

ARMedicalSketch: Exploring 3D Sketching for Medical Image Using True 2D-3D Interlinked Visualization and Interaction

Nan Zhang , Tianqi Huang , and Hongen Liao , *Senior Member, IEEE*

Abstract—In traditional clinical practice, doctors often have to deal with 3D information based on 2D-displayed medical images. There is a considerable mismatch between the 2D and 3D dimensions in image interaction during clinical diagnosis, making image manipulation challenging and time-consuming. In this study, we explored 3D sketching for medical images using true 2D-3D interlinked visualization and interaction, presenting a novel AR environment named ARMedicalSketch. It supports image display enhancement preprocessing and 3D interaction tasks for original 3D medical images. Our interaction interface, based on 3D autostereoscopic display technology, provides both floating 3D display and 2D tablet display while enabling glasses-free visualization. We presented a method of 2D-3D interlinked visualization and interaction, employing synchronized projection visualization and a virtual synchronized interactive plane to establish an integrated relationship between 2D and 3D displays. Additionally, we utilized gesture sensors and a 2D touch tablet to capture the user’s hand information for convenient interaction. We constructed the prototype and conducted a user study involving 23 students and 2 clinical experts. The controlled study compared our proposed system with a 2D display prototype, showing enhanced efficiency in interacting with medical images while maintaining 2D interaction accuracy, particularly in tasks involving strong 3D spatial correlation. In the future, we aim to further enhance the interaction precision and application scenarios of ARMedicalSketch.

Index Terms—3D aerial display, augmented reality, autostereoscopic visualization, human-computer interaction.

Manuscript received 24 October 2023; revised 7 May 2024 and 2 July 2024; accepted 15 July 2024. Date of publication 2 August 2024; date of current version 18 September 2024. This work was supported in part by the National Natural Science Foundation of China under Grant U22A2051, Grant 82027807, and Grant 62101305, in part by the National Key Research and Development Program of China under Grant 2022YFC2405200, in part by the Intelligent Healthcare Tsinghua University under Grant 2022ZLB001, and in part by the Tsinghua-Foshan Innovation Special Fund under Grant 2021THFS0104. This article was recommended by Associate Editor S. Cao. (Corresponding author: Hongen Liao.)

Nan Zhang was with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China. He is now with the Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: zhangnan@ia.ac.cn).

Tianqi Huang is with the School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China (e-mail: huangtianqi@sjtu.edu.cn).

Hongen Liao is with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China, and also with the School of Biomedical Engineering and the Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: liao@tsinghua.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/THMS.2024.3432735>.

Digital Object Identifier 10.1109/THMS.2024.3432735

I. INTRODUCTION

IN CLINICAL practice, how medical images are interacted with significantly affects the accuracy and efficiency of image operations for treatment, which is crucial for both healthcare providers and patients. The prevalent method in clinics involves using 2D flat displays, operated through conventional devices such as a mouse and keyboard. Physicians often perform complex 3D tasks using these 2D-rendered medical images, such as segmenting 3D lesions, planning 3D surgical paths, and conducting 3D minimally invasive procedures. However, dealing with medical images in a 3D format through 2D-based visualization and interaction presents challenges. Complex 3D structures of human tissues are not adequately captured without 3D depth perception, hampering effective observation. Tasks such as surgical path planning based on 3D images are difficult in a 2D setting, and switching between display layers is inconvenient and less precise. To address these issues, we propose using medical 3D sketching [1] for 3D interactions with medical images. This approach is critical but challenging. It helps overcome the limitations of 2D-based methods, enhancing both the precision of medical image operations and the quality of patient care.

As computer technology advances, exploration and research into 3D sketching within virtual reality (VR) [2], [3], [4] and augmented reality (AR) [5] have gained momentum. Some research based on 2D display [6] can also achieve 3D sketching, which is not the focus of this study. The demand for 3D sketching is extensive, particularly in various design fields involving manual work. Compared with 2D sketching, 3D spatial sketching offers a clear advantage, providing a “what you see is what you get” human-computer interaction (HCI) experience. Nevertheless, achieving accurate and efficient 3D sketching is a challenging task. Two primary difficulties in achieving precision and efficiency in 3D sketching involve ensuring users possess a correct perception of 3D space and the ability to interactively translate their mental 3D sketches effectively [7].

In the realm of VR/AR-based 3D sketching [4], head-mounted displays (HMDs) [3], [8] are a significant approach for imparting 3D spatial perception, offering advantages such as real-time rendering and high-quality graphics. However, several drawbacks, such as inducing discomfort like nausea, vomiting, and vertigo, have yet to be effectively addressed [9], [10]. Given the demanding workload of clinical experts, there is a compelling

need for a 3D visualization method that does not rely on HMDs. Naked-eye-based methods emerge as a viable alternative, encompassing autostereoscopic [11], [12], [13], holographic [14], [15], and volumetric [16], [17] displays. Among them, the 3D autostereoscopic display stands out as a full-color 3D visualization technique supporting real-time interaction. Zhang et al. [11] proposed a floating autostereoscopic display system with in-situ interaction, leveraging a novel light-field symmetrical multiview-point rendering algorithm and incremental gestures. It is noteworthy, however, that limited attention has been devoted to 3D autostereoscopic display-based 3D sketching methods and systems, warranting further exploration in this domain.

Based on the considerations mentioned earlier, this study aims to explore 3D sketching for medical images using naked-eye-based visualization. We propose ARMedicalSketch, a novel medical 3D sketching environment featuring true 2D-3D inter-linked visualization and interaction. The main contributions of this work are as follows.

- 1) The definition of a design space for 3D medical image sketching, as well as the presentation of interaction metaphors arising from the integration of naked-eye-based 2D and 3D visualization and interaction with a floating 3D autostereoscopic display and tablet.
- 2) The concept and implementation of ARMedicalSketch involve a sketching application that combines a user-controlled tablet and floating 3D display, enabling seamless content control through hand gestures. It integrates 3D mid-air sketching with 2D surface sketching for mutual responsiveness to changes.
- 3) First insights regarding usage patterns of ARMedicalSketch, based on a one-hour-long usability walkthrough study with twenty-three participants. Comparative user experiments demonstrate greater efficiency in 3D-based interaction compared with 2D.

II. RELATED WORK

Our work is strongly influenced by previous work on HMD-based 3D sketching VR/AR environments and medical image sketching studies.

A. VR/AR 3D Sketching

3D sketching holds a significant role in HCI [4]. As early as 1995, a VR-based 3D sketch drawing system was introduced [3], marking an early milestone in the development of VR-based 3D geometry creation and manipulation tools. This system facilitated the creation and manipulation of 3D tube and wire geometries. While sketching systems based on predefined surface shapes serve practical purposes in specific scenarios, they can limit user creativity. Freeform 3D space sketching, on the other hand, demands a more sophisticated HCI design. In recent years, HMD-based sketching systems have emerged [18], [19], [20], [21] to address this challenge. Common approaches include constraining input dimensions and providing visual guidance to improve sketch alignment with users' intentions.

