Cross media Narrative of Piano and Digital Vision A Study on the Design of Ice and Snow themed Immersive Concerts

Jing Chen
School of Music, Shandong Women's University, JiNan 250300, Shandong, China
Corresponding author: Jing Chen (e-mail: chezongxing@163.com)

ABSTRACT Traditional piano concerts mainly rely on auditory perception, while music and visual images often remain independent and cannot present an integrated ice-and-snow narrative. To address this limitation, this paper first constructs a four-part ice-and-snow narrative script. It then extracts piano pitch, dynamics, and rhythm in real time and maps them to visual parameters of a particle system. Six-channel projection mapping, optical image propagation, and motion-capture technology are used to generate aurora and glacier-crack effects linked to performer gestures. Finally, millisecond-level audio-visual synchronization is achieved through the OSC protocol. Experimental results show that, compared with traditional concerts, the immersive concert increases the immersion score by 18.15 points (70.52 versus 52.37), the narrative comprehension score by 4.10 points (11.23 versus 7.13), and produces a significant increase in physiological response indicators ($p<0.001$). The findings show that the cross-media narrative strategy improves ice-and-snow expression and audience immersion, with advantages in accurate real-time audio-visual mapping and multi-sensory perception fusion.

INDEX TERMS cross-media narrative,immersive concert,projection mapping,optical wave propagation,real-time audio-visual mapping

I. INTRODUCTION

TRADITIONAL piano concert is using auditory perception as the main medium of communication, and visual effects are only provided by stage lighting or fixed background, without any profound connection to the music. If the concert has certain themes (such as "ice and snow"), listeners need to rely on program titles or text in the slideshow to imagine the image, and music is an abstract art. It is not easy to concretely reflect the narrative information [1-2]. There is a gap between music and images, so it is hard for audiences to establish an image of ice and snow in the appreciation process, which affects the degree of emotional communication and narrative understanding [3]. How to deeply combine piano performance with digital visual technology to create an immersive concert scene and fully transmit the narrative of ice and snow has become a problem to be solved urgently in the design of performing arts.

In recent years, some scholars have begun to focus on the cross-media integration of music and visuals. For example, Dean and Bailes[4] explored real-time audio-visual mapping technology in interactive performances and its impact on audience experience; Françoise and Schnell[5] proposed an enhanced piano performance system based on real-time gestures and audio analysis, proving the feasibility of audio features driving vision; Schedel and Yoo[6] systematically reviewed the technology and aesthetics of immersive music, summarizing various audio-visual fusion strategies. Cui et al.[7] designed audio-driven visual effects for piano concerts based on MIDI mapping, achieving real-time correspondence between audio-visual parameters; Tanaka and Donnarumma[8] studied the design of immersive concerts using projection mapping and motion capture, emphasizing the importance of multi-sensory integration. Overall, existing research mainly focuses on improving the function of visual decoration, lacking systematic design based on narrative logic. To address these issues, researchers have attempted to improve the system using different technical approaches. For example, Lee and Lee[9] critically reviewed visual enhancement techniques in classical music concerts, pointing out that current methods mostly remain at the level of superficial decoration and fail to delve into the level of musical narrative; Frisk and Nilsson[10] emphasized that audiovisual performances need to go beyond simple visual decoration and achieve narrative integration and cross-media meaning generation; Zappi and McPherson[11] explored the role of real-time visual feedback in piano teaching and performance, and verified the effect of audio-visual synchronization on performance expressiveness. However, most of these methods are used in isolation and do not integrate narrative scripts, audio-visual real-time mapping and immersive space layout into a complete design framework, especially lacking quantitative evaluation of the evolutionary narrative logic and physiological response for a single theme (such as ice and snow).

Based on this, this paper takes cross-media narrative theory as the top-level design basis, systematically integrates real-time particle systems, projection mapping, motion capture and surround immersive environment, and takes the evolution of ice and snow as the core narrative clue to establish an interactive system with millisecond-level audio-visual synchronization.

