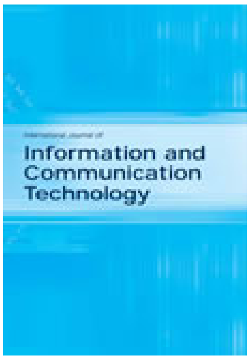


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Harmonising code and composition: computational design strategies for multimedia classical concerts

Shanshan Li

DOI: 10.1504/IJICT.2025.10072179

Article History:

Received:	21 April 2025
Last revised:	14 May 2025
Accepted:	15 May 2025
Published online:	20 July 2025

Harmonising code and composition: computational design strategies for multimedia classical concerts

Shanshan Li

Huainan Normal University,
Huainan, Anhui, 232038, China
Email: lss13675547827@163.com

Abstract: There are growing intersections between classical music performance, its tradition, acoustic fidelity, and advanced computational technologies, creating immersive multimedia concert experiences. We discuss how classical concert design is advanced by integrating real-time audio processing, machine learning models, generative visual engines, and audience interaction frameworks. A modular architecture was developed and tested through two case studies (Beethoven 360° and Bach Rewired), demonstrating two separate use cases of artificial intelligence, real-time rendering, and biofeedback systems. Sensor-based audience input has been used to learn about the adaptive performance elements, and musical structure was mapped to visual output using deep learning models. Results showed that audience immersion increased from an average of 5.4 to 6.8, while system latency was less than 25 ms. The findings suggest that computational multimedia systems might improve the classical concert experience by bringing new ways for emotional expression, structural clarity, or participatory design while maintaining the integrity of the repertoire.

Keywords: multimedia concert design; computational music systems; artificial intelligence in music; real-time audio-visual synchronisation; generative visuals; audience interaction; machine learning in performance arts.

Reference to this paper should be made as follows: Li, S. (2025) ‘Harmonising code and composition: computational design strategies for multimedia classical concerts’, *Int. J. Information and Communication Technology*, Vol. 26, No. 27, pp.67–91.

Biographical notes: Shanshan Li is affiliated with Huainan Normal University, Huainan, Anhui, China. She contributes to academic teaching and research at the institution. Her professional work likely involves education or scientific development in her field of expertise.

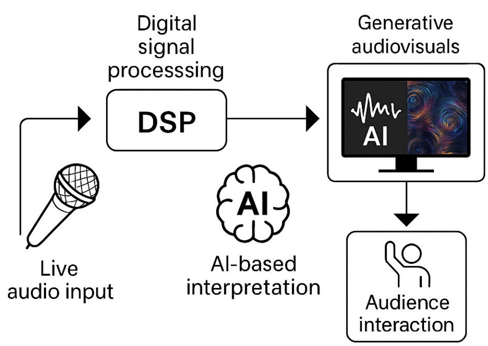
1 Introduction

A classical music tradition has contained acoustic purity, ritualised presentation, and historical preservation. Concert halls aimed at the maximum natural resonance of strings and winds, passive, reverent audiences, and no more visual engagement with the performers than their gestural nuance (Cairoli and Agostinelli, 2024; Vickers, 2021). But, in the 21st century, as classical performance reaches 21st audiences exposed to interactive and multimedia experiences in immersive theatre, digital installations, gaming environments, and augmented realities, classical performance is being re-imagined digitally in the context of computational enhancement.

Over the past years, artists and institutions have begun to manipulate concerts from passive time-based events into immersive multimedia experiences by applying real-time data processing, artificial intelligence (AI), generative visual systems, and sensor-driven feedback systems (Tromp, 2025; Bloomberg, 2020; Jagatheesaperumal et al., 2021). The motivation for this evolution stems from changing cultural norms, declining attendance at the traditional location, and the availability of affordable, accessible digital tools that can process things in real-time (McHaney, 2023; Dede, 1996; Sherron and Boettcher, 1997). In today's world, and especially with younger audiences, one looks for a performance that is not just musically profound but also visually and experientially.

There are, however, no precedents for these developments. Before the power of today's computing could make it scalable, pioneers like Iannis Xenakis or Karlheinz Stockhausen imagined a synthesis of sound, space, and light (Humięcka-Jakubowska, 2019; May, 2020). Today, however, systems use powerful graphical rendering engines (e.g., unity or unreal engine) combined with digital signal processing (DSP) and AI that can interpret, predict, and visually represent musical features on the fly in Figure 1. The modern computational concert system is a pipeline that includes live audio capture, feature extraction, machine learning inference, and multi-sensory rendering outputs augmented with audience layers.

