Displays (2024)

Contents lists available at ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa

Painting++: Human-Computer Collaborative Painting in VR with Multisensory Interaction

Zhuoshu Li[a], Pei Chen[a,b,*], Hongbo Zhang[a], Yexinrui Wu[a], Xuanhui Liu[c], Lingyun Sun[a]

[a]*College of Computer Science and Technology, Zhejiang University, Hangzhou, 310027, China*
[b]*Zhejiang-Singapore Innovation and AI Joint Research Lab, Hangzhou, 310027, China*
[c]*Hangzhou City University, Hangzhou, 310015, China*

ARTICLE INFO

Communicated by

Keywords:
Virtual Reality
Multisensory interaction
Human-AI collaboration
VR painting
Human-computer collaboration

ABSTRACT

Extensive human-computer collaborative painting systems have been explored to assist people in ideating and creating the content of paintings. Recently, the popularity of VR has opened up the novel possibility to facilitate painting by providing an immersive virtual environment. To create immersion, previous VR painting systems mainly focused on simulating real painting environments. However, our preliminary study revealed that the lack of multisensory interaction in existing VR painting applications detracted people from the fully immersive experience. Grounded on the findings of the preliminary study, we developed Painting++, a human-computer collaborative painting system that synthetically integrates visual, audio, haptic, and smell feedback to enhance the sense of immersion. Particularly, the content generation ability of AI is also adopted to assist users in obtaining completed stereo paintings from rough sketches. Painting++ has been evaluated in a comparative study with two mainstream VR painting applications (N=18). The results illustrate that multisensory interaction and AI assistance make Painting++ succeed in creating highly immersive painting experience. Our work further reveals distinct VR painting systems' impact on painting behavior and demonstrates the multiple roles that Painting++ plays in supporting human-computer collaborative painting.

1. Introduction

Painting serves as a potent medium for individuals to explore and express their thoughts through visual representation [1]. With the advent of computer technology, painting has transitioned into the digital era. While the traditional role of the computer was to provide digital painting tools, the recent advancements in Artificial Intelligence (AI) have elevated computers to the status of collaborators in the painting process [2, 3]. For instance, Creative Sketching Partner effectively stimulated ideation by generating sketches similar to user input [4], and DuetDraw facilitated real-time collaborative painting between the user and AI [5].

While AI predominantly assists in producing the content of paintings, Virtual Reality (VR) as a computational collaborator could support painting by providing an immersive virtual environment (VE). In this paper, we define human-computer collaboration as the mutual engagement of human and computers in achieving a shared goal through coordinated efforts [6]. In the context of painting, computational collaborators include not only those who paint together with people like AI does, but also those who provide novel stimuli to induce creativity [7] as VR does. Studies have demonstrated that immersive VR could alleviate stress [8], improve task performance [9], and promote creative thinking [10]. To create immersion, existing VR paint-

*Corresponding author: chenpei@zju.edu.cn

http://dx.doi.org/10.1016/j.displa.xxxx.xx.xxx
Received ; Received in final form ; Accepted ; Available online
0141-9382/

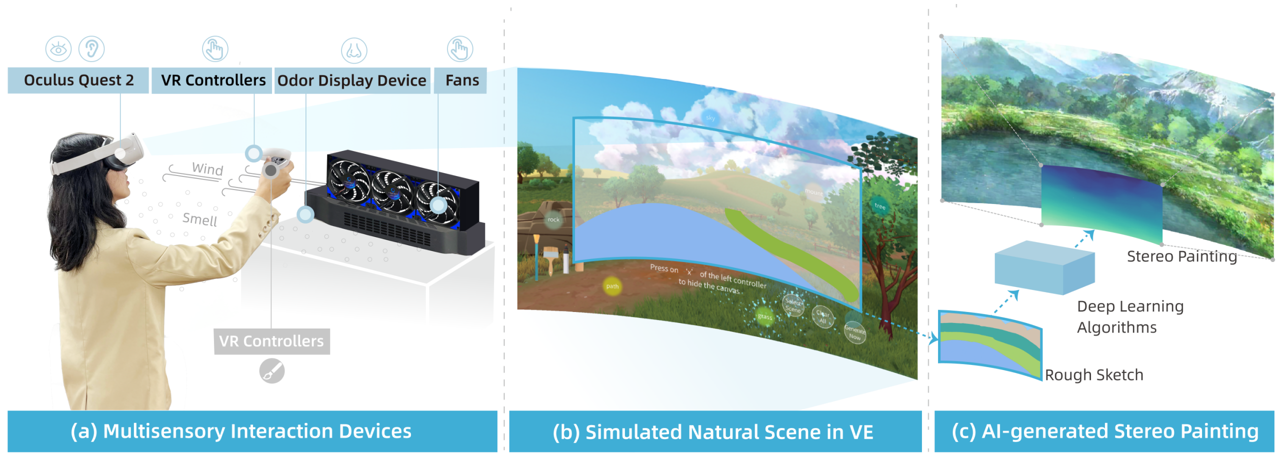

Fig. 1. Painting++ is a human-computer collaborative painting system based on multisensory VR and AI. (a) The system adopts various devices to provide visual, audio, haptic, and smell feedback, so as to create immersion for users. (b) The simulated natural scene immerses users in VE and fosters their creativity. (c) AI's capability is adopted to help users obtain stereo paintings from casual 2D sketches.

ing systems such as Painting VR [11] and Vermillion [12] transported users to a virtual painting studio, where they dove into the painting process as authentically as they would in the physical world.

Despite the existing research, the potential of immersive VR in facilitating painting has not been fully exploited. According to prior research, immersion augmented by multisensory interaction can effectively activate people's multiple brain areas and physical senses [13, 14, 15], which is conducive to promoting creative activities [16]. For example, earlier work combined vision, audition, touch, and olfaction in VE to aid students in exploring creative ideas [17]. However, in the realm of existing VR painting systems, the incorporation of multisensory interactions remains limited, impeding users from experiencing painting in its fullest immersive potential.

In this context, we aim to explore how to encompass multisensory interaction within VR painting systems for enhancing immersion and facilitating idea expression. To achieve this, we first conducted a preliminary study involving two commercial VR painting applications (Painting VR and Tilt Brush) with 12 participants. The preliminary study aimed to deepen our understanding of existing VR painting applications and explore strategies for integrating multisensory interaction into VR painting. Our findings revealed four design implications for an immersive painting system, including sensory interaction with scenes, external painting assistance, scenes for inspiration, and simplified painting interaction.

Grounded on the above findings, we designed and developed Painting++, a human-computer collaborative painting system that provides users with an immersive VR painting experience through multisensory interaction. As depicted in Fig. 1, the system seamlessly integrated visual, audio, haptic, and smell feedback to immerse people in the natural scene and invigorate their creativity. In addition, we applied a conditional generative adversarial network (GAN) [18] to generate high-quality stereo cartoon paintings based on people's rough 2D sketches. To validate the usability of Painting++ and whether it effectively provides immersion, we conducted an evaluation involving 18 participants. The subjective results indicated that Painting++ outperformed the existing VR painting applications in most usability factors, largely due to the AI assistance it provides. Our multisensory system not only inspired participants but also facilitated a deeply immersive painting experience. We further utilized behavior coding for objective evaluation, revealing Painting++'s advantages in usability and immersion based on participants' behavior pattern and analyzing how distinct VR painting applications influenced participants' painting behavior. The main contributions of this paper are as follows:

- We performed an empirical analysis of existing VR painting applications, which provides implications for the design of an immersive VR painting system.
- We presented a human-computer collaborative system named Painting++, which allows people to paint collaboratively with multisensory natural scenes and AI. Painting++ immerses users in painting and helps them produce pleasant paintings.
- We conducted an evaluation that demonstrated the effectiveness of multisensory interaction in supporting immersive painting experience and revealed the design opportunity to incorporate AI into VR painting.
- We summarized the multiple roles that Painting++ played in painting as well as the advantages and limitations of our system, which could hopefully inspire future research on human-computer collaborative painting.

2. RELATED WORK

2.1. Human-computer Collaborative Painting

Recently, human-computer collaborative painting systems have been extensively explored. Based on previous research that delved into supporting human-computer collaborative activities [19, 20], this section reviews various collaboration modes in existing painting systems.

A branch of research seeks to inspire users by showcasing supportive examples. For instance, systems such as IdeaExpander presented sketches related to the design task as inspiration [21], Creative Sketching Partner applied deep learning models to generate inspiring sketches according to users' input [4], and the work of Tholander and Jonsson generated concept illustrations to inspire kids [22]. Furthermore, ImageCascade provided images and generated color palettes to support ideation [23]. Beyond virtual painting collaborators, embodied robots such as Cobbie helped designers ideate iteratively by drawing creative sketches on paper [24].

Another type of systems created and improved the content of paintings in collaboration with users. Several computational collaborators could complete paintings with users in real time [25, 5]. For example, Drawing Apprentice analyzed users' input using object recognition algorithms and responded on a shared canvas [26]. For individuals not skilled at painting, systems such as StoryDrawer [27] applied text-to-image generation models to transform users' descriptive words into paintings, and Artverse supported painting generation through users' action, vision, and linguistic interaction [3]. Focusing on AI-based visual art creation, Shahriar further summarized existing human-AI collaborative painting systems [28]. Prior work also assisted users in colorizing paintings by analyzing color combinations [29, 30] or target styles [31], while systems like ShipShape [32] attempted to beautify and smoothen users' strokes.

Meanwhile, the recent popularity of VR has transformed human-computer collaborative painting. Instead of contributing to the content of paintings, VR supports painting by creating an immersive and inspiring VE. As indicated by prior work, immersive VR helps users reach a special state where they are able to dive into the creative process [33]. However, the potential of immersive VR in empowering painting has not been thoroughly explored yet. Therefore, we seek to further harness the immersive experience brought by VR to facilitate painting.

2.2. Multisensory Interaction in Immersive VR

VR provides users with a computer-simulated synthetic environment, opening up new possibilities for fields such as games [34], sports [35], and education [36]. Recently, more VR-related research paid attention to immersion enhancement [37, 38], and providing sensory cues has been verified effective in improving immersion [39].

Stereoscopic vision using head-mounted displays (HMDs) [40] and spatial audio based on space rendering [41, 42] have been widely used in VR. Besides, the touch and olfaction augmentation methods have been extensively explored. Haptic feedback enhances physical presence, emotional scale, and user interaction [43], have been integrated with VR to enrich gaming interactions [44], improve the design expression of visual artists [45], and enhance the perception of environment by simulating haptic sensations such as vibration and wind [46, 47]. In addition, smell feedback has been proven to be related to users' attention and immersion [48], commonly delivered through fans [49, 50], tubes [51], or wearable devices [52, 53]. Adaptive smell feedback has also been explored to enhance sensory engagement within interactive experiences [54].

Beyond providing a single sense channel, recent research has suggested the approach of achieving full immersion in VR by synthesizing multiple senses [55]. Multisensory interaction can activate the brain [13, 15] and facilitate bodily engagement [14]. Taking advantage of this, multisensory interaction has been incorporated into many scenarios, such as creating memorable experiences in tourism and marketing [56], increasing the expressiveness of visual art [57], and enhancing situated learning in education [38]. Multisensory interaction is not a simple superposition of several senses. Instead, special attention should be paid to the coordination and consistency of different senses [58]. For example, Season Traveler customized scents and haptic feedback for different natural scenes [54]. However, while existing practices have consciously incorporated haptic feedback in painting through VR controllers, physical proxies such as pens and tablets [59, 60], and wearable gloves [61], we barely see existing research integrate multiple sensory channels to enhance immersion in the field of VR painting. This gap motivates us to develop a multisensory VR system that facilitates immersive painting experience.