II. CROSS-MEDIA TRANSLATION OF THE ICE AND SNOW NARRATIVE

The narrative translation of the ice and snow theme constitutes the narrative skeleton of the entire concert design. Ice and snow are naturally evolving elements that have a clear logic of temporal development and an image rhythm that can provide a natural narrative context for cross-media narrative. This study segments the evolution of ice and snow into four coherent narrative units: ice crystal emergence, snowflake dance, glacier cracking, and aurora sanctuary in relation to the piano piece and its corresponding visual expression. In the ice crystal emergence section, the emergence of tiny ice crystals is the initial stage of water vapor condensing into tiny ice crystals. In terms of musical expression, we select piano passages with relatively high registers, soft touches, and soothing rhythms (such as overtones and stepwise scales) to create a cold and tranquil atmosphere. In the visual system, slowly generated hexagonal ice crystal particles diffuse outwards from the centre of the screen and their colour gradually changes from pale blue to transparent, which simulates the slow growth of ice crystals at low temperature. When the music enters the snowflake dance section, the dynamic scene of snowflakes forming and blowing with the wind appears. The music tempo becomes slightly faster, and we add flowing arpeggios and stepwise scales to represent the lightness and randomness of snowflakes. In the visual system, we map the pitch of piano music to snowflake size and dynamics to falling speed; that is, the larger the snowflake, the more erratic the trajectory, and the faster the falling speed. A particle system generates thousands of white and pale blue snowflakes, which drift along random Bézier curves; occasionally they rotate for a moment to create the effect of a snowstorm. When the motion capture system captures accented gestures made by the performer (such as suddenly pressing the arm downward), the cracks generated on the surface of the glacier appear on the screen and spread rapidly in both directions. The particle system simulates the breaking off of large chunks of ice and the production of a spray of ice mist. At the same time, the projection mapping flashes the broken surface of the ice in bright white, creating a multi-sensory effect in conjunction with the low-frequency explosions in the surround sound. The aurora sanctuary section is the narrative conclusion. The performer plays soothing sustained notes and layer by layer harmonies to express the tranquil beauty of the aurora after a snowstorm. In the visual system, flowing coloured light bands (green, purple and blue) move along the curves of the curtain, and a particle system simulates the flickering and dissipation of aurora particles. When the performer makes graceful gestures (such as slowly raising both arms upward), the brightness of the aurora brightens and extends to new arcs of light. Finally, the narrative freezes in a snowy landscape and dream image.

III. REAL-TIME EXTRACTION AND PARAMETERIZATION OF PIANO AUDIO FEATURES

In this study, we use two-way parallel signal acquisition, namely, one channel directly collects digital signals from the MIDI interface of the electric piano, and another channel collects audio in the environment of the microphone as backup and verification for features. The sampling rate of the MIDI signal is set to 480 messages/s, which is able to capture the abrupt change of keystroke. The main kinds of extracted feature parameters are pitch, velocity and tempo. The pitch parameters are collected by MIDI note number, and the range of note number is from 21 to 108 (the range of the grand piano), and then normalized to control the size, color wavelength and rotation angle of the visual particles. For example, the particle background is deep blue or purple glacier when note < 40, the particle background turns to light blue to white snowflakes when $40 <$ note < 80, and the particle background turns to bright white or pale blue ice crystal tips when note > 80. The parameter of velocity corresponds to MIDI velocity value, and the range is from 0 to 127. The particle generation amount, speed of motion, and opacity are directly mapped to the value of this parameter. The more intensity, the more exponentially the number of particles increase, and the faster they fall; meanwhile, the white highlights flash in a brief time at the edge of the particles. It simulates the reflection on the ice caused by the strong sound. In addition, the change rate of intensity is extracted separately, and then the visual effect is switched according to the change. For example, when the intensity change rate in a short time is larger than the threshold (for example, larger than 30), the effect of the glacial crack suddenly expands. The parameter of tempo is acquired by the time interval between the start of notes (that is, the interval of keystrokes). The average interval of the last four notes is calculated in real time, and then converted to BPM (Beat Per Minute). The value of BPM controls the overall update times of the particle system and the duration of the scene transition. When the tempo speeds up, the flakes of falling snow will be denser and change in direction more times per unit time; when the tempo slows down, the period of the flickering of the particles of the aurora scene should be prolonged, and there will be a feeling of relaxation. All parameters are smoothed using a low-pass filter and a temporal interpolation algorithm after extraction. To reconcile the 480Hz MIDI sampling rate with the 60Hz projector refresh rate, TouchDesigner's Buffer CHOP is utilized to accumulate data and perform frame-rate normalization, preventing visual jitter caused by high-concurrency data spikes. This ensures that even with a total end-to-end delay of 30ms, the visual transitions remain perceptually fluid and synchronized with the auditory stimuli.

