# *Vanti*: A Novel Interaction Design for Immersive VR Walking Tours

Xiaojie Wang
School of Film and Television Arts
Xi'an Jiaotong-Liverpool University, Suzhou, China

Yue Li *
School of Advanced Technology
Xi'an Jiaotong-Liverpool University, Suzhou, China


Figure 1: Demonstration of *Vanti*: (a) push the controller joystick to move forward in the virtual tour (play the video clip); pull the joystick to move backwards (rewind the video clip); (b) select the next route to continue the virtual tour (switch to the video clip matching the user's selection); (c) The overhead view of the virtual tour path, consisting of five short routes (five video clips).

**ABSTRACT**

Virtual Reality (VR) videos have become increasingly accessible with the advancement of immersive technologies. Viewing immersive VR videos using Head-Mounted Displays (HMDs) demonstrates significant advantages in the field of view and immersion compared to traditional video viewing on 2D screens. However, current interaction methods of video playback in VR are mainly inherited from traditional 2D video players, using timeline with panels and icons overlaid on the video. However, these interactions may not be as intuitive and easy to use in VR, as users are immersed in a spatial enclosure that extends far beyond a 2D menu. This research proposes *Vanti (Vanishing Timeline)*, a novel interaction design that combines video playback with locomotion in VR walking tours. The technique eliminates traditional timeline and button playback controls, transforming the passive viewing experience of traditional immersive VR videos into an active exploration of the real-world environments in the virtual tour. By mapping the steering locomotion with the video timeline, users can control their direction of movement and steer in VR to play (or rewind) the recorded video. Reducing video attributes of the immersive VR video improves users' freedom of movements in the surrounding virtual scenes. Our evaluation study with 25 participants showed that *Vanti* significantly improved user presence, agency and engagement in their interactions with immersive VR walking tours.

**Index Terms:** Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

## 1 INTRODUCTION

The emergence of 360-degree video technology marks a significant shift in audience engagement, enabling users to control viewing angles independently from playback. This immersive experience allows exploration from multiple angles, enhancing engagement and digital storytelling. Initially, 360-degree content was consumed on 2D devices like monitors and mobile screens via platforms like YouTube [13]. VR Head-Mounted Displays (HMDs) further enhance this by using head motion for navigation, aligning with human proprioception and vestibular systems [15]. However, most viewing experiences remain passive. The inherited playback controls from traditional 2D interfaces limit presence and immersion in VR.

360-degree images and videos have been used for virtual tours in cultural heritage [5], tourism [6], storytelling [2], and educational activities [12]. Various interaction techniques have been proposed for immersive 360-degree videos. For example, Argyriou et al. [1] recommended using steady, slow movement to simulate real walking when implementing 360-degree video for immersive tours. Recent studies proposed recommendations for the design of timelines for VR 360-degree videos [14], and suggested the use of alternative interaction techniques like hand gestures and head orientation [10]. Our examination of previous works showed that most video playback methods rely on 2D or 3D UIs similar to traditional screen-based viewing. Users need to point to the UI panel to drag the timeline and click buttons for playback control in VR. Despite some alternative designs, there is a gap in seamlessly integrating video playback as a navigation method within VR environments. There is an opportunity for innovation to bridge 360-video viewing and VR locomotion, enhancing the immersive experience of 360-degree videos.

This study proposes the concept of *Vanishing Timeline (Vanti)*, rethinking traditional timeline-based interaction techniques to create a dynamic and engaging user experience for immersive VR walking tours. *Vanti* moves beyond standard controls to an interactive approach using joystick movements to explore virtual environments. The motivation for the design of *Vanti* was not to improve the video-viewing experience, but to utilize the 360-degree video playing as an alternative for navigating a virtual environment, taking advantage of its relatively cheap production cost compared to 3D reconstruction. Our evaluation showed that *Vanti* enhances presence, agency, and engagement by allowing users to navigate 360-degree content similarly to real-world body movements.

