

VRContour: Bringing Contour Delineations of Medical Structures Into Virtual Reality

Chen Chen^{1*} Matin Yarmand¹ Varun Singh¹

Michael V. Sherer² James D. Murphy² Yang Zhang³ Nadir Weibel¹

¹Computer Science and Engineering and ²School of Medicine, University of California San Diego, La Jolla, CA, United States

³Electrical and Computer Engineering, University of California Los Angeles, Los Angeles, CA, United States

ABSTRACT

Contouring is an indispensable step in Radiotherapy (RT) treatment planning. However, today's contouring software is constrained to only work with a 2D display, which is less intuitive and requires high task loads. Virtual Reality (VR) has shown great potential in various specialties of healthcare and health sciences education due to the unique advantages of intuitive and natural interactions in immersive spaces. VR-based radiation oncology integration has also been advocated as a target healthcare application, allowing providers to directly interact with 3D medical structures. We present *VRContour* and investigate how to effectively bring contouring for radiation oncology into VR. Through an autobiographical iterative design, we defined three design spaces focused on contouring in VR with the support of a tracked tablet and VR stylus, and investigating dimensionality for information consumption and input (either 2D or 2D + 3D). Through a within-subject study ($n = 8$), we found that visualizations of 3D medical structures significantly increase precision, and reduce mental load, frustration, as well as overall contouring effort. Participants also agreed with the benefits of using such metaphors for learning purposes.

Keywords: Human–Computer Interactions for Health, Radiotherapy (RT) Treatment Planning, VR in Health & Medical Applications

Index Terms: Applied computing—Healthcare information systems;

1 INTRODUCTION

Radiotherapy (RT) treatment is indispensable in cancer management. Contouring is a critical step in the RT treatment planning workflow where oncologists need to identify and outline malignant tumors and/or healthy organs from a stack of medical images (Fig. 1). Inaccurate contouring could lead to systematic errors throughout the entire treatment course, leading to missing the tumors or overtreating the healthy tissues [62], and could cause increased risks of toxicity, tumor recurrence and even death [59]. Oncology software tools, *e.g.*, Eclipse [4], have been introduced to simplify the contouring workflow: oncologists can identify and segment key structures by analyzing and delineating outlines on a stack of 2D scans using a general 2D desktop display (Fig. 1b–d). This is sometimes not intuitive, and mentally reconstructing 3D physiological structures could introduce significant cognitive load while analyzing 2D images [36].

3D immersive experiences in Virtual Reality (VR) have unlocked a variety of applications in healthcare (*e.g.*, [34]) and workspaces (*e.g.*, [23]). Because of its unique advantages of intuitive and natural interaction, VR has shown great potential in different specialties of healthcare and health science education [13]. “Radiation oncology integration” has also been advocated as a target

healthcare application for providers in VR [19, 38, 57]. Prior research has evaluated the affordances of VR while being applied to visualize RT treatment plans [46, 53] and enhance understanding of dose distributions [47]. However, the delineation of contours — the initial phase of the workflow (Fig. 1a) — has not been explored. A few recent works (*e.g.*, [1, 11, 24]) attempted to bring radiation planning procedures into VR. However, the lack of ways to enable precise drawing of contours hinders the practical uses. While Williams *et al.* [57] showed that 83% of oncologist participants found it useful to use VR for visualizing and contouring medical images, the kinds of merits that VR could offer are still unclear.

In this work, we present *VRContour* and investigate *how to bring radiation oncology contouring into head-mounted VR by leveraging the merits introduced by 3D User Interfaces (UI)*. Since contouring is a challenging and highly specialized task in oncology, we spent more than nine months with domain experts and applied a user-centered design approach to consider their practical needs (*e.g.*, precision and lead time). Through an iterative and autobiographical design process with healthcare domain experts at different experience levels (*incl.* a MD student, a resident and an attending physician), we decided to use a tracked tablet and Logitech VR stylus [7] for its high drawing precision, and defined three design spaces by considering the dimensionality (*i.e.*, either 2D or 2D + 3D) of *information consumption* (*i.e.*, 2D: only visualizing cutting-planes on 2D tablet; 2D + 3D: visualizing both cutting-planes on 2D tablet and 3D medical structure) and *information input* (*i.e.*, 2D: only support contouring on 2D tablet with VR stylus; 2D + 3D: support contouring on both 2D tablet and direct drawing in 3D). We populated essential components and metaphors (*e.g.*, menu designs in immersive environment, 3D object manipulations, and inter-slice interpolations) based on prior works and through our iterative design.

We built a proof-of-concept prototype to evaluate three design spaces on the HTC Vive Eye Pro [6]. A separate prototype used to emulate today's contouring on a desktop-based 2D display was also implemented on the same headset as a baseline reference. To minimize the impact of prior experience with 2D image interpretation, we intentionally involved only junior MD students and other health science students (*e.g.*, pre-med) who have been trained in anatomy, but without extensive exposure to 2D medical images. Through a within-subject study ($n = 8$), we showed how 3D visualizations could increase contouring precision, although the capabilities of direct contouring in 3D did not reduce lead time. Our user experience evaluations suggested that *VRContour* leads to reductions in mental load, frustration, and overall effort when introducing 3D visualizations, and all participants agreed that contouring in 3D is promising for learning purposes. We believe our work will benefit and direct future efforts on designing VR-based oncology software.

*corresponding author, e-mail: chenchen@ucsd.edu

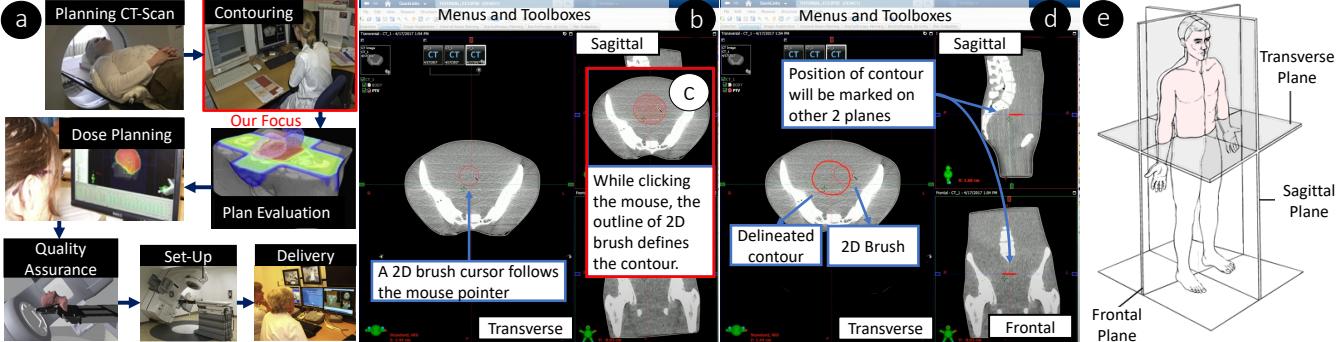


Figure 1: (a) Workflow of today's RT treatment [19]. (b – d) Contouring using Eclipse [4]; For delineating contours, a 2D brush is used (b); while clicking the mouse, an area is colored (c); upon releasing the mouse, the contour is defined by the colored area (d). (e) Three planes of motions.

2 RELATED WORK

2.1 Contouring in Radiotherapy (RT) Treatment Planning

RT treatment aims to deliver high doses of radiation to a tumor while limiting the dose that is received by the surrounding healthy tissue, which can be categorized into seven steps (Fig. 1a) [19]. Providers need to interpret CT scans (MRI and PET if needed) to mentally reconstruct a “virtual anatomy” and precisely locate and segment tumors as well as surrounding healthy tissues on each slice. Oncology software, *e.g.*, Eclipse [4], is widely used in today's clinics to streamline this process, allowing oncologists to easily align the scans and draw outlines with mouse and/or stylus. Figures 1b–d show an example of the Eclipse [4] interface, where oncologists can locate and contour the targets by rapidly switching slices across the preferred plane of motion (Figures 1e) using scroll wheels. A 2D brush is commonly used as the delineation tool for drawing and refining contours (Fig. 1b–d) [56]. Advanced features, *e.g.*, inter-slice interpolations [15, 66] are also used to reduce the number of slices to be manually contoured, leading to less lead time.

Despite these tools, contouring is still a lengthy process, as the oncologists need to examine all slices, each requiring manually distinguishing the tumoral volume from surrounding healthy tissues [28]. While Artificial Intelligence (AI) assisted contouring can effectively empower clinicians beyond their current practices [62], their performance is considered less favourable than manual approaches. Additionally, the limited availability of training datasets could impact the scalability and inclusive design of AI-based contouring. Ramkumar *et al.* [49] investigated the possibilities of integrating novel interaction modalities into the today's contouring workflow. They found that tablet–PC, with multi-touch and gesture-based input are the optimal ones, and could be integrated into the existing workflow for enhancing contouring efficacy.