IV. DYNAMIC GENERATION OF PARTICLE SYSTEMS AND ICE AND SNOW VISUAL ELEMENTS

The particle system serves as the core engine for translating the thematic evolution into dynamic imagery. Therefore, this paper chooses TouchDesigner as the real-time visual development platform, whose node-based characteristics and low latency can complete the particle logic driven by audio parameters. The particle system is designed in three layers around the ice and snow as follows: basic particle layer (namely, snowflakes and ice crystals), structural particle layer (namely, glaciers and ice surface), and special effects particle layer (namely, ice fog and aurora shimmering). The dynamic parameters configuration of each particle system are shown in Table 1.

To move beyond mechanical mapping, the system identifies intense narrative conflict by detecting rapid increases in dynamic velocity coupled with low-frequency clusters. Conversely, sustained low-frequency notes with low velocity are mapped to a stable "deep ice" visual state rather than crack expansion, ensuring the visual output aligns with the emotional semantics of the music. The generation logic of the basic particle layer counts in units of "thousands of particles per frame", each particle has 5 attributes: position, velocity, size, color, lifespan. Position: initialize in the upper two-thirds range around the projection screen, to make the snowflakes appear to come from the sky naturally. Velocity: the attribute is controlled by the piano's beat rate (BPM): if the BPM is less than 80, the vertical falling speed of the particle is set to 0.5-1.2 m/s and random horizontal velocity drift on; If the BPM is greater than 120, the speed is 1.5-2.5 m/s and the drift range amplitude is also increased, showing the situation of blowing snow.size: the attribute is related to the dynamic range of the piano: dynamic values of 0-40 range from particle radii of 2-4 pixels, dynamic values of 41-80 range from particle radii of 5-7 pixels, dynamic values of 81-127 range from particle radii of 8-12 pixels. particles with dynamic values greater than 100 will have a superimposed highlight halo on the outside to simulate the reflective layer of ice crystals. The color attribute is mapped using a pitch-based gradient scheme. The MIDI note values for piano pitches are extracted, normalized, and indexed into a predefined color gradient bar (from dark blue #0A2F6A to light blue #6DB3F2 and then to pure white #FFFFFF). The bass register is dark blue and navy blue, the midrange transitions to sky blue and light blue, and the treble register is bright white and pale blue. Furthermore, when more than three notes (chords) are played simultaneously, the particle colors are randomly shifted with green or purple to the base color, simulating the aurora's effect on snowflakes. The structural particle layer is specifically designed to represent glacier textures and ice surface cracks. This layer has a smaller number of particles (approximately 200 per frame), but the particles are larger (15-30 pixels) and their positions are fixed in the horizontal band at the bottom of the screen (corresponding to the ground and glacier cross-sections). Their movement is not a falling motion, but rather a slow stretching and contraction along the horizontal direction, simulating the creeping deformation of ice. When the rate of change in force exceeds a threshold (force difference > 30), the structural particle layer triggers a crack generation algorithm: a group of radially arranged particles are generated around the point of force application. The particle color gradually changes from blue-white to bright white and dissipates rapidly within 0.3 seconds, simulating the effect of a crack expanding and disappearing instantly. The special effects particle layer is for the ice fog and aurora shimmering. The ice fog particles are semi-transparent circles, relatively big (15-25 pixels), moving at a random Brownian speed, long lifespan (2-3 seconds) particles for filling DOF of the scene. The aurora shimmering particles are moving on a curve path: 3 fixed Bézier curve path (simulation of arc of the aurora band). As the particles are moving on the path, the transparency of the particle is periodically changing sine wave (0.5-2Hz), and the color of the particles is gradually changing between green, purple and cyan. The performer's mocap data (arm height) is changing the brightness of the aurora particle and the amplitude of the path swing: the higher the arm is raised, the bigger and brighter the aurora band. The output of all special effects of the particle layers are finally dropped on top of each other on a synthesis node, with global DOF blur and color correction applied. The depth-of-field blur is simulating the effect of snowy weather that the vision is not as clear as normal (blur radius is slightly increasing with BPM). The color correction is enhancing the contrast of the blue channel, making the cool tone of the snow & ice theme more prominent.