2.3. Creativity Work and Painting Experience in VR

Augmented by multisensory interaction, immersive VR has a positive effect on individual's creativity [16]. In VR Rehearse & Perform, musicians were transported to rehearsal rooms with pleasant music in VR, which facilitated their creative improvisation [62]. In addition, prior work indicated that a VE combining vision, audition, touch, and olfaction helped students explore creative ideas [17] and natural scenes with sound effects foster creativity in brainstorming and design [63].

As a form of art, painting also bares several unique advantages when combined with VR. First, VE provided in VR allows users to paint without restrictions of place or time. For example, VR can reproduce reality and even create novel perceptual dimensions that have never existed [64, 65, 66]. Second, VR expands the painting dimensions to 3D space, enabling individuals to create 3D objects directly through 2D strokes. Similar to Tilt Brush [67], some studies also supported users to create 3D artworks in VE [68, 69, 61, 70, 71]. Furthermore, painting in VR brings enjoyment to users owing to the unique interactive details. For instance, previous work adopted gesture-based input [61] and transformed local scenic images into colors [72] for digital artwork. Additionally, a VR-embedded tutorial system was developed to enhance users' understanding of three-dimensional painting details [73].

While the aforementioned studies demonstrate VR's value in enriching painting experience, addressing existing limitations is crucial for better collaboration between users and VR painting systems. First, users find it challenging to precisely control stroke shapes in 3D space [69, 74]. To alleviate the difficulty of 3D mid-air painting, previous work has explored solutions such as providing a physical drawing interface [75], optimizing the input algorithm [59, 76], and using tactile feedback to enhance rendering accuracy [73, 77]. Another potential solution is to beautify strokes or paintings with AI assistance [78, 79], similar to what Calliope achieved, enabling designers to explore and manipulate generative design solutions in real time manner [80]. Despite this, these systems primarily focused on the aesthetic quality of the painting output, seldom addressing subjective aspects such as users' sense of immersion and their holistic experience during the painting process. Secondly, as stated in Section 2.2, there is ample opportunity for existing VR painting systems to enhance immersion by integrating multiple sensory channels.

3. PRELIMINARY STUDY

3.1. Methods

To enhance users' immersion during human-computer collaborative painting in VR, the preliminary study was guided by two research questions: (1) what contributes to and detracts from users' immersion in existing VR painting applications? and (2) how to integrate multisensory interaction in VR painting systems for immersive experience?

3.1.1. Participants

We recruited participants via social media and employed a pre-screen based on applicants' interest in painting. Twelve participants (P1-P12; 6 males and 6 females; aged 21-26 years old) who were interested in painting took part in the preliminary study. Seven participants had prior experience using the VR headset while the others did not. All participants signed informed consent forms and were compensated monetarily.

3.1.2. Procedure

During the study, each participant used Painting VR and Tilt Brush respectively on an Oculus Quest 2 headset. Painting VR provides a vast virtual studio with various painting tools and 2D canvases. Tilt Brush allows users to paint in 3D space with dynamic brushes. These applications were chosen because (1) they are among the most popular commercial VR painting applications with relatively comprehensive functionality; (2) their efforts to create immersion for users align with our goals; (3) their target users include both painting enthusiasts and expert painters, similar to our participant profile.

For each application, we introduced its main functions and interactions, and participants were given 10 minutes to paint with the application. Participants were encouraged to freely explore the functionality of the application, and they could ask questions anytime if they had trouble using the application or the VR headset. Participants were specifically instructed to consider their collaboration with the system and their sense of immersion while using the applications. Afterward, participants were asked to describe their experience of using the application with examples. Follow-up questions were asked based on their responses.

After participants used two applications, a more comprehensive semi-structured interview was conducted. The main questions include: (1) Can you compare the immersion levels of the two applications? (2) What multisensory feedback do you expect from VR painting applications (e.g., haptic and smell feedback)? (3) What additional support do you wish existing applications would offer? These questions were adjusted flexibly based on participants' responses.

3.1.3. Thematic Analysis

Two coders (co-authors of this paper) analyzed the interview results through a hybrid process of deductive and inductive thematic analysis [81]. The recordings of the interview were transcribed into text, and the two coders repeatedly read the transcripts to acquaint themselves with the interview data. With the two research questions serving as predefined themes, the coders independently assigned different statements to these themes and manually developed a set of initial codes. The coding process strictly pertained to the semantic content of the interview data instead of the coders' assumptions. For example, the statement "*I think it was very close to the real painting scene*" (P5) was coded as "*realistic scenes*". Then, the two coders resolved disagreements by discussion and finally obtained eight codes: *realistic scenes*, *visual effect*, *expanded painting dimensions*, *complex interactions*, *restricted scenes*, *lack of sensory feedback*, *painting process*, and *painting environment*.

3.2. Findings

This section first summarizes the factors related to immersion, then introduces findings regarding the integration of multisensory interaction in VR painting systems, and finally, we draw conclusions about the design implications for an immersive VR painting system. We support our discussion with illustrative quotes, where "P1" represents quotes from the first participant.

3.2.1. Factors Concerning Immersion

For factors enhancing immersion, participants appreciated the *realistic scenes* of Painting VR and the *visual effect* of Tilt Brush. Participants felt that Painting VR "*simulated the painting scene in real life*" (P5, P11), which gave them "*the feeling of being there*" (P2, P10) and made them feel like artists. In Painting VR, participants could select brushes and mix colors as they did in real life. For Tilt Brush, "*the vivid brushes inspired me to create wonderful paintings*" (P9). In the infinity 3D space of Tilt Brush, participants could "*walk around to enjoy my painting from different views*" (P6).

However, the *expanded painting dimensions* in VR presented challenges for participants to paint concentratedly and decreased immersion. While P2 appreciated the novelty of painting in VR because it subverted the traditional methods, some participants found"*it was struggling to control strokes in 3D space*" (P3) as they were not accustomed to 3D mid-air painting. In this context, participants tended to think about "*how to obtain the shapes I wanted*" (P12) instead of immersing themselves in the painting process. Despite being able to paint on a 2D canvas in Painting VR, some participants still found it difficult to create satisfactory paintings. They felt that painting with VR controllers lacked the precision of using pens and brushes. Furthermore, three participants mentioned that they hoped to "*receive assistance from external sources, such as remote collaborative partners or the assistance of generative AI.*"(P3, P10, P11).

Moreover, the *lack of sensory feedback* made participants feel like "*painting on a browser instead of in VR*" (P3), and "*I could not interact with the environment or my painting*"

(P6). Although Painting VR provided haptic feedback when the brush touched the virtual canvas, P11 thought that "*it was not enough for me to immerse in the painting process. I felt that there was almost no audio feedback.*" For Tilt Brush, participants noted that the absence of feedback corresponding to the input operation negatively influenced immersion. P6 stated that "*although I was painting in a 3D space, I could not gain the feeling of presence through other senses such as touch*". Participants also expressed disappointment in the lack of sensory feedback from the scene.

In addition, other factors such as *complex interactions* and *restricted scenes* also slightly decreased immersion. In Painting VR, the process of selecting brushes aims to replicate the real-world experience of picking up a brush. However, as highlighted by P4 and P10, mastering this interaction took time. Furthermore, according to P1 and P6, "*the brushes usually fell accidentally and could not be picked up again*", leading to frustration among some participants. Regarding the scene, participants thought that "*current VR painting applications offered relatively limited scenes*" (P3, P8), which restricted their inspiration. Specifically, P11 mentioned that "*when I put on the VR headset and faced with the blank space in Tilt Brush, I struggled to decide what to paint*". P2 also suggested that "*it would be nice to have a river or other natural scenes in front of me when I was painting*".

To summarize, the findings were consistent with our assumption that the *lack of sensory feedback* somewhat reduced immersion while also revealed additional factors influencing immersion. Existing VR painting systems showcased how *realistic scenes* and *visual effect* brought immersion, which could be also integrated into our system. To further increase immersion, we have to assist users in overcoming challenges posed by the *expanded painting dimensions* and *complex interactions* through providing more painting assistance or simplifying the interaction. Specifically, inspired by various existing painting systems leveraging AI, we could incorporate AI's generative ability into VR to collaboratively create paintings with users. Additionally, the scenes in VR painting systems are supposed to inspire users.

3.2.2. Approaches for Integrating Multisensory Interaction

When it comes to integrating multisensory interaction in VR painting, participants expressed a desire for augmented sensory feedback during the *painting process*. Meanwhile, some participants longed for richer sensory stimuli from the *painting environment*. For the painting process, P2 mentioned that "*when I painted with the snow brush, I expected there to feel the sensation of snow or the sound of snow falling*". In this regard, Painting VR provided richer feedback, with participants noting that the haptic and audio feedback during painting made them feel like they were painting on a real canvas (P7, P12). Furthermore, participants desired to interact with the scene through different senses. While painting in VR, they not only conveyed their artistic ideas but also "*released pressure in the virtual environment*" (P4, P10, P11). Compared with painting in the real world, P12 figured out that VE shielded them from external distractions, and "*if there were some sound from the environment, it might be more immersive than painting on a pad*". Similarly, P3 stated that "*having sound and natural scents would create a more relaxing atmosphere*".

3.2.3. Implications for System Design

Based on the above findings, we derive the implications for the system design as follows:

- Sensory interaction with scenes. Users desire the feeling of painting from life in VR, necessitating the system to facilitate interaction with the surroundings and offer multisensory feedback. (I1)
- External painting assistance. Users require external sources to refine their painting effects in VR and enhance the aesthetic appeal of their artwork. (I2)
- Scenes for inspiration. The prefabricated scenes provided by the system should be consistent with the content of the painting, from which users can gain painting inspiration. For instance, providing urban scenes when painting streetscapes. (I3)
- Simplified painting interaction. Simple and intuitive painting interaction is necessary for users to concentrate more on painting. (I4)

4. DESIGN AND IMPLEMENTATION

4.1. Design Consideration

In the preliminary study, participants painted on 2D canvas (2D painting) in Painting VR and painted in 3D space (3D painting) in Tilt Brush. In fact, the debate surrounding whether to apply 2D or 3D painting in VR has been ongoing.

Taking advantage of the expanded painting dimensions in VR, Tilt Brush facilitates the creation of stereo paintings and 3D models. However, participants also noted challenges such as difficulty in locating brushes and strokes, as well as the lack of a clear sense of distance between their painting and themselves. This finding aligns with previous study [75, 82], which indicated the increased complexity of 3D aerial painting in VR compared to 2D painting.

Compared to 3D painting, 2D painting aligns more with people's cognitive processes and habits. During the preliminary study, most participants viewed 2D painting as advantageous due to its lower learning curve and higher painting accuracy. However, those who preferred Tilt Brush mentioned that 2D painting might limit their ability to create 3D objects in VR. Additionally, some participants expressed that 2D painting lacked novelty in the setting of VR.

When designing Painting++, we carefully considered these factors and adopted 2D painting. We assume that getting accustomed to the subverted way of 3D painting may take time for users. More importantly, while VR offers a three-dimensional space, not all activities necessarily require a shift to 3D forms. A promising solution is to ensure people's performance on 2D tasks while integrating 3D interaction [83, 84]. Therefore, we plan to retain 2D painting functionality in VR while also simulating various natural scenes and enabling the creation of stereo paintings to provide a 3D experience.

