmentation and its clinical implications: virtual reality, desktop screen and artificial intelligence. *Expert Systems with Applications* 239 (2024), 122275.

[28] Ryan Gunther, Rick Kazman, and Carolyn MacGregor. 2004. Using 3D sound as a navigational aid in virtual environments. *Behaviour & Information Technology* 23, 6 (2004), 435–446.

[29] Vincent Hayward and Karon E MacLean. 2007. Do it yourself haptics: part I. *IEEE Robotics & Automation Magazine* 14, 4 (2007), 88–104.

[30] Thomas Hermann and Helge Ritter. 2004. Sound and meaning in auditory data display. *Proc. IEEE* 92, 4 (2004), 730–741.

[31] Luis J Hernández-Sierra, Marco O Mendoza-Gutiérrez, and Aldo R Mejía-Rodríguez. 2020. Design of a haptic system for medical image manipulation using augmented reality. *Health and technology* 10, 2 (2020), 447–451.

[32] Takeo Igarashi, Naoyuki Shono, Taichi Kin, and Toki Saito. 2016. Interactive volume segmentation with threshold field painting. In *Proceedings of the 29th annual symposium on user interface software and technology*. 403–413.

[33] Zachary Jones, Simon Drouin, and Marta Kersten-Oertel. 2025. Exploring interaction paradigms for segmenting medical images in virtual reality. *International Journal of Computer Assisted Radiology and Surgery* (2025), 1–9.

[34] Daniel Jönsson, Erik Sundén, Anders Ynnerman, and Timo Ropinski. 2014. A survey of volumetric illumination techniques for interactive volume rendering. In *Computer Graphics Forum*, Vol. 33. Wiley Online Library, 27–51.

[35] Daisuke Kawahara and Yasushi Nagata. 2021. T1-weighted and T2-weighted MRI image synthesis with convolutional generative adversarial networks. *reports of practical Oncology and radiotherapy* 26, 1 (2021), 35–42.

[36] Jiazhou Liu, Barrett Ens, Arnaud Prouzeau, Jim Smiley, Isobel Kara Nixon, Sarah Goodwin, and Tim Dwyer. 2023. Datadancing: An exploration of the design space for visualisation view management for 3d surfaces and spaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[37] Karon E MacLean. 2000. Designing with haptic feedback. In *Proceedings 2000 icra. millennium conference. ieee international conference on robotics and automation. symposia proceedings (cat. no. 00ch37065)*, Vol. 1. IEEE, 783–788.

[38] Chiara Magliaro, Alejandro L Callara, Nicola Vanello, and Arti Ahluwalia. 2017. A manual segmentation tool for three-dimensional neuron datasets. *Frontiers in neuroinformatics* 11 (2017), 36.

[39] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H Thomas. 2018. *Immersive analytics*. Vol. 11190. Springer.

[40] Sasan Matinfar, Shervin Dehghani, Mehrdad Salehi, Michael Sommersperger, Navid Navab, Koorosh Faridpooya, Merle Fairhurst, and Nassir Navab. 2025. From tissue to sound: A new paradigm for medical sonic interaction design. *Medical Image Analysis* 103 (2025), 103571.

[41] Syed T Mubarrat, Suman K Chowdhury, and Antonio S Fernandes. 2024. Evaluating visual-spatiotemporal co-registration of a physics-based virtual reality haptic interface. *IEEE Access* 12 (2024), 57017–57032.

[42] Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. 2022. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes* 15, 1 (June 2022). https://doi.org/10.1186/s13104-022-06096-y

[43] Mary K. Ngo and Charles Spence. 2010. Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, Perception, & Psychophysics* 72, 6 (2010), 1654–1664. https://doi.org/10.3758/APP.72.6.1654

[44] Donald A Norman. 1988. *The Design of Everyday Things*. Springer.

[45] Ingela Nyström, Johan Nysjö, and Filip Malmberg. 2011. Visualization and haptics for interactive medical image analysis: Image segmentation in cranio-maxillofacial surgery planning. In *International Visual Informatics Conference*. Springer, 1–12.

[46] Dinesh D Patil and Sonal G Deore. 2013. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing* 2, 1 (2013), 22–27.

[47] Duc-Minh Pham and Wolfgang Stuerzlinger. 2019. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. 1–11.

[48] Vahid Pooryousef, Ross Brown, and Selen Turkay. 2019. Shape recognition and selection in medical volume visualisation with haptic gloves. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*. 433–436.

[49] Bernhard Preim and Charl P Botha. 2013. *Visual computing for medicine: theory, algorithms, and applications*. Newnes.