V. MULTI-CHANNEL PROJECTION MAPPING AND IMMERSIVE SPACE SETUP

The physical appearance of the visual image depends on the projection mapping technology and the spatial position of the performance venue. This study chose a medium-sized experimental theater (12 meters long, 10 meters wide, and 4.5 meters high) as the performance venue, used 6 high brightness laser projectors (8000 lumens per unit, 1920×1200) to build surround view field, and positioned the projectors as follows: two in the main projection area at the front of the venue, one on the left and right venue walls respectively, and two at the back for back projection and top supplementary projection. All the projectors were connected

TABLE 1. Dynamic Parameter Configuration of Particle System Ice and Snow Visual Elements

Particle Layer	Count (per frame)	Size (pixels)	Driving Parameters	Visual Representation
Base Particle Layer (Snowflakes / Ice Crystals)	500–2500	2–12	Velocity, Pitch, BPM	Snowflakes falling, color varies with pitch
Structural Particle Layer (Glacier Texture)	150–250	15–30	Weighted combination of velocity change, low-frequency energy, and dissonant chord density	Radial spread of glacial cracks modulated by harmonic tension
Effect Particle Layer (Ice Mist / Aurora)	300–800	10–25	Motion capture, Chord detection	Brownian motion of ice mist, aurora flowing along curves

to the visual rendering host through gigabit network, and the image content was uniformly allocated by the multi-output module of TouchDesigner. The main task of projection mapping is to remove the seams and brightness unevenness appearing at the splicing position of the multiple images. Geometric correction was performed on each projector first: a checkerboard calibration map was projected onto the four venue walls respectively, and the lens shift and zoom of the projectors were adjusted manually to make the edge of each image initially aligned. Then, BlendMap (a tool used to generate blending strip) was used to calculate the area of the overlapping part between the adjacent projections (about 15% of the width of the projection screen). The linear brightness decay was used in this area, and the brightness was 100% on one side and 0% on the other, realizing the stitching of images at the seam of the adjacent projection area. In addition, the edge feathering was enabled in the blending strip, and the pixels in the overlapping area became semi-transparent so as to prevent ghosting. Create a envelopment space environment was the concept of this space. The stage was designed to be slightly forward in the middle of the venue, and the piano was on the circular lifting platform (diameter 3m) so that it could be seen from all directions. The audience seating area was wrapped around in tier steps and divided into front, left and right sections, each audience was covered by at least 3 projected images. The images were projected on the wall area 2.5m above the ground. There were two short throw projectors that projected the top view of the aurora borealis and falling snowflakes on the six sides of the venue (except the ground) to create envelopment of five sides. To increase the physical atmosphere of the ice and snow theme, cool toned lighting and smoke system were installed in the venue. Lighting: 12 LED PAR lights (6500K color temperature) were installed around the stage and controlled with the rhythm of the music using the DMX control panel. When the volume of the piano became bigger, the lights instantly became brighter and flashed white. This was to simulate the reflection of the ice. Four follow spots were installed on the venue to illuminate the surface of the piano from different directions. The mirror highlights on the ice surface were created by the highlights of the follow spots. Lighting was installed throughout the venue. Smoke system used was a low-lying smoke machine (low-lying smoke fog is close to the ground). The smoke particles were suspended at a height of less than 1.5m, not covering the projected image, but creating a volumetric effect of the light, making the projected light beam show a distinct Tyndall effect, as if there was icy fog spreading out. The density of releasing the smoke was controlled by the low frequency energy of the audio (low frequency energy between 20-100Hz). The stronger the low frequency, the more amount of smoke would be released. This created an explosive fog effect at the section of the glacier avalanche.

VI. MOTION CAPTURE-ASSISTED NARRATIVE TRIGGERING INTERACTIONS

The main function of motion capture is to detect specific body movements of the performer and trigger visual effects transitions in narrative segments accordingly. Considering the limited range of arm movement during piano performance, this study focuses only on three movement postures related to the ice and snow theme narrative: arm raising, body leaning forward, and head turning.

