https://doi.org/10.2352/EI.2026.38.13.ERVR-194

Sublime VR: The Influence of Auditory Presentation

Yoshihiro BANCHI, Kosuke AOKI, Ryosuke SUZUKI, Takashi KAWAI; Department of Intermedia Art and Science, School of Fundamental Science and Engineering, Waseda University; Tokyo Japan.
Tomohiro ISHIZU; Department of Psychology, Graduate School of Psychology, Kansai University; Osaka, Japan.

Abstract

This study aims to clarify the role of sound in evoking the sublime experience within a virtual reality (VR) environment. The sublime is a complex emotion combining awe and fear, arising from vast objects or overwhelming forces. VR is considered an effective medium for safely inducing this experience. However, existing research has predominantly focused on visual factors, and the influence of auditory stimuli—essential for immersion—remains insufficiently explored. In this study, a 3D 360° video of a volcanic crater was presented via a head-mounted display (HMD) under three auditory conditions: Silent, Normal (natural environmental sound), and Reverbed (processed sound). We evaluated the experience using subjective measures (Awe Experience Scale) and objective physiological measures (Electrodermal Activity, pupil diameter, gaze data). The results demonstrated that the presence of sound significantly amplified the sublime experience across both subjective and objective indices. Specifically, the Normal condition showed high integration with visual information, eliciting the strongest emotional arousal and significant pupil dilation. Conversely, while the Reverbed condition induced spatial exploratory behavior (gaze dispersion), it caused a sense of incongruence between sight and sound, tending to lower the quality of the experience. These findings suggest that audio-visual congruence is critical in designing sublime experiences in VR.

Introduction

The sublime is a traditional aesthetic category referring to the complex emotion of awe and reverence mingled with fear, experienced when confronting vast objects, overwhelming power, or ambiguity. Unlike the simple pleasure associated with "beauty," the sublime involves a negative aspect of threat or terror. Recently, this experience has attracted attention in psychology as "awe," due to its potential to diminish the individual's self-focus ("small self") and promote prosocial behavior[1]. Virtual Reality (VR) technology, particularly Head-Mounted Displays (HMDs), is an extremely effective tool for inducing this sublime experience[2]. HMDs block external stimuli, allowing users to immersively experience overwhelming natural phenomena (such as massive storms or volcanic eruptions) that would be dangerous in the real world, while maintaining the "safe distance" required for the sublime to arise[3].

Most prior research on the sublime has focused on visual stimuli such as paintings and photographs[4, 5]. While the authors' previous research investigated visual factors in VR, such as Field of View (FOV) and stereoscopic presentation (3D/2D), the influence of auditory factors remained unexamined[6]. Although sound is essential for immersion and presence, existing research on sound and emotion has largely focused on music stimuli. The role of "environmental sounds" in dynamic VR environments remains largely unexplored.

This study focuses on "environmental sound" as a constituent element of the VR sublime experience. Specifically, we test the hypothesis that "reverb" processing, which adds spatial resonance, is effective for perceiving "vastness," a key component of the sublime. The objective is to multimodally investigate how the presence of sound (Silent vs. Sound) and the quality of sound (Normal vs. Reverbed) affect the sublime experience using subjective and physiological indices.

Method

Stimuli

The video content was created using the game engine Unity (ver. 2021.2.8f1). A 3D 360° VR environment was constructed using the Unity Asset Store package [7] "L.V.E (Lava & Volcano Environment)," depicting a scene surrounded by volcanoes approximately 60m in height. To evoke the feelings of "overwhelming scale" and "threat" associated with the sublime, the