[50] Arnaud Prouzeau, Maxime Cordeil, Clement Robin, Barrett Ens, Bruce H Thomas, and Tim Dwyer. 2019. Scaptics and highlight-planes: Immersive interaction techniques for finding occluded features in 3d scatterplots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[51] Hanna-Riikka Rantamaa, Jari Kangas, Sriram Kishore Kumar, Helena Mehtonen, Jorma Järnstedt, and Roope Raisamo. 2023. Comparison of a vr stylus with a controller, hand tracking, and a mouse for object manipulation and medical marking tasks in virtual reality. *Applied Sciences* 13, 4 (2023), 2251.

[52] Tom Ritchey. 1998. Fritz Zwicky, morphologie and policy analysis. In *16th EURO conference on operational analysis, Brussels*.

[53] Renaud Ronsse, Veerle Puttemans, James P Coxon, Daniel J Goble, Johan Wagemans, Nicole Wenderoth, and Stephan P Swinnen. 2011. Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. *Cerebral cortex* 21, 6 (2011), 1283–1294.

[54] Louis B Rosenberg. 1992. The use of virtual fixtures as perceptual overlays to enhance operator performance in remote environments. *Air force material command* (1992), 1–42.

[55] Kenneth Salisbury, Francois Conti, and Federico Barbagli. 2004. Haptic rendering: introductory concepts. *IEEE computer graphics and applications* 24, 2 (2004), 24–32.

[56] Laura Schütz, Sasan Matinfar, Gideon Schafroth, Navid Navab, Merle Fairhurst, Arthur Wagner, Benedikt Wiestler, Ulrich Eck, and Nassir Navab. 2024. A Framework for Multimodal Medical Image Interaction. *IEEE Transactions on Visualization and Computer Graphics* (2024).

[57] Afshin Shoeibi, Marjane Khodatars, Mahboobeh Jafari, Navid Ghassemi, Delaram Sadeghi, Parisa Moridian, Ali Khadem, Roohallah Alizadehsani, Sadiq Hussain, Assef Zare, et al. 2024. Automated detection and forecasting of covid-19 using deep learning techniques: A review. *Neurocomputing* (2024), 127317.

[58] Krit Somkantha, Nipon Theera-Umpon, and Sansanee Auephanwiriyakul. 2010. Boundary detection in medical images using edge following algorithm based on intensity gradient and texture gradient features. *IEEE transactions on biomedical engineering* 58, 3 (2010), 567–573.

[59] Maurício Sousa, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge, and Daniel Simões Lopes. 2017. Vrrrroom: Virtual reality for radiologists in the reading room. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 4057–4062.

[60] The AVRO Lab. 2018. DICOM VR: Virtual Reality Contouring System for Medical Imaging. https://avrolab.bwh.harvard.edu/. Accessed: 2025-06-05.

[61] Arnau Valls-Esteve, Núria Adell-Gómez, Albert Pasten, Ignasi Barber, Josep Munuera, and Lucas Krauel. 2023. Exploring the potential of three-dimensional imaging, printing, and modeling in pediatric surgical oncology: a new era of precision surgery. *Children* 10, 5 (2023), 832.

[62] Erik Van der Burg, Christian NL Olivers, Adelbert W Bronkhorst, and Jan Theeuwes. 2008. Audiovisual events capture attention: Evidence from temporal order judgments. *Journal of vision* 8, 5 (2008), 2–2.

[63] Erik Van der Burg, Christian NL Olivers, Adelbert W Bronkhorst, and Jan Theeuwes. 2008. Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance* 34, 5 (2008), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

[64] Bruce N Walker and Michael A Nees. 2011. Theory of sonification. *The sonification handbook* 1 (2011), 9–39.

[65] Song Wang, Dong Zhu, Hao Yu, and Yadong Wu. 2020. Immersive wysiwyg (what you see is what you get) volume visualization. In *2020 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 166–170.

[66] Adam S Williams, Jason Garcia, and Francisco Ortega. 2020. Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3479–3489.

[67] Koon-Pong Wong. 2005. Medical image segmentation: methods and applications in functional imaging. In *Handbook of Biomedical Image Analysis: Volume II: Segmentation Models Part B*. Springer, 111–182.

[68] Nan Zhang, Tianqi Huang, and Hongen Liao. 2024. ARMedicalSketch: Exploring 3D Sketching for Medical Image Using True 2D-3D Interlinked Visualization and Interaction. *IEEE Transactions on Human-Machine Systems* (2024).

[69] Yichi Zhang, Qingcheng Liao, Le Ding, and Jicong Zhang. 2022. Bridging 2D and 3D segmentation networks for computation-efficient volumetric medical image segmentation: An empirical study of 2.5 D solutions. *Computerized Medical Imaging and Graphics* 99 (2022), 102088.