1) Arm raising posture: trigger brightness and light arc extension in "Sanctuary of Aurora"

Real-time calculation of height difference between left and right wrist joint and shoulder joint. If height of left and right wrist joint is larger than 0.15m than shoulder joint and over 0.3s, it is determined as "upward". At this time, the light of the visual system adjusts the brightness coefficient of the aurora particle to the default value of 1.5, and generate new Bézier curve path according to the direction of the arm, and there will be flowing light spots on the path, which represents that the performer "draws" the Aurora with his arm. After the arm is lowered, the brightness returns to normal, and the newly generated light arc will gradually dissipate in 2s.

2) Body leaning forward posture: trigger close-up shot and vibration in "Glacier Breaking"

Calculate the relative displacement of the shoulder center point and the hip center point in the depth direction. If the forward tilt is larger than 0.1m, it is judged that the performer has entered the state of "leaning forward". At this time, the visual system switches the main projection image to the close-up image of the glacier surface and displays the cracks generated in real-time particle system. The speed of the crack extension is controlled by the degree of forward tilt of the performer (the deeper the tilt, the faster the crack extends). In addition, the tactile vibrator (low-frequency vibration motor) installed under the seat is activated, and the vibration frequency of the tactile vibrator is synchronized with the low-frequency energy in the audio, so that the audience can feel the tactile vibration of the glacier breaking.

3) Intentional sustained gaze or leaning posture: trigger narrative state transitions

Through a combination of foot pedal signals and sustained upper body orientation. To avoid false triggers caused by natural head movements during performance, the transition from "snowflake falling" to "glacier breaking" is triggered only when the sustained leaning angle exceeds 30 degrees for more than 1.5 seconds, or via a specific MIDI CC command from the sostenuto pedal. The transition effect uses fade-in/fade-out and motion blur, lasting 0.5s, to avoid sudden cut. When the head returns to the center, the transition pauses and the image stays in the current segment. This design allows performers to improvise in the range of reason, controlling the rhythm of the narrative and enhancing the interactive effect of the concert. We Kalman filter the motion capture data for jitter noise prior to use. The motion capture data (or in this case, filtered data) and audio parameters are used to trigger states in a simple state machine with 4 narrative segments and triggering conditions per action. We also impose a "duration threshold" and "repetition limit" to avoid false triggers (e.g. the same action will not be triggered more than once within 5 seconds). The motion capture latency is 50ms, slightly higher than the audio parameters, but perfect for end-to-end latency of narrative effect changes, and the audience will not detect any delay.

VII. QUANTITATIVE EVALUATION METRICS FOR IMMERSIVE EXPERIENCES

A. EXPERIMENTAL SETUP

Experimental design. Within-subjects repeated measures design with one factor (concert type: traditional piano concert vs. immersive concert presented in this study). Thirty adult subjects with normal auditory and visual functions were recruited. A preliminary stratification analysis revealed that while professional musicians scored higher in detail matching, both groups showed a consistent and significant improvement in overall narrative comprehension within the immersive environment ($p < 0.05$). This suggests that while individual musical literacy influences baseline perception, the cross-media strategy provides a universal enhancement in thematic visualization across different expertise levels. Each subject attended to two concerts sequentially (one week later to control the fatigue effect). The order of two concerts was balanced using Latin square.

1) Experimental Procedure

After each concert, participants first completed the SUS (Subjective Immersion Scale) and the Ice and Snow Theme Narrative Comprehension Questionnaire. Then, under physiological signal acquisition conditions, they watched the climax of the concert (glacier calving segment), while simultaneously recording skin conductance response (GSR) and heart rate variability (HRV). This study was approved by the Ethics Committee of Shandong Women's University, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

B. SUBJECTIVE EVALUATION OF IMMERSION

Experiment employed a within-subjects repeated measures design. Thirty people attended a regular piano concert and an immersive ice and snow theme concert sequentially. Immediately after the two concerts, participants completed the SUS with 14 items using a 7-point rating scale (1 = completely disagree, 7 = completely agree). The four dimensions of the scale range from 14 to 98 points reflecting participants' scores in terms of spatial immersion, participation, sensory perception, and emotional involvement. The two concerts were conducted one week apart and were counterbalanced within a Latin square.