*Corresponding author: yue.li@xjtlu.edu.cn

2771-1110/24/$31.00 ©2024 IEEE
DOI 10.1109/ISMAR-Adjunct64951.2024.00145

Although there were instances of video playback interaction using a joystick, our contribution lies in being the first to combine video playback interaction control with navigation. This approach enhances the user's active exploration interest, mitigates the inherent video attributes of 360-degree videos, and allows users to become more immersed in the virtual environment. The significance of this research lies in its potential to redefine interaction design for 360-degree video in VR. By integrating video playback control with locomotion in VR, it paves the way for more intuitive and immersive experiences. This design can transform passive viewing into active exploration, with implications for entertainment, education, training, and cultural content.

## 2 Design of *Vanti* for Immersive Walking Tours

Based on the identified research gap, we propose *Vanti*, a novel interaction design for immersive VR walking tours. Previous work showed that users prefer to be able to directly control their movement [9]. Therefore, instead of playing the video automatically, *Vanti* requires the user to control its play by 'moving' in the virtual environment. This approach transforms the 360-degree video experience from a passive viewing activity to an active exploration of a virtual environment. By allowing users to control video playback and rewind using the joystick on their handheld controller, reminiscent of actual movement in a 3D virtual environment, it supports a sense of presence and engagement. Users could perceive themselves as navigating through a virtual space that they can freely explore. This design eliminates the traditional button-clicking operation for video playback, replacing it with a more intuitive mapping of video playback to the user's direct movements in VR.

**Video Clip Preparation** Preparing video materials is crucial for this interaction design to effectively simulate natural human movement in VR. To minimize the risk of simulator sickness, it is essential to ensure that the camera used to capture the 360-degree video maintains stability throughout the recording process [8]. In addition, the walking speed should be set to a comfortable and realistic pace that mimics natural human movement. For the experimental system development, the 360-degree videos were shot using Insta360 X2. The video resolution is $3008 \times 1504$, and the refresh rate is 60 frames per second. The videos were recorded during off-peak hours to avoid disturbances during shooting. The video content included an indoor video clip and four outdoor video clips (see Figure 1c). Each video clip lasted 15.8 seconds on average ($SD = 3.54$).

**Navigation Control and Specification** We mapped the joystick control commonly used in VR with the movements in VR for video playback control. For the implementation, we used a Meta Quest 3 HMD with two hand-held controllers. Here we follow Zielasko and Weissker's guideline on steering locomotion [16] to specify the *direction*, *speed*, and *rotation* in the design of *Vanti*. The *directions* are derived from the thumbstick mounted on tracked controller. Pushing the joystick forward, which corresponds to moving forward in the virtual scene, controls the video playback in the forward direction. Similarly, pulling the joystick backward, mirroring backward movement in the virtual scene, is mapped to video playback in reverse (rewind) direction (see Figure 1a). Based on previous research [4], we set up an average *speed* of 1.8 steps/s. The movement is indicated by the $y$ axis of the `PrimaryThumbstick`, with a value range of $(-1,1)$. When $y > 0.5$, it signals the user's intent to push the joystick forward, initiating forward video playback. Conversely, when the $y < -0.5$, it indicates the user's intention to pull the joystick backward, triggering backward video playback. To prevent unintended actions, a dead zone is implemented for the values between -0.5 and 0.5, ensuring that slight movements within this range do not trigger any video playback actions. During the tour, users can change their viewing *orientation* at all times. The head *rotation* was used to detect the facing direction, so that the forward and backward movements were mapped with users' orientation. No timeline or UI panel was displayed during the virtual tour (video playback). When a route reaches the end (a video ends playing), buttons for route selection will appear, allowing the user to choose the next route (video clip). Figure 1 illustrate the interaction design and a schematic view of the route for the virtual tour.

## 3 User Study

**Traditional Timeline Design** We implemented a traditional timeline design for a baseline comparison. It includes a 2D UI panel that is always displayed in front of the user's view, with a timeline and play/pause buttons on it. Such form factor (a menu bar and control buttons) is common and familiar to users. Given that virtual tours require frequent actions of moving (play) and stopping (pause), a constantly displayed menu allowed users to pause or continue at any time without the additional step of activating the menu. This helped maintain consistency for comparison - neither the menu-based approach nor *Vanti* required users to activate the control. The same speed of 1.8 step/s was used for this condition. Whenever users reach the end of a route (the end of a video), two additional buttons will appear, allowing the user to choose to turn/continue to the next route (the next video clip) or return to the original route (rewind). Users can use the controller to point and select the direction buttons (see Figure 1b). They can also navigate to different points in the video by clicking on different positions along the timeline or dragging the timeline slider.