One of the challenges in rendering 3D space is managing the high degree of interaction freedom. Virtual 2D interactive

surfaces or the use of 2D tablets as input devices have shown promise in addressing this challenge. For instance, Arora et al. [22] developed a hybrid interactive drawing system based on HMDs. Their hybrid mode combines the precision of 2D plane interaction with the immersive qualities of 3D air interaction, offering users greater flexibility. Drey et al. [4] expanded on this concept with "VRSketchIn," enhancing 2D and 3D hybrid interactions. The system features six-degrees-of-freedom interaction and real-time tracking for input devices (pens and tablets), earning recognition from artists, product designers, and other creative professionals. Jiang et al. [23] proposed an immersive 3D drawing system called "HandPainter," wherein users wear HMDs and interact using gloves to draw on their palms. User studies have demonstrated that this system enables more accurate spatial interaction compared with drawing on a tablet or using a 3D controller.

Moreover, some studies focus on aiding users in reproducing target sketches by providing visual guidance or stroke optimization. Tsang et al. [24] introduced an image-guided, pen-based suggestive interface for sketching 3D wireframe models. This environment combines 3D sketching with gestures, concept images, and suggestions, resulting in a compelling and effective design tool. Yu et al. [25] proposed a method for curve and surface drawing in an immersive environment based on HMD technology. Their 3D drawing optimization framework incorporates functions such as automatic curve network connection and patch prediction, enhancing the reasonableness and aesthetics of the output compared with user-input three-dimensional spatial paths.

While HMD-based VR/AR sketching research has yielded valuable interaction solutions, there remains an unmet need for a naked-eye-based 3D sketching system, free from the constraints of glasses. This work aims to address that gap and contribute to the field.

B. Medical Image Sketching

In the medical context, research on 3D sketching remains relatively scarce. Relevant studies can be categorized into two main areas: education, and clinical diagnosis and treatment.

In the field of medical education, 3D sketching aids users in rapidly grasping complex anatomical structures and physiological dynamics. For instance, Saalfeld et al. [26] introduced a semi-immersive 3D user interface that allows users to sketch intricate vascular structures and vessel pathologies by drawing centerlines in 3D. A qualitative evaluation affirmed the practicality and user-friendliness of their approach. Subsequently, they presented computer graphic techniques [27] to enhance and facilitate the creation and annotation of complex sketches, resulting in more lucid, expressive, and comprehensible resources for medical teaching, treatment planning, and patient education.

In contrast to the realm of medical education, 3D sketching in clinical diagnosis and treatment is a demanding endeavor characterized by the need for efficiency and precision. The efficiency and precision of these tasks often have a direct bearing on subsequent diagnosis and treatment outcomes. Therefore, within the domain of medical space mapping, there is a heightened

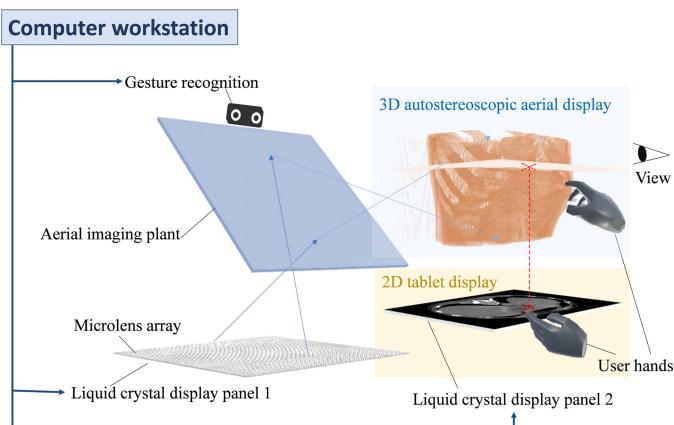


Fig. 1. System schematic.

emphasis on optimizing interaction efficiency and precision. Heckel et al. [28] proposed sketch-based editing tools for tumor segmentation in 3D medical images using a 2D interface. Ijiri and Yokota [29] introduced an innovative system for the intuitive and efficient refinement of medical volume segmentation by modifying multiple curved contours, supporting both topological manipulations and contour shape modifications. Johnson et al. [30] explored VR and hybrid 2D/3D sketch-based interfaces and visualizations for medical image segmentation, highlighting the growing demand for 3D sketching in this domain.

Human medical images inherently comprise 3D data, and many manual processing tasks in medical image analysis can be effectively summarized as 3D sketching. Therefore, a comprehensive platform capable of supporting medical images from raw data to 3D sketch processing is needed. Such a platform could encompass features such as anatomical feature markers and surgical pathway planning, presenting a valuable contribution to the field of medical image analysis.

III. OVERVIEW OF THE ARMEDICALSKETCH

Despite numerous studies having been conducted in the domains of VR/AR 3D sketching and medical image sketching, systems enabling 3D spatial sketching through naked-eye displays have been rarely investigated. In this study, we present ARMedicalSketch (Fig. 1), a novel medical 3D sketching environment that supports genuine 2D and 3D displays and free interaction based on the user's hand. The displayed content in 2D and 3D is dynamically linked in real-time, with the content in one dimension synchronously adjusting whenever a user interacts with the other dimension.

By combining naked-eye floating 3D and 2D images, our AR medical sketching environment serves as an effective and precise entry point for handling 3D sketching tasks in medical imaging. The environment consists of three parts (Fig. 2): medical image pro-processing (Section IV-A), naked-eye-based 2D and 3D display (Section IV-B), and interlinked visualization and interaction (Section IV-C–D).

The system employs 3D aerial autostereoscopic visualization technology to provide true 3D visualization, enabling real-time

display of 3D sketch schemes of medical images. Given the precision challenges inherent in interactions within environments with high spatial freedom, we utilize a 2D tablet as a reliable input mechanism. Additionally, to ensure users accurately perceive and position themselves relative to both the 3D and 2D displays, we have designed a virtual synchronized interactive plane. The 3D-human interaction offers intuitiveness and 2D-human interaction ensures precision. Our 2D-3D interlinked mode helps bridge the visualization and interaction advantages present in these two distinct dimensions.

The combination of naked-eye-based 2D and 3D displays with freehand interaction facilitates an easy-to-perceive and engaging 3D medical sketching environment. It allows multiple users to view without glasses and engage effortlessly. The user's hand in the prototype can be replaced with any required input device. For instance, an optical tracking pen can be used for higher input precision, and a force feedback device can be employed for perceiving the tactile sense of medical anatomical structures, and so forth.

IV. PROOF OF CONCEPT IMPLEMENTATION

A. Medical Image Pro-Processing

In our system, a specialized interface for medical image data has been developed. The workflow involves reading and reorganizing the original medical images, transforming them into 3D texture data optimized for use in the AR environment. This 3D texture data is then utilized in the AR environment for both 2D slice display and 3D volume rendering. To enhance medical image visualization, we have implemented tailored enhancements based on the 3D texture data.

In 2D display, an approach based on histogram normalization using 3D volumetric data was adopted to address challenges posed by the limited grayscale range and low image contrast often observed in medical images. This method effectively distributes voxel grayscale values from the medical 3D texture data uniformly within the entire grayscale range of 0–1, significantly enhancing image clarity. Each voxel value (with $dx : dy : dz = 1 : 1 : 1$, representing the equal spatial resolution in the x , y , and z dimensions resulting from preprocessing) is considered as a pixel. When statistically analyzing the grayscale range, the entire 3D volume data is considered as a collective, as opposed to individual 2D planes. Similarly, during grayscale transformation, the alteration and distribution of grayscale values are computed using the complete 3D volume data as a unified set. Fig. 3 illustrates the effects of histogram normalization based on 3D volumetric data, enhancing medical image interpretation.