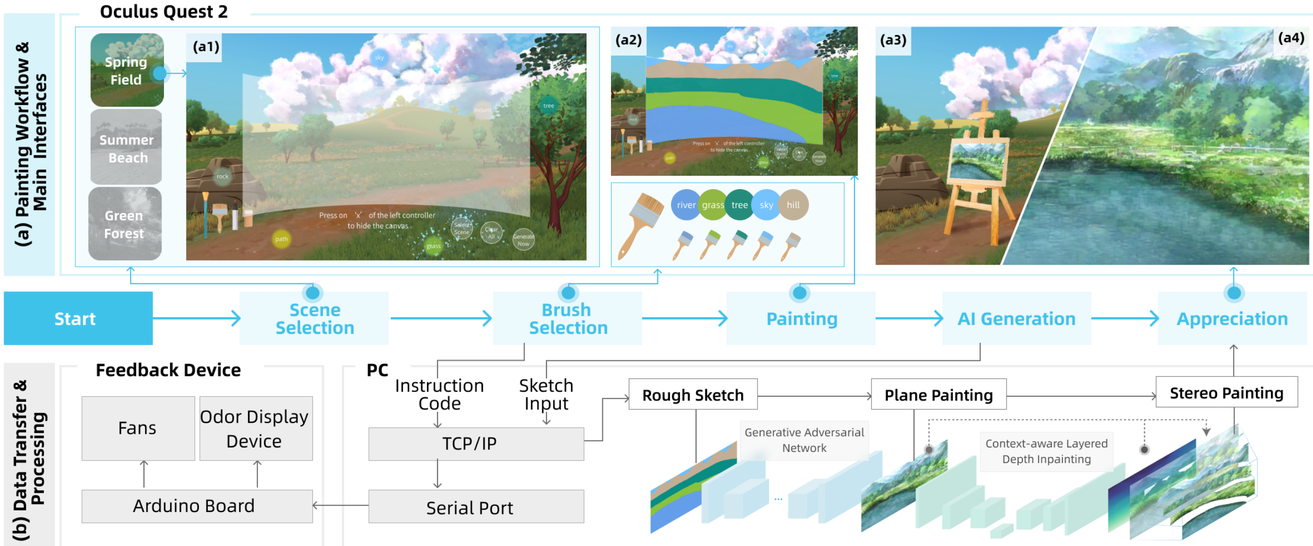

Fig. 2. (a) The painting workflow and main interfaces of Painting++. (b) The workflow of data transfer and data processing.

4.2. System Design

Motivated by the findings from the preliminary study, we designed a human-computer collaborative VR painting system named Painting++. It transports users to natural scenes, allowing them to create landscape paintings. The system provides visual, audio, haptic, and smell feedback during the whole painting workflow to produce a sense of immersion. Meanwhile, it incorporates deep learning algorithms to generate vivid stereo images from user input sketches intelligently, which greatly reduces the difficulty of painting in VR.

4.2.1. Painting Workflow

The painting workflow and main interfaces are illustrated in Fig. 2(a). Users start by coming to a natural scene with a large, translucent 2D canvas (see Fig. 2(a1)). We provide three prefabricated scenes including the field, forest, and beach, and users can freely switch between them. Each scene features several natural objects, including the tree, grass, stone, river, sky, hill, and path. Users can hide the canvas and walk around to enjoy the natural scene as an inspiration for landscape painting (I3).

We provide three brush types with different thicknesses to accommodate various painting preferences (see Fig. 2(a2)). Upon selecting a brush, users can pick colors from the natural scene. For instance, if a user intends to paint stones, they can point to a stone in the scene, and the brush color will adjust to a suitable gray shade for drawing stones (I4). This approach encourages users to interact with the natural scene, sparking creativity and motivation for painting (I1).

4.2.2. Multisensory Interaction Design

Our system incorporates multisensory feedback throughout the painting process, with different forms of feedback triggered by user behaviors. Specifically, to promote users' interaction with the natural scene, our system provides sensory feedback according to the selected object (I1). For instance, when users interact with grass, they experience the scent of freshly cut grass; when pointing to a tree, they will hear the sound of wind blowing and smell the woody fragrance. In addition, during the painting process, when the brush touches the canvas in VR, users receive haptic feedback through the controller, simulating the tactile sensation of brush and physical surface contact. Users can also continuously hear the sound corresponding to the object they are currently painting.

4.2.3. AI-assisted Painting Generation

To address the challenge of painting in VR, we seek to incorporate AI to enable collaborative painting (I2). Utilizing AI algorithms (detailed in Section 4.3), users can sketch out the semantic layout of their painting (see Fig. 2(a2)) to convey their ideas to AI, and then AI generates a cartoon stereo image with rich textures automatically. This approach empowers users, even those without formal training, to create engaging artworks in our system. We provide two ways for users to in depth appreciate the completed paintings. The first is to present the painting in the current scene (see Fig. 2(a3)), simulating a real-life painting experience where users can walk forth and back to observe it from different angles. Another way is a full-screen display (see Fig. 2(a4)) that allows users to observe the details of the painting and appreciate the scenery in it.

4.3. Implementation

Our system was developed using Unity 3D and compatible with the Oculus Quest 2 VR headset. To minimize physical strain on users, we primarily utilize desktop equipment to provide sensory feedback. Table 1 details the feedback of different objects and the employed devices. Audio feedback is delivered through the built-in headset speakers. The system leverages spatial audio technology provided by the Oculus Spatializer to enhance the immersive experience. Haptic feedback is achieved using controllers and three fans. The controllers vibrate during user interactions with the painting interface (e.g., selecting, painting), while the fans simulate wind sensations in

Table 1. Feedback of each object and the adopted feedback device. (A: Audio feedback, H: Haptic feedback, S: Smell feedback)

Object	Feedback	Device
Tree	Sound of wind blowing (A)	Headset
	Feeling of wind (H)	Fans
	Woody fragrance (S)	Odor display device
Grass	Odor of mown grass (S)	Odor display device
Stone	Sound of rolling stones (A)	Headset
River	Sound of river flowing (A)	Headset
	Damp smell of water (S)	Odor display device
Sky	Warm feeling (H)	Fans
	Smell of sunshine (S)	Odor display device
Hill	Birdsong (A)	Headset
	Woody fragrance (S)	Odor display device
Path	Footstep (A)	Headset

the natural scenes. Each fan has a 12cm blade capable of rotating at a maximum speed of 1000 revolutions per minute (RPM). The fans are controlled via a PWM (Pulse Width Modulation) signal, allowing precise control over their speed. Smell feedback is facilitated by a multi-channel odor display device that can emit distinct scents through separate tubes. The device uses piezoelectric pumps to control the flow of scented micro airflows, ensuring accurate and immediate delivery of smells to the user's nose. The system is designed to prevent odor mixing by employing a series of valves and airflow channels that direct the scents to the user while isolating each scent stream.

The AI generation of stereo images involves two stages, taking approximately 15 seconds in total. The first stage translates the user sketch into a plane cartoon image using a GAN model trained on a dataset consisting of 7,234 cartoon images with corresponding semantic annotations [85]. This model includes a painting producer for image-to-image translation and an edge synthesizer to produce an edge map based on the input sketch, providing details for painting generation. After training, the GAN model can generate cartoon images with appropriate textures based on the semantics and layouts of unseen user sketches. In the second stage, a context-aware layered depth inpainting algorithm converts these images into stereo images. It first utilizes the pre-trained MegaDepth [86] to produce a depth map for an input image and then learns multi-layer representations to synthesize novel views, ultimately producing the final stereo image [87]. The above algorithms run on a PC (Intel Core i9-11900K 3.50 GHz CPU, 64.0G RAM, NVIDIA GeForce RTX 3090 GPU) using Windows 10 OS.

Fig. 2(b) depicts the workflow of data transfer and processing. Information exchange between Oculus Quest 2 and the PC is realized via TCP/IP. All sensory feedback devices are connected to the PC through serial ports using a standard baud rate of 115200 bps, within a local area network (LAN) environment. When a user selects a color (representing a natural object), Oculus Quest 2 sends the corresponding instruction code to the PC via TCP/IP. The instruction code, encoded in JSON format, is processed by a custom-built parser developed in C within the Unity environment. The PC then utilizes an Arduino microcontroller to control the sensory devices. To ensure consistent multisensory feedback, the PC synchronizes all actions using a real-time clock and a feedback control loop, which monitors the status of each device and adjusts the commands dynamically. When the painting is completed, users can click the "AI Generation" button, sending a high-priority TCP/IP packet from the Oculus Quest 2 to the PC. The PC processes the sketch using the aforementioned algorithmsand the generated stereo painting is sent back to Oculus Quest 2 for users to view in VR.

5. EVALUATION

We conducted the evaluation with three main goals: (1) Comparing the usability and immersion of Painting++ with other mainstream VR painting applications (Section 6.1.1- 6.1.2). (2) Exploring the influence of multisensory interaction on immersion and participants' satisfaction of each provided sense (Section 6.1.3). (3) Gaining insights into Painting++'s impact on users' painting behavior and its roles in human-computer collaborative painting (Section 6.2- 7.1).

5.1. Participants

Based on the same pre-screening standard as the preliminary study, we recruited 18 participants (P13-P30, 8 males and 10 females, aged 20 to 27 years old) via social media. None of them overlapped with participants from the preliminary study. All participants were interested in painting, and 11 participants were skilled painters. Two-thirds of them had experience using VR headsets, while the others did not. Each participant confirmed the informed consent and received compensation.

5.2. Procedure

The evaluation procedure is outlined in Fig. 3. During the study, each participant wore an Oculus Quest 2 headset to use three VR painting applications (i.e., Painting VR, Tilt Brush, and Painting++) in shuffled orders based on the balanced Latin square to avoid order effects. The setting of the evaluation apparatus is shown in Fig. 4. Participants received brief instructions for each application and had 10 minutes to explore its features

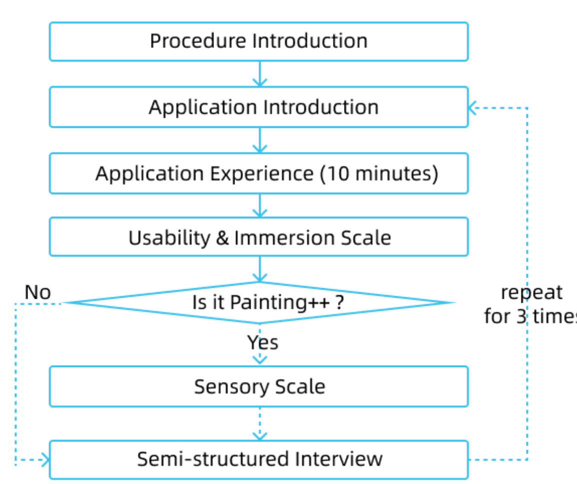

Fig. 3. The procedure of the evaluation.

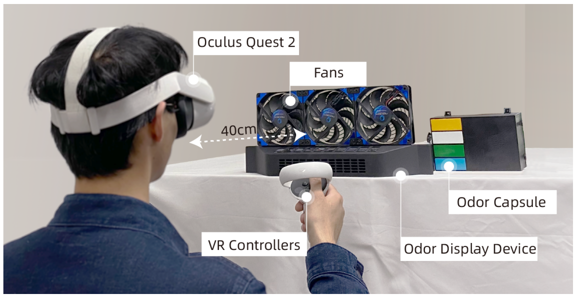

Fig. 4. The setting of the evaluation apparatus.

Table 2. Statements of factors in the usability and immersion scale.

Factor	Statement
Ease of use	Whether it takes physical or mental effort to use the system for painting.
Flexibility	Whether the system supports users to paint what they want freely.
Functionality	Whether the system could provide the expected function for painting.
Effectiveness	Whether the quality of the painting outcome is satisfying.
CF	Whether users could take the initiative to interact with the system.
DF	Whether the system allows users to intently paint without being distracted by external factors.
SF	Whether the system appropriately encompasses multisensory feedback.
RF	Whether the painting experience is similar to that in real world.

without thematic restrictions, and the whole process was videotaped. Participants could ask if they encountered any trouble using the application or the VR headset. After that, participants needed to fill out a seven-point Likert scale related to the usability and immersion of the application.