**Procedure** A within-subjects study was conducted to compare the use of *Vanti* and traditional timeline-based interactions in four aspects of user experience metrics: simulator sickness, agency, presence, and engagement. The experimental task for the participants was to simulate a point-to-point tour movement in the university, i.e., go to a classroom. Simulator sickness was measured using the Simulator Sickness Questionnaire (SSQ) [7]. Experience measures about agency and presence were adapted from previous validated questionnaires [3, 11]. We assessed cognitive, emotional and behavioral engagement. The three categories were based on the conceptualization of engagement in education theories. Participants rated their experience on a 7-point Likert scale (7 = strongly agree). This was followed by an interview to further understand the feelings and suggestions of the users who participated in the experiment. After the experiment data collection, we used IBM SPSS Statistics for data analysis. This study was identified as low-risk research and approved by our University Ethics Committee. Figure 2 shows the experimental procedure.

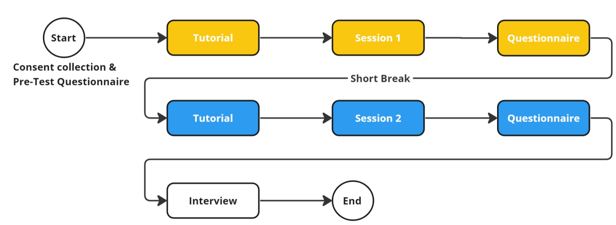
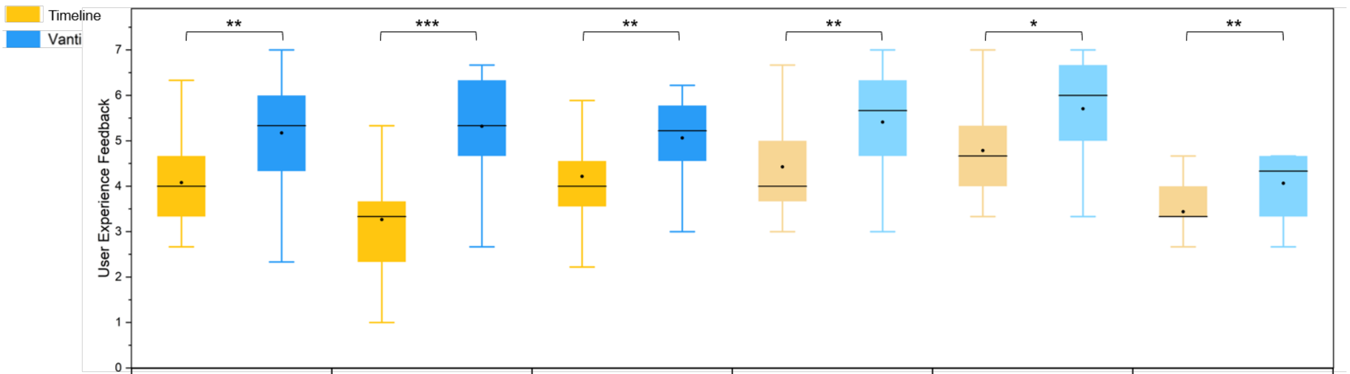



Figure 2: Flow chart showing the procedure of the experiment.