Contouring on today's 2D interface is also error-prone, especially for clinicians with less experience or in lower tier medical centers [14, 29, 31]. This is because it requires experience for oncologists to mentally reconstruct the medical structure for target localization by looking at a stack of planar scans, the only support provided by existing tools. Dowsett *et al.* [28] showed that more than 30% of errors are introduced while contouring spinal canal and lungs using a mouse and lightpen on 2D displays. Due to the difficulties of mastering contouring skills, existing research has proposed accessible learning supports using web-based [10, 52] and mobile applications [61]. While enhancing contouring training might be helpful, we believe that the new immersive 3D visualization technology can potentially help oncologists simplify and streamline the contouring procedure, and therefore might lower the barriers and efforts of contouring procedures in the long term.

2.2 Precise Drawing in VR

Enabling oncologists to engage in precise VR delineations is challenging. While most existing works focus on drawing for 3D modeling, many design components are transferable to our context.

Early studies have investigated the issues of drawing inaccuracies in 3D [50]. Fitzmaurice *et al.* [33] indicated how orientations of pen and paper could affect the 2D drawing performance. Arora *et al.* [17] showed how the lack of drawing surface support is the major cause of inaccuracies, which also depends on the orientations of drawing surface. Cannavò *et al.* [22] and Chen *et al.* [25] demonstrated the merits of using stylus for 2D and 3D drawing respectively in VR while being compared to traditional VR controller.

With the aforementioned VR based drawing evaluations, several existing works proposed drawing tools by integrating 2D tablets for general designers. For example, SymbiosisSketch [16] proposed a 3D drawing application by combining the interaction metaphors of mid-air drawing and surface drawing, as as to create detailed 3D designs with head-mounted AR. They showcased how designers could create a 3D canvas with mid-air stroke, and subsequently refine and add details with 2D tablet.

While focusing on designing contouring applications for oncologists, our work borrowed several design ideas from these works. However, due to the uniqueness of contouring requirements and the needs of practicing oncologists, directly transferring the techniques from applications facing general users to contouring applications does not necessarily lead to a usable system by radiation oncologists. We therefore co-designed with domain experts iteratively to better understand our system requirements (Sec. 3).

2.3 XR in Healthcare and Health Science Educations

Extended reality (XR) has been advocated as the next frontier for future healthcare and health science education [18]. Existing works have explored the affordances of immersive 3D visualization of medical structures and their real-world applications. Zhang *et al.* [63] built a server-aided 3D DICOM viewer using mobile Augmented Reality (AR), with real-time rendering, interactions and full transfer functions editing. However the lack of contouring support limits its use for RT treatment planning. Anatomy Studio [65] proposed an AR tool for virtual dissection that combines tablets with stylus to assist anatomists by easing manual tracing and exploring cryosection images. While supporting manual segmentation, the limited resolution and the non-transparent isosurface rendering prevents anatomists from exploring finer grained structures.

As for health sciences education, Nicholson *et al.* [45] have demonstrated significant merits of using VR for anatomy education, with an average of 18% of score improvement. Ma *et al.* [40] designed an AR applications by overlaying the CT dataset on top of real-world participants, and demonstrated 86.1% approval in terms of educational value. Boejen *et al.* [19, 53] summarized how VR could benefit today's RT training by enabling students to simulate

	Motivations & Needs Analysis [N#]	Insights for Interactions and Interface Design [D#]
Iteration 1 (P1, P2, P3) Initial Needs-Finding	[N1] 3D visualization of medical structure could be beneficial to understand clinical target volume; [N2] Considerations of learning of 3D user interfaces is critical;	[D1] Directly contouring in 3D can be useful, but fine-grained control is still needed;
Iteration 2 (P2, P3) Initial Prototype Design & Evaluations	[N3] Today's doctors are well trained for, and thus getting used to 2D medical image interpretations. Radically switching to 3D might be a too large of a change; [N4] 3D visualizations of medical structure might be beneficial for contouring education;	[D2] Aggregate all UI elements in the same control panel could be overwhelming; [D3] Placing cutting-planes in different orientation could be helpful, yet 3 basic planes of motion are what domain experts are familiar with; [D4] Visualizations of 3D medical structures along with cutting planes might be helpful for better leveraging the features of inter-slice interpolation; [D5] Transparent medical structure might be useful;
Iteration 3 (P2, P3) Prototype Enhancement	[N5] Mapping between 2D and 3D slices might be beneficial;	[D6] 3D volume should be visualized with different rendering parameters; [D7] Brush size should be adjustable based on the size of contouring structures;
Iteration 4 (P2) Re-Evaluations	[N6] Requiring domain expert participants to tune rendering parameters for better 3D visualization could be less intuitive;	[D8] Users should be able to fine tune the slice indices, due to their frequent use in practical contouring procedures;

Figure 2: Insights and takeaways from preliminary interviews and iterative design.

and train clinical situations without interfering with practical clinical workflow and without the risks of making errors.

In terms of RT treatment planning, Patel *et al.* [46] showed how the stereoscopic visualization could foster the understanding of spatial relationships between anatomy and calculated dose distributions. Williams *et al.* [57] suggested that 3D visualization could be advantageous when the definition of the clinical target volume is based on a 3D spread of disease along anatomical features with intricate topology. Such volumes could be difficult to observe and evaluate as they usually do not lie neatly in any of the motion of planes [57]. DICOM VR [1,57] showed a system that allows oncologists to manually segment the target directly in 3D space. Early results of using this system show that while the contouring time seems to be reduced compared to 2D display based software, the drawing precision might be sacrificed to some extent [57]. Although 83% of oncologists found DICOM VR useful [57], the underlying reasons are still obscure. In addition, the lack of physical surface support and use of VR controllers naturally hinders precise delineations while refining outlines on a 2D virtual plane [17]. We aim to understand how VR can benefit contouring, and how to design such a system by fully exploiting the affordances that this medium offers.

3 PRELIMINARY EXPLORATORY AND ITERATIVE DESIGN

We used an autobiographical approach and user-centered design, to explore and then implement as well as refine our prototype *iteratively* with domain experts [24, 43]. All studies (*incl.* Sec. 5) have been approved by the Institutional Review Boards (IRB).

3.1 Participants

Our participants include a third year medical student (P1, F), a senior resident fellow (P2, M), and an attending oncologist (P3, M). P2 and P3 have extensive experience teaching contouring, and all participants are affiliated to the UC San Diego Health System. No participant had prior VR experience. As today's contouring on 2D display is highly dependent on familiarity and understanding of 2D radiology images, the results from P2 and P3 are highly affected by their past experience regarding interpretations of medical scans and 2D contouring workflow. To minimize such impacts we therefore included P1 intentionally, who is familiar with human anatomy, but not with 2D medical image interpretations.

3.2 Procedure

Our preliminary design consisted of four iterations. Each interview (or co-design workshop) lasted for 30–90 minutes ($M = 53$ min., $SD = 30$ min.). Thematic analysis was used for data interpretation.

[I1] Iteration 1 – We first conducted three remote semi-structured interviews with all participants [60]. We explained basic VR concepts by showing example VR based medical image visualization applications, to help them brainstorm conceptually. We focused on two guiding questions: **(i)** “*What are the benefits that VR might*

bring to today's contouring?” and **(ii)** “*What are the features that might be useful to augment during contouring procedure?*”

[I2] Iteration 2 – We then built a simple prototype based on [I1], where a tracked tablet and Logitech VR stylus was used as the input tool (Fig. 5). Participants could draw contours on a cutting-plane rendered on a tracked tablet using a 2D brush metaphor (Fig. 4a – d), which is consistent with today's 2D contouring (Fig. 1b – d). We also designed a 3D brush metaphor by extending the existing 2D brush, which is essentially a semi-transparent sphere attached to the tip of the VR pen. Participants could then directly draw a colored volume inside 3D medical structures using a 3D brush (Fig. 4e–f). All control menus were attached to the tracked tablet. Section 4 outlines the details of our design. P2 and P3 were invited to a co-design workshop where they could try out the system; P1 was excluded from this second iteration due to the lack of familiarity with practical contouring procedures.

[I3] Iteration 3 – We improved the system design based on lessons learnt from [I2], and organized remote interviews with P2 and P3. The goal of the session was to understand the feasibility and setbacks of the provisional design.

[I4] Iteration 4 – We finally invited P2 to experience the enhanced system. We also discussed extensively with P2 to ensure domain experts participants could inspect the target structure in both 2D and 3D through our rendering algorithm. P3 was not involved during [I4] due to clinical duties.

3.3 Findings

Fig. 2 shows key takeaways during each iteration. [N#] and [D#] denote *the motivation & Needs analysis*, reflecting experts' thoughts of how 3D contouring might be beneficial to future oncology diagnosis and educations, as well as suggestions on *interaction and interface Design*. The lessons learnt are summarized into three areas:

Inspecting Medical Images in 3D – All participants recognized the challenges of learning how to analyze 2D medical images, especially for small organs (“*things like vasculature can be harder to identify.*” [P1]). Participants explained the approach to locate and understand a specific target from a set of medical images: “*You get a stack [of planar medical images]. And then you can go [back and forth], and change the different [cutting plane] to see where you are. [...] We don't get to see only one image. You can go up and down and then kind of orient yourself in.*” [P1]. Key benefits include the “*overall view*” capabilities for helping determining the boundary between tumors and surrounding healthy tissues — “[...] often you're trying to decide where does that cancer stop? [...] Being able to visualize everything in 3D could help you see those borders better.” [P2]— and having flexibility for placing and orienting cutting-planes for inspecting target as well as inter-slice interpolations [P2, P3].