The usability scale evaluated aspects including the ease of use, flexibility, functionality, and effectiveness of applications, while the immersion scale estimated users' immersion in terms of four factors: control factor (CF), distraction factor (DF), sensory factor (SF), and realism factor (RF). Each factor contained several underlying questions, which were distilled from widely-used questionnaires [88, 89] and tailored to the context of VR painting. For each factor, we calculated the average score of all questions for the quantitative analysis. Table 2 provides brief descriptions of all factors, while detailed questions can be found in the Appendix. In the following semi-structured interview, participants were asked to describe their experience using the application and provide specific instances. Follow-up questions were posed based on their responses. Notably, for Painting++, participants completed an additional seven-point Likert scale to rank each sensory channel's impact on immersion and report their satisfaction with each sense.

5.3. Analysis

For subjective measures, a Shapiro-Wilk test [90] indicated a significant departure from normality distribution for most scores, so we applied the Friedman test to compare the three

Table 3. Coded types of behavior and the corresponding sub-behaviors.

Behavior	Introduction	Sub-behavior
Observation	Look around and no actual operation	Observe environment (OE), observe painting (OP)
Selection	Select something and prepare for painting	Select brush (SB), select color (SC), select eraser (SE), select other tools (SO)
Painting	Create or change strokes	Paint (P), erase (E)

VR painting applications with a significant level of 5% (0.05). When significant differences were found, we conducted pairwise comparisons to determine the variations between each pair. To explain participants' quantitative feedback, interview recordings were transcribed into text. One co-author assigned each statement to one or more factors in the scales. For example, the statement "*the audio feedback was so realistic that it gave me a sense of presence*" (P13) was assigned to both SF and RF. Subsequently, another co-author reviewed the assignment and resolved disagreements by discussing with the first co-author. Furthermore, we incorporated participants' feedback to comprehensively compare the interaction and technical details of the three systems from various aspects, including painting interaction, painting environment, auxiliary functionality, and multisensory interactions. We also analyzed how these aspects impact factors including usability and immersion.

In addition, we applied behavior coding to analyze the performance of Painting++ in an objective way. Specifically, we coded the screen recordings to delve into participants' behavior throughout the whole painting process, deepening our understanding of the systems' usability and immersion, as well as users' collaboration with different VR painting systems. To begin with, two co-authors independently watched the recordings and collaboratively established the coding rules. Following discussions, three primary types of behavior were determined: observation, selection, and painting, each encompassing specific sub-behaviors (see Table 3). Subsequently, one co-author annotated all video recordings based on the behavior types. The other co-author went through the annotations and refined them together with the first co-author. Based on the annotated recording, we further (1) quantified data, such as the percentage of each behavior and the duration/frequency of each sub-behavior, to alternatively reveal the usability and immersion; (2) analyzed participants' behavior sequences and identified self-correcting behaviors. In our analysis, self-correcting behavior referred to users repeatedly performing selection actions without engaging in actual painting, which could also be considered as an indicator of usability [91]. Behavior sequences were also visualized to help understand participants' behavior patterns when interacting with different systems.

6. RESULTS

6.1. Subjective Results

6.1.1. Usability

The quantitative analysis (see Fig. 5) revealed significant differences in ease of use ($\chi^2(2) = 24.889$, $p < .001$), functionality ($\chi^2(2) = 21.556$, $p < .001$) and effectiveness ($\chi^2(2) = 15.483$, $p < .001$). Post hoc comparisons indicated that Painting++ and Painting VR had significant differences in ease of use and functionality, and the effectiveness of Painting++ significantly surpassed both Painting VR and Tilt Brush. The flexibility of the three applications had no significant difference ($\chi^2(2) = 2.338$, $p = .311$).

The simplified interaction of Painting++ made it significantly outperform Painting VR in ease of use. In Painting++, participants could directly point to the color or brush and select it, which was "*easy for VR novice to get started with painting*" (P21). In contrast, participants found the interaction of Painting VR "*so realistic that it took me much effort to pick up the brush*" (P28) or "*wash the brush*" (P29). However, both Painting++ and Painting VR had the issue of misoperation, such as accidental brush switching (e.g., P22, P25). In this aspect, participants thought the interaction of Tilt Brush was simple and intuitive, although they needed some time to adapt to get used to 3D mid-air painting. Additionally, more than half of the participants preferred painting on the 2D canvas in Painting++, as it aligned better with their cognition and habits, according to both skilled painters and novices.

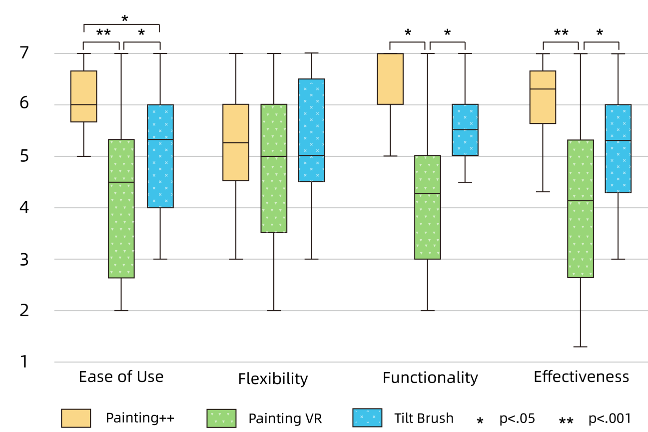

Fig. 5. Results of the usability scale.

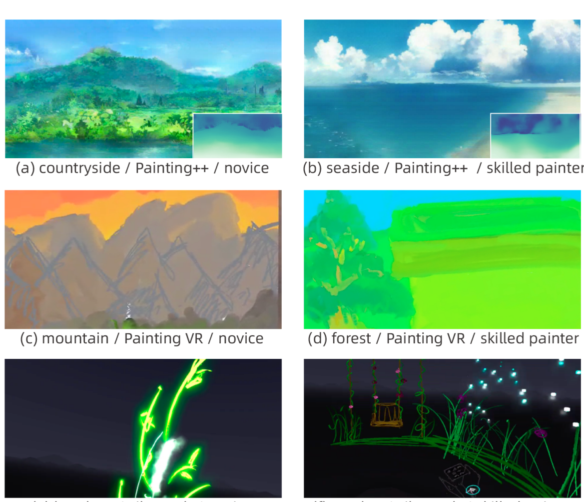

Fig. 6. Paintings created by participants in Painting++, Painting VR, and Tilt Brush. The caption below each painting shows the theme of the painting / the used application / expertise of the participant.

Regarding flexibility, feedback from participants indicated the need for greater flexibility in all three applications. For Painting++, participants noted that the limited color options somewhat restricted their creative freedom, as expressed by P13: "*I could only choose the color from the given scene, and it did not allow me to draw a dog or a cat*" (P13). Apart from that, P20 and P28 reported difficulty in obtaining desired colors in Painting VR by simply mixing the provided colors. In comparison, as stated by P22 and P26, Tilt Brush provided a wider range of brushes, scenes, and auxiliary tools. However, "*its painting style was somewhat limited to the cyber style*" (P30).

Participants praised the functionality and effectiveness of Painting++ due to its AI assistance. We present several paintings created by participants in the three applications in Fig. 6 to illustrate the quality of paintings. Both novices and skilled painters achieved high-quality paintings in Painting++ (Figs. 6(a) and 6(b)). Conversely, even skilled painters faced challenges completing paintings in Painting VR and Tilt Brush (Figs. 6(c)- 6(f)). On the question "*I can create better paintings in this application than in reality*", Painting++ scored an average of 5.72 ($SD = 1.28$), while Painting VR and Tilt Brush scored 3.67 ($SD = 1.73$) and 5.33 ($SD = 1.37$), respectively. P16 and P23 mentioned that "*for an unprofessional painting lover like me, the collaboration with AI helped me obtain a better painting with vivid details*". In particular, P20 described that "*AI enriched the details of my painting, and I could see the leaves of the trees*". Regarding Painting VR, participants thought that "*it only provided the tools for painting, but did not give me any further assistance in painting*" (P21). As a result, some participants found the painting outcome of Painting VR somewhat disappointing. As for Tilt Brush, P16 and P18 appreciated its captivating brush effects that enhanced their paintings to some extent, although creating ideal shapes and items in 3D space remained challenging.

6.1.2. Immersion

For immersion, the quantitative results indicated significant differences in DF ($\chi^2(2) = 8.100$, $p = .017$), SF ($\chi^2(2) = 14.282$, $p = .001$), and RF ($\chi^2(2) = 14.812$, $p = .001$). Since the immersion of Painting VR and Tilt Brush has been analyzed in Section 3.2, this section mainly focuses on participants' feedback on Painting++ (see Fig. 7).

Although there was no significant difference in CF, Painting++ had the highest mean score among the three applications. For the question "*How much were you able to control the effect of painting?*", most participants agreed that Painting++'s AI generation capability guaranteed the quality of their paintings. In addition, interaction with the natural scene and multisensory feedback contributed to a higher score for the question "*How responsive was the environment to actions that you initiated (or performed)?*". However, due to the limited brush options in Painting++, participants felt they had more freedom to control the painting content in Tilt Brush and Painting VR. P29 also mentioned that he wanted to manipulate the AI generation results in Painting++.

Regarding DF, the simplified interaction of Painting++ effectively immersed participants in painting. They found the system helped them with tasks such as selecting the suitable color for natural objects, allowing them to concentrate on painting rather than the mechanisms. In addition, the audio feedback of Painting++ prevented distractions from real-wold noise (e.g., P25), contributing to higher DF scores. However, the average DF score was relatively low among the four factors. Qualitative feedback revealed that some participants were not used to the VR equipment (e.g., P22), and the heavy HMD distracted them from painting. Additionally, participants need to wait about 15 seconds before obtaining the final painting. During this waiting period, some participants may be distracted by events occurring in the real world.

Participants' positive feedback on SF also indicated that the multisensory interaction of Painting++ profoundly shapes the immersive painting experience. Generally, participants agreed that the multiple sensory channels integrated into Painting++ vastly enriched their sense of immersion. P21 further pointed out that "*including multisensory interaction into the painting process made Painting++ distinct from other VR painting applications*". Participants also appreciated that in Painting VR, "*I felt the touch and heard the sound when I was painting*" (P17, P19), and spatial audio feedback in Tilt Brush surprised P16. Participants' detailed feedback on each sensory channel of Painting++ will be introduced Section 6.1.3.

The RF of Painting++ was positively influenced by realistic scenes and multisensory interaction. Participants agreed that the natural scenes in Painting++ authentically reproduced the real sceneries, including the beach, forest, and field. In contrast, the scenes in Tilt Brush leaned toward a more futuristic style. It could be noticed that some questions of RF overlapped with those of SF, indicating that multisensory feedback not only enriched participants' sensory experience but also increased the realism of VE.

6.1.3. Multisensory Interaction

In terms of participants' feedback on different sensory channels, we observed that for most participants, visual feedback remained predominant (see Fig. 8(a)). Surprisingly, two participants thought that haptic feedback had the most significant influence on their immersion.