**Participants** There were 25 participants (12 males, 13 females) aged from 18 to 40 years old, who voluntarily signed up to participate in the experiment. 80% of the participants were aged between 22-26 years old, with an average of 24.6 years old ($SD = 3.92$). The participants were mainly students from a local university. 44% of participants were undergraduate students and 52% were graduate students. There are 13 participants used *Vanti* and 12 used the traditional timeline menu in the first round of the test. Regarding




| | Presence | | Agency | | Engagement | | Cognitive Engagement | | Emotional Engagement | | Behavioural Engagement | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 4.1 | 5.2 | 3.3 | 5.3 | 4.2 | 5.1 | 4.4 | 5.4 | 4.8 | 5.7 | 3.4 | 4.1 |
| SD | 1.3 | 1.2 | 1.4 | 1.1 | 0.9 | 0.8 | 1.1 | 1.1 | 1.2 | 1.1 | 0.8 | 0.7 |

Figure 3: Box plots and tables showing participants' self-reported experience of perceived presence, agency, and engagement.

participants' experience with VR technology, 60% of the participants ($n = 15$) reported that they used VR equipment very rarely, 24% of the participants ($n = 6$) used VR equipment rarely and only 16% of participants ($n = 4$) used VR equipment occasionally.

## 4 Results

We used IBM SPSS Statistics for the data analysis. The normality of the data distribution was evaluated using the Shapiro-Wilk test. The SSQ data deviated from the normal distribution ($p < 0.05$). Presence, agency, and engagement data appeared to be normally distributed. Thus, we conducted Wilcoxon signed-rank test for the analysis of simulator sickness, and paired-samples $t$ test for the analysis of presence, agency, and engagement.

**Simulator Sickness** Wilcoxon signed-rank test showed no significant difference in the total severity of simulator sickness between the traditional timeline interaction design ($M = 20.3, SD = 22.8$) and *Vanti* ($M = 21.2, SD = 25.5$), $z = -0.327, p = .744$.

**Presence** Paired-samples $t$ test showed a statistically significant difference in perceived presence, $t(24) = -3.03, p = 0.006$. Participants perceived significantly greater presence using *Vanti* than the traditional timeline interaction design (see Figure 3).

**Agency** Paired-samples $t$ test showed a statistically significant difference in perceived sense of agency, $t(24) = -5.921, p < 0.001$. Participants perceived significantly greater agency using *Vanti* than the traditional timeline interaction design (see Figure 3).

**Engagement** Paired-samples $t$ test showed a statistically significant difference in audience engagement, $t(24) = -5.921, p = 0.003$. Participants perceived significantly greater agency using *Vanti* than the traditional timeline interaction design (see Figure 3). Specifically, significant differences were shown in cognitive engagement ($t(24) = -3.501, p = 0.002$), emotional engagement ($t(24) = -2.619, p = 0.015$), and behavioral engagement ($t(24) = -3.389, p = 0.002$). Participants perceived significantly greater engagement in all three dimensions using *Vanti* than the traditional timeline design.

**Time Recording** Analysis of the system log showed that participants spent an average time of 237.04 seconds ($SD = 55.62$) using *Vanti* and 198.52 seconds ($SD = 45.13$) using the traditional approach. The time spent using *Vanti* was greater than using the traditioanl timeline design.

**Interview** Qualitative data were obtained through direct observation and audio-recorded interviews. About 80% of participants ($n = 20$) preferred the *Vanti* design, citing a higher sense of control and immersion. For example, P2 stated, *'this kind of interaction is mapped with my movements, making me feel more like I am moving in a real environment rather than just watching a video'*. Five participants preferred the traditional design due to its simplicity and familiarity with traditional video playback. Many users ($n = 10$) suggested adding auxiliary navigation tools such as minimaps to improve spatial awareness. P14 recommended determining movement directions based on body or head rotation. P19 mentioned that connections between video clips were stiff, and P16 wanted smoother transitions to maintain immersion. P17 suggested increasing video resolution and quality for a better experience. These comments pointed out directions of improvements to the future design of techniques for immersive VR walking tours.