Annotating & Contouring Medical Images in 3D – All participants, during I2 and I4, agreed that directly drawing in 3D could be helpful for identifying hard tissues (e.g., bones), yet not for the soft ones (e.g., blood vessels). P3 emphasized the importance of preserving the

Figure 3: Design of VRContour. (a) Explorations of the design space by considering the dimension of information input and consumption. (b – i) Final prototype of four conditions generated by our preliminary autobiographical and iterative design process.

features of 2D contouring, yet using 3D as an assistive visualization technique: “*as you’re drawing in 2D, you can see a 3D rendering of a structure [...] that actually could be a really interesting concept and might improve contouring accuracy. To really see the structure, like a prostate cancer, being able to see the course of the rectum, see where the prostate sits, see how the seminal vesicles go over [...] and kind of understand that anatomy, and actually see that 3D, that actually might result in more accurate contouring, especially for people that are just learning.*” P2 further extrapolated the usefulness of mapping 3D annotations back to 2D slices: *[It would be nice to] have like a bright star, and you can put it somewhere. And then when you switch over [to 2D] and start scrolling up and down, [you see] a star on the excellent position where you put it.* [P2]

Learning about 3D UI and Interactions – During [I1], P2 and P3 took on average 10 minutes to get used to the prototype. They believed such learning effort should *not* be problematic for general practicing oncologists (e.g., “*I don’t think it would take that long to get used to pushing the buttons*” [P2]). Participants also emphasized the importance of leveraging the skills of inspecting and contouring medical slices on 2D displays that oncologists already have in today’s practice: “*we work with 2D images, but have to think in 3D. It’s a little bit of paradox.*” [P3], “*No doctor will ever see it as 3D, we’re always scrolling a plane [...] but it’s always on 2D.*” [P2].

3.4 Discussions and Design Space Explorations

We finally defined a design space, leading to three evaluation conditions + a baseline condition for emulating today's 2D contouring (Fig. 3a). The prototypes of the four conditions are demonstrated in Fig. 3b–j. We re-use the concept of information flow as the dimensions of our morphological analysis [51]. Specifically, the *dimension of information consumption* indicate the ways that the users consume the output information from VR (e.g., inspecting scans), and the *dimension of information input* refers to how users outlines contours inside VR (Fig. 3a). While each dimension can be either 2D or 2D + 3D, the dimension for information consumption should not be lower than the one of the information input. Based on [N3] and [D1], we also intentionally excluded any design that would enable contouring *purely* on 3D, as mid-air drawing in VR makes it challenging to be precise due to the lack of physical support and perception of depth [17]. The four conditions are summarized below:

[C1] Contour on a 2D desktop display (Baseline) – We built a prototype to emulate today’s contouring on 2D desktop display using a keyboard and mouse (Fig. 3b+f). Since enhancing the comfort of the VR headset is beyond our scope, we re-created this scenario to establish a baseline, such that the confounding factors introduced by comfort of VR headset could be minimized. While wearing the VR headset, oncologists need to go through each transverse slices to mentally understand the medical structures, followed by outlining key structures on each slices. During [I4], we confirmed with P2 that the prototype could sufficiently emulate today’s 2D contouring.

[C2] Contour on a 2D tablet without visualizations of 3D medical structure – C2 leverages the idea of direct interaction where

the drawn stroke is aligned with the physical tip of the VR pen (Fig. 3c+g). Participants are required to inspect and contour on 2D slices rendered on a tracked tablet, using a VR stylus. No 3D visualizations were provided in this condition.

[C3] Contour on a 2D tablet with visualizations of 3D medical structure – C3 offers a 3D rendered medical structure hovering in the mid-air (Fig. 3d+h). This, hypothetically, could reduce the cognitive task load for thinking in 3D while inspecting different cutting-planes in 2D. As oncologists are not allowed to draw in 3D, we consider the dimensions for information input as 2D only.

[C4] Contour in the 3D medical structure and on tablet – Participants are expected to contour on both 3D and 2D (Fig. 3i+j). Besides the hypothesized benefits in C3, we extrapolated that the direct 3D contour could reduce lead time, due the potential elimination of slice-wise drawing on the 2D interface. After 3D contouring, oncologists could continue refine the outlines with pen and tablet.

4 SYSTEM DESIGN

4.1 Volumetric Rendering

Although investigating optimal rendering techniques is beyond our scope, existing works (e.g., [63]) have shown that there is no single algorithm that could be applied to every 3D task (also confirmed in [D4]). While considering the needs of transparent rendering [D5], we decided on using a ray-casting approach that treats the 3D medical structure as a transparent object that could transmit, emit and absorb light (like a cloud). Although Direct Volume Rendering (DVR) and Maximum Intensity Projection (MIP) are commonly used approach to render transparent volumes, MIP has a well known limitation regarding depth perception (e.g., ambiguous front and back relations) [41, 44]. We thus decided on using DVR. As we focus on understanding the merits that the 3D UI could offer to contouring, and the tuning of rendering parameters is important [D6] yet beyond our scope, we therefore pre-adjusted the rendering parameters for each evaluation task before final evaluation (see Step 1 below and Fig. 7). This process was done during [I4] with P2. The implementations of rendering pipelines include three steps:

Step 1: Data Prepossessing & Sampling – We represent the 3D medical structure using a 3D array, by stacking all DICOM slices. To emulate practical contouring, we retain the original resolution of each slices without either upsampling or downsampling. The physical space between voxels over each dimensions are determined by the scanning machine. We used min–max normalization to scale the density at each voxels to $[0, 1]$ and the resultant view of 3D visualizations are called “*full density range view*”. During [I4], we adjusted such range (*i.e.*, vary the lower and upper density bound) to make sure that the target structures are visible (*i.e.*, “*fine density range view*”, see also Fig. 7). The ray, cast from camera, would then sample the volume into 256 equal–distance steps that were chosen based on visual effects and system performance.

Step 2: Classification – Classification refers to the step of using transfer functions to assign each discrete step with color and opacity.

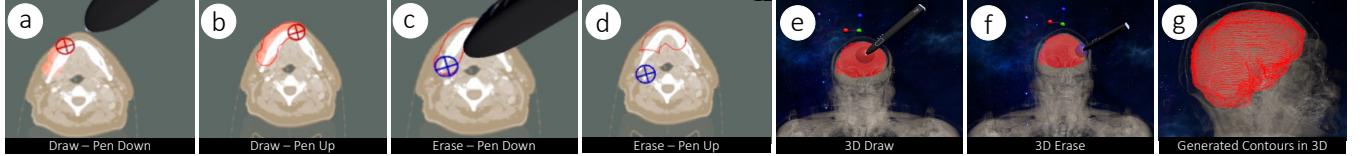


Figure 4: Drawing/erasing contours using brush metaphors. (a–b) The contour is automatically generated after the VR pen is lifted from the tablet; (c–d) the new contour is re-generated upon the completion of erasing; (e–f) contours, visualized as a volume, could also be drawn/erased directly inside 3D volume; (g) generated contours in 3D (side view).

For simplicity, we only consider the simplest 1D transfer function to map the normalized (or adjusted) density to a RGBA vector.

Step 3: Composition – Finally, we aim to generate a composited value by blending the RGBA at each step along the ray, which would be eventually rendered on the 2D projection. We used the *over* operator [48] to compute the final RGBA vector. The algorithm for front-to-back traversal could be referred to equations (1–2), where \mathbf{c}_i and α_i refer to the color and opacity at i^{th} step ($i \in \{1, 2, 3, \dots, 256\}$).

$$\mathbf{c}_i = \mathbf{c}_{i-1} + \mathbf{c}_i \alpha_i (1 - \alpha_i) \quad (1)$$

$$\alpha_i = \alpha_{i-1} + \alpha_i (1 - \alpha_i) \quad (2)$$

More details in terms of volumetric rendering techniques are outlined by Lichtenbelt *et al.* [39].

4.2 Interaction Techniques

Pen + Tablet Interactions and Supporting Precise Contouring – We decided to use the Logitech VR stylus [7] as input tool in both 3D and 2D contouring for its high precision, which meets the need for precision drawing in a virtual environment [17] (Fig. 5a). We used a tracked physical board, with same size as a 10.2" iPad, since having a physical support should increase the drawing precision [17] (Fig. 5a–b). Participants could scale the size of the virtual tablet to inspect the image at higher resolution (Fig. 5c–d), in which case the boundary of the physical tablet are highlighted to help visually identify the drawing area (Fig. 5d–e).