In this study, the 3D volume rendering is implemented using the raycasting algorithm. During the rendering process, we sample and synthesize from the front to the back of the image using the substitution.

$$Color_i = (1 - Color_{i-1} \cdot a) * Color_i + Color_{i-1} \cdot$$

Here, $Color_i$ represents the color of the sample point, and $Color_i$ is the cumulative displayed color at position i . To improve the rendering effect, we incorporate three rendering parameters, I , D , and K , into the sampling process, defined by

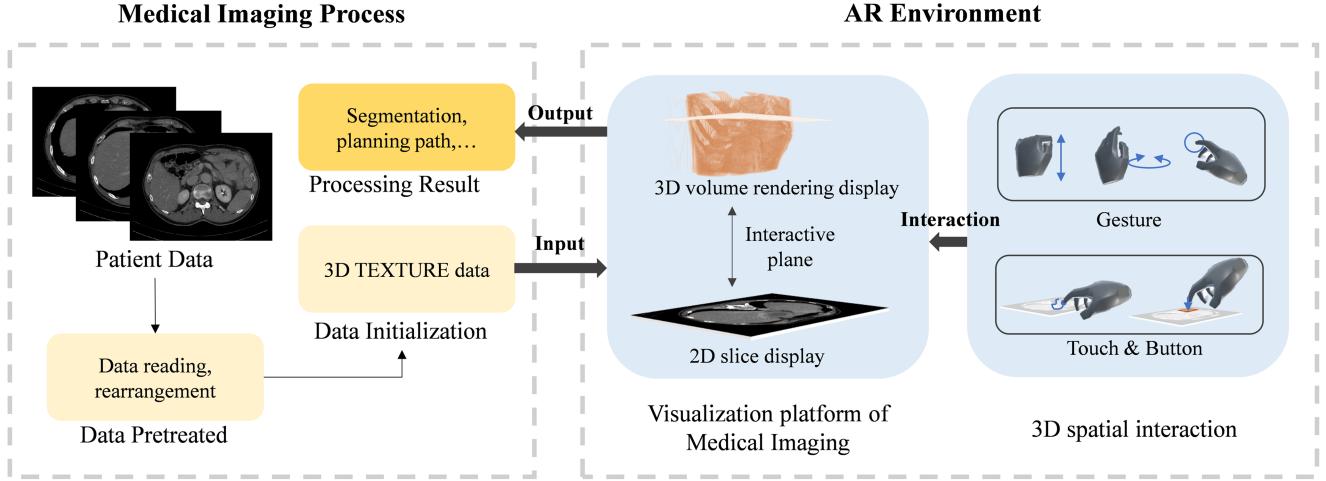


Fig. 2. Overview of our approach.

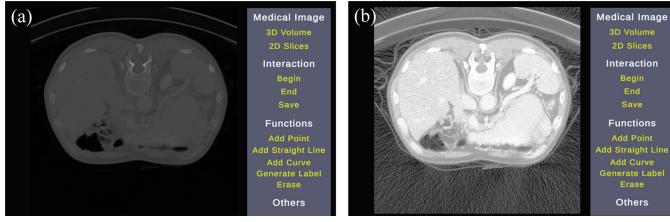


Fig. 3. Histogram normalization enhancement based on three-dimensional volume data. (a) Original medical image. (b) Processed medical image.

the substitutions.

$$Color_i.a = D \times \text{pow}(Color_i.a, K)$$

$$Color_i.rgb = I \times Color_i.rgb \times Color_i.a,$$

where I adjusts the color intensity of the sample point, D determines the sampling depth, and K fine-tunes the contrast between colors at the sample point. Voxel values are mapped to a pseudocolor palette, allowing for selective enhancement of specific aspects of the volume rendering display. Moreover, we integrate boundary enhancement into the rendering process to accurately showcase the range of data. Fig. 4(a) displays the original volume rendering of the medical image, while Fig. 4(b)–(d) demonstrates the rendering outcomes after adjusting rendering parameters and applying pseudocolor mapping. The data border can be tailored for display, including adjustments to transparency, color, and width.

B. True 2D and 3D Visualization Without Glasses

The system employs glasses-free naked-eye 2D tablet and 3D floating display technology. For authentic 2D visualization, a touch-enabled 2D display tablet is utilized. In this section, our primary focus is elucidating the methodology for realistic 3D visualization. We employ the floating autostereoscopic display technology, as depicted in Fig. 5. Initially, the system calculates the real-time capture of 3D parallaxes within the virtual AR environment using a virtual camera array, resulting in the generation of an elemental image array (EIA) through pixel resampling.

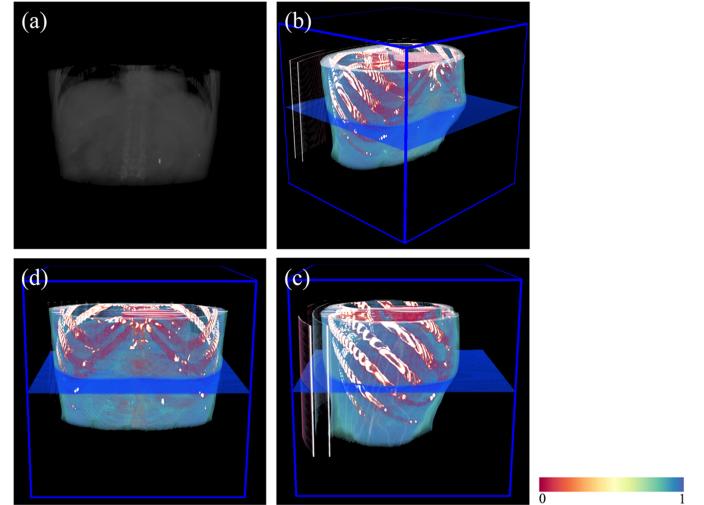


Fig. 4. Three-dimensional volume rendering enhancement visualization. (a) Original medical image volume rendering. (b)–(d) Rendering results after parameter adjustments and pseudocolor mapping.

Subsequently, the EIA is showcased on a liquid crystal display (LCD), and the images are converged into a 3D light field through a lens array. To achieve spatial interaction, we project the 3D parallaxes into the air using an aerial imaging plate (AIP). AIP can project incident light into the air at identical angles. For a comprehensive understanding of the AIP's structure and internal light path, we refer to [11]. A light-field symmetrical multiview rendering algorithm is employed to rectify viewpoint inversion during pixel resampling, ensuring the correct multiview parallax of the 3D parallax in the air. The viewing angle of the floating 3D display is calculated as follows:

$$\varphi = 2 \times \arctan(pitch/2gap)$$

where $pitch$ represents the diameter of a single lens, and gap denotes the separation between the lens array and LCD. The advantage of the AIP-based floating 3D display is to present undistorted proportional real images. Consequently, by