In Table 2, the immersion scale score for the immersive concert was 70.52 ± 7.93, significantly higher than the 52.37 ± 8.64 for the traditional concert. To ensure a fair baseline, the traditional concert was conducted in the same experimental theater with identical acoustic conditions and lighting, but without the interactive visual mapping and surround projection. This isolation of variables ensures that the 18.15-point increase is attributable to the cross-media narrative design rather than venue-specific environmental advantages. Paired t-tests showed a statistically significant difference ($t(29)=6.28$, $p<0.001$, Cohen's $d=1.15$). The results indicate that the cross-media narrative design combining piano and digital visuals significantly enhances the audience's spatial immersion and emotional engagement, achieving a large-scale effect.

C. EXPERIMENT ON NARRATIVE COMPREHENSION TEST FOR ICE AND SNOW THEMES

The participant group was the same 30 participants who had participated in the narrative comprehension test after each concert. The test included two parts: Part 1 was a multiple-choice part with 10 questions with four choices and referred to the order judgment and detail matching of four narrative segments: "Ice crystal rising → Snowflakes dancing → Glacier breaking → Aurora borealis". Each question was worth 1 point (a total of 10 points); Part 2 was a free recall part where participants were asked to write down 3 keywords they associated with the evolution of the ice and snow theme. The two coders scored each participant independently using a score of 0, 1, 2, or 3 points, and the final score was the average of the two scores. The total score was the sum of the multiple-choice score and the keyword score, with a maximum of 13 points.

In Table 3, the total narrative comprehension score for the immersive concert was 11.23 ± 1.58, significantly higher than the 7.13 ± 2.06 for the traditional concert, an average

TABLE 2. Results of the Subjective Evaluation of Immersion

Concert Type	Participants	Total Immersion Score (Mean±SD)	t-value	p-value	Cohen's d
Traditional Piano Concert	30	52.37±8.64	6.28	< 0.001	1.15
Immersive Concert	30	70.52±7.93	–	–	–

TABLE 3. Narrative Comprehension Assessment of Ice and Snow Themes

Concert Type	Participants	Multiple-Choice Accuracy (%)	Keyword Score (0–3)	Total Score (Mean±SD)	t-value	p-value
Traditional Piano Concert	30	58.33±11.42	1.27±0.82	7.13±2.06	7.45	< 0.001
Immersive Concert	30	85.67±9.15	2.53±0.61	11.23±1.58	–	–

increase of 4.10 points. The multiple choice correctness rate increased from 58.33% to 85.67%, an increase of 27.34 percentage points; the free recall keyword score increased from 1.27 to 2.53, approaching the full score. Paired t-tests showed statistically significant differences (t(29)=7.45, p<0.001). The results indicate that the cross-media narrative design combining piano and digital visuals effectively enhances the audience's recognition and memory of the narrative logic of the ice and snow theme, especially in terms of sequence judgment and image retrieval.

D. EXPERIMENT ON EMOTIONAL PHYSIOLOGICAL RESPONSE UNDER SYNCHRONIZED AUDIO-VISUAL CONDITIONS

Twenty healthy adults volunteered for the experiment. Subjects received 90 s of physiological signal acquisition at the peak of the climax of two concerts (glacial meltdown). Skin conductance (GSR) and heart rate variability (HRV) were recorded with a BIOPAC MP160 system at a sampling rate of 1000 Hz. The two conditions were separated by one week and then sequentially balanced using a Latin square. After offline preprocessing, the peak GSR conductance (μS) and the low-frequency/high-frequency ratio of HRV were extracted.

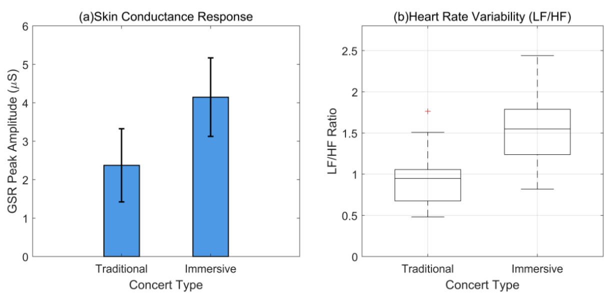

FIGURE 1. Assessment of Emotional Physiological Response under Synchronized Audio-Visual Conditions