Figure 1 Conceptual architecture of a computational concert system integrating live audio input, AI-driven interpretation, and generative audio-visual output (see online version for colours)



Unlike stylised concert formats, which depend on static staging and lighting to support the purely aural narrative, computational systems create dynamic, reactive environments. Thus, these systems can highlight formal structure (e.g., sonata form or fugue entries), react emotionally via colour and motion, and engage the audience via interactive devices (Fitzgerald, 2004; Kanellos, 2024; Filimowicz, 2022). Table 1 compares fundamental aspects of the design of traditional concerts to those augmented by computational methods, and in particular demonstrates a fundamental shift from a fixed presentation to an adaptive, data-driven performance design.

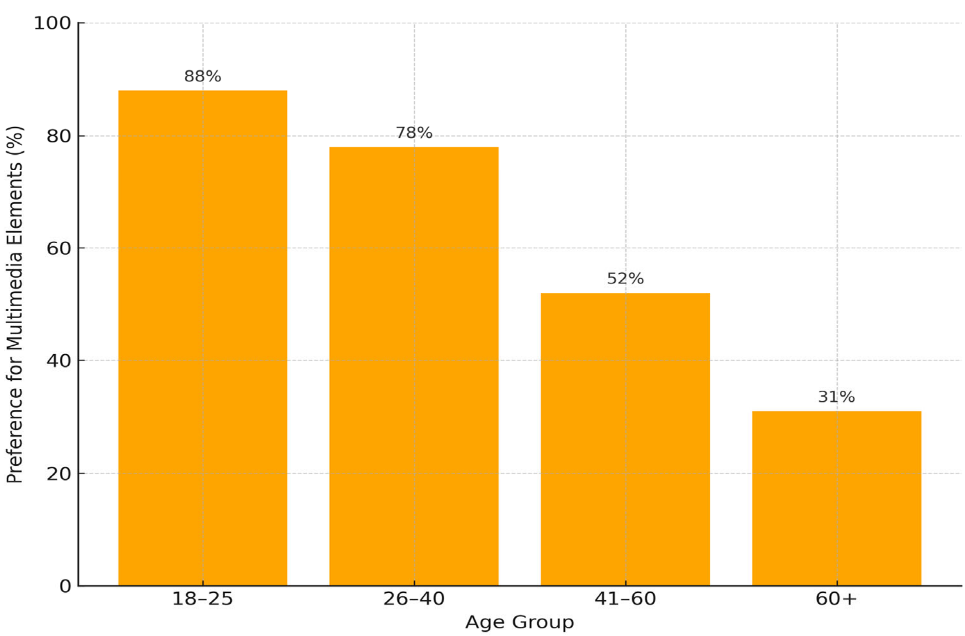
It has many motivations. The ecosystem of these multimedia augmentation tools (AI toolkits, OSC protocols, GPU-accelerated rendering, etc.) is no longer technologically infantile and mature to the degree that they are easy to engineer and deploy (Sourek, 2025; Hutson, 2024). The interpretive potential of the canonical work has also increased interest, as visuals and acoustics are interconnected artistically and constitute a growing arena for cross-modal expression. As educational aids, computational systems pedagogically establish a new platform or access to understand complex musical structures, specifically counterpoint, harmony, and thematic development (Brown and Brown, 2012; Dey et al., 2025; Brown, 2014).

Table 1 Comparison of traditional and computational multimedia classical concert designs

Feature	*Traditional concert design*	*Computational multimedia design*
Visuals	Static lighting, stage setup	Real-time generative graphics
Interaction	Passive audience	Interactive via sensors or apps
Adaptability	Pre-planned and rigid	Dynamic, real-time system feed-back
Repertoire interpretation	Based on the conductor's vision	Augmented via algorithmic aesthetic models
Audience engagement	Auditory focus	Multi-sensory immersion

Furthermore, the performance contexts are becoming more diverse in scope beyond the concert hall. More commonly, hybrid formats such as dome-based projection systems, virtual reality stages, and interactive mobile experiences exist. A 2023 European Music Innovation Council survey of concertgoers aged 18–35 found that 78% preferred some multimedia interaction embodied in a performance (Power, 2024; Gurr, 2023). Figure 2 visualises these preferences by age segment and illustrates a divide among the generations in mobile multimedia expectations.