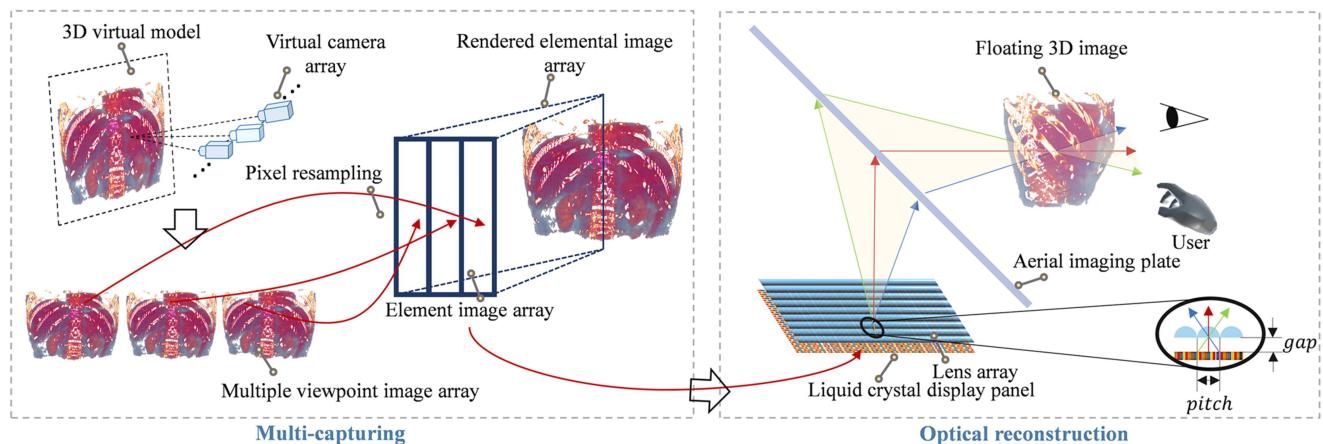


Fig. 5. Rendering and display procedure of the floating 3D image, encompassing multiviewpoint image capture and subsequent 3D image reconstruction through a lens array.

calibrating the physical position in the real world, we establish a vertical projection relationship between 3D and 2D displays.

Our proposed naked-eye visualization approach eliminates the need for wearable devices, offering an accessible and user-friendly experience. The integration of 2D and 3D displays enhances the precision of signal input through the tablet, while the 3D display facilitates easily accessible 3D depth perception.

C. 2D-3D Interlinked Visualization

Enhanced visualization of medical 3D texture data is utilized in our proposed platform for both the 2D tablet and a floating 3D display. In our research, we provide a 2D-3D interlinked visualization method to seamlessly integrate 2D projection and 3D spatial representation. It ensures a close correlation between the content displayed in 2D and the floating 3D projection. The 2D display serves as a projection, offering an accurate and easy-to-use display for users, while the floating 3D display brings the content to life in a spatial context, providing comprehensive and immersive understanding.

To achieve the integration, we adopt a parameter-based control system and a virtual synchronized interactive plane that governs the content displayed in both 2D and 3D. This means that the same set of parameters is employed to regulate visualization in both dimensions, ensuring synchronicity and cohesiveness between the 2D projections and the 3D spatial representations. At the same time, the interactive plane will be presented in the floating 3D display image, representing the orientation of the 2D plane in the floating 3D display.

In the volume rendering process, 3D medical images are visualized as a whole in space. Consequently, establishing a correct spatial overlay relationship between an independent virtual interaction plane and the 3D medical images becomes nontrivial. A virtual plane can typically either be entirely displayed in front of or behind the volume-rendered medical images, but fails to accurately represent the spatial pose relationship between them. Therefore, we adopted a technique grounded in volume rendering graphics pipelines to cohesively integrate the virtual interaction plane with the 3D medical images. We utilize the fragment shader program to compute the coordinates of the

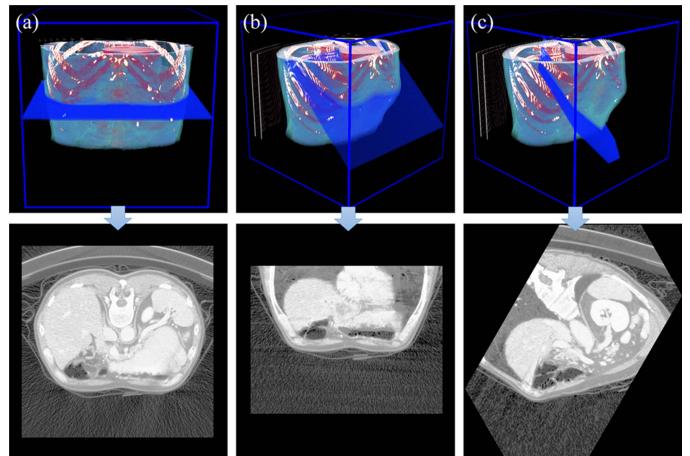


Fig. 6. 2D-3D synchronized projection. (a) Horizontal plane. (b) and (c) Free angle plane. First row: 3D volume rendering; second row: corresponding 2D displays.

virtual plane based on model-space coordinates. During the rendering process, we assign the voxel values to the coordinates of the virtual plane and incorporate them into the final output screen color values using a front-to-back raycasting rendering approach. This integration places the interactive plane and the 3D medical images under a unified rendering framework, allowing the depiction of their accurate tangential relationship, which is pivotal in achieving seamless synchronization and spatial coherence.

The control system includes three core parameters N_{dir} , N_{slice} and N_{pos} . N_{dir} is a 3D vector representing the normal vector of the synchronized interactive plane, N_{slice} is the distance from the center point to the interactive plane in 3D medical image volume data, and N_{pos} is the real-time interaction point location of the user on the medical image. These parameters control critical aspects such as orientation, position, and appearance, allowing for a seamless transition between 2D and 3D, enhancing the overall user experience and comprehension of the medical data. Fig. 6 illustrates the results of 2D-3D interlinked visualization. By adjusting the parameters N_{dir} and N_{slice} , changes in the 3D volume rendering image and the

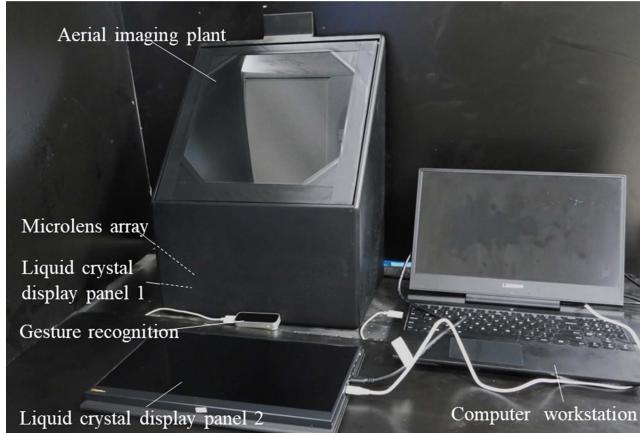


Fig. 7. Setup of the prototype.

interactive plane within it can be observed. Clearly, a greater set of parameters can be configured, such as those used to store user interaction results.

Our 2D-3D interlinked visualization method leverages the synergy between 2D projection and 3D spatial representation, unifying them under a common set of parameters. This innovative approach not only enriches the visualization experience but also maximizes the potential for effective interpretation and interaction with complex medical data.

D. 2D-3D Interlinked Interaction With Hand and Tablet

To facilitate seamless HCI within the environment, we employ a fully user-centric interaction mode, leveraging the user's hands extensively. This interaction is made possible through a sophisticated integration of sensory inputs from two critical sources: the 2D tablet and gesture sensors.

The user's hands serve as the focal point of interaction, acting as conduits for input into our system. The interaction process is orchestrated through a combination of sensory data obtained from the 2D tablet and gesture sensors. The 2D tablet, equipped with touch-sensitive capabilities, forms a pivotal component of this dynamic interplay, enabling precise and intuitive touch-based inputs from the user. Complementing the 2D tablet, gesture sensors play a crucial role in enriching the interaction dynamics. The sensors capture the nuances of hand gestures and the positioning of fingertips. By amalgamating the touch-based inputs from the 2D tablet and the hand data from gesture sensors, our system hopes to achieve a comprehensive and sophisticated understanding of the user's intent and actions.