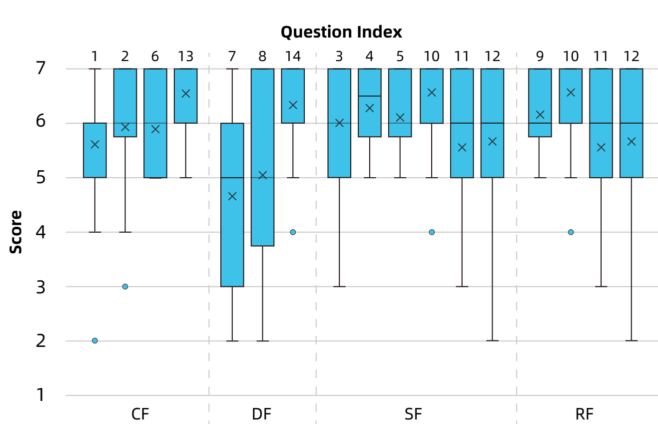

Fig. 7. Painting++'s results of the immersion scale.

Fig. 8(b) depicts the detailed results of sensory satisfaction. All participants appreciated the visual effect of Painting++ (M = 6.39, SD = 0.68). P20 specifically described that "*when I look around, I found a vast ocean behind me, which inspired me to add an island to my painting*". Besides, P19 and P25 agreed that the natural scene broadened their views and evoked feelings of painting outdoors. Moreover, participants put forward expectations for improved visual effects of the brush. For instance, P27 desired textured strokes, and P23 wished for color overlay effects.

As for audio feedback, participants' average response was 6.50 (SD = 0.83), with more than half of them expressing strong satisfaction with it. For instance, P17 mentioned that "*the sounds, especially the singing of birds, could bring me to the natural scene in a short time and increased my willingness to paint*". The audio feedback also brought P22 the illusion of being outdoors and reminded P27 of painting experience in natural settings.

The average score for haptic feedback was 6.11 (SD = 0.87). Participants noted that the simulated wind from VE made them feel comfortable and relaxed. P19 further mentioned, "*I could feel my hair fluttering with the wind*". The vibration from the VR controllers provided intuitive feedback, as described by P18, "*once I painted on the canvas, I could get the haptic feedback*". However, while some participants felt the haptic feedback distinctly, others found it hard to recognize. Meanwhile, P20 observed that the haptic feedback was not always noticeable enough and expected the system to provide feedback when the brush was lifted from the canvas.

The average score of smell feedback was relatively low (M = 5.39, SD = 1.25). Most participants appreciated the involvement of smell feedback, with P14 describing that "*the smell made me think of the immersive cinema in Disneyland*". It was generally agreed that smell feedback could assist participants in choosing their desired brush color (e.g., P21). However, participants also found that some scents needed to be stronger, as they were difficult to distinguish.

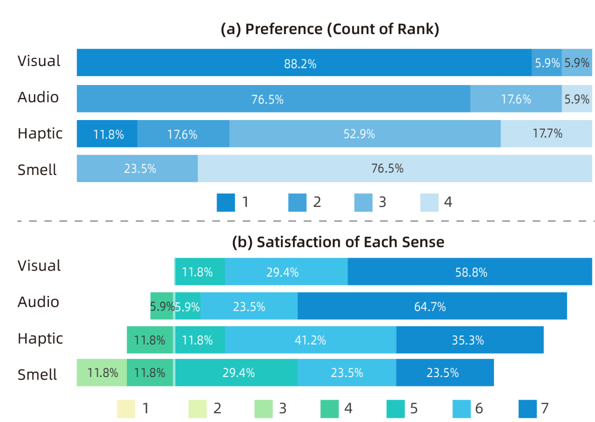

Fig. 8. (a) Participants' preference for different senses. (b) Participants' satisfaction of each sense.

6.1.4. System Comparison

Despite sharing similar core functionalities, target users, and design objectives, Painting VR and Tilt Brush differed from Painting++ in terms of interaction and technical details. These differences could impact the above factors including usability and immersion. Therefore, we further summarized the primary differences among the three systems (see Table 4), and made a comprehensive comparison based on participants' feedback.

Painting Interaction. While Painting VR and Tilt Brush adopted embodied interaction and menu-based interaction separately, Painting++ combined these two ways of interaction to maximize their respective strengths. Embodied interaction refers to how people interact with digital technologies through their physical bodies [92]. The incorporation of embodied interaction reproduces users' real-life experience and facilitates the involvement of one's body, increasing the sense of presence for participants. However, while embodied interaction of creating strokes (Painting++ and Painting VR) was mentioned to increase usability by 8 participants, employing embodied interaction (Painting VR) to select tools decreased usability (stated by 14 participants). Meanwhile, 12 participants agreed that menu-based selection (Painting++ and Tilt Brush) positively influenced system usability.

Painting Environment. Environment is a critical component of a VR painting system, which is typically characterized by its realism and type. For technical implementation, all three systems used Unity to build their environments but differed in their choices regarding realism and type. Realism of environment refers to the extent to which a VE convincingly simulates a real-world environment [93], and 13 participants thought that the realistic environment in Painting++ contributed to its higher immersion. In contrast, Tilt Brush offered an abstract environment without any realistic objects, and participants were unable to draw inspiration directly from this environment. The types of VE mainly include neutral environment (e.g., plain background), indoor environment (e.g., studio, gallery), and outdoor environment (e.g., forest, street), and the neutral environment in Tilt Brush was considered to decrease immersion (suggested by 6 participants).

Auxiliary Functionality. With auxiliary functionality from VR painting systems, users could possibly obtain more satisfying works than painting alone. In Painting VR and Tilt Brush, some basic auxiliary functionalities such as rulers (to help users draw straight lines) and mirrors (to copy a symmetrical pattern from the original painting) were available. As for Painting++, we incorporated GAN-based AI generation to help users obtain rich-textured output by outlining the layout of painting, and 11 participants appreciated it for increased usability.

Multisensory Interaction. Painting VR and Tilt Brush primarily relied on the built-in equipment in Oculus to provide multisensory feedback, whereas Painting++ utilized additional hardware for more comprehensive haptic and smell feedback. According to feedback, 13 participants felt that the multisensory feedback provided by Painting++ enhanced immersion, while 8 participants mentioned that the insufficient sensory feedback in the other two systems negatively impacted immersion.

6.2. Objective Results

6.2.1. Quantitative Metrics of Behaviors

Observation. Based on the behavior coding, the average percentage of observation in Painting++ (M = 42.74%) exceeded that of Painting VR (M = 7.94%) and Tilt Brush (M = 3.84%). By analyzing the average duration and count of observed behaviors (see Fig.9), we found that when using Painting++, participants' average duration of observing the environment and painting far exceeded that of the other two systems, while the frequency of these behaviors did not differ significantly.

Selection. The time proportion participants spent on selection in Painting VR (M = 53.56%) and Tilt Brush (M = 50.01%) far exceeded Painting++ (M = 17.42%). Evident findings include:

Duration of selecting brush. In Painting VR, the average brush selection time ($M = 10.80s$) far exceeded Painting++ ($M = 4.99s$) and Tilt Brush ($M = 3.67s$), which somewhat illustrated Painting VR's weakness in usability. Specifically, the longest time for one brush selection took $68s$ in Painting VR. P28 attributed this extended brush selection process to the highly realistic embodied interaction in Painting VR, which might enhance immersion but meanwhile diminish usability.

Frequency of selecting brush. When using Tilt Brush (M = 29.30), participants' brush selection frequency obviously surpassed that of Painting++ (M = 3.42) and Painting VR (M = 9.92). This was primarily due to Tilt Brush's variety of brushes,

Table 4. Differences in interaction and technical details of Painting++, Painting VR, and Tilt Brush.

		Painting++	Painting VR	Tilt Brush
Painting Interaction		Menu-based interaction, embodied interaction	Embodied interaction	Menu-based interaction
Painting Environment		Realistic, natural environment modeling with Unity	Realistic indoor painting studio modeling with Unity	Abstract neutral environment created by panoramic pictures
Auxiliary Functionality		GAN-based AI generation, scene switching	Ruler, browser	Ruler, mirror, teleport, zoom in and zoom out
Multisensory Interaction	Audio	Seven types of simulated natural sound (Oculus' built-in speaker)	Simulated sound of physical contact between brush and canvas (Oculus' built-in speaker)	Magic sound effects for different brushes (Oculus' built-in speaker)
	Haptic	Haptic feedback to simulate physical contact (Oculus), simulate natural feeling (fans)	Haptic feedback to simulate physical contact (Oculus)	Haptic feedback (Oculus)
	Smell	Simulated smell of natural environment (a multi-channel odor display device)	-	-

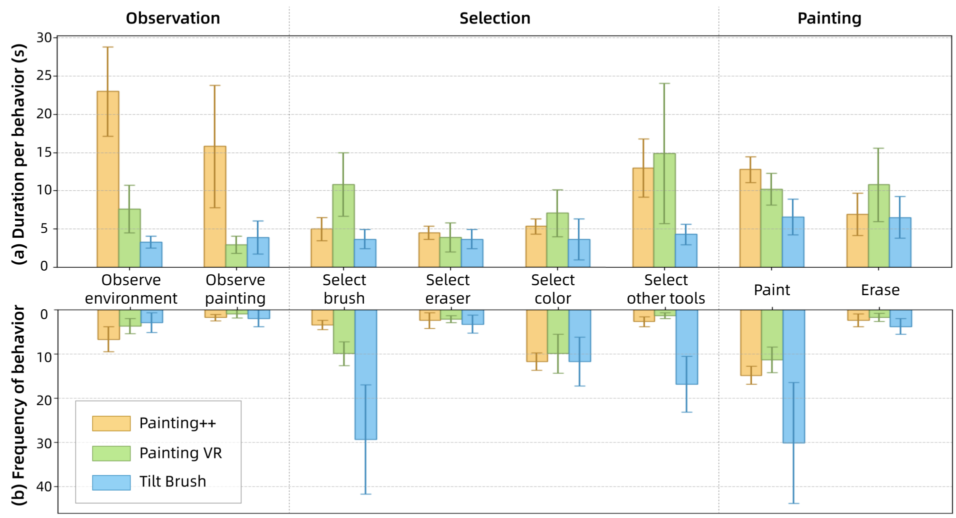

Fig. 9. Quantitative metrics of behaviors, including (a) average duration per behavior and (b) average count of each behavior.

and participants tended to "*switch the brush frequently to explore the visual effects of different brushes*" (P15). Such variety provided by Tilt Brush enhanced the freedom of painting, but it also resulted in more self-correcting behaviors that reflected lower usability (see Section 6.2.2). Besides, the frequency of brush selections in Painting++ was the lowest because it offered fewer brush types compared to the other two systems.

Duration and frequency of selecting other tools. Participants demonstrated frequent and rapid selection of other tools ($M_{duration} = 4.29s, M_{frequency} = 16.80$) when using Tilt Brush, but an opposite trend in Painting++ ($M_{duration} = 13.30s, M_{frequency} = 2.67$) and Painting VR ($M_{duration} = 1.33s, M_{frequency} = 14.90$). In Painting++, participants mainly used the scene-switching and AI generation tools, with the latter resulting in a longer duration due to a technical bottleneck. In Painting VR, participants primarily utilized other tools to move their positions in the virtual environment, but the unnatural interaction design of this feature caused many participants (e.g., P21 and P30) to spend some time adapting to the current view and overcoming dizziness after moving, reflecting its lower usability. As for Tilt Brush, participants tended to switch different auxiliary tools to understand their actual functions. This revealed that Tilt Brush lacked a suitable tutorial for beginners, and the relatively high learning curve of these tools was considered a factor decreasing usability.

Painting. Regarding the time proportion spent on actual painting, there was no obvious difference among Painting++ ($M = 39.84\%$), Painting VR ($M = 38.50\%$), and Tilt Brush ($M = 46.15\%$). However, the average duration of individual painting instances in Tilt Brush ($M = 6.59s$) was shorter than in Painting++ ($M = 12.79s$) and Painting VR ($M = 10.20s$). Participants noted that the virtual canvas and embodied painting interaction in Painting++ and Painting VR allowed them to '*concentrate more on the painting process*", possibly explaining the variance in average duration and frequency. Notably, the average time of erasing in Painting VR ($M = 10.8s$) was higher than in Painting++ ($M = 6.94s$) and Tilt Brush ($M = 6.52s$), possibly due to the frequent falling of the eraser, thus negatively influencing the usability of Painting VR.