Cutting–Planes and Cross–Dimension Contouring – We found that participants enjoyed the capabilities of viewing and contouring in 3D, yet they still wanted to preserve the capability of contouring on the 2D surface, due to its merits of precision and ease of observing detailed structures. Additionally, 2D contouring leverages existing skills of 2D medical images interpretation ([N3], [D1]). Therefore, while designing C4, we proposed a cross–dimension contour workflow, by combining fast and direct contouring in 3D and precise 2D contouring on a tracked 2D tablet. This process includes three aspects: (a) Upon finding the target, oncologists first perform contouring *directly* in the 3D volume using a 3D brush (as the design in [I2], see Sec. 3.2 and Fig. 4e–f). The brush–based contour would result in a 3D colored volume. (b) Oncologists would

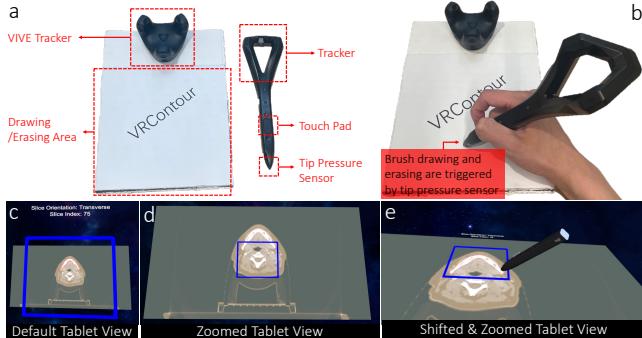


Figure 5: Contour on 2D. (a) Tracked tablet and VR pen; (b) real-world scene using VR pen and a tablet; (c) a blue frame indicating the physical boundary of the tablet; (d) the cutting planes could be zoomed in for gaining a detailed view; (e) oncologists can only draw inside the blue frame due to the physical surface support.

then select a cutting–plane with one of three orientations (Fig. 1e). We did not include an arbitrary orientation, as understanding cross–sectional slices with arbitrary orientation is challenging for domain experts [D3]. The contours on each slices were generated based on the selected slicer orientation (Fig. 4g). (c) Oncologists would finally inspect and revise contours mirrored on the tablet (Fig. 4a–d). The inter–slice interpolation will be applied as needed to reduce the lead time (see Sec. 5.5).

Transforming 3D Medical Structures – During C3 and C4, participants were required to see the 3D medical structures while contouring, as well as to draw directly inside the 3D volume. It is therefore critical to design efficient ways for manipulating 3D medical structures inside VR. To address this, we revisited and revised the design techniques of workspace translation and scaling in SymbiosisSketch [16], which was evaluated among six design professionals. After initiating the transforming mode, two rays are used to determine the new position where the 3D volume will be placed. The new position is decided by the midpoint of the line segment whose length is the shortest between two rays (Fig. 6a). To avoid visual clutter, a 3D cursor is used to indicate the center of the newly placed 3D volume. A second touchpad click allows the 3D volume to be moved into the target place (Fig. 6b). Oncologists can rotate the volume by clicking–and–dragging the near–hand anchor (Fig. 6c). The quaternion of the near–hand anchor is then mapped to the 3D medical structure.

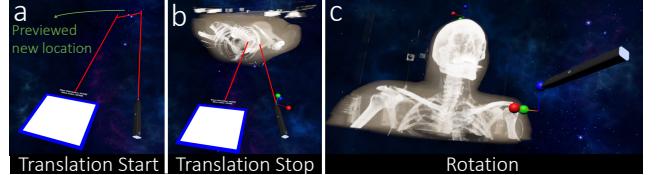


Figure 6: Transforming 3D medical structure. (a) Two rays from the VR pen and tablet are initiated after translation start; (b) after the translation stop, the 3D medical structures are anchored to the new place; (c) during rotation mode, the quaternion of the near–hand anchor is projected onto the 3D medical structure.

Support of Coarse– and Fine–grained Tuning of Parameters – Finding optimal parameters has been identified as a critical factor that affects contour delineations (*e.g.*, locating the slices, changing the tablet and brush size, see [D7], [D8]). For coarsely tuning parameters, oncologists can use the scrollbar shown in front of them. They can also click on the pen’s touchpad for fine–grained parameter adjustments. Although we enabled both tuning methods for all parameters during [I2], oncologists suggested only adding fine–grained tuning to the *slice selections* due to its frequent uses during contouring.

Spatially Distributed Control Panels – While only retaining essential features of general computer–supported contouring applications, we found that it was challenging for participants to promptly locate target UI elements if aggregated onto only one panel [D2]. Therefore, we decided to dispatch UI elements onto three sub–control panels for C2 – C4. Each panel has elements with similar features: *tablet and volume control*, *stylus control*, and *interpolation control*. Through this design, oncologists are able to re–organize the UI layout and/or hide specific panels as needed. Participants were satisfied with these design enhancement during [I4].

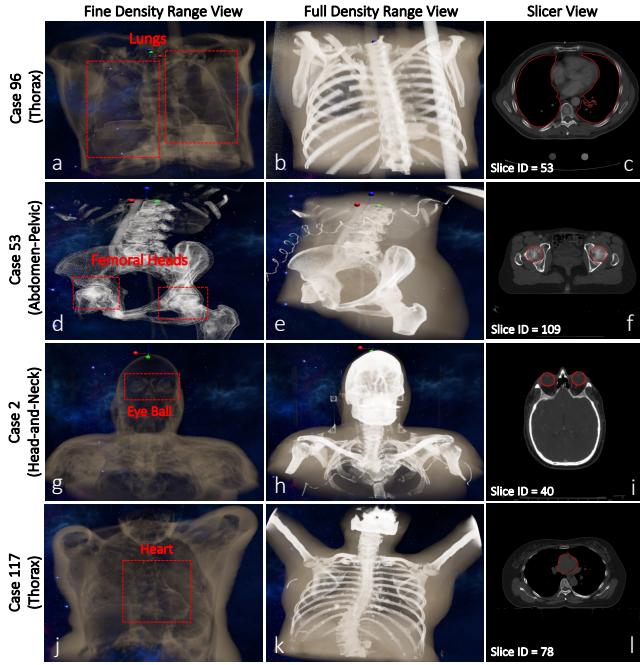


Figure 7: Contour tasks: (a – c) lungs from the thorax structure; (d – f) femoral heads from the abdomen–pelvic structure; (g – i) eye balls from the head–neck structure; (j – l) heart from the thorax structure. Red contours in c, f, i, and l indicate expert contours. Displayed data are extracted from eContour [10], cases 96, 53, 2, and 117.

5 STUDY METHODS

5.1 Participants

We recruited 8 participants (2 females and 6 males, age: $M = 25.6$, $SD = 1.2$) who have taken human anatomy related curricula, but who do *not* yet had formal training on 2D medical image interpretation. Our rationale behind this decision is based on two factors; *First*, since we are evaluating the affordances of a specific modality, extensive past learning experience of 2D medical image interpretation could affect UX evaluations and system preferences; *Second*, participants taking human anatomy classes are able to identify and contour basic structures, even though clinical contouring is performed by senior residents and attending physicians. To accommodate participants’ anatomy knowledge, we intentionally selected fundamental structures with domain experts and made sure that the task was feasible for any students who have taken anatomy classes. During our pre-study phase (see Fig. 8), participants overall *agreed* on a 5-point Likert scale that they were able to identify key structures from 3D anatomy ($MED = 4$), yet *disagreed* that they could interpret 2D medical images ($MED = 2.5$). Any participant that was involved with our design work (Sec. 3) was excluded from this evaluation.

5.2 Contouring Tasks

We decided to use the dataset from eContour [52], which contains 119 anonymized cases with expert contouring guidance. Since investigating 3D rendering is out of our scope, we decided on four structures from four different patients with oncology educators that could be well visualized with our rendering pipeline (see Sec. 4.1). Fig. 7 demonstrates the 3D views of the selected tasks (with fine and full density ranges), along with one example transverse cutting-plane, as well as an expert contour (noted by the red outlines).

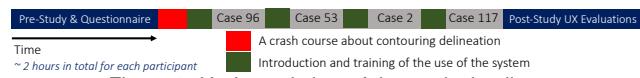


Figure 8: Horizontal view of the study timeline.

5.3 Procedures

We structured our within-subject study into three phases (see Fig. 8).

Phase 1: Pre-Study Orientations & Questionnaire – Participants were first asked to complete a pre-study questionnaire, aiming to collect demographic information, VR experience, and the skills for identifying key structures in 3D anatomy and from 2D scans. Participants were then introduced to the concept of contouring and inter-slice interpolation.

Phase 2: VR Data Collection – We aim to evaluate participants’ contouring performance of four conditions (see Fig. 3). To avoid the confounding factors of learning experience, a Latin square design [26] was used to counterbalance the order of conditions. This means that the contouring task (see Fig. 7) performed by each participants using a specific condition would be varied and balanced across participants. Participants were given opportunities before each session to familiarize themselves with the system.

Phase 3: Post-Study UX Evaluations – Participants were finally asked to fill out a questionnaire, and think-aloud while recording their responses. System Usability Scale (SUS) [21] and NASA TLX [35] were used to evaluate their perceived usability and task load. Participants were also asked to rate the usability of using all four conditions as tools to *learn* contouring, on a 5-point Likert scale. Brief semi-structured interviews were also conducted focusing on the merits and setbacks of seeing and directly drawing inside 3D medical structures.