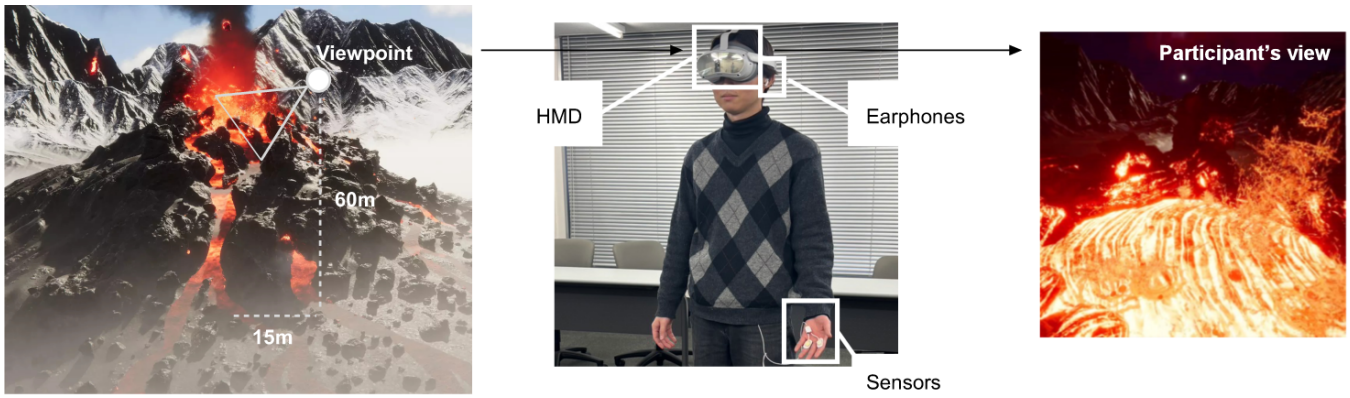


Figure 1. *Experimental Layout*

participant's viewpoint (camera position) was set at a close distance, approximately 15m from the erupting crater. Additionally, to improve the accuracy of gaze data alignment, a fixation cross video was inserted before the start of the main video (approx. 2 minutes).

The audio source was derived from environmental sounds extracted from a YouTube video filming an actual volcano via drone [8]. A specific section (3:00–4:00) where drone movement noise was minimal and the distance to the crater remained constant was sampled to create the following two audio types:

- Normal: The sampled environmental sounds (roaring, eruption) were used without processing. The mean sound pressure level was 79.1 dB.
- Reverbed: To simulate spatial expanse and "vastness," the sound was processed using a Schroeder Reverb algorithm implemented in Python. This simulated the reverberation of a vast space similar to a concert hall. The processed mean sound pressure level was 81.7 dB.

The video was presented using a Pico4 Enterprise HMD. Audio was presented via AirPods Pro with active noise canceling enabled. Physiological data were acquired using a biosignalsplux sensor.

Conditions

The following three auditory conditions were established. To control for order effects, the presentation order of conditions was counterbalanced across participants.

1. Silent: Video only (earphones were worn to maintain consistent physical sensation across conditions).
2. Normal: Unprocessed volcanic environmental sound presented.
3. Reverbed: Reverb-processed environmental sound presented.

Measures

The Japanese version of the Awe Experience Scale (AWE-S) was used to measure six factors: Time perception, Self-Diminishment, Connectedness, Perceived Vastness, Physical Sensations, and Need for Accommodation. Detailed introspective reports were conducted after the experiment.

Physiological indices included Electrodermal Activity (EDA) measured via finger sensors (analyzing Skin Conductance Level [SCL] and Response [SCR]). Additionally, pupil diameter and gaze behavior were recorded using the HMD's built-in eye tracker. To account for individual physiological differences, objective indices were standardized using baseline data collected during a resting state at the beginning of the experiment.

Procedure

Fifteen university students in their 20s (13 males, 2 females) participated in the experiment. Participants first received an explanation of the experiment and provided informed consent, followed by an attribute survey. Participants then donned the HMD and controllers were used to adjust Inter-Pupillary Distance and calibrate eye tracking. Subsequently, the HMD was temporarily removed to attach the EDA sensor and earphones, then re-equipped. A 90-second baseline measurement was conducted while viewing a neutral VR scene in a relaxed standing position. This was followed by the main trials. To improve gaze data accuracy, fixation cross was presented before each video. Participants then viewed the volcano video under the three auditory conditions (Silent, Normal, Reverbed) in a random order. Participants stood and were free to look around during the video. After each condition, participants removed the HMD/earphones and completed the AWE-S questionnaire. After all trials, a post-experiment interview was conducted to gather qualitative feedback on the video and sound.

Results

AWE-S

A one-way ANOVA on the AWE-S factors revealed significant differences between conditions. Multiple comparisons showed that compared to the Silent condition, both Normal and Reverbed conditions yielded significantly higher scores for F2 (Self-Diminishment), F3 (Connectedness), F4 (Vastness), F5 (Physiological), and F6 (Accommodation) ($p<.05$). Notably, for the core sublime factors of Vastness (F4) and Physiological changes (F5), there was an approximate 2-point difference between Silent and sound conditions. While there was no statistically significant difference between Normal and Reverbed conditions, Normal consistently scored higher across all factors.