6.2.2. Behavior Sequences

During the behavior coding of participants' video recordings, we found distinct degrees of self-correcting behaviors when using different systems. Based on our predefined behavior codes, recognized typical sequences of self-correcting behaviors include:

- P-SE-SB-P: Participants selected the eraser but switched to a brush without erasing anything.
- P-SB(-SB)-SB-P: Participants consecutively switched brushes, with some brushes not being used.
- P-SC(-SC)-SC-P: Participants consecutively switched colors, with some colors not being used.

According to the statistical results, the average number of self-correcting behaviors in Painting++ ($M = 2.08, SD = 0.76$) was lower than in Painting VR ($M = 2.75, SD = 1.09$) and Tilt Brush ($M = 6.42, SD = 2.63$). Based on participants' feedback, we believed this difference was mainly because that: (1) lower usability led to frequent misoperations and self-corrections, and (2) systems did not provide sufficient inspiration or guidance to help participants decide which tool to use. Overall, according to our observation of self-correcting behaviors, participants interacted more smoothly with Painting++, demonstrating higher usability for our system.

To gain a more comprehensive understanding of differences in participants' behavior among the three VR painting systems, we further compared the visualized behavior sequences. Examples of behavior sequences for P14 and P22 are illustrated in Fig. 10, aiding our discussion on these differences. For Painting++, participants typically observed the environment first before selecting a brush. The observation process helped them make rapid and correct selections, which might also explain the fewer self-correcting behaviors in Painting++. In addition,

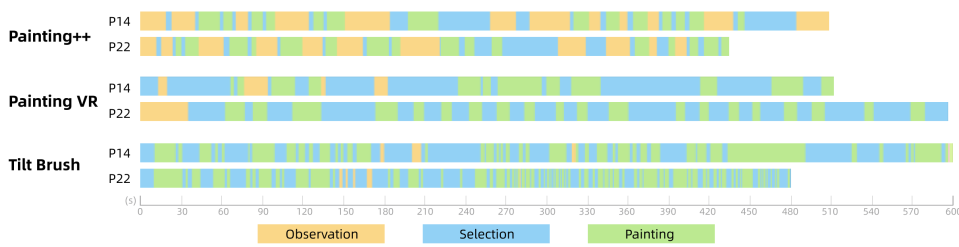

Fig. 10. Example of behavior sequences in three VR painting systems.

while in Painting++ participants made observation throughout the entire painting process, in Painting VR and Tilt Brush, they primarily observed at the beginning only. These differences were likely influenced by the distinct environmental setups. In Painting++, participants were able to draw inspiration from the natural environment surrounding them (as demonstrated by P24), thereby leading to longer and more frequent observations.

7. DISCUSSION

7.1. Roles of Painting++ in Human-computer Collaborative Painting

In Painting++, we aimed to enhance the immersion of VR painting and extend the roles that a computational collaborator could play in collaborative painting. To gain a deeper understanding of the collaboration between Painting++ and users, we outline the roles of Painting++ based on participants' feedback.

7.1.1. Painting++ as a Painting Assistant

Our system leveraged AI's generation capabilities to produce vibrant cartoon paintings from rough sketches and converted plane paintings into stereo paintings. The utilization of AI assists users in achieving more satisfying paintings in VR compared to painting alone. In our preliminary study, both skilled painters and novices reported that painting in VR posed more challenges than traditional tools, hindering their ability to convey intentions and create aesthetically pleasing works. These challenges stem from factors including the limited controller flexibility and demanding 3D spatial imagination requirements. Concurrently, as AI continues to advance, collaborative creative systems involving AI are becoming prevalent [94]. Especially, AI enables untrained users to participate in creative activities without additional training or specialization [24]. Capitalizing on this, Painting++ demonstrates the potential of integrating AI capabilities into the VR experience. The AI-generated paintings made users excited that they could "*easily create personalized virtual worlds in VR*" (P30).

7.1.2. Painting++ as an Immersion Creator

By incorporating appropriate multisensory interactions within simulated scenes, Painting++ deepened users' immersion in painting and rejuvenated their mental states. Our findings (see Section 6.1.3) highlighted that Painting++'s multisensory interactions prevented participants' distractions and facilitated deeper concentration on their inner thoughts. For example, P15 remarked that "*the audio feedback was like white noise and the breeze from fans made me feel like I was getting a facial massage*". Prior research also proved that multisensory interaction could effectively engage one's body [14]. Our study, along with existing literature, suggests that combining painting with multisensory interactions can effectively relax the mind and enhance the immersive experience. The created immersion could hopefully activate users' brains [13, 15] and bring users to a creative state [33].

7.1.3. Painting++ as an Inspiration Trigger

Natural scenes augmented by multisensory interaction have the potential to stimulate users' inspiration and creativity. Participants particularly noted that audio and smell feedback activated their willingness to create, while visual feedback provided inspiration for the layout, theme, and details of the painting. Previous studies have indicated that true creativity extends beyond mere conceptual imagination; it must manifest physically or sensorially [17], as seen in real sensory feedback. From this aspect, multisensory interaction further enhances users' understanding and perception of VE by allowing them to actually hear, smell, and touch the scene, thus contributing to the inspiration. Moreover, advancements in virtual and augmented reality technologies have expanded people's perceptual boundaries, allowing for the creation of novel perceptual dimensions that have never existed [64], which further contributes to triggering creative inspiration.

7.2. Potential Value of VR Painting

7.2.1. Escaping from Physical Limitations in Simulated VE

Based on our findings, a significant benefit of VR painting systems is their ability to liberate users from the constraints of the physical world. When painting outdoors, individuals obtain diverse experiences and rejuvenate their minds by engaging with nature. By harnessing VR technology, users can immerse themselves in their preferred painting settings within VE and interact with it through multisensory feedback. For instance, they can paint in a virtual countryside or seaside setting anytime. According to the qualitative feedback on RF, participants acknowledged that the natural scenes like the beach, forest, and field in Painting++ crafted a vibrant virtual realm parallel to reality, transporting them to "*an imaginary world*" (P20).

7.2.2. Creating Content in VR with AI Assistance

Our study delved into the collaboration between people and AI to create content in VR, which is expected to promote the development of VR applications [95]. Recently, numerous VR applications have garnered public attention. Providing content such as diversified scenes is necessary to prevent users from monotony in VR. However, crafting and rendering 3D content remains time-consuming and complex, even for highly skilled artisans. With the advancements of AI, certain existing models such as Stable Diffusion [96] and DALL·E [97] can swiftly generate high-quality images, and research like 3DALL·E [98], DreamFusion [99], and Dreambooth3D [100] further explored the generation of 3D objects. While existing research already supports text-based 3D generation, content creation through human-AI collaborative painting leaves more room for users to customize the shape and layout of the outcomes.In our evaluation, one participant mentioned that "*the painting generated in VR could be used as scenes for animations or games*" (P24). Future endeavors could delve deeper into human-AI collaboration, integrating cutting-edge AI algorithms to facilitate user-driven content customization across various VR applications.

7.3. Limitations and Future Work

Here we present some issues worthy of further discussion, and consider several limitations of current Painting++ that warrant future research efforts.

When designing Painting++, we employed a self-developed AI algorithm for generating stereo paintings, considering two aspects. First, we chose a GAN-based model rather than large models like Stable Diffusion, because these models mainly rely on text-based prompts to produce images, which to some extent diminishes the interaction of painting activities. Furthermore, our primary goal is to preliminarily validate the feasibility and value of AI generation in VR painting, for which our self-developed AI model is sufficient. Building upon our validation, the adopted algorithm could potentially be replaced by other models with better performance in the future. Second, we chose to further transfer the generated 2D paintings to stereo ones because the approach could provide a more consistent experience in the three-dimensional space of VR without increasing the complexity of painting.

To validate Painting++, we compared it with Painting VR and Tilt Brush, but these two applications differed somewhat in functionality and painting styles from our system. Ideally, we should compare Painting++ with another painting system that provides multisensory interaction and AI assistance. However, no VR painting application has the same functionality as Painting++. Nonetheless, our evaluation primarily aimed to validate the usability and immersion of Painting++, which aligned with the design goals of Painting VR and Tilt Brush. Another factor that may influence the evaluation results is AI's impact on the expressiveness of the created paintings. We acknowledge that AI assistance can enhance painting expressiveness, leading to the potential unfairness of the comparison between our system and the other two. However, the design goal of Painting++ and the primary focus of our evaluation are concerned with the enhancement of the overall painting experience through human-computer collaborative painting, rather than the impact on the expressiveness of the painting results.

Regarding the evaluation, another potential limitation is the lack of an objective benchmark test suite for evaluating Painting++, primarily due to the current unavailability of mature and applicable suites. Additionally, the core design goal of Painting++ is to enhance immersion in VR painting, which is inherently a subjective experience. Nevertheless, future work could make efforts to develop objective evaluation standards for VR painting experiences, such as assessments based on behavior or electroencephalography (EEG) signals.

Although our system effectively brings immersion to VR painting, the evaluation results imply that multisensory feedback is insufficient in some cases. We speculate that this is because the feedback devices used in Painting++ have a relatively limited scope of effect. We adopted desktop equipment to provide sensory feedback to reduce users' physical load, but this approach may have restricted the scope of effects. To address this issue, lightweight wearable feedback devices could be considered as alternatives. Additionally, adopting more advanced equipment such as intelligent gloves and ultrasonic simulators could enrich the tactile experience.

Another limitation of Painting++ lies in its relatively lower functional flexibility compared to other commercial VR painting applications. While the simplified painting interaction and external painting assistance of Painting++ reduce the difficulty of painting in VR, they may limit the proactiveness of professional users, particularly those well-versed in using VR painting applications. Despite this, Painting++ could still benefit professional users by creating immersion and triggering inspiration. Apart from that, participants have expressed a desire for Painting++ to provide more available colors and richer brushes. Some participants also wish to make additional adjustments, such as modifying layouts or textures in the generated paintings. We have to clarify that this paper focuses on exploring the collaboration between human and the computer in VR painting, rather than developing a fully functional commercial product. However, we will continually improve the functions based on users' painting habits to enhance functional flexibility.

Finally, some specific limitations of Painting++ can be improved in a short term: (1) Optimize the data transfer approach between PC and Oculus Quest 2 to reduce users' waiting time. (2) Add timely guidance during painting to avoid misoperations. (3) Allow users to import their own sceneries and models to address the issue of restricted scenes.

8. CONCLUSION

In this paper, we designed and developed Painting++, a system that enabled human-computer collaborative painting in VR with multisensory interaction. Specifically, we provided natural painting scenes and integrated visual, audio, haptic, and smell feedback to enhance immersion in VR painting. We also adopted AI to collaborate with users in creating stereo paintings from rough sketches. The evaluation results suggested that Painting++ outperformed existing VR painting applications in terms of usability owing to the assistance from AI. In addition, it effectively delivers a highly immersive painting experience for users. Our work elaborates on the potential of multisensory interaction and AI capability in facilitating painting. We expect that Painting++ will provide inspiration for the design of future human-computer collaborative painting systems.

Acknowledgments

This work was supported by the National Key R&D Program of China (2022YFB3303301).

References

[1] E.Y.L. Do, VR Sketchpad, in: Computer Aided Architectural Design Futures 2001, Springer Netherlands, 2001, pp. 161–172.