Our 2D-3D interlinked interaction approach leverages the user's hands for intuitive interaction. The integration of sensory inputs from the 2D tablet and gesture sensors offers a cohesive, enriched interaction experience, contributing to effective interpretation and interaction with intricate medical data.

V. IMPLEMENTATION AND EVALUATION

A. System Architecture and Technical Details

The proposed prototype is realized through Unity 2020.3.8. Considering the hardware construction, the prototype (Fig. 7) is

TABLE I
CONFIGURATION PARAMETERS OF THE PROTOTYPE

Equipment	Parameters	Specification
Liquid crystal display panel 1	Size	344 mm × 194 mm
	Resolution	3840 × 2160
	PPI	282
Liquid crystal display panel 2	Size	344 mm × 194 mm
	Resolution	1920 × 1080
	PPI	141
Lenticular lens array	Lens pitch	0.2115 mm
	Gap	1.052 mm
Gesture recognition device	Equipment	Leap Motion
	Aerial imaging plate	Size
Computer workstation	Size	303 mm × 298 mm
	CPU	CORE i7-9750H
	GPU	GeForce GTX 1660Ti

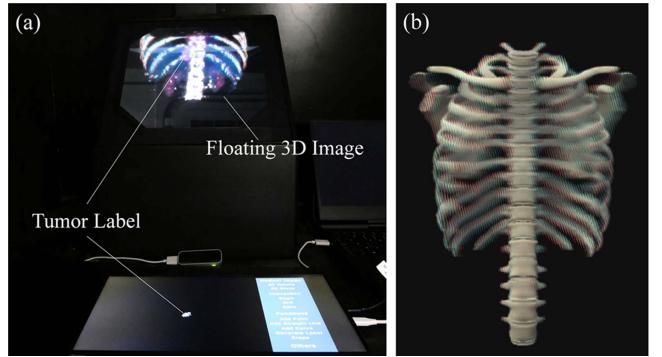


Fig. 8. (a) 2D-3D interlinked visualization. (b) Rendered EIA imaging of floating 3D display.

configured with the specifications in Table I. The prototype contains two LCD panels, an MLA, an AIP, a gesture recognition, and a computer workstation. LCD 1 is utilized for the floating 3D display and LCD 2 for 2D display and touch.

Fig. 8(a) shows the 2D-3D interlinked visualization with a floating 3D image and the corresponding tumor image in 2D tablet display. We used a skeleton model to demonstrate the 3D floating horizontal parallax. Fig. 8(b) shows a rendered EIA of the skeleton model for the floating 3D display, and Fig. 9 shows the 3D aerial autostereoscopic display with horizontal parallax.

In the prototype interaction experiment, users can utilize gestures to modify the display interaction plane, resulting in changes in both 2D and 3D images [Fig. 10(a) and (b)]. Furthermore, when users draw the tumor boundary on the touchscreen, the floating 3D display synchronously presents the drawn contours [Fig. 10(c) and (d)]. The rendering rate exceeded 110 frames per second. More control gestures can be designed to meet a variety of medical sketching objectives.

B. Application Walkthroughs

We demonstrate the 3D sketching capabilities of our proposed system through tumor segmentation and tumor puncture path planning.

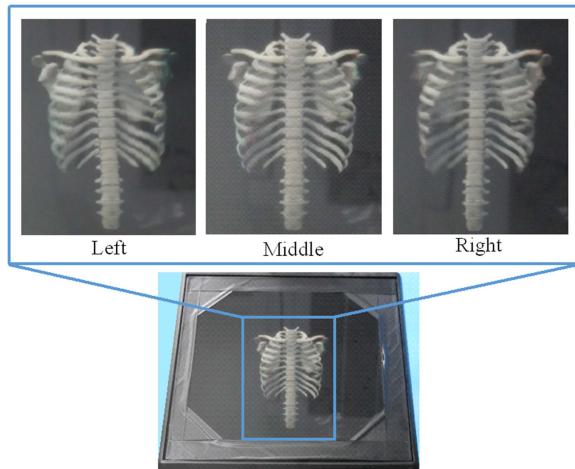


Fig. 9. 3D aerial autostereoscopic display with parallax effect.

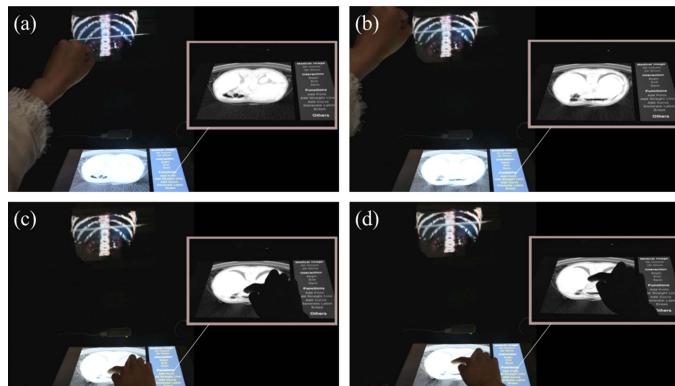


Fig. 10. User prototype interaction experiment. (a) and (b) Gestures alter 2D display content. (c) and (d) Touch gestures draw the tumor boundary.

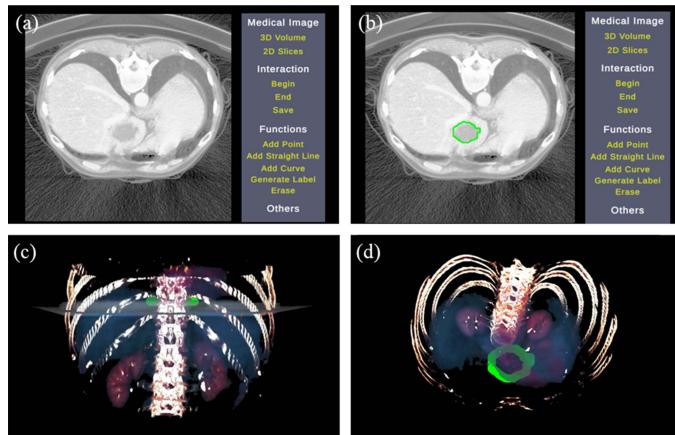


Fig. 11. Synchronized effect of tumor boundary sketching. (a) 2D drawing interface. (b) Touch drawing of tumor boundary. (c) and (d) Bold 3D spatial boundary.