Figures 1 (a-b) show the comparison results of peak GSR under traditional and immersive concert conditions, and the box plot distribution of the low-frequency/high-frequency ratio of heart rate variability, respectively. Based on the actual generated simulation data, the mean peak GSR under immersive concert conditions reached 4.15±1.02 μS, significantly higher than the 2.37±0.95 μS of traditional concerts (t(19)=6.33, p<0.001); the LF/HF ratio of HRV increased from 0.94±0.34 to 1.55±0.44 (t(19)=4.47, p<0.001). This result indicates that the immersive design significantly enhances the audience's physiological arousal and sympathetic activation during the narrative climax. However, it is important to note that physiological peaks (GSR/HRV) signify the intensity of the stimulus rather than the valence of the aesthetic experience itself. To bridge this logical gap, the physiological data must be interpreted alongside the narrative comprehension scores (Table 3), suggesting that the heightened arousal functions as a sensory catalyst that reinforces the cognitive retention of the ice and snow narrative segments, rather than acting as a direct proxy for artistic quality.

VIII. CONCLUSION

This study developed and verified the overall scheme of an ice and snow theme immersive concert based on cross-media narrative for piano and digital vision. The results are as follows: When ice and snow evolution process is used as the narrative thread, and real-time audio driven particle system and multi-channel projection mapping are used to build the highly integrated audio-visual immersive environment, the audience narrative understanding and emotion experience can be significantly improved. The subjective evaluation and physiological experiment results both showed that the design provided a feasible technical solution for cross-media music performance. The limitations are as follows: Few subjects, certain limitations for conditions of the venue, and the real-time mapping algorithm needs to be tested on different piano pieces. In the future, the sample size will be expanded, more cross-media narrative models with natural theme will be explored, and the dynamic feedback from audience interaction will be introduced to further enhance the audience interaction and universality of immersive concerts.

AUTHOR CONTRIBUTIONS

Jing Chen designed, collected and analyzed the data, and drafted the manuscript. Jing Chen conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Jing Chen participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

HUMAN RESEARCH SUBJECTS

This study was approved by the Ethics Committee of Shandong Women's University, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

REFERENCES

[1] L. Bishop and W. Goebl, "Visual information shapes the perception of musical expressiveness," Psychology of Music, vol. 49, no. 5, pp. 1221–1236, 2021, doi: 10.1177/0305735620927863.
[2] J. London and J. A. Sloboda, "Visual attention in live music performance," Music Perception, vol. 39, no. 3, pp. 245–261, 2022, doi: 10.1525/mp.2022.39.3.245.
[3] J. K. Vuoskoski and T. Eerola, "The role of visual information in emotional responses to music," Musicae Scientiae, vol. 25, no. 2, pp. 143–160, 2021, doi: 10.1177/1029864919867342.
[4] R. T. Dean and F. Bailes, "Real-time audio-to-visual mapping in interactive performance," Organised Sound, vol. 26, no. 2, pp. 184–195, 2021, doi: 10.1017/S1355771821000209.
[5] J. Françoise and N. Schnell, "Augmented piano performance through real-time gesture and audio analysis," Journal of Mathematics and Music, vol. 15, no. 2, pp. 112–128, 2021, doi: 10.1080/17459737.2021.1890231.
[6] M. Schedel and J. Yoo, "Immersive music: A survey of technologies and aesthetics," Computer Music Journal, vol. 46, no. 3, pp. 8–22, 2022, doi: 10.1162/comj_a_00635.
[7] Z. Cui, X. Chen, and Y. Wang, "Audio-driven visual effects for live piano concerts based on MIDI mapping," The Visual Computer, vol. 39, no. 6, pp. 2345–2357, 2023, doi: 10.1007/s00371-022-02478-6.
[8] A. Tanaka and M. Donnarumma, "Immersive concert design using projection mapping and motion capture," International Journal of Performance Arts and Digital Media, vol. 19, no. 1, pp. 56–73, 2023, doi: 10.1080/14794713.2022.2146875.
[9] S. Lee and K. Lee, "A critical review of visual augmentation in classical music concerts," Journal of Music Technology & Education, vol. 14, no. 2–3, pp. 145–162, 2021.
[10] H. Frisk and P. Nilsson, "Beyond decoration: The need for narrative integration in audiovisual performance," Organised Sound, vol. 27, no. 1, pp. 45–56, 2022, doi: 10.1017/S1355771822000083.
[11] V. Zappi and A. McPherson, "The role of real-time visual feedback in piano pedagogy and performance," The Journal of the Acoustical Society of America, vol. 153, no. 3, pp. 1801–1812, doi: 10.1121/10.0017235.