5.4 Quantitative Measures

Consensus Measurement – To evaluate the quality of participants’ contour, we compute the Dice Similarity Coefficient (DSC) [64] between participants’ and expert’s contours. Equation (3) shows the computations of DSC, where X_i and Y_i refer to the contour masks of expert’s and user’s contour at i^{th} slice.

$$DSC = \frac{2 \sum_{i=0}^n |X_i \cap Y_i|}{\sum_{i=0}^n |X_i| + \sum_{i=0}^n |Y_i|} \quad (3)$$

Temporal Domain Evaluations – We measure the time for initial explorations and overall Task–Completion–Time (TCT). The *time for initial exploration* refers to the duration *after* the spatial UI being anchored and *before* the first stroke being drawn. During this step, participants need to identify and localize the majority of the target. On the other hand, the *overall TCT* refers to the duration that participants take for completing the entire contouring procedures.

Gaze Analysis – We analyzed the gaze direction to analyze *where did participant focus on during a particular fixation instant?*. We first filter the frames with the gaze angular speed $\geq 150^\circ/\text{s}$, which are usually considered as saccade [20]. We then compute the % of attention that participants allocated during each sliding window on *tablet* (2D) and *volume* (3D), quantified by the % of frames when gaze collides with tablets and volume, respectively. Since the overall TCT varies among individuals, we used the % of the completion progress as the independent variable. For simplicity, 1% of overall TCT is used as the width of sliding window, and the % of attention was computed upon every unit increment of task completion progress. Gaze analysis are only used to analyze C3 and C4, as only 2D interfaces are included in C1 and C2.

5.5 Implementation

We implemented all testbeds on HTC VIVE Pro Eye [6] using Unity (v2020.3.6f1) and SteamVR (v1.21.2). Tobii XR SDK (v3.0.1.179) [12] was used to track gaze direction. The rendering ran on a gaming laptop (32GB RAM, 12-core CPU, and a RTX 3080 GPU). Inter-slice interpolation was also integrated [15, 66]. For C1, we used a keyboard [5] and mouse [8] as the input device, to emulate contouring on today’s 2D display. For C2 – C4, we used a Logitech VR Pen [7] and a tracked tablet as the input device (see Sec. 4.2). we

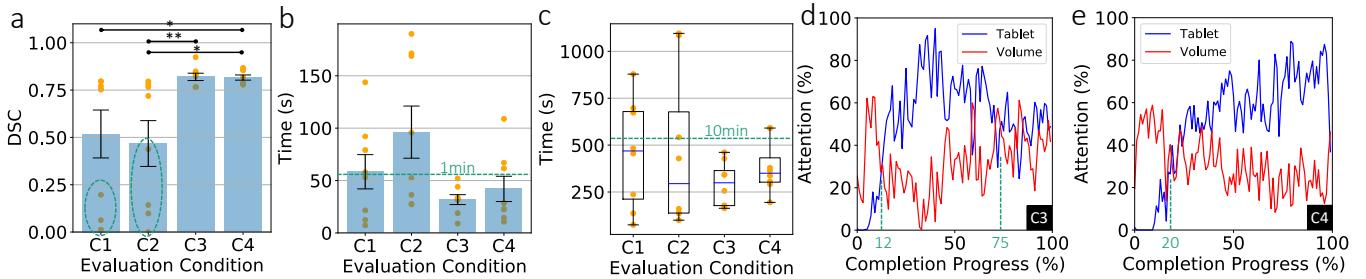


Figure 9: Quantitative results ($* = p < .05$, $** = p < .01$, $*** = p < .001$) for DSC (a), time for initial explorations (b), and overall TCT (c) measured on four conditions; standard error is used to visualize the error bar and the overlaid yellow scatters indicate individual measurements. Panels (d) and (e) show the average of attention (%) that participants allocated to tablet and 3D volume while contouring using C3 and C4 respectively.

used three SteamVR Base Station (v2.0) [2] for tracking purposes. As investigations of tracking performance is out of our scope, we have adjusted the placements of base station and workspace and ensure that the pen and tablet are tracked during at least 99% of the frames (see Appendix A). We implemented a vertex and fragment shader based on [9], for rendering cutting-planes and 3D volumes. A compute shader was implemented for coloring each pixel/voxel. For C3 and C4, participants’ drawn contours would be mirrored on 3D medical structure while drawing on tablet, and vice versa.

6 RESULTS

6.1 Quantitative Measurements

Figure 9 shows quantitative results. Repeated Measures Analysis of Variance (RM-ANOVA) were used to analyze various measures ($\alpha = .05$). We used Aligned Rank Transformation (ART) [58] for those that failed to pass the normality check, and ART procedures for multi-factor contrast test for post-hoc test, with Bonferroni corrections [32].

Overall DSC – We first show that changing the evaluation condition (C1, C2, C3, or C4) might change the overall Dice Similarity Coefficient (DSC) ($F_{3,28} = 5.79$, $p = .005$, $\eta^2 = 0.45$). The subsequent post-hoc test demonstrates a statistical significance between C1 and C4 ($p = .04$), C2 and C4 ($p = .04$), as well as C2 and C3 ($p = .02$). By looking at the overlaid scatter plot in Fig. 9a, we observed that three and four participants failed to locate the target structure in C1 and C2 (*without* 3D visualizations), leading to $< 50\%$ of DSC ($M_1 = 51.80\%$, $M_2 = 46.76\%$), and higher standard deviations compared to their counterpart ($SD_1 = 13.65\%$, $SD_2 = 12.10\%$ vs. $SD_3 = 1.93\%$, $SD_4 = 1.37\%$). All participants could successfully identify the majority of the targets with C3 and C4 (*with* 3D visualization), leading to $> 75\%$ of DSC ($M_3 = 81.98\%$, $M_4 = 81.61\%$). It is worth mentioning that the DSC measured from our study should not be considered as a reference for clinical purposes (*e.g.*, usually $> 80\%$ of DSCs were measured in contouring of cervical cancer [55]), as all participants are pre-med or junior MD students without practical contouring experience. Additionally, tracking inaccuracies could also hinder the consensus measurements.

Time for Initial Exploration – RM-ANOVA ($F_{3,28} = 3.03$, $p = .04$, $\eta^2 = 0.25$) shows a marginal p -value, but large effect size, implying that our results show an effect, but are likely limited by the small sample size. No statistical significance was detected during post-hoc test. All participants located the target using C3 within 1 min.

Overall TCT – No statistical significance was detected when it comes to overall Task-Completion-Time (TCT). To illustrate our results we therefore visualize the boxplot shown in Fig. 9c. Results indicate all participants were able to finish the instructed contour tasks within 10 min with C3. While it seems that the capability of direct contouring in 3D could accelerate the overall task by minimizing the procedures of delineating outlines, more time was spent for refining the outlines, compensating for the merits brought by the novel affordance given by direct 3D drawing.

Gaze Analysis – Fig. 9d–e approximate the average of % of attention that participants allocated on tablet (2D) and volume (3D) while progressing towards task completion. With C3, participants heavily relied on the 3D volume for initial explorations (\sim first 12% of progress). When task completion progress reached more than 75%, both tablet and 3D volume were used for finalizing the contouring. With C4, participants mainly used the 3D volume during initial explorations and 3D contours (\sim first 20%).

6.2 Post-Study UX Evaluations

We report here on participants’ responses to SUS (Fig. 10a), task load (Fig. 10b–g), and level of agreement of *usefulness as a learning tool* (Fig. 10h). We used Sauro’s [3] and Hart *et al.*’s [35] recommendations to interpret SUS and task load scores respectively. Overall, participants rated C3 (*with* 3D visualization, *without* 3D contouring) as the most usable prototype ($MED_3 = 87.5$ vs. $MED_4 = 78.8$, $MED_2 = 72.5$, $MED_1 = 71.3$). “Excellent” is considered for SUS ≥ 80.8 [3, 21]. Four takeaways are generated through thematic analysis:

Seeing 3D Medical Structures – All participants appreciated the merits of seeing the 3D structure inside VR, and think “*the 3D was very useful*” [P5]. With the 3D visualization (C3 and C4), more participants reported less mental load (Fig. 10b) and higher performance (Fig. 10e). First, participants felt that 3D visualizations are useful for understanding the structure (“*when I just look at the slides, I’m not sure if that’s the target or not, but if I see it in 3D, I’m more confident*.” [P3]), reduce mental load (“*It helps to have a bigger picture, and it takes off that cognitive load to actually imagine the 3D representation, which is a very high load*.” [P2]), and locating targets [P5]. P7 pointed out that the 3D visualization addresses one of the hardest challenges in anatomy class: “*I think that the hardest part of anatomy is trying to look at pictures on a 2D surface, and then make them 3D, which is like exactly what this is accomplishing*” [P7]. P3 and P7 emphasized that the 3D medical structure helped them tremendously on deciding the boundary of the target structures (*e.g.*, heart and lungs). Some participants appreciated the transparent volumetric rendering: “*we use an app called Complete Anatomy, but I like [the transparent rendering] a lot better. You can see more detail and get depth. It’s really hard to get depth for non-transparent rendering*.” [P7]. Second, with 3D structures, participants could better leverage the inter-slice interpolation by deciding the slice position where the rapid structure changes occurred (*e.g.*, “*it will help in locating the cutting-plane*.” [P5], “*you can see the plane that you put in your virtual cadaver, and it helps feel where the rapid changes occur*.” [P7]).