[2] J. Koch, A. Lucero, L. Hegemann, A. Oulasvirta, May ai? design ideation with cooperative contextual bandits, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–12.

[3] C. Guo, Y. Dou, T. Bai, X. Dai, C. Wang, Y. Wen, Artverse: A paradigm for parallel human–machine collaborative painting creation in metaverses, IEEE Transactions on Systems, Man, and Cybernetics: Systems 53 (2023) 2200–2208.

[4] P. Karimi, J. Rezwana, S. Siddiqui, M.L. Maher, N. Dehbozorgi, Creative sketching partner: an analysis of human-ai co-creativity, in: Proceedings of the 25th International Conference on Intelligent User Interfaces, 2020, pp. 221–230.

[5] C. Oh, J. Song, J. Choi, S. Kim, S. Lee, B. Suh, I lead, you help but only with enough details: Understanding user experience of co-creation with artificial intelligence, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 2018, pp. 1–13.

[6] S. Barbara, M. Celebrate, Collaboration (1989).

[7] A. Kantosalo, H. Toivonen, Modes for creative human-computer collaboration: Alternating and task-divided co-creativity, in: Proceedings of the seventh international conference on computational creativity, 2016, pp. 77–84.

[8] R.R. Feinberg, U. Lakshmi, M.J. Golino, R.I. Arriaga, ZenVR: Design Evaluation of a Virtual Reality Learning System for Meditation, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22, Association for Computing Machinery, 2022, pp. 1–15.

[9] K. Gruchalla, Immersive well-path editing: investigating the added value of immersion, in: IEEE Virtual Reality 2004, 2004, pp. 157–164. ISSN: 1087-8270.

[10] J. Nelson, J. Guegan, "i'd like to be under the sea": Contextual cues in virtual environments influence the orientation of idea generation, Computers in Human Behavior 90 (2019) 93–102.

[11] Painting VR, PaintingVR, `https://www.paintingvr.xyz/`, 2019.

[12] T. van den Berge, Vermillion: Oil Painting Simulation In Virtual Reality: A new tool for digital artists offering the analog control of traditional painting with the benefits of a virtual environment., in: ACM SIGGRAPH 2021 Immersive Pavilion, ACM, 2021, pp. 1–2.

[13] M.A. Meredith, B.E. Stein, Interactions Among Converging Sensory Inputs in the Superior Colliculus, Science 221 (1983) 389–391. Publisher: American Association for the Advancement of Science.

[14] E. Panagiotopoulou, M.L. Filippetti, M. Tsakiris, A. Fotopoulou, Affective Touch Enhances Self-Face Recognition During Multisensory Integration, Scientific Reports 7 (2017) 12883. Number: 1 Publisher: Nature Publishing Group.

[15] M. Marucci, G. Di Flumeri, G. Borghini, N. Sciaraffa, M. Scandola, E.F. Pavone, F. Babiloni, V. Betti, P. Aricò, The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence, Scientific Reports 11 (2021) 4831. Number: 1 Publisher: Nature Publishing Group.

[16] X. Yang, L. Lin, P.Y. Cheng, X. Yang, Y. Ren, Y.M. Huang, Examining creativity through a virtual reality support system, Educational Technology Research and Development 66 (2018) 1231–1254.

[17] K.W. Lau, P.Y. Lee, The use of virtual reality for creating unusual environmental stimulation to motivate students to explore creative ideas, Interactive Learning Environments 23 (2015) 3–18.

[18] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784 (2014).

[19] R. Zhang, N.J. McNeese, G. Freeman, G. Musick, "an ideal human": Expectations of ai teammates in human-ai teaming, Proc. ACM Hum.-Comput. Interact. 4 (2021).

[20] J. Rezwana, M.L. Maher, Designing creative ai partners with cofi: A framework for modeling interaction in human-ai co-creative systems, ACM Transactions on Computer-Human Interaction (2022).

[21] H.C. Wang, D. Cosley, S.R. Fussell, Idea expander: supporting group brainstorming with conversationally triggered visual thinking stimuli, in: Proceedings of the 2010 ACM conference on Computer supported cooperative work, 2010, pp. 103–106.

[22] J. Tholander, M. Jonsson, Design ideation with ai - sketching, thinking and talking with generative machine learning models, in: Proceedings of the 2023 ACM Designing Interactive Systems Conference, DIS '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 1930–1940.

[23] J. Koch, N. Taffin, M. Beaudouin-Lafon, M. Laine, A. Lucero, W.E. Mackay, Imagesense: An intelligent collaborative ideation tool to support diverse human-computer partnerships, Proc. ACM Hum.-Comput. Interact. 4 (2020).

[24] Y. Lin, J. Guo, Y. Chen, C. Yao, F. Ying, It is your turn: collaborative ideation with a co-creative robot through sketch, in: Proceedings of the 2020 CHI conference on human factors in computing systems, 2020, pp. 1–14.

[25] J.E. Fan, M. Dinculescu, D. Ha, Collabdraw: an environment for collaborative sketching with an artificial agent, in: Proceedings of the 2019 on Creativity and Cognition, 2019, pp. 556–561.

[26] N. Davis, C.P. Hsiao, K.Y. Singh, L. Li, S. Moningi, B. Magerko, Drawing apprentice: An enactive co-creative agent for artistic collaboration, in: Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition, 2015, pp. 185–186.

[27] C. Zhang, C. Yao, J. Wu, W. Lin, L. Liu, G. Yan, F. Ying, Storydrawer: A child–ai collaborative drawing system to support children's creative visual storytelling, in: CHI Conference on Human Factors in Computing Systems, 2022, pp. 1–15.

[28] S. Shahriar, Gan computers generate arts? a survey on visual arts, music, and literary text generation using generative adversarial network, Displays 73 (2022) 102237.

[29] E. Kim, J. Hong, H. Lee, M. Ko, Colorbo: Envisioned mandala coloringthrough human-ai collaboration, in: 27th International Conference on Intelligent User Interfaces, 2022, pp. 15–26.

[30] S. BV, J. Patel, A. Naik, Y.P. Butala, S. Sharma, N. Chhaya, Towards enabling synchronous digital creative collaboration: Codifying conflicts in co-coloring, in: CHI Conference on Human Factors in Computing Systems Extended Abstracts, 2022, pp. 1–7.

[31] F. Wu, S.W. Hsiao, P. Lu, An aigc-empowered methodology to product color matching design, Displays 81 (2024) 102623.

[32] J. Fišer, P. Asente, D. Sỳkora, Shipshape: a drawing beautification assistant, in: Proceedings of the workshop on Sketch-Based Interfaces and Modeling, 2015, pp. 49–57.

[33] J. Amores, J. Lanier, HoloARt: Painting with Holograms in Mixed Reality, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM, 2017, pp. 421–424.

[34] S. Jung, Y. Wu, R. McKee, R.W. Lindeman, All Shook Up: The Impact of Floor Vibration in Symmetric and Asymmetric Immersive Multi-user VR Gaming Experiences, in: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2022, pp. 737–745. ISSN: 2642-5254.

[35] M. Bonfert, S. Lemke, R. Porzel, R. Malaka, Kicking in Virtual Reality: The Influence of Foot Visibility on the Shooting Experience and Accuracy, in: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2022, pp. 711–718. ISSN: 2642-5254.

[36] H. Cecotti, Z. Day-Scott, L. Huisinga, L. Gordo-Pelaez, Virtual Reality for Immersive Learning in Art History, in: 2020 6th International Conference of the Immersive Learning Research Network (iLRN), 2020, pp. 16–23.

[37] P.H. Han, Y.S. Chen, C.E. Hsieh, H.C. Wang, Y.P. Hung, Hapmosphere: Simulating the Weathers for Walking Around in Immersive Environment with Haptics Feedback, in: 2019 IEEE World Haptics Conference (WHC), IEEE, 2019, pp. 247–252.

[38] K.A. Mills, A. Brown, Immersive virtual reality (VR) for digital media making: transmediation is key, Learning, Media and Technology 47 (2022) 179–200.

[39] E. Kruijff, C. Trepkowski, R.W. Lindeman, The effect of vibration and low-frequency audio on full-body haptic sensations, in: Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology, ACM, 2015, pp. 194–194.

[40] University of National and World Economy, Sofia, Bulgaria, D. Velev, P. Zlateva, Virtual Reality Challenges in Education and Training, International Journal of Learning and Teaching (2017).

[41] J.H. Wang, D.L. James, KleinPAT: optimal mode conflation for time-domain precomputation of acoustic transfer, ACM Transactions on Graphics 38 (2019) 122:1–122:12.

[42] M. Narbutt, S. O'Leary, A. Allen, J. Skoglund, A. Hines, Streaming VR for immersion: Quality aspects of compressed spatial audio, in: 2017 23rd International Conference on Virtual System & Multimedia (VSMM), 2017, pp. 1–6. ISSN: 2474-1485.

[43] E. Bouzbib, G. Bailly, S. Haliyo, P. Frey, "Can I Touch This?": Survey of Virtual Reality Interactions via Haptic Solutions, in: 32e Conférence Francophone sur l'Interaction Homme-Machine, 2021, pp. 1–16. ArXiv:2101.11278 [cs].

[44] T. Singhal, O. Schneider, Juicy Haptic Design: Vibrotactile Embellishments Can Improve Player Experience in Games, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, 2021, pp. 1–11.

[45] M. Azh, S. Zhao, S. Subramanian, Investigating Expressive Tactile Interaction Design in Artistic Graphical Representations, ACM Transactions on Computer-Human Interaction 23 (2016) 1–47.

[46] P.H. Han, Y.S. Chen, K.C. Lee, H.C. Wang, C.E. Hsieh, J.C. Hsiao, C.H. Chou, Y.P. Hung, Haptic around: multiple tactile sensations for immersive environment and interaction in virtual reality, in: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, ACM, 2018, pp. 1–10.

[47] Y. Singhal, H. Wang, H. Gil, J.R. Kim, Mid-Air Thermo-Tactile Feedback using Ultrasound Haptic Display, in: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology, ACM, 2021, pp. 1–11.

[48] S. Zou, X. Hu, Y. Ban, S. Warisawa, Simulating Olfactory Cocktail Party Effect in VR: A Multi-odor Display Approach Based on Attention, in: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2022, pp. 474–482.

[49] Y. Yanagida, S. Kawato, H. Noma, N. Tetsutani, A. Tomono, A nose-tracked, personal olfactory display, in: ACM SIGGRAPH 2003 Sketches & Applications, SIGGRAPH '03, Association for Computing Machinery, 2003, p. 1.

[50] H. Matsukura, T. Yoneda, H. Ishida, Fragrant multimedia display system: Presenting odor distribution on display screen, in: 2012 IEEE SENSORS, 2012, pp. 1–4. ISSN: 1930-0395.

[51] T. Yamada, S. Yokoyama, T. Tanikawa, K. Hirota, M. Hirose, Wearable Olfactory Display: Using Odor in Outdoor Environment, in: IEEE Virtual Reality Conference (VR 2006), 2006, pp. 199–206. ISSN: 2375-5334.

[52] D. Dobbelstein, S. Herrdum, E. Rukzio, inScent: a wearable olfactory display as an amplification for mobile notifications, in: Proceedings of the 2017 ACM International Symposium on Wearable Computers, ISWC '17, Association for Computing Machinery, 2017, pp. 130–137.

[53] Y. Wang, J. Amores, P. Maes, On-Face Olfactory Interfaces, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20, Association for Computing Machinery, 2020, pp. 1–9.

[54] N. Ranasinghe, P. Jain, N. Thi Ngoc Tram, K.C.R. Koh, D. Tolley, S. Karwita, L. Lien-Ya, Y. Liangkun, K. Shamaiah, C. Eason Wai Tung, C.C. Yen, E.Y.L. Do, Season Traveller: Multisensory Narration for Enhancing the Virtual Reality Experience, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, 2018, pp. 1–13.