[70] Ezequiel R Zorzal, Maurício Sousa, Daniel Mendes, Rafael Kuffner dos Anjos, Daniel Medeiros, Soraia Figueiredo Paulo, Pedro Rodrigues, José João Mendes, Vincent Delmas, Jean-Francois Uhl, et al. 2019. Anatomy studio: a tool for virtual dissection through augmented 3D reconstruction. *Computers & Graphics* 85 (2019), 74–84.

List of existing designs in our design space

1. Zorzal, Ezequiel R., Maurício Sousa, Daniel Mendes, Rafael Kuffner dos Anjos, Daniel Medeiros, Soraia Figueiredo Paulo, Pedro Rodrigues, et al. 2019. Anatomy studio: A tool for virtual dissection through augmented 3D reconstruction. *Computers & Graphics* 85 (2019), 74–84. https://doi.org/10.1016/j.cag.2019.09.006 [**Anatomy Studio**]

2. **Takeo Igarashi, Naoyuki Shono, Taichi Kin, and Toki Saito. 2016.** Interactive volume segmentation with threshold field painting. *In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16).* ACM, New York, NY, USA, 403–413. https://doi.org/10.1145/2984511.2984537 [**Threshold Field Painting**]

3. Chen, Chen, Matin Yarmand, Varun Singh, Michael V. Sherer, James D. Murphy, Yang Zhang, and Nadir Weibel. 2022. VRContour: Bringing contour delineations of medical structures into virtual reality. *In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '22).* IEEE, 64–73, doi: 10.1109/ISMAR55827.2022.00020 [**VRContour**]

4. Ferreira, Pedro M., Teresa Mendonça, Jorge Rozeira, and Paula Rocha. 2012. An annotation tool for dermoscopic image segmentation. *In Proceedings of the 1st International Workshop on Visual Interfaces for Ground Truth Collection in Computer Vision Applications (VIGTA '12).* ACM, New York, NY, USA, Article 1, 1–6. https://doi.org/10.1145/2304496.2304501 [**Annotation tool**]

5. Wang, Song, Dong Zhu, Hao Yu, and Yadong Wu. 2020. Immersive wysiwyg (what you see is what you get) volume visualization. *In Proceedings of the 2020 IEEE Pacific Visualization Symposium (PacificVis '20).* IEEE, 166–170, doi: 10.1109/PacificVis48177.2020.1001 [**Immersive WYSIWYG**]

6. Schütz, Laura, Sasan Matinfar, Gideon Schafroth, Navab Navid, Merle Fairhurst, Arthur Wagner, Benedikt Wiestler, Ulrich Eck, and Nassir Navab. 2024. A Framework for Multimodal Medical Image Interaction. *IEEE Transactions on Visualization and Computer Graphics* (TVCG), September 10, 2024, doi: 10.1109/TVCG.2024.3456163 [**MMII**]

7. Sousa, Maurício, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge, and Daniel Simões Lopes. 2017. VRRRRoom: Virtual Reality for

Radiologists in the Reading Room. *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI ’17)*. Association for Computing Machinery, New York, NY, USA, 4057–4062. https://doi.org/10.1145/3025453.3025566 [**VRRRRoom**]

8. Zhang, Nan, Tianqi Huang, and Hongen Liao. 2024. ARMedicalSketch: Exploring 3D Sketching for Medical Image Using True 2D–3D Interlinked Visualization and Interaction. *IEEE Transactions on Human-Machine Systems* (Aug. 2024), doi: 10.1109/THMS.2024.3432735 [**ARMedicalSketch**]

9. The AVRO Lab. 2018. DICOM VR: Virtual Reality Contouring System for Medical Imaging. *Brigham and Women’s Hospital, Harvard Medical School*. Retrieved June 5, 2025 from https://avrolab.bwh.harvard.edu/ [**DICOM VR**]

10. Belec, Jason, Justin Sutherland, Matthew Volpini, Kawan Rakhra, Dal Granville, Dan La Russa, Teresa Flaxman, Eduardo P. de Oliveira, Edward P. Glikstein, Marc P. Dos Santos, and Jason W. Werier. 2024. A pilot clinical and technical validation of an immersive virtual reality platform for 3D anatomical modeling and contouring in support of surgical and radiation oncology treatment planning.

Journal of Imaging Informatics in Medicine, 2024, 1–16. https://doi.org/10.1007/s10278-024-01048-3 [**Elucis**]

11. Nyström, Ingela, Johan Nysjö, and Filip Malmberg. 2011. Visualization and haptics for interactive medical image analysis: Image segmentation in cranio-maxillofacial surgery planning. In *Proceedings of the International Visual Informatics Conference (IVIC 2011)*. Springer, Berlin, Heidelberg, 1–12. https://doi.org/10.1007/978-3-642-25191-7_1 [**CMF Haptics**]