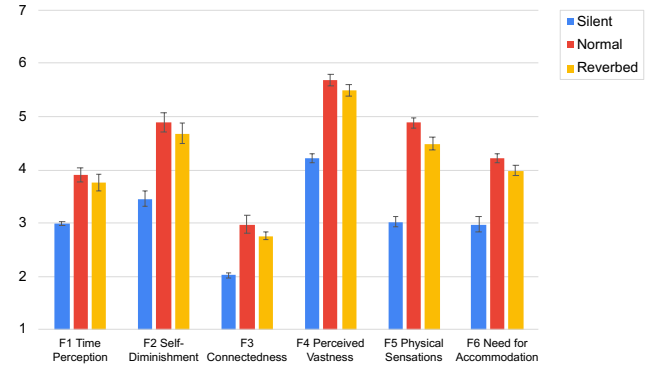


Figure 2. *AWE-S Questionnaire Results*

Pupil Diameter

Pupil diameter (standardized mean) showed significant differences across all measurement intervals (20s bins) ($p<.05$). The magnitude followed the order Normal > Reverbed > Silent. Normal sound elicited the largest dilation, followed by Reverbed, with Silent eliciting the least.

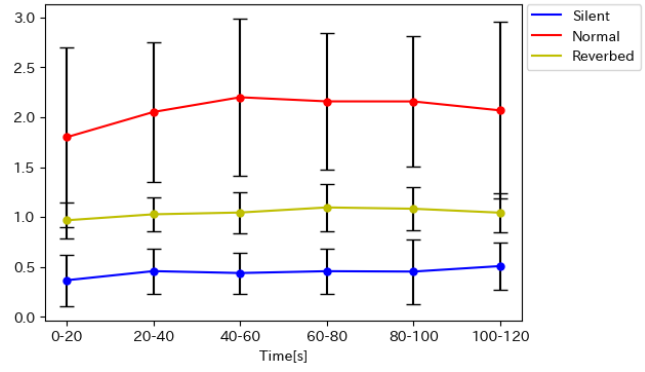


Figure 3. *Pupil Diameter over Time (20-second intervals)*

EDA

No significant differences were found in SCL or SCR count between conditions. However, the variance of the EDA amplitude was larger in the Reverbed condition. This suggests individual differences in reaction to the reverb, with some participants experiencing cognitive load or dissonance.

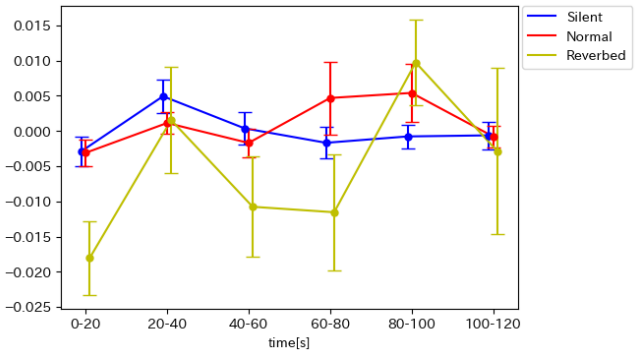


Figure 4. *EDA over Time (20-second intervals)*

Gaze Data

Heatmap analysis showed a general trend of gaze concentration on the crater. However, comparisons revealed that in the Reverbed condition, fixation points were less concentrated on the crater, with a tendency for participants to explore the broader spatial environment.

Discussion

In this study, we investigated the influence of sound on the sublime experience in VR using subjective evaluation and multiple physiological indices. The results confirmed that the presence or absence of auditory stimuli consistently caused differences in subjective evaluations of the sublime and pupil diameter changes.

Regarding the AWE-S results, both the Normal and Reverbed conditions showed significantly higher scores than the Silent condition for factors F2 (Self-diminishment), F3 (Connectedness), F4 (Vastness), F5 (Physical sensations), and F6 (Accommodation). This suggests that the addition of environmental sound complements physical and emotional aspects that are difficult to evoke via visual information alone, effectively reinforcing core sublime elements such as "Vastness" and "Physical sensations".