[55] S. Jung, A.L. Wood, S. Hoermann, P.L. Abhayawardhana, R.W. Lindeman, The Impact of Multi-sensory Stimuli on Confidence Levels for Perceptual-cognitive Tasks in VR, in: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2020, pp. 463–472.

[56] J. Martins, R. Gonçalves, F. Branco, L. Barbosa, M. Melo, M. Bessa, A multisensory virtual experience model for thematic tourism: A Port wine tourism application proposal, Journal of Destination Marketing & Management 6 (2017) 103–109.

[57] D.B. Faustino, S. Gabriele, R. Ibrahim, A.L. Theus, A. Girouard, SensArt Demo: A Multisensory Prototype for Engaging with Visual Art, in: Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, ACM, 2017, pp. 462–465.

[58] G. Bernal, N. Hidalgo, C. Russomanno, P. Maes, Galea: A physiological sensing system for behavioral research in Virtual Environments, in: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2022, pp. 66–76.

[59] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, E. Rukzio, VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, ACM, 2020, pp. 1–14.

[60] H. Elsayed, M.D. Barrera Machuca, C. Schaarschmidt, K. Marky, F. Müller, J. Riemann, A. Matviienko, M. Schmitz, M. Weigel, M. Mühlhäuser, VRSketchPen: Unconstrained Haptic Assistance for Sketching in Virtual 3D Environments, in: 26th ACM Symposium on Virtual Reality Software and Technology, ACM, 2020, pp. 1–11.

[61] Y. Jiang, C. Zhang, H. Fu, A. Cannavò, F. Lamberti, H.Y.K. Lau, W. Wang, HandPainter - 3D Sketching in VR with Hand-based Physical Proxy, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, 2021, pp. 1–13.

[62] S. Ppali, V. Lalioti, B. Branch, C.S. Ang, A.J. Thomas, B.S. Wohl, A. Covaci, Keep the VRhythm going: A musician-centred study investigating how Virtual Reality can support creative musical practice, in: CHI Conference on Human Factors in Computing Systems, ACM, 2022, pp. 1–19.

[63] S. Fleury, P. Blanchard, S. Richir, A study of the effects of a natural virtual environment on creativity during a product design activity, Thinking Skills and Creativity 40 (2021) 100828.

[64] M.P. Jeon, P. Fishwick, Special Issue on Arts, Aesthetics, and Performance in Telepresence: Guest Editors' Introduction: Homo Ludens in Virtual Environments, Presence: Teleoperators and Virtual Environments 26 (2017) iii–vii.

[65] M. Jeon, R. Fiebrink, E.A. Edmonds, D. Herath, From rituals to magic: Interactive art and HCI of the past, present, and future, International Journal of Human-Computer Studies 131 (2019) 108–119.

[66] R. Zender, P. Sander, M. Weise, M. Mulders, U. Lucke, M. Kerres, Handlevr: Action-oriented learning in a vr painting simulator, in: Emerging Technologies for Education: 4th International Symposium, SETE 2019, Held in Conjunction with ICWL 2019, Magdeburg, Germany, September 23–25, 2019, Revised Selected Papers 4, Springer, 2020, pp. 46–51.

[67] Alphabet Inc, Tilt Brush by Google, `https://www.tiltbrush.com/`, 2019.

[68] E. Rosales, C. Araújo, J. Rodriguez, N. Vining, D. Yoon, A. Sheffer, AdaptiBrush: adaptive general and predictable VR ribbon brush, ACM Transactions on Graphics 40 (2021) 1–15.

[69] R. Arora, R.H. Kazi, F. Anderson, T. Grossman, K. Singh, G. Fitzmaurice, Experimental Evaluation of Sketching on Surfaces in VR, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, 2017, pp. 5643–5654.

[70] T. Wang, X. Qian, F. He, K. Ramani, LightPaintAR: Assist Light Painting Photography with Augmented Reality, in: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, 2021, pp. 1–6.

[71] M. Kusunoki, R. Furuhama, R. Toshima, H. Mori, H. Xie, T.Y. Wang, T. Yuizono, T. Sato, K. Miyata, Multibrush: 3d brush painting using multiple viewpoints in virtual reality, in: 2023 9th International Conference on Virtual Reality (ICVR), 2023, pp. 481–486.

[72] I. Druzetic, F. Büntig, C. Vogel, A. Treskunov, M. Bertram, C. Geiger, Photo Sprayer: A VR Application for Digital Art Creation, in: Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, ACM, 2018, pp. 537–543.

[73] B. Thoravi Kumaravel, C. Nguyen, S. DiVerdi, B. Hartmann, TutoriVR: A Video-Based Tutorial System for Design Applications in Virtual Reality, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, 2019, pp. 1–12.

[74] N. Chaniaud, S. Fleury, B. Poussard, O. Christmann, T. Guitter, S. Richir, Is virtual reality so user-friendly for non-designers in early design activities? comparing skills needed to traditional sketching versus virtual reality sketching, Design Science 9 (2023) e28.

[75] P. Panda, C. Ho, D. Ham, Morphaces: Exploring Morphable Surfaces for Tangible Sketching in VR, in: Creativity and Cognition, ACM, 2021, pp. 1–11.

[76] Z. Zhang, J.P. Schulze, VirtualForce: Simulating writing and sketching on a 2D-surface in virtual reality, Electronic Imaging 34 (2022) 297–1–297–5.

[77] A.R. Fender, T. Roberts, T. Luong, C. Holz, Infinitepaint: Painting in virtual reality with passive haptics using wet brushes and a physical proxy canvas, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, Association for Computing

Machinery, New York, NY, USA, 2023.
[78] M.D. Barrera Machuca, P. Asente, J. Lu, B. Kim, W. Stuerzlinger, Multiplanes: Assisted Freehand VR Drawing, in: Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology, ACM, 2017, pp. 1–3.
[79] W. Li, Synthesizing 3D VR sketch using generative adversarial neural network, in: Proceedings of the 2023 7th International Conference on Big Data and Internet of Things, BDIOT '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 122–128. doi:10.1145/3617695.3617723.
[80] J. Urban Davis, F. Anderson, M. Stroetzel, T. Grossman, G. Fitzmaurice, Designing co-creative AI for virtual environments, in: Proceedings of the 13th Conference on Creativity and Cognition, 2021, pp. 1–11.
[81] J. Swain, A Hybrid Approach to Thematic Analysis in Qualitative Research: Using a Practical Example, SAGE Publications Ltd, 2018.
[82] L.E. Ramsier, Evaluating the usability and user experience of a virtual reality painting application (2019) 84.
[83] R.P. Darken, R. Durost, Mixed-dimension interaction in virtual environments, in: Proceedings of the ACM symposium on Virtual reality software and technology, 2005, pp. 38–45.
[84] J. Wang, R. Lindeman, Coordinated hybrid virtual environments: Seamless interaction contexts for effective virtual reality, Computers & Graphics 48 (2015) 71–83.
[85] L. Sun, P. Chen, W. Xiang, P. Chen, W.y. Gao, K.j. Zhang, SmartPaint: a co-creative drawing system based on generative adversarial networks, Frontiers of Information Technology & Electronic Engineering 20 (2019) 1644–1656.
[86] Z. Li, N. Snavely, MegaDepth: Learning Single-View Depth Prediction from Internet Photos, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 2041–2050. ISSN: 2575-7075.
[87] M.L. Shih, S.Y. Su, J. Kopf, J.B. Huang, 3D Photography Using Context-Aware Layered Depth Inpainting, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8025–8035. ISSN: 2575-7075.
[88] Roy S.Kalawsky, VRUSE—a computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems, Applied Ergonomics 30 (1999) 11–25. Publisher: Elsevier.
[89] B.G. Witmer, M.J. Singer, Measuring Presence in Virtual Environments: A Presence Questionnaire, Presence: Teleoperators and Virtual Environments 7 (1998) 225–240.
[90] S.S. Shapiro, M.B. Wilk, An analysis of variance test for normality (complete samples), Biometrika 52 (1965) 591–611. Publisher: Oxford Academic.
[91] C. Altin Gumussoy, A. Pekpazar, M. Esengun, A.E. Bayraktaroglu, G. Ince, Usability evaluation of tv interfaces: Subjective evaluation vs. objective evaluation, International Journal of Human–Computer Interaction 38 (2022) 661–679.
[92] P. Dourish, Where the action is: the foundations of embodied interaction, MIT press, 2001.
[93] M. Slater, S. Wilbur, A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments, Presence: Teleoperators & Virtual Environments 6 (1997) 603–616.
[94] M. Guzdial, N. Liao, J. Chen, S.Y. Chen, S. Shah, V. Shah, J. Reno, G. Smith, M.O. Riedl, Friend, collaborator, student, manager: How design of an ai-driven game level editor affects creators, in: Proceedings of the 2019 CHI conference on human factors in computing systems, 2019, pp. 1–13.
[95] M. Wang, X.Q. Lyu, Y.J. Li, F.L. Zhang, VR content creation and exploration with deep learning: A survey, Computational Visual Media 6 (2020) 3–28.
[96] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Müller, J. Penna, R. Rombach, Sdxl: Improving latent diffusion models for high-resolution image synthesis, arXiv preprint arXiv:2307.01952 (2023).
[97] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, I. Sutskever, Zero-shot text-to-image generation, in: International conference on machine learning, Pmlr, 2021, pp. 8821–8831.
[98] V. Liu, J. Vermeulen, G. Fitzmaurice, J. Matejka, 3dall-e: Integrating text-to-image ai in 3d design workflows, in: Proceedings of the 2023 ACM designing interactive systems conference, 2023, pp. 1955–1977.
[99] B. Poole, A. Jain, J.T. Barron, B. Mildenhall, Dreamfusion: Text-to-3d using 2d diffusion, arXiv preprint arXiv:2209.14988 (2022).
[100] A. Raj, S. Kaza, B. Poole, M. Niemeyer, N. Ruiz, B. Mildenhall, S. Zada, K. Aberman, M. Rubinstein, J. Barron, et al., Dreambooth3d: Subject-driven text-to-3d generation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 2349–2359.

Appendix

Appendix A. Questions in the Scales

Table 1 presents questions in the usability scale and the corresponding factor of each question. Table 2 presents questions in the immersion scale and the related factors of each question.

Table 1. All questions in the usability scale.

Question	Factor
I can easily paint whatever I want with this application.	Ease of use
I can freely express my ideas through painting in this application.	Flexibility
The functions provided by this application can help me to paint.	Functionality
I am satisfied with the painting created in this application.	Effectiveness
I do not need any extra assistance when using this application.	Ease of use
This application will not restrict the content of my painting.	Flexibility
I am well aware of all the functions provided by this application.	Functionality
I can create better paintings in this application than in reality.	Effectiveness
I can easily modify my painting in this application.	Ease of use
The sensory feedback from the application facilitates my painting process.	Flexibility

Table 2. All questions in the immersion scale.

Question	Factor
How much were you able to control the effect of painting?	CF
How responsive was the environment to actions that you initiated (or performed)?	CF
How completely were all of your senses engaged?	SF
How much did the visual aspects of the environment involve you?	SF
How much did the auditory aspects of the environment involve you?	SF
How natural was the mechanism which controlled movement of painting?	CF
How aware were you of events occurring in the real world around you?	DF
How aware were you of your display and control devices?	DF
How much did your experiences in the virtual environment seem consistent with your real-world experiences?	RF
How well could you identify sounds?	RF SF
How well could you localize sounds?	RF SF
How well could you actively survey or search the virtual environment using touch?	RF SF
How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?	CF
How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?	DF