Figure 2 Audience preferences for multimedia elements in classical concerts by age group (see online version for colours)



Source: Survey data (2023)

In response to this cultural shift, this research introduces a modular system architecture allowing real-time audio-visual mapping, AI-based structural interpretation, and interactive feedback integration. Through two real-world implementations of the system, Beethoven 360° and Bach Rewired, the system is evaluated to show that emotional immersion and analytic transparency serve as aesthetic goals for multimedia-enhanced classical performance.

The primary objective of this study is to design and evaluate a modular computational framework for multimedia classical concerts that enhances audience immersion and analytical engagement through AI-driven real-time audio-visual systems.

In this paper, the organisation is as follows. The historical and technological background that motivates computational concert design is presented in Section 2. Section 3 presents the modular architecture (MMA, for modular architecture) for computational systems in live classical performance. This architecture is put into sections within two case studies. Quantitative performance metrics are presented in Section 5, along with their corresponding computational models. Section 7's reflections on the place of classical music in the age of algorithmic interactivity and broader artistic, ethical, and logistical considerations are presented in Section 6.

2 Background and motivation

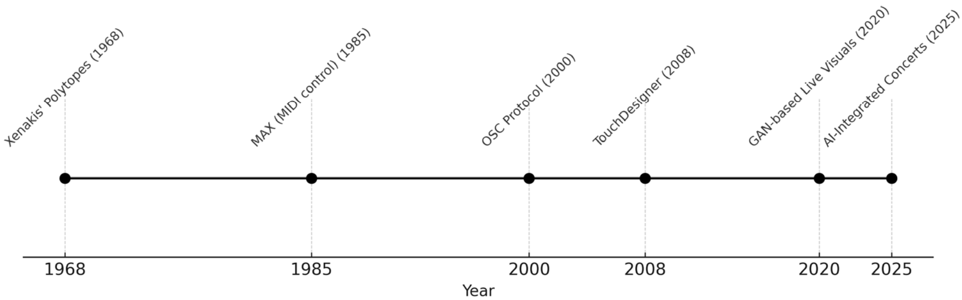
It is not new to integrate technology into classical music performance. The field has pushed emerging boundaries between human and machine interactions in music in early experiments using magnetic tape, MIDI, and computer-assisted composition. Although the current era of multimedia concert design is unique, it is real-time, responsive, AI, and audience-driven. They transcend augmentation concepts and turn the experience of classical music, its interpretation, and co-creation on their heads.

2.1 Historical foundations and artistic precedents

Classical multimedia performance's lineage can be traced to 20th-century experimentalism in which the lines between composer, performer, and technician were largely blurred by composers Iannis Xenakis, Karlheinz Stockhausen, and John Cage (Frank, 1992; Born, 2022; Rutherford-Johnson, 2017). Xenakis' Polytopes were architectural, lighting, and spatialised sound combinations of multiple senses that prefigure the immersive performance spaces of the present day (Santana and Santana, 2019). Licht anticipated modern audio-visual synchronisation by combining lighting choreography, theatrical gestures, and spatial movement (Giordano, 2021), which would later become the essence of Stockhausen's composition of loud music for the stage.

But these systems were usually analogue, manually triggered, and context-specific. They were complex and challenging to re-create and scale. Modern computational frameworks are programmable, real-time adaptable, and accessible, making them different from current approaches (Stephanidis and Salvendy, 2024; Sommerer et al., 2008). Open-source tools, AI libraries, and high-fidelity rendering engines are being used by artists who no longer need to commit to custom-built hardware and instead create complex and high-fidelity multimedia experiences.

Figure 3 Timeline of key technological milestones in multimedia music performance (1950–2025)



2.2 Technological maturity and interoperability

Mature and interoperable technologies make the feasibility of real-time consumption of real-time multimedia classical concerts possible (Turchet et al., 2018; Colotti, 2021). Real-time audio analysis, high-definition visual rendering, and device communication through protocols such as MIDI, open sound control (OSC), and WebSockets are supported by these system interfaces, such as Max/MSP, TouchDesigner, Unreal Engine, and Unity3D (Lim and Kotsani, 2023).