## 5 Discussion and Conclusion

In this paper, we presented *Vanti*, a novel interaction design for immersive VR walking tours that transformed the passive linear watching experience to an active interactive exploration activity. The evaluation results showed its significant effect on increasing the sense of presence, agency and engagement of user experience. *Vanti* mitigates the use of 2D UIs in an immersive spatial enclosure and enhances immersion and interactivity, simulating real-world walking tours in VR. Compared to the traditional timeline approach, the technique requires more efforts in the preparation and processing of high-quality video content. However, it contributes positively to the user experience, and requires much lower production cost compared to the reconstruction of a realistic environment using laser scanning and photogrammetry technologies. Our work offers ideas for improving VR walking tour experiences, and validates the feasibility of a vanishing timeline, expanding immersive VR applications and offering cost-effective, user-centered solutions. Still, it should be acknowledged that the current study examined a specific setup of a VR walking tour. Future work is needed to validate the applicability of the technique to other types of VR videos. The technique is promising to be extended to a wide range of scenarios that are walking-based, such as heritage sites, museums and galleries, and campus tours. Based on idea presented in this paper, future research can explore the application and usage scenarios for presenting a wider variety of immersive video content.

## ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (62207022). We thank our participants for their time and the reviewers for their comments and suggestions.

## REFERENCES

[1] L. Argyriou, D. Economou, and V. Bouki. Design methodology for 360 immersive video applications: the case study of a cultural heritage virtual tour. *Personal and Ubiquitous Computing*, 24(6):843–859, 2020.

[2] L. Bassbouss, S. Steglich, and I. Fritzsch. Interactive 360° video and storytelling tool. In *2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT)*, pp. 113–117. IEEE, 2019.

[3] J. Bergström, J. Knibbe, H. Pohl, and K. Hornbæk. Sense of Agency and User Experience: Is There a Link? *ACM Trans. Comput.-Hum. Interact.*, 29(4), mar 2022. doi: 10.1145/3490493

[4] N. Biswas and S. Bhattacharya. Finding a range of perceived natural visual walking speed for stationary travelling techniques in VR. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 209–211, 2021. doi: 10.1109/ISMAR-Adjunct54149.2021.00050

[5] S. Cai, E. Ch'ng, and Y. Li. A comparison of the capacities of VR and 360-degree video for coordinating memory in the experience of cultural heritage. In *2018 3rd Digital Heritage International Congress (DigitalHERITAGE)*, pp. 1–4. IEEE, 2018.

[6] V. De Luca, G. Marcantonio, M. C. Barba, and L. T. De Paolis. A Virtual Tour for the Promotion of Tourism of the City of Bari. *Information*, 13(7), 2022. doi: 10.3390/info13070339

[7] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[8] S. Kim, S. Lee, and Y. M. Ro. Estimating vr sickness caused by camera shake in vr videography. In *2020 IEEE International Conference on Image Processing (ICIP)*, pp. 3433–3437, 2020. doi: 10.1109/ICIP40778.2020.9190721

[9] Y. Li, P. Tennent, and S. Cobb. Appropriate control methods for mobile virtual exhibitions. In *VR Technologies in Cultural Heritage: First International Conference, VRTCH 2018, Brasov, Romania, May 29–30, 2018, Revised Selected Papers 1*, pp. 165–183. Springer, 2019.

[10] T. Pakkanen, J. Hakulinen, T. Jokela, I. Rakkolainen, J. Kangas, P. Piippo, R. Raisamo, and M. Salmimaa. Interaction with webvr 360 video player: Comparing three interaction paradigms. In *2017 IEEE Virtual Reality (VR)*, pp. 279–280. IEEE, 2017.

[11] M. Slater et al. Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence: teleoperators and virtual environments*, 8(5):560–565, 1999.

[12] D. Veliz, R. Ccori, and L. Alfaro. Division and composition of 360° videos for the generation of personalized tours for tourism, hotel industry and education based on immersive VR. In *2023 IEEE World Engineering Education Conference (EDUNINE)*, pp. 01–06. IEEE, 2023.

[13] Wired. You can now watch and upload 360-degree videos on YouTube, 2015.

[14] M. Yu and C. Bi. Timeline Exploration in 360° Video. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*, SUI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3607822.3614533

[15] C. Zhou, Z. Li, and Y. Liu. A Measurement Study of Oculus 360 Degree Video Streaming. In *Proceedings of the 8th ACM on Multimedia Systems Conference*, MMSys'17, p. 27–37. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3083187.3083190

[16] D. Zielasko and T. Weissker. Stay Vigilant: The Threat of a Replication Crisis in VR Locomotion Research. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, VRST '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3611659.3615697