Drawing Contours in 3D – Participants appreciated the capabilities of direct drawing in 3D while attempting to locate the target (*e.g.*, “*A huge part of the work gets done way faster*” [P2], “*I like the mapping that map my annotation in 3D to 2D*” [P6]). Some participants also appreciated the precision offered by the VR pen (“*[the pen] is intuitive for precise hand control*” [P3]). However, we also observed participants expressing slightly higher frustration by comparing C3

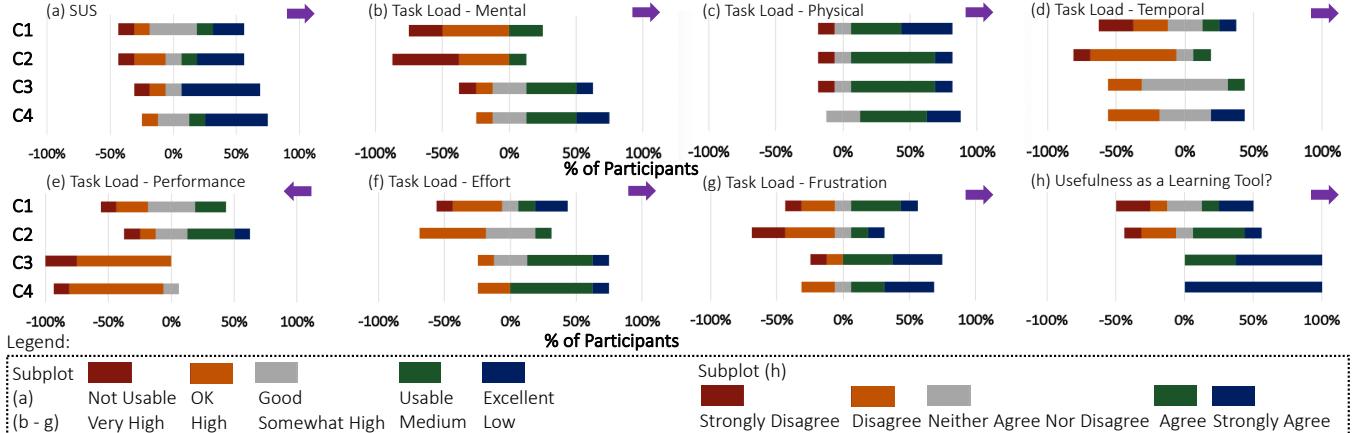


Figure 10: Responses to post-study surveys. Purple arrow (top right) indicates the direction where a *good* system should be designed toward.

and C4 (*MED*: 12.5% vs. 20%); potential reasons include difficulties of drawing precisely in 3D (e.g., *“It was very hard for me to be precise when drawing in 3D, I didn’t see where I was starting”* [P2]), lack of capabilities for inspecting inside 3D volumes (e.g., *“You cannot look inside the body. When you look inside the body, you cannot see very clearly where you are. [...] I think a better thing is as your pen is approaching, in the 3D model, you’re looking at a slice of the 3D model, so that you know exactly where you are”* [P3]), and small size (e.g., *“if you could blow it up more, to be able to get finer precision, that would make it a lot easier”* [P7]).

Learning Cost – Echoing [N2] and [N3], some participants also mentioned that *“medical school students mostly never used VR”* [P3]. While trained to use the system before each session, all participants found it not hard to learn and get familiar with system operations (e.g., *“a week for learning probably would be sufficient”* [P8]). This echoes the fact that seven participants *agree* or *strongly agree* with the statement that *“I would imagine that most people would learn to use this system very quickly”* in SUS responses.

Contouring Education – All participants *agreed* or *strongly agreed* that seeing and drawing in 3D is extremely helpful for learning contouring (e.g., *“if you could put this into like, even our anatomy labs, that’d be amazing”* [P7]). While comparing the experience of using C3 and C4, P1 outlined additional setbacks of learning anatomy with cadavers, including the unpleasant smell of formaldehyde, the mental stress triggered by a real body, and the reduced accessibility and resulting practicing opportunities. P7 appreciated the transparent rendering while comparing it with *Complete Anatomy* [42], a mobile application designed for anatomy learning purpose, because of the capability of *“seeing more details and getting depth information”*.

7 DISCUSSION AND FUTURE WORK

7.1 Seeing and Contouring in 3D

Our results show that there is great potential for radiation oncology and contouring when integrating 3D visualizations into existing contouring workflows through VR. While our focus is on RT treatment planning, many findings are more general and could be transferred to other similar applications that require annotations on 3D volumes (e.g., annotations during surgical planning [34]).

Since some of our participants felt that it was less intuitive and overall harder to directly draw in 3D – mostly because of lack of depth perception and fine grained control – one natural follow up would be to *investigate the design of novel interaction techniques where participants can draw precisely in 3D in these contexts*.

Another direction that future work might consider is the inherent challenges raised by interacting with 3D medical structures. While the metaphor of pen + tablet could offer precise drawing, some participants believe that they are less intuitive when it comes to 3D

object manipulations. Future investigations of interactions with 3D volume might be needed to uncover more insights.

7.2 Re-thinking the Contouring Training Curriculum

Contouring is a challenging skill to master, and it is typically trained during residency programs. While interpreting 2D scans is usually difficult, it is a necessary prerequisite for learning contouring. However, Dmytriw *et al.* [27] surprisingly showed that < 10% of students are *very* or *completely* confident in interpreting basic radiology images. Although additional training during residency program might address such gaps, many residents might not have that opportunity, especially in less developed areas [30, 37, 54]. With VRContour we unveil an alternative solution – bringing the contouring workflow into a VR-based 3D UI – that could possibly help radiation oncologists (and others) to massively scale contouring education.

Although participants recognized that getting familiar with VR-Contour is not challenging, we also found that half participants did not have extensive experience on using VR (rated as *“only tried it out once or twice before”* in our pre-study questionnaire), and exhibited less experience with exploring 3D user interfaces. One future direction is to rethink and modernize today’s oncology educations by fully integrating VR as part of future curricula.

7.3 Limitations

The limitation of our work can be summarized into three areas. *First*, our results are based on 8 participants, and the small sample size together with the lack of background/expertise/preference diversity might affect some results. While we only considered participants without extensive experience in interpreting 2D scans to minimize the confounding factors of learning experience, such participants will also not have prior contouring experience, limiting the application of our results to different levels of contouring proficiency. *Second*, since examining participants’ anatomy knowledge was beyond our scope, our study was based on only four cases extracted from eContour [52] during [I4], with the goal to have baseline scenarios that would be familiar for any students with basic anatomy training. However, practical contouring usually involves tumors and finer-grained anatomy, which might require prior experience on oncology and pathology, which would have been challenging for current participants. Future work might introduce more realistic tasks, and include residents and attending physicians as part of the evaluation. *Finally*, we intentionally excluded procedures for adjusting volumetric rendering parameters from our study. Instead, we pre-tuned parameters and offered *full* and *fine* density range view for participants to inspect the target and overall view of the 3D volume. However, adjusting such parameters is an inevitable step in practical contouring, and it might be considered in future work.

8 CONCLUSION

We presented *VRContour*, and investigated how to bring radiation oncology contouring into VR. We explored and defined three design spaces by considering the dimensionality for consuming and inputting information. In a within – subject study (n = 8), we found that the inclusion of a 3D visualization could increase contour precision and reduce mental load, frustration, and overall effort, as well as introduce important benefits for learning and educational purposes. We believe that our work will benefit future research in designing VR-based medical software tools for radiation oncology and beyond.