Conversely, no statistically significant difference was found between the Normal and Reverbed conditions in AWE-S scores. However, the Normal condition consistently showed higher mean values across all factors, and an order of magnitude was confirmed in pupil diameter: Normal > Reverbed > Silent. Additionally, in the Reverbed condition, a tendency for larger variance in EDA amplitude was observed. Introspective reports included comments such as "highly processed" and "the sound and video felt like separate things". These results indicate a potential divergence in participant reactions, suggesting that for some participants, the auditory spatial representation via reverb was perceived as an incongruence (uncanniness) with the visual information, potentially suppressing emotional arousal.

Gaze data analysis revealed that fixation points in the Reverbed condition were not concentrated around the crater but tended to disperse throughout the space. This result indicates that reverb processing functioned as a cue for "spatial vastness," influencing the allocation of visual attention. However, the fact that this exploratory behavior did not necessarily lead to an increase in subjective sublimity or physiological arousal is inferred to be influenced by factors such as the aforementioned "incongruence".

In summary, while sound influences the intensity and modality of the VR sublime experience, it is suggested that its effect depends not merely on adding auditory vastness, but on the relationship with visual information and individual differences.

Conclusion

This study examined the effects of auditory stimuli on the VR sublime experience using subjective and physiological measures, yielding the following findings:

• The presentation of environmental sound significantly increased subjective evaluations related to the sublime experience and pupil diameter changes compared to the silent condition.

• Regarding the influence of sound quality, the Normal condition using natural environmental sound elicited the highest responses in AWE-S factors and pupil diameter. Conversely, while the Reverbed condition with reverb processing showed a tendency for gaze to disperse throughout the space, the subjective sublime experience and physiological arousal (pupil diameter) did not surpass the Normal condition.

These results indicate that sound is a potentially important factor in altering the intensity of the sublime experience in VR space. In particular, while reverb processing functions to induce spatial gaze behavior, it did not necessarily result in increased intensity of the sublime experience under the conditions of this experiment. This suggests that auditory spatial effects can have complex impacts on the quality of the experience depending on their combination with visual information.

However, this study is limited by a small sample size and specific auditory conditions involving particular environmental sounds and reverb processing. Future research should more precisely manipulate sound source types and physical characteristics of spatial audio to examine in detail the conditions under which sound reinforces or alters the sublime experience.

References

[1] Piff, P. K., Dietze, P., Feinberg, M., Stancato, D. M., Keltner, D., "Awe, the small self, and prosocial behavior," Journal of Personality and Social Psychology, vol. 108, no. 6, pp. 883–899, 2015.

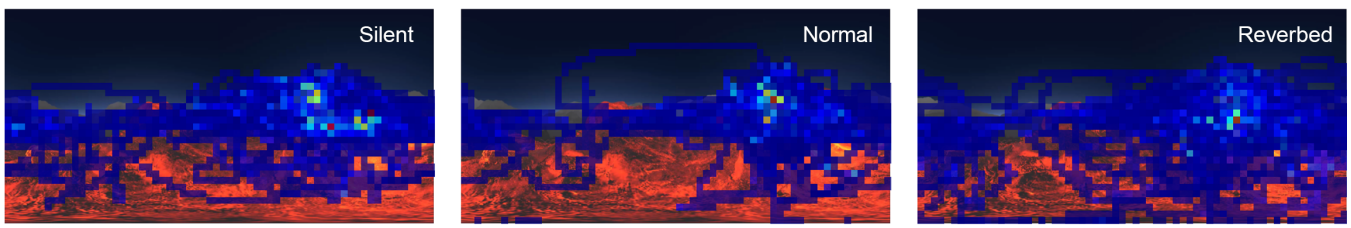


Figure 5. *Gaze Heatmap Results*

[2] Chirico, A., Yaden, D. B., Riva, G., Gaggioli, A., "The potential of virtual reality for the investigation of awe," Frontiers in Psychology, vol. 7, 2016.

[3] Pelowski, M., Hur, Y.-J., Cotter, K. N., Ishizu, T., Christensen, A. P., Leder, H., McManus, I. C., "Quantifying the if, the when, and the what of the sublime: A survey and latent class analysis of incidence, emotions, and distinct varieties of personal sublime experiences," Psychology of Aesthetics, Creativity, and the Arts, vol. 15, no. 2, pp. 216–240, 2021.

[4] Ishizu, T. & Zeki, S., "A neurobiological enquiry into the origins of our experience of the sublime and beautiful," Front. Hum. Neurosci. 8:891, 2014.