1) *Scenario 1. Tumor Segmentation:* Fig. 11 illustrates the synchronous effect of tumor boundary sketching. Users adjust the synchronous interaction interface to the appropriate tumor position using their left hand and then draw the tumor boundary on the 2D tablet using their right hand. Simultaneously, the drawn boundary appears in the volume rendering image of

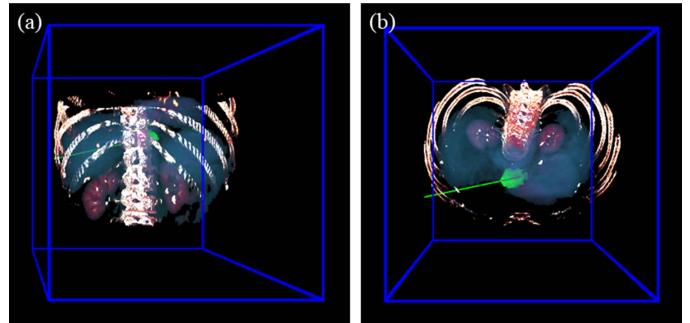


Fig. 12. Tumor puncture pathing view in (a) frontal and (b) top-down perspectives.

the 3D autostereoscopic aerial display (bold green outline). After completing the drawing of a tumor boundary, a search is conducted within the medical 3D volume data based on the drawn boundary. Using the tumor voxel values for growth, a 3D segmented tumor label is generated, and then a 3D surface model of the label is generated based on the Marching Cubes algorithm.

2) *Scenario 2. Path Planning:* Fig. 12 showcases the tumor puncture path sketching effect. Users adjust the synchronous interaction interface to the suitable tumor puncture position using the left hand and then mark the puncture starting position on the 2D touchscreen using the right hand. Then, the left-hand interacts to change the synchronous interaction plane N_{slice} value. A real-time puncture path is generated in the 3D volume rendering, synchronously changing with the N_{slice} , and when the user marks the second puncture termination position on the 2D tablet, the final puncture path is determined and generated. The coordinates of the completed sketching points and lines can be adjusted according to the user's preference. The provided sketching system assists users in achieving the desired drawing results.

C. User Study

1) *Purpose:* The control test evaluated the following two parameters. First, training effectiveness: a) determining whether the system improved the sketching efficiency of medical images; b) determining whether the system improved the accuracy of sketching medical images. Second, system performance: a) determining the system's usability and participant satisfaction, and b) determining the degree of psychological burden on participants caused by system use.

2) *Participants:* Twenty-three biomedical engineering students volunteered to participate in the user study and completed the consent form. Despite having a basic medical knowledge foundation, participants had no actual expertise in organ diagnosis of medical images. Participants ranged in age from 21 to 33 years (average = 25 years) and included 13 males and 10 females. None of the participants in the study were involved in the system's creation. Three participants were assigned to test the prototype, and based on their feedback, we established the system's learning curve process time in procedures. The remaining twenty participants were randomly assigned to an experimental group or a control group, each with ten students.

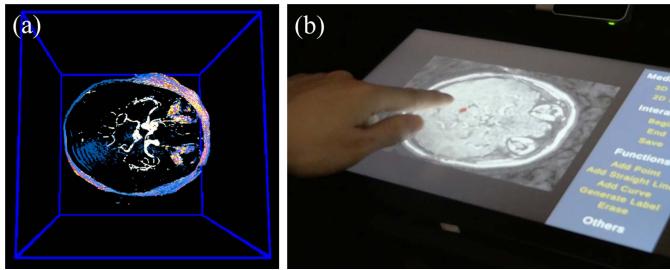


Fig. 13. Labeling experiment of Willis's circle. (a) Willis's circle enhanced rendering through the floating 3D display. (b) Users marked anatomical points through the 2D tablet.

Willis's circle, recognized as one of the complex anatomical structures at the base of the human brain, is the intersection of several important arteries. We collected five MR images of Willis's circle, split them into a comparable sample set, and minimized differences in position. Similarly, we collected and processed ten healthy liver CT images.

3) *Procedures*: We set up two prototypes. Our proposed prototype was Prototype 1, which featured a naked-eye floating 3D display alongside a 2D tablet. Prototype 2 replaced the floating 3D display with a standard 2D flat display for visualizing 3D medical models rendered using a single virtual camera, and also along with a 2D tablet. Different prototypes were utilized by different groups.

The experiment comprised five stages: (1) introduction, (2) initial testing, (3) learning, (4) follow-up testing, and (5) questionnaire period. Initially, around 20 min were allocated for each participant to be introduced to their prototype and learn about the structure of Willis's circle and liver imaging region. Subsequently, volunteers underwent two tasks for the first time. Task 1 involved selecting any region within the liver images, while Task 2, a more challenging test, focused on Willis's circle (Fig. 13), where volunteers were required to mark the anterior communicating artery and the top of the basilar artery by interacting with their prototype. Participants then spent 20 min learning about their prototype, during which they were encouraged to engage freely with the system during the learning process. Following this learning phase, volunteers underwent the same two tests for a second time. We documented the time and marked locations.

Fig. 14 presents the results, including mean scores, standard deviations, and significance across prototypes for interaction duration and accuracy across two tests. In the accuracy assessment, the precision of the liver experiment was measured by users' ability to identify points within the liver region, while the precision of Willis's experiment was gauged by positional deviation distances. A critical p-value of 0.05 is used for the hypothesis test to determine whether the interprototype difference is significant, and significant findings are depicted in Fig. 14 asterisks. ANOVA tests were used to obtain p-values. Prototype 1 exhibited higher completion efficiency in Task 2.

Participants were also asked to complete two questionnaires: (1) assessing system usability using the System Usability Scale (SUS) [31]; (2) evaluating workload with the NASA-TLX assessment tool [32], gauging cognitive load with mental demand, physical demand, temporal demand, performance, effort, and

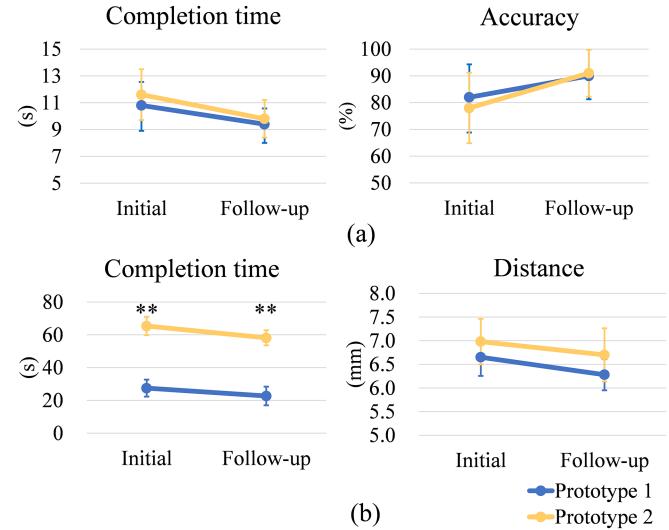


Fig. 14. Results of the user study on training effectiveness. (a) Task 1 involved identifying liver regions. (b) Task 2 focused on labeling key points of Willis's circle. (**): $p < 0.01$.

TABLE II
RESULTS FOR THE QUESTIONNAIRES: MEAN (INTERQUARTILE RANGE)

Questionnaires	Prototype1	Prototype2	Significance
System usability	81.6(4)	76.8(2)	** ($p < 0.01$)
Workload	53.2(8)	50.7(5)	

frustration. For the SUS questionnaire, responses were recorded on a five-point Likert scale and converted to a usability score ranging from 0 to 100, where higher SUS scores indicate better usability. In the workload assessment, participants rated each dimension on a scale from 0 to 10, and provided pairwise comparisons to determine their relative importance as weights. Overall workload scores, ranging from 0 to 100, were calculated based on these ratings, with lower scores indicating lower workload levels.

Table II shows two questionnaires' results on system performance, including the score mean, interquartile range, and significance levels (one-way ANOVAs). There was significance in the SUS questionnaire.