REFERENCES

- [1] DICOM VR: A New Way of Viewing Volumetric Medical Imaging and Planning Targeted Radiation Treatment in Virtual Reality. <http://www.dicomvr.com>. Accessed: 08-08-2022.
- [2] SteamVR Base Station 2.0. <https://www.vive.com/us/accessory/base-station2/>. Accessed: 08-09-2022.
- [3] 5 Ways to Interpret a SUS Score. <https://measuringu.com/interpret-sus-score>, 2018. Accessed: 08-07-2022.
- [4] Eclipse: Fast, Precise Planning for Advanced Cancer Care. <https://www.varian.com/products/radiotherapy/treatment-planning/eclipse>, 2021. Accessed: 08-08-2022.
- [5] Hiwings Mechanical Gaming Keyboard. https://www.amazon.com/gp/product/B095SDPH6V/ref=ppx_yo_dt_b_asin_title_o06_s00?ie=UTF8&psc=1, 2021. Accessed: 08-07-2022.
- [6] HTC VIVE Pro. <https://www.vive.com/us/product/vive-pro>, 2021. Accessed: 08-07-2022.
- [7] Logitech VR Ink. <https://www.logitech.com/en-us/promo/vr-ink.html>, 2021. Accessed: 08-07-2022.
- [8] Logitech Wired Mouse. <https://www.amazon.com/Logitech-Wired-Mouse-M90-Black/dp/B003D8ZT0C>, 2021. Accessed: 08-07-2022.
- [9] Unity Volumetric Rendering. <https://github.com/mlavik1/UnityVolumeRendering>, 2021. Accessed: 08-07-2022.
- [10] eContour: Interactive Contouring Guidelines for the Busy Radiation Oncology Professional. <https://econtour.org>, 2022. Accessed: 08-07-2022.
- [11] Elucis – The Future of Medical Modelling. <https://www.realizemed.com/elucis>, 2022. Accessed: 08-07-2022.
- [12] HTC VIVE Pro Eye Development Guide. <https://vr.tobii.com/sdk/develop/unity/getting-started/vive-pro-eye>, 2022. Accessed: 08-08-2022.
- [13] The Benefits of Virtual Reality in Healthcare. <https://www.healthcare.digital/single-post/2017/11/02/the-benefits-of-virtual-reality-in-healthcare>, 2022. Accessed: 08-06-2022.
- [14] R. A. Abrams, K. A. Winter, W. F. Regine, H. Safran, J. P. Hoffman, R. Lustig, A. A. Konski, A. B. Benson, J. S. Macdonald, T. A. Rich, and C. G. Willett. Failure to Adhere to Protocol Specified Radiation Therapy Guidelines Was Associated With Decreased Survival in RTOG 9704 – A Phase III Trial of Adjuvant Chemotherapy and Chemoradiotherapy for Patients with Resected Adenocarcinoma of the Pancreas. *International Journal of Radiation Oncology Biology Physics*, 82(2):809–816, 2012. doi: 10.1016/j.ijrobp.2010.11.039
- [15] A. B. Albu, T. Beugeling, and D. Laurendeau. A Morphology-Based Approach for Interslice Interpolation of Anatomical Slices From Volumetric Images. *IEEE Transactions on Biomedical Engineering*, 55(8):2022–2038, 2008. doi: 10.1109/tbme.2008.921158
- [16] R. Arora, R. Habib Kazi, T. Grossman, G. Fitzmaurice, and K. Singh. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, pp. 1–15. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173759
- [17] R. Arora, R. H. Kazi, F. Anderson, T. Grossman, K. Singh, and G. Fitzmaurice. Experimental Evaluation of Sketching on Surfaces in VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI ’17, pp. 5643–5654. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025474
- [18] S. Balasubramanian. The Next Frontier For Healthcare: Augmented Reality, Virtual Reality, And The Metaverse, Dec 2021. Accessed: 08-07-2022.
- [19] A. Boejen and C. Grau. Virtual Reality in Radiation Therapy Training. *Surgical Oncology*, 20(3):185–188, 2011. Special Issue: Education for Cancer Surgeons. doi: 10.1016/j.suronc.2010.07.004
- [20] E. Britannica. Sensory Reception: Human Vision: Structure and Function of the Human Eye. *Encyclopedia Britannica*, 27:179, 1987.
- [21] J. Brooke. SUS – A Quick and Dirty Usability Scale. *Usability Evaluation in Industry*, 189(194):4–7, 1996.
- [22] A. Cannavò, D. Calandra, A. Kehoe, and F. Lamberti. Evaluating Consumer Interaction Interfaces for 3D Sketching in Virtual Reality. In *International Conference on ArtsIT, Interactivity and Game Creation*, pp. 291–306. Springer, 2021. doi: 10.1007/978-3-030-73426-8-17
- [23] C. Chen, K. Sun, and X. Zhang. ExGsense: Toward Facial Gesture Sensing with a Sparse Near-Eye Sensor Array. In *Proceedings of the 20th International Conference on Information Processing in Sensor Networks (Co-Located with CPS-IoT Week 2021)*, IPSN ’21, pp. 222–237. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3412382.3458268
- [24] C. Chen, M. Yarmand, V. Singh, M. V. Sherer, J. D. Murphy, Y. Zhang, and N. Weibel. Exploring Needs and Design Opportunities for Virtual Reality-based Contour Delineations of Medical Structures. In *Companion of the 2022 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS ’22 Companion, pp. 19–25. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3531706.3536456
- [25] C. Chen, M. Yarmand, Z. Xu, V. Singh, Y. Zhang, and N. Weibel. 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)*, 2022.
- [26] F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media, 2005.
- [27] A. A. Dmytriw, P. S. Mok, N. Gorelik, J. Kavanaugh, and P. Brown. Radiology in the Undergraduate Medical Curriculum: Too Little, Too Late? *Medical Science Educator*, 25(3):223–227, 2015. doi: 10.1007/s40670-015-0130-x
- [28] R. J. Dowsett, J. M. Galvin, E. Cheng, R. Smith, R. Epperson, R. Harris, G. Henze, M. Needham, R. Payne, M. A. Peterson, A. L. Skinner, and A. Reynolds. Contouring Structures for 3-Dimensional Treatment Planning. *International Journal of Radiation Oncology – Biology – Physics (IJROBP)*, 22(5):1083–1088, 1992. doi: 10.1016/0360-3016(92)90812-v
- [29] E. Dühmke, J. Franklin, M. Pfreundschuh, S. Sehlen, N. Willich, U. Rühl, R.-P. Müller, P. Lukas, A. Atzinger, U. Paulus, B. Lathan, U. Rüffer, M. Sieber, J. Wolf, A. Engert, A. Georgii, S. Staar, R. Herrmann, M. Beykirch, H. Kirchner, A. Emminger, R. Greil, E. Fritsch, P. Koch, A. Drochert, O. Brosteanu, D. Hasenclever, M. Loeffler, and V. Diehl. Low-Dose Radiation Is Sufficient for the Noninvolved Extended-Field Treatment in Favorable Early-Stage Hodgkin’s Disease: Long-Term Results of a Randomized Trial of Radiotherapy Alone. *Journal of Clinical Oncology*, 19(11):2905–2914, 2001. doi: 10.1200/jco.2001.19.11.2905
- [30] P. Eansor, M. E. Norris, L. A. D’Souza, G. S. Bauman, Z. Kassam, E. Leung, A. C. Nichols, M. Sharma, K. Y. Tay, V. Velker, A. Warner, K. E. Willmore, D. A. Palma, and N. Campbell. Development, Implementation, and Initial Participant Feedback of an Online Anatomy and Radiology Contouring Bootcamp in Radiation Oncology. *Journal of Medical Education and Curricular Development*, 8:1 – 9, 2021. doi: 10.1177/23821205211037756
- [31] A. Eisbruch, J. Harris, A. S. Garden, C. K. S. Chao, W. Straube, P. M. Harari, G. Sanguineti, C. U. Jones, W. R. Bosch, and K. K. Ang. Multi-Institutional Trial of Accelerated Hypofractionated Intensity-Modulated Radiation Therapy for Early-Stage Oropharyngeal Cancer (RTOG 00-22). *International Journal of Radiation Oncology – Biology – Physics (IJROBP)*, 76(5):1333–1338, 2010. doi: 10.1016/j.ijrobp.2009.04.011