[5] Chirico A, Clewis RR, Yaden DB, Gaggioli A, "Nature versus art as elicitors of the sublime: A virtual reality study," PLOS ONE, 16(3): e0233628, 2021.

[6] Yoshihiro Banchi, Taisei Tsukahara, Tomohiro Ishizu, Takashi Kawai, "Effects of Stereoscopic Representations in Sublime Experiences Induced by Immersive VR" in Electronic Imaging, pp 343-1 - 343-5, 2025.

[7] Unity Technologies, "L-V-E-2023 Lava & Volcano Environment 2023", Unity Asset Store, https://assetstore.unity.com/packages/tools/terrain/l-v-e-2023-lava-volcano-environment-2023-253816?locale=ja-

[8] Traveller In The Whole World, "ICELAND VOLCANO REAL SOUND! CLOSE APROACH NEAR THE CRATER EDGE IN FULL ERUPTION MODE!," YouTube, https://www.youtube.com/watch?v=ie3gt9ftizA

Author Biography

Yoshihiro Banchi is an Assistant Professor at Waseda University, Japan. He received Ph.D., M.A. from Waseda University, in 2020, 2018. His research focuses ergonomics and data science on psyco-physiological effects in advanced technology, e.g. VR, XR, self-driving car.

Kosuke Aoki is a Master's Student at Waseda University, Japan. He received B.A. in Human Sciences from Waseda University, in 2024. His research focuses on psycho-physiological effects of audiovisual stimulation in mixed reality, particularly on perceived speed and crossmodal perception during locomotion.

Ryosuke Suzuki is a Master's Student at Waseda University, Japan. He received B.A. in Human Sciences from Waseda University, in 2024. His research focuses on analyzing audiovisual experiences using smartphones and spatial sound.

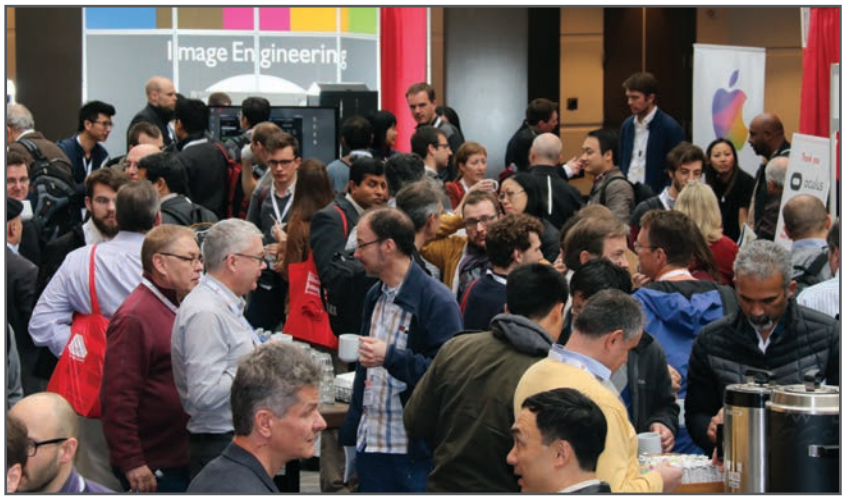
Tomohiro Ishizu is a Professor at Kansai University, Japan. Using experimental psychology, neuroimaging methods (functional MRI, MEG, EEG, NIRS), brain stimulation (tDCS), and neuroaesthetic research, we conduct 1) research on cognitive and brain functions related to aesthetic appraisal and artistic activities, 2) research on perception and techniques used in visual arts, and 3) research on the clinical and social application of empirical aesthetics science.

Takashi Kawai is a Professor at Waseda University, Japan. He received Ph.D., M.A., B.A. in Human Sciences from Waseda University, in 1998, 1995, 1993. His research focuses ergonomics in immersion technologies, e.g. 3D, VR, XR. He is a Certified Professional Ergonomist. Currently he is in charging of Japan Committee Chair of Advanced Imaging Society, Executive Committee of International Ergonomics Association, and Conference Chair of Stereoscopic Displays and Application.

JOIN US AT THE NEXT EI!

electronic IMAGING

Imaging across applications . . . Where industry and academia meet!



• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

ADVANCING IMAGING SCIENCE 75 1947 IS&T 2022 imaging.org