4) *Clinical Expert Evaluation*: We invited an orthopedic surgeon with seven years of clinical experience and a neurosurgeon with three years to provide expert feedback after using our system. Both acknowledged that the glasses-free 3D spatial HCI technology provides an intuitive, immersive visualization experience, enhancing comprehension and evaluation of intricate anatomical structures for improving precision and safety in medical imaging tasks. The integration of a 2D touch-screen tablet with 3D floating image projection enables seamless annotation and manipulation on a 2D plane, with corresponding 3D effects, facilitating flexible surgical path planning and procedural guidance. The platform's versatility across various minimally invasive surgical tasks, including tumor resection and endovascular interventions, offers a novel adjunctive tool to optimize surgical efficiency and success rates.

However, for more intricate surgical procedures, further improvements and optimizations may be necessary to align with

practical requirements, particularly in adapting to interpatient variations and addressing unforeseen challenges across clinical scenarios and specialties. Limitations in the clarity and stability of the floating 3D images could affect the observation and interpretation of fine details and structures. Recommendations include improving the quality and clarity of the floating 3D images to ensure precise visualization, and fortifying system stability and real-time responsiveness to facilitate relevant support.

VI. DISCUSSION

Overall, our prototype was well-received by the volunteers and clinical experts. Based on the results of the user experiments, Prototype 1 demonstrated superior performance and user acceptance. In Fig. 14, the completion time for Task 2 indicated a significant improvement in efficiency for Prototype 1 when handling challenging 3D medical tasks. In Table II, the results of the SUS revealed that users appreciated Prototype 1 for its system usability and demonstrated a high level of acceptance. Results from both the initial and follow-up tests showed a slight improvement in task completion efficiency and accuracy, demonstrating the usability of the proposed prototypes.

User experiments confirm that the proposed true 2D-3D interlinked visualization and interaction efficiently facilitates interactive tasks based on 3D medical images. Fig. 14 shows that the combined display mode of floating 3D visualization and projected 2D display significantly improves user interaction efficiency compared with traditional 2D displays when dealing with complex 3D image tasks (such as Task 2). For medical image tasks with weak 3D spatial correlation (such as Task 1), the efficiency of Prototype 1 and traditional 2D devices are similar. These findings align with previous research in related fields. Studies in surgical applications [33] and anatomical education [34] have consistently shown that 3D visualization improves performance in 3D image recognition tasks. Additionally, a comparative study [35] of 3D and 2D camera-based laparoscopic procedures revealed that interactions facilitated by 3D visualization significantly reduced error rates. Table II further indicates that spatial gesture interaction based on 3D autostereoscopic aerial display is acceptable and does not significantly increase user workload. Therefore, the proposed naked-eye 2D-3D visual interaction interface is conducive to assisting volunteers in rapid spatial localization.

One of the innovative aspects of our work lies in the fusion visualization of true 2D and 3D. In Fig. 14, the interaction accuracy of users with Prototypes 1 and 2 is similar and does not significantly differ. When it comes to 3D spatial tasks, Prototype 1 shows a slightly more accurate representation, which is speculated to be due to the spatial perception provided by the floating 3D autostereoscopic display, facilitating the user's visual judgment. The result of the SUS in Table II suggests that introducing floating 3D visualization in the naked-eye display significantly enhances the usability level of the system. These results indicate that our proposed true 2D-3D visualization interface provides a high-precision 3D spatial interaction method, achieving an interaction accuracy level comparable to that of a 2D interface. This result is promising, achieving precise 3D

spatial interaction through the projected visual association of 2D and 3D visualizations [22].

Leveraging the strengths in visualization and interaction, the integration of ARMedicalSketch into medical environments holds considerable promise. As an innovative HCI system developed for medical use, it provides users with glasses-free 3D spatial perception and an intuitive, precise mode of interaction. The system serves diverse diagnostic, therapeutic, and educational purposes by customizing interactive interfaces to suit various clinical scenarios. For instance, in neurosurgery or interventional cardiology, ARMedicalSketch enhances diagnostic capabilities by visualizing intricate case images, facilitates segmentation and planning of 3D radiation targets in radiology, and aids in explaining lesion presentations to patients and their families, thereby enhancing comprehension of treatment strategies and surgical procedures. Additionally, the naked-eye 3D floating display and gesture interaction module can be combined independently to form a fully floating, noncontact 3D HCI system, extending ARMedicalSketch's potential application demands to the operating room, including path planning display and intraoperative navigation.

The system is compatible with standard DICOM medical image formats for seamless integration into existing systems within medical institutions, facilitating easy access and manipulation of patient medical images by healthcare professionals. While introducing ARMedicalSketch into clinical settings requires careful considerations, it will strictly adhere to relevant regulatory standards and undergo validation to ensure continuous compliance. This includes robust data security, stringent user privacy protection, and system stability measures. Regarding hardware deployment, it is implemented in sealed facilities with integrated tablets, facilitating strict adherence to disinfection requirements in line with healthcare institution hygiene standards. Furthermore, ongoing plans for feedback iteration mechanisms are in place to better align with the needs and preferences of clinical users. Ultimately, as a medical device, ARMedicalSketch will need to comply with relevant regulatory standards and regulations in the future to ensure its safety, efficacy, and compliance.

Our research addresses the gap in naked-eye 2D-3D fusion visualization and interaction. However, this work still has several limitations. There is room for improvement in the interaction accuracy of the prototype. In this study, we captured interaction information through users' hands, but the positioning information at the fingertip or obtained by finger touch is relatively coarse. Future plans include designing a spatial pen with 3D optical tracking and 2D tablet touch functions, replacing hand touch while retaining quick gesture control. Additionally, we recruited 23 volunteers for this study, and aim to expand the participant pool for a comprehensive evaluation, comparing our system with HMDs, traditional 2D flat displays, and freehand handheld tablets in 3D medical sketching in future work.

VII. CONCLUSION

In this article, we explored the realm of 3D sketching for medical images by leveraging the potential of true 2D-3D interlinked visualization and interaction, and presenting a novel medical

3D sketching AR environment named ARMMedicalSketch. In clinical practice, 3D interaction based on 2D medical image displays is challenging due to a mismatch between dimensions, and current 3D drawing methods with HMDs limit prolonged use for medical professionals. Our ARMMedicalSketch presents a new approach by offering an interaction interface devoid of the need for glasses, presenting a seamless fusion of both floating 3D display and 2D tablet display capabilities. This amalgamation empowers users to interact with 3D medical imaging tasks without the encumbrance of wearing glasses. We introduced a method of 2D-3D interlinked visualization and interaction, employing synchronized projection visualization and a virtual synchronized interactive plane to establish an integrated relationship between 2D and 3D displays. The outcomes gleaned from our user study unveil a promising trajectory. When compared with traditional 2D interaction interfaces, our true 2D-3D system significantly amplifies the efficiency of 3D interactive tasks related to medical image visualization while preserving precision and accuracy.

ARMMedicalSketch provides a novel glasses-free 3D spatial interaction method tailored specifically for interactive tasks in 3D medical images. In the future, our aim is to refine the precision of interactions, expand the spectrum of application scenarios for ARMMedicalSketch, and even introduce machine learning methods, such as automated feature detection and segmentation, as well as connect to large-scale language models, to improve the intelligence of the system. This will unlock its full potential in medical imaging and beyond.

ACKNOWLEDGMENT

The authors would like to thank Z. Liao and J. Liu for helping design and print the prototype frame.