- [32] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, p. 754–768. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474784
- [33] G. W. Fitzmaurice, R. Balakrishnan, G. Kurtenbach, and B. Buxton. An Exploration into Supporting Artwork Orientation in the User Interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, p. 167–174. Association for Computing Machinery, New York, NY, USA, 1999. doi: 10.1145/302979.303033
- [34] D. Gasques, J. G. Johnson, T. Sharkey, Y. Feng, R. Wang, Z. R. Xu, E. Zavala, Y. Zhang, W. Xie, X. Zhang, K. Davis, M. Yip, and N. Weibel. ARTEMIS: A Collaborative Mixed-Reality System for Immersive Surgical Telementoring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445576
- [35] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock and N. Meshkati, eds., *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139–183. North-Holland, 1988. doi: 10.1016/S0166-4115(08)62386-9
- [36] L. He, A. Guayaquil-Sosa, and T. McGraw. Medical image atlas interaction in virtual reality. In *Immersive Analytics Workshop*, 2017.
- [37] N. F. Holt. Medical Students Need More Radiology Education. *Academic Medicine*, 76(1):1, 2001.
- [38] W. Jin, B. Birkhead, B. Perez, and S. Hoffe. Augmented and Virtual Reality: Exploring a Future Role in Radiation Oncology Education and Training. *Applied Radiation Oncology*, 6:13–20, 2017.
- [39] B. Lichtenbelt, R. Crane, and S. Naqvi. *Introduction to Volume Rendering*. Prentice-Hall, Inc., USA, 1998.
- [40] M. Ma, P. Fallavollita, I. Seelbach, A. M. Von Der Heide, E. Euler, J. Waschke, and N. Navab. Personalized Augmented Reality for Anatomy Education. *Clinical Anatomy*, 29(4):446–453, 2016. doi: 10.1002/ca.22675
- [41] A. S. Mady and S. A. El-Seoud. An Overview of Volume Rendering Techniques for Medical Imaging. 05 2020.
- [42] S. K. Motsinger. Complete Anatomy. *Journal of the Medical Library Association*, 108(1):155–157, 2020. doi: 10.5195/jmla.2019.853
- [43] C. Neustaeder and P. Sengers. Autobiographical Design in HCI Research: Designing and Learning through Use-It-Yourself. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, p. 514–523. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2317956.2318034
- [44] D. R. Ney, E. K. Fishman, D. Magid, and R. A. Drebin. Volumetric Rendering of Computed Tomography Data: Principles and Techniques. *IEEE Computer Graphics and Applications*, 10(2):24–32, 1990. doi: 10.1109/38.50670
- [45] D. T. Nicholson, C. Chalk, W. R. J. Funnell, and S. J. Daniel. Can Virtual Reality Improve Anatomy Education? A Randomised Controlled Study of A Computer-Generated Three-Dimensional Anatomical Ear Model. *Medical Education*, 40(11):1081–1087, 2006. doi: 10.1111/j.1365-2929.2006.02611.x
- [46] D. Patel, L. P. Muren, A. Mehus, Y. Kvinnslund, D. M. Ulvang, and K. P. Villanger. A Virtual Reality Solution for Evaluation of Radiotherapy Plans. *Radiotherapy and Oncology*, 82(2):218–221, 2007. doi: 10.1016/j.radonc.2006.11.024
- [47] R. Phillips, J. W. Ward, L. Page, C. Grau, A. Bojen, J. Hall, K. Nielsen, V. Nordenstoft, and A. W. Beavis. Virtual Reality Training for Radiotherapy Becomes a Reality. *Studies in Health Technology and Informatics*, 132:366, 2008.
- [48] T. Porter and T. Duff. Compositing Digital Images. In *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pp. 253–259, 1984. doi: 10.1145/964965.808606
- [49] A. Ramkumar, E. Varga, W. J. Niessen, and A. Freudenthal. Exploring Input Devices for Contouring in External Radiotherapy. *Innovative Imaging to Improve Radiotherapy Treatments*, 1:29–34, 2013.
- [50] R. Schmidt, A. Khan, G. Kurtenbach, and K. Singh. On Expert Performance in 3D Curve-Drawing Tasks. In *Proceedings of the 6th Eurographics Symposium on Sketch-Based Interfaces and Modeling*, SBIM '09, p. 133–140. Association for Computing Machinery, New York, NY, USA, 2009. doi: 10.1145/1572741.1572765
- [51] A. Sears and J. A. Jacko. *Human-Computer Interaction Fundamentals*. CRC press, 2009.
- [52] M. V. Sherer, D. Lin, K. Puri, N. Panjwani, Z. Zhang, J. D. Murphy, and E. F. Gillespie. Development and Usage of eContour, a Novel, Three-Dimensional, Image-Based Web Site to Facilitate Access to Contouring Guidelines at the Point of Care. *JCO Clinical Cancer Informatics*, 3(3):1–9, 2019. doi: 10.1200/cci.19.00041
- [53] T. Su, W. Sung, C. Jiang, S. Sun, and C. Wu. The Development of a VR-Based Treatment Planning System for Oncology. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 6104–6107, 2005. doi: 10.1109/embc.2005.1615886
- [54] K. Sura, J. W. Lischalk, J. Leckie, J. S. Welsh, A. J. Mundt, and E. Fernandez. Webinar-Based Contouring Education for Residents. *Journal of the American College of Radiology*, 14(8):1074–1079.e3, 2017. doi: 10.1016/j.jacr.2017.03.013
- [55] A. N. Viswanathan, B. Erickson, D. K. Gaffney, S. Beriwal, S. K. Bhatia, O. Lee Burnett, D. P. D'Souza, N. Patil, M. G. Haddock, A. Jhingran, E. L. Jones, C. A. Kunos, L. J. Lee, L. L. Lin, N. A. Mayr, I. Petersen, P. Petric, L. Portelance, W. Small, J. B. Strauss, K. Tournamchai, A. H. Wolfson, C. M. Yashar, and W. Bosch. Comparison and Consensus Guidelines for Delineation of Clinical Target Volume for CT- and MR-Based Brachytherapy in Locally Advanced Cervical Cancer. *International Journal of Radiation Oncology – Biology – Physics (IJROBP)*, 90(2):320–328, 2014. doi: 10.1016/j.ijrobp.2014.06.005
- [56] M. Werner-Wasik, A. D. Nelson, W. Choi, Y. Arai, P. F. Faulhaber, P. Kang, F. D. Almeida, Y. Xiao, N. Ohri, K. D. Brockway, J. W. Piper, and A. S. Nelson. What is the Best Way to Contour Lung Tumors on PET Scans? Multiobserver validation of a gradient-Based Method Using a NSCLC Digital PET Phantom. *International Journal of Radiation Oncology – Biology – Physics (IJROBP)*, 82(3):1164–1171, 2012. doi: 10.1016/j.ijrobp.2010.12.055
- [57] C. L. Williams and K. A. Kovtun. The Future of Virtual Reality in Radiation Oncology. *International Journal of Radiation Oncology – Biology – Physics (IJROBP)*, 102(4):1162–1164, 2018. doi: 10.1016/j.ijrobp.2018.07.2011
- [58] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, p. 143–146. Association for Computing Machinery, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963
- [59] E. J. Wuthrich, Q. Zhang, M. Machtay, D. I. Rosenthal, P. F. Nguyen-Tan, A. Fortin, C. L. Silverman, A. Raben, H. E. Kim, E. M. Horwitz, N. E. Read, J. Harris, Q. Wu, Q.-T. Le, and M. L. Gillison. Institutional Clinical Trial Accrual Volume and Survival of Patients with Head and Neck Cancer. *Journal of Clinical Oncology*, 33(2):156, 2015. doi: 10.1200/jco.2014.56.5218
- [60] M. Yarmand, C. Chen, D. Gasques, J. D. Murphy, and N. Weibel. *Facilitating Remote Design Thinking Workshops in Healthcare: The Case of Contouring in Radiation Oncology*. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411763.3443445
- [61] M. Yarmand, M. Sherer, C. Chen, L. Hernandez, N. Weibel, and J. D. Murphy. Evaluating Accuracy, Completion Time, and Usability of Everyday Touch Devices for Contouring. *International Journal of Radiation Oncology, Biology, Physics*, 2022.
- [62] H. Zhai, X. Yang, J. Xue, C. Lavender, T. Ye, J.-B. Li, L. Xu, L. Lin, W. Cao, Y. Sun, et al. Radiation Oncologists' Perceptions of Adopting an Artificial Intelligence-Assisted Contouring Technology: Model Development and Questionnaire Study. *Journal of Medical Internet Research*, 23(9):e27122, 2021. doi: 10.2196/27122
- [63] M. Zhang and J. P. Schulze. Server-Aided 3D DICOM Viewer for Mobile Platforms. *Electronic Imaging*, 2021(13):179–1, 2021. doi: 10.2352/issn.2470-1173.2021.13.evr-179
- [64] A. Zijdenbos, B. Dawant, R. Margolin, and A. Palmer. Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. *IEEE Transactions on Medical Imaging*, 13(4):716–724, 1994.

doi: 10.1109/42.363096

- [65] E. R. Zorza, M. Sousa, D. Mendes, R. Kuffner dos Anjos, D. Medeiros, S. F. Paulo, P. Rodrigues, J. J. Mendes, V. Delmas, J.-F. Uhl, J. Mogorrón, J. A. Jorge, and D. S. Lopes. Anatomy Studio: A Tool for Virtual Dissection through Augmented 3D Reconstruction. *Computers & Graphics*, 85:74–84, 2019. doi: 10.1016/j.cag.2019.09.006
- [66] D. Zukic, J. Vicory, M. McCormick, L. Wisse, G. Gerig, P. Yushkevich, and S. Aylward. ND Morphological Contour Interpolation. *The Insight Journal*, pp. 1–8, 2016.

APPENDIX A EXPERIMENTAL SETUPS

This section provides supplementary material in addition to Sec. 5.5, specifically in terms of describing our experimental workspace.

While investigating tracking accuracy of pen and tablet is out of our scope, it is still important to maintain reasonable tracking performance for conducting all evaluation conditions. Fig. A1 demonstrate the setups of the experimental office space, as well as the final placements of three SteamVR base stations [2].

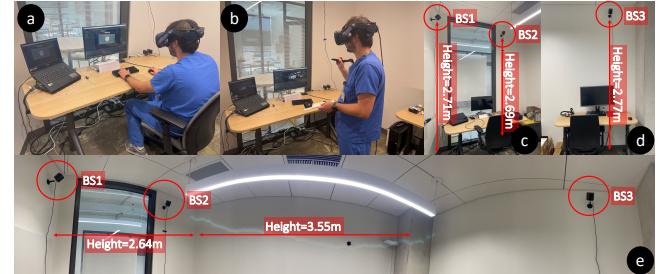


Figure A1: Experimental setups. (a – b) Participant could sit or stand while performing designated contouring tasks. (c) The base station mounted at the front of the room. (d) The base stations are mounted at the back of the room. (e) A panoramic view of the experimental office.