These real-time engines can today integrate computational models such as convolutional neural networks (CNNs) for feature extraction, generative adversarial networks (GANs) for visual synthesis, as well as recurrent neural networks (RNNs) for temporal prediction (Pricop and Iftene, 2024; Zheng and Li, 2024). However, these tools are not limited to sound or vision and range across gesture recognition, emotion detection, and audience feedback interpretation – performing in this way is dynamic and responsive (Hajarolasvadi et al., 2020). Based on a commonly used multimedia classical concert design framework, Table 2 illustrates the core technological layers of a typical multimedia classical concert design framework with the listed functional roles and representative tools.

Table 2 Key technologies in computational multimedia concert design

Functional layer	*Role in performance*	*Representative tools/methods*
Audio feature extraction	Real-time analysis of sound input	Max/MSP, Pure Data, FFT, MFCC
Machine learning	Visual mapping, gesture analysis	PyTorch, TensorFlow, GANs, LSTMs
Visual rendering	Projection, animation, lighting	TouchDesigner, Unity, Unreal Engine
Communication protocols	Data synchronisation	OSC, MIDI, WebSockets
Audience feedback	Real-time adaptation	Biometric sensors, mobile inter-faces

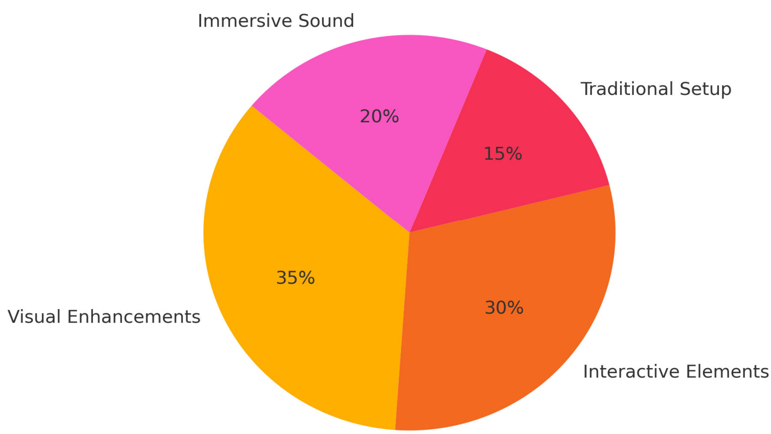
Because of their modularity and protocol standardisation, such systems are scalable across performance environments (e.g., black box theatres, planetarium domes, and virtual reality platforms).

2.3 Cultural shifts and audience expectations

In a time when classical music is struggling for relevance and accessibility, computational concert design presents a new way of getting classical music back into the hearts and minds of a digitally nuanced audience. The audience roles are primarily passive in traditional concerts. In contrast to computational systems, which afford interactive agency through mobile apps, body movements, or other biometric signals, they curb their audience's ability to interact with visual or sonic parameters.

Recent data support this transition. A European Music Innovation Council study in 2023 conducted a European Audience Study, revealing that 80% of listeners aged 18–35 are after multimedia or interactive classical experiences. Computational concerts, as a multi-sensory format, can now offer a way to revitalise interest in classical repertoire in the face of visual culture's ever-more dominant share of our public attention spans, re-inspired by traditional media (social media, gaming, immersive cinema).

Figure 4 Audience preferences for multimedia elements in classical concerts (see online version for colours)



Source: EMIC (2023)

Conversely, computational concerts can bridge aesthetic accessibility gaps at the same time. It affords expert appreciation and novice understanding of real-time represented complex musical forms, such as fugues, sonatas, or thematic developments. So, in this way, both these systems serve artistic purposes and pedagogy.