REFERENCES

- [1] D. Gasques, J. G. Johnson, T. Sharkey, and N. Weibel, "What you sketch is what you get: Quick and easy augmented reality prototyping with PintAR," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–6.
- [2] N. Zhang, H. Wang, T. Huang, X. Zhang, and H. Liao, "A VR environment for human anatomical variation education: Modeling, visualization and interaction," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 391–403, 2024.
- [3] M. F. Deering, "HoloSketch: A virtual reality sketching/animation tool," *ACM Trans. Comput.-Human Interact.*, vol. 2, no. 3, pp. 220–238, 1995.
- [4] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio, "VRS-sketchin: Exploring the design space of pen and tablet interaction for 3D sketching in virtual reality," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–14.
- [5] J. H. Shuhaiher, "Augmented reality in surgery," *Arch. Surg.*, vol. 139, no. 2, pp. 170–174, 2004.
- [6] L. Besançon, M. Sereno, L. Yu, M. Ammi, and T. Isenberg, "Hybrid touch/tangible spatial 3D data selection," *Comput. Graph. Forum*, vol. 38, no. 3, pp. 553–567, 2019.
- [7] M. D. B. Machuca, W. Stuerzlinger, and P. Asente, "Smart3DGuides: Making unconstrained immersive 3D drawing more accurate," in *Proc. 25th ACM Symp. Virtual Reality Softw. Technol.*, 2019, pp. 1–13.
- [8] C. R. Spitzer and C. Spitzer, Eds., *Digital Avionics Handbook*, 1st ed., CRC Press, 2000. [Online]. Available: <https://doi.org/10.1201/9781420036879>
- [9] Y. Gao, Y. Liu, D. Cheng, Y. Wang, and S. Optoelectronics, "A review on development of head mounted display," *J. Comput.-Aided Des. Comput. Graph.*, vol. 28, no. 6, pp. 896–904, 2016.
- [10] H. Hua, "Enabling focus cues in head-mounted displays," *Proc. IEEE*, vol. 105, no. 5, pp. 805–824, May 2017.
- [11] N. Zhang, T. Huang, X. Zhang, and H. Liao, "35.1: A novel in-situ interactive 3D floating autostereoscopic display system with aerial imaging plate," in *Proc. SID Symp. Dig. Tech. Papers*, 2021, vol. 52, pp. 244–247.
- [12] H. Liao, T. Inomata, I. Sakuma, and T. Dohi, "3-D augmented reality for MRI-guided surgery using integral videography autostereoscopic image overlay," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1476–1486, Jun. 2010.
- [13] N. Zhang, T. Huang, X. Zhang, C. Hu, and H. Liao, "Omnidirectional 3D autostereoscopic aerial display with continuous parallax," *J. Opt. Soc. Amer. A*, vol. 39, no. 5, pp. 782–792, 2022.
- [14] A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–16, 2017.
- [15] T. Inoue and Y. Takaki, "Table screen 360-degree holographic display using circular viewing-zone scanning," *Opt. Express*, vol. 23, no. 5, pp. 6533–6542, 2015.
- [16] D. L. MacFarlane, "Volumetric three-dimensional display," *Appl. Opt.*, vol. 33, no. 31, pp. 7453–7457, 1994.
- [17] R. Balakrishnan, G. W. Fitzmaurice, and G. Kurtenbach, "User interfaces for volumetric displays," *Computer*, vol. 34, no. 3, pp. 37–45, 2001.
- [18] D. Keefe, R. Zeleznik, and D. Laidlaw, "Drawing on air: Input techniques for controlled 3D line illustration," *IEEE Trans. Visual. Comput. Graph.*, vol. 13, no. 5, pp. 1067–1081, Sep./Oct. 2007.
- [19] S.-H. Bae, R. Balakrishnan, and K. Singh, "ILoveSketch: As-natural-as-possible sketching system for creating 3D curve models," in *Proc. 21st Annu. ACM Symp. User Interface Softw. Technol.*, 2008, pp. 151–160.
- [20] R. C. Zeleznik, K. P. Herndon, and J. F. Hughes, "SKETCH: An interface for sketching 3D scenes," in *Proc. ACM SIGGRAPH 2006 Courses*, 2006, Art. no. 9-es.
- [21] D. F. Keefe, D. A. Feliz, T. Moscovich, D. H. Laidlaw, and J. J. LaViola Jr, "CavePainting: A fully immersive 3D artistic medium and interactive experience," in *Proc. Symp. Interactive 3D Graph.*, 2001, pp. 85–93.
- [22] 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 *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2018, pp. 1–15.
- [23] Y. Jiang et al., "Handpainter-3D sketching in VR with hand-based physical proxy," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2021, pp. 1–13.
- [24] S. Tsang, R. Balakrishnan, K. Singh, and A. Ranjan, "A suggestive interface for image guided 3D sketching," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2004, pp. 591–598.
- [25] E. Yu, R. Arora, T. Stanko, J. A. Bærentzen, K. Singh, and A. Bousseau, "Cassie: Curve and surface sketching in immersive environments," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2021, pp. 1–14.
- [26] P. Saalfeld, A. Stojnic, B. Preim, and S. Oeltze-Jafra, "Semi-immersive 3D sketching of vascular structures for medical education," in *Proc. VCBM/MedViz*, 2016, pp. 123–132.
- [27] P. Saalfeld, S. Oeltze-Jafra, S. Saalfeld, U. Preim, O. Beuing, and B. Preim, "Sketching and annotating vascular structures to support medical teaching, treatment planning and patient education," in *Proc. Eurographics (Dirk Bartz Prize)*, 2017, pp. 5–8.
- [28] F. Heckel, J. H. Moltz, C. Tietjen, and H. K. Hahn, "Sketch-based editing tools for tumour segmentation in 3D medical images," *Comput. Graph. Forum*, vol. 32, no. 8, pp. 144–157, 2013.
- [29] T. Ijiri and H. Yokota, "Contour-based interface for refining volume segmentation," *Comput. Graph. Forum*, vol. 29, no. 7, pp. 2153–2160, 2010.
- [30] S. Johnson, B. Jackson, B. Tourek, M. Molina, A. G. Erdman, and D. F. Keefe, "Immersive analytics for medicine: Hybrid 2D/3D sketch-based interfaces for annotating medical data and designing medical devices," in *Proc. ACM Companion Interactive Surfaces Spaces*, 2016, pp. 107–113.
- [31] J. Brooke, "SUS: A quick and dirty usability scale," in *Usability Evaluation Industry*, P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, Ed. London, U.K.: CRC Press, 1996, pp. 4–7.
- [32] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Advances in Psychology*, vol. 52. Amsterdam, The Netherlands: Elsevier, 1988, pp. 139–183.
- [33] C. Lin et al., "Three-dimensional visualization technology used in pancreatic surgery: A valuable tool for surgical trainees," *J. Gastrointestinal Surg.*, vol. 24, no. 4, pp. 866–873, 2020.
- [34] T. Itamiya et al., "A novel anatomy education method using a spatial reality display capable of stereoscopic imaging with the naked eye," *Appl. Sci.*, vol. 11, no. 16, 2021, Art. no. 7323.
- [35] S.-H. Kong et al., "Comparison of two-and three-dimensional camera systems in laparoscopic performance: A novel 3D system with one camera," *Surg. Endoscopy*, vol. 24, pp. 1132–1143, 2010.