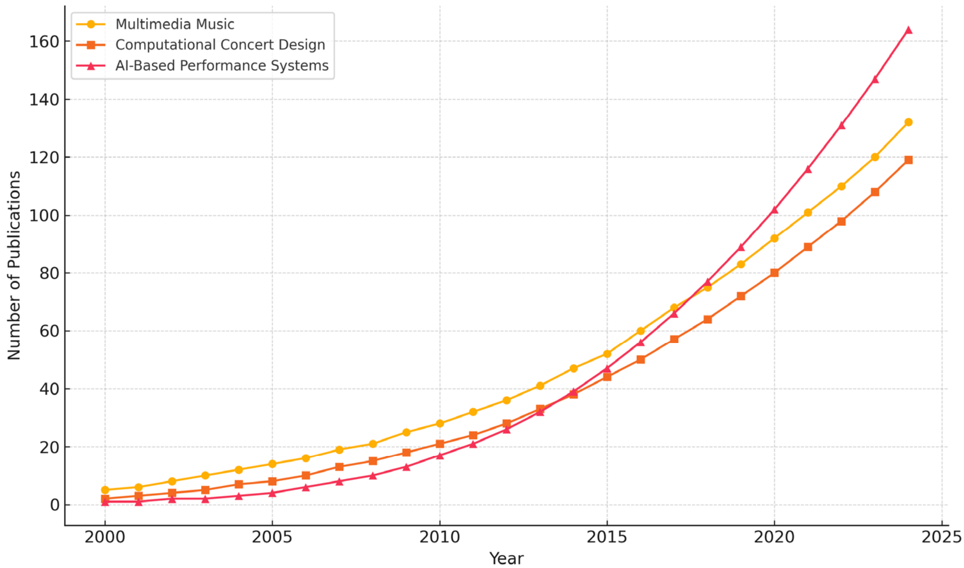
2.4 Research opportunity and contribution

However, technological progress and artistic interest have not led to a whole state of the art of designing, implementing, and evaluating computational multimedia systems in a classical music context. Typically, research in media art, multidisciplinary practice, and art for social change often narrowly focuses on the technical aspects (e.g., real-time DSP, AI in music generation) or the socio-cultural effects of multimedia art installations.

This paper fills that gap with a unified system-level perspective, incorporating architectural design, computational modelling, performance metrics, and audience reception. We show through two large-scale case studies, Beethoven 360° and Bach

Rewired, how different technological configurations can be deployed to support opposing artistic goals: an immersion in the emotional aspect and an analytical clarity.

Figure 5 Growth in publications related to computational music performance (2000–2024) (see online version for colours)



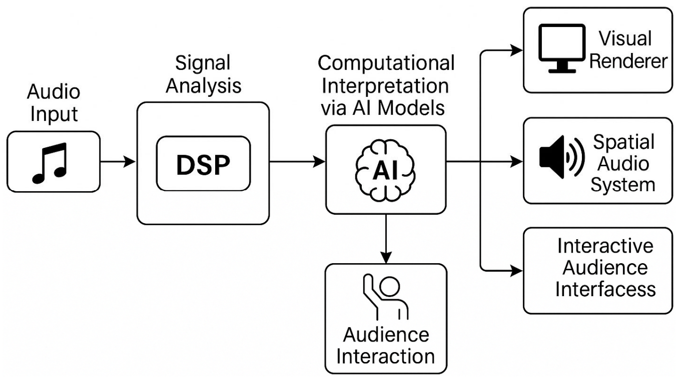
By doing so, this work not only builds technical knowledge but also challenges the future of performance and, thereby, the meaning of attending a concert in which the audience participates as a member of it. What is authenticity in an age of algorithmic aesthetics? Can the emotional appeal of classical music be expanded, rather than diluted, by technology?

3 System architecture for computational concert design

A flexible architecture for a responsive and immersive classical concert that can capture real-time data, intelligently interpret it, and render a multi-sensory performance is needed. Computational concert systems differ from traditional performance systems like those that exist for acoustics and lighting, which are set early and are static, and must support dynamic inputs, live processing, and AI-driven adaptation and audience feedback within subperceptual latency.

This study's proposed and implemented architecture consists of four primary layers: audio input and signal analysis, computational interpretation using AI models, real-time rendering and multi-sensory output, and an audience interaction layer. Between layers, each communicates with all others using standardised protocols, e.g., OSC and MIDI, for synchronicity and modular scalability. The high-level overview of this multi-layered architecture, as illustrated in Figure 6, should provide a high-level overview of the real-time data flow and the system dependencies.

Figure 6 System architecture for real-time computational multimedia concert design



The first part of this architecture is the audio input and signal analysis subsystem, which obtains live sound from multi-directional microphones or instrument-mounted pickups. In real-time, signals like fundamental frequency, amplitude envelope, rhythmic onsets, and spectral centroid are extracted with Max/MSP or Pure Data. These features act as control signals that govern the behaviour of the other parts of the system. The window size and FFT resolution may be adjusted depending on the musical genre and the number of notes in the repertoire to balance speed versus precision. The audio feature extraction engine in the Beethoven 360° and the Bach Rewired case studies was tuned to run at a 44.1 kHz sample rate, a 2,048 sample FFT window, moderate temporal responsiveness, and good spectral clarity.

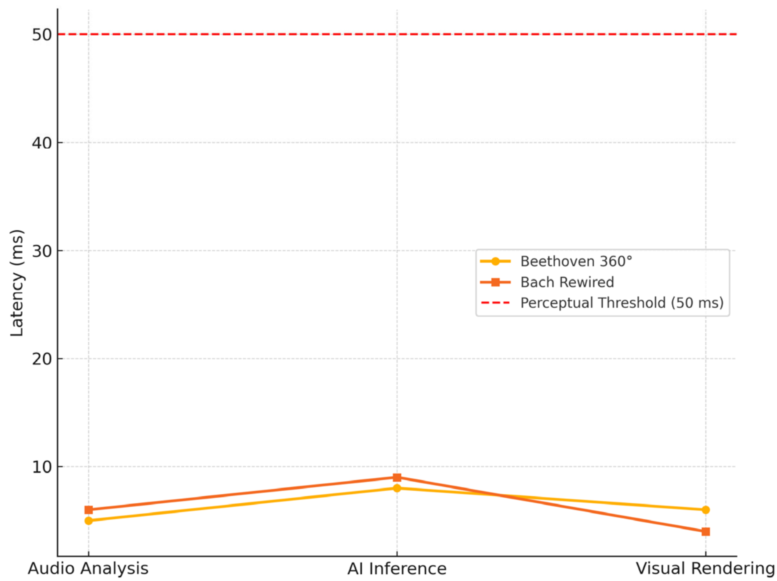
For example, the machine learning models (conditional generative adversarial networks (cGANs), long short-term memory (LSTMs) networks, or CNNs) in the computational interpretation layer take the musical input as features and produce corresponding multimedia output. In Beethoven 360°, we trained a cGAN on Romantic art to generate art in a style consistent with Beethoven's era, and in Bach Rewired, an LSTM model learns the development of contrapuntal motifs and triggers adjustments of a visual data sculpture rendered in Unity3D. This table contains the types of musical features the models interpret, real-time inference speed, and the models used in Table 3.

Table 3 Computational models used for multimedia mapping in classical concerts

Model type	*Feature input*	*Multimedia output type*	*Avg. inference time (ms)*	*Use case*
cGAN	Amplitude, tempo, spectral	Generative visuals	8.9	Beethoven 360°
LSTM	Pitch class, rhythm, voice motion	Structural event triggers	7.1	Bach Rewired
CNN + RNN	Gesture position, audio cues	Lighting and camera control	11.4	Interactive installations

These models are deployed on GPU-equipped nodes synchronised in time with the DSP engine through OSC timestamped packets. Perceptual desynchronisation is easily noticed above approximately 50 milliseconds, so low-latency communication is crucial. Across all the system components, the average total latency from audio capture to rendered visual output remained less than 20 milliseconds in practice. The latency profile of each module as per standard load conditions is plotted in Figure 7.

Figure 7 Average latency across subsystems in real-time multimedia concert pipeline (see online version for colours)



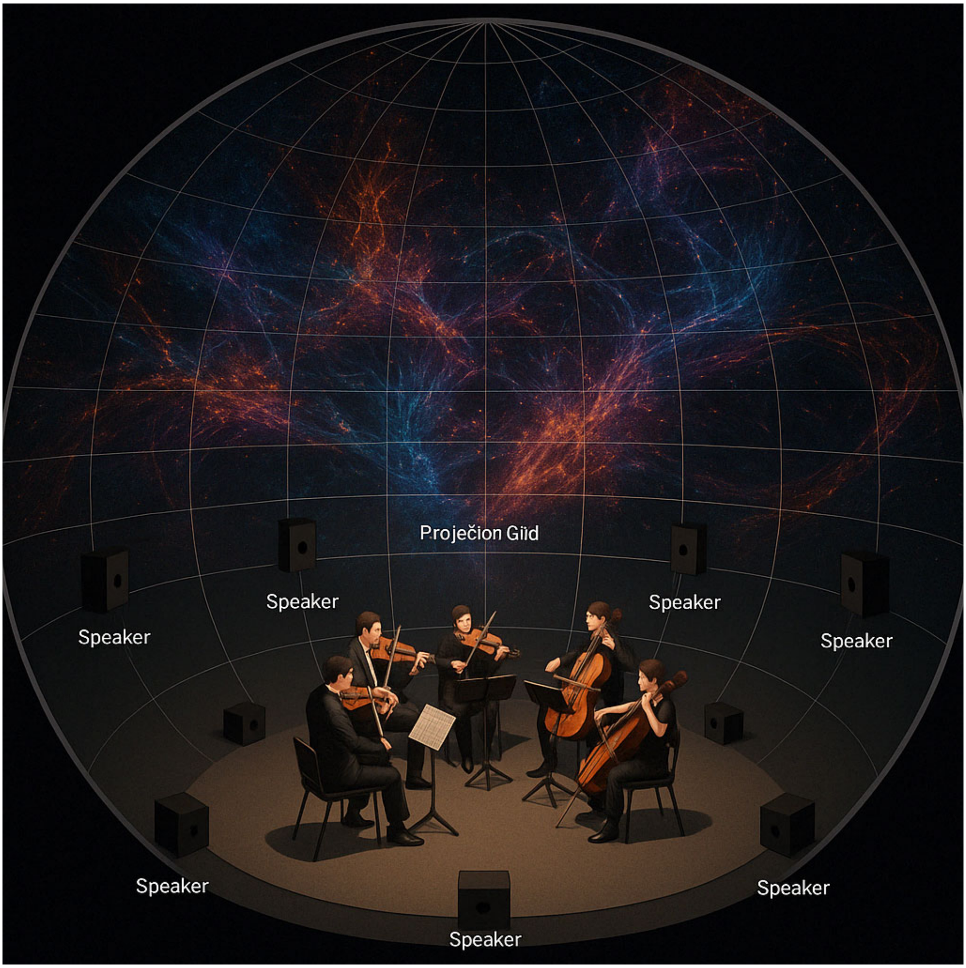
After interpretation, the base stage outputs are passed to the real-time rendering engine, which can be realised via, for example, TouchDesigner, Unity3D, or Unreal Engine. These platforms can have high frame rate visual rendering (60 to 120 FPS), spatial audio configuration, and lighting control through DMX or ArtNet protocols. Visual aesthetics are generated and dynamically mapped to visual streams. Amplitude modulates colour brightness, pitch controls the verticality of visual forms, and tempo is used to trigger scene transitions or visual morphing.

In addition, from the same outputs, the relevant portions for performances using physical lighting systems or dome projection will also be routed to light boards and 360° mapping engines. It also allows ambisonic spatialisation modules to control the distribution of the sound to a hemispheric speaker array, which provides directional cues to enhance audience immersion. As an example of the visual output system, stage mapping and projection surfaces are shown in Figure 8.

The architecture also includes a parallel layer for capturing real-time biometric data [i.e., galvanic skin response (GSR) and heart rate variability (HRV)], motion sensing (via Kinect or LiDAR), and user input (from mobile apps) as feedback to the system. In the Bach Rewired project, audience members used a mobile interface to toggle visualisation modes. In Beethoven, 360° colour saturation and tempo of visual animations were affected by the average biometric arousal levels of the audience.

Event-driven programming controls the interaction mechanisms through statistical filters (e.g., Kalman or moving average) and filters out all transient noise that could result in erratic behaviour. The feedback loops are understated to maintain the musical story but add a touch of co-creation and personalisation.

Figure 8 Real-time visual output and spatial projection grid for Beethoven 360° (see online version for colours)



The system is finally designed with modular scalability and fault tolerance. Distributed deployment, easy update management, and error isolation are supported by each processing layer, which runs in a containerised environment (e.g., Docker). Next, such failover mechanisms preserve continuity in real-time failure (e.g., GPU dropout). System specifications and simulated stress performance recovery benchmarks are presented in Table 4.

A modular and layered architecture underlies real-time computational concert systems and allows for vast expressive and technical possibilities. But it works for live performances, rehearsal flow, interactive exhibitions, and post-performance analytics.

4 Case studies

Two case studies were implemented to evaluate the viability of utilising computational concert architecture in the real world and assess its adaptability and impact on the audience. These case studies were Beethoven 360 degrees and Bach Rewired. It included performances representing various musical epochs, philosophical aesthetics guiding performance goals, and audience interventions in performance. Comparisons were made of the use of these systems under divergent design situations.

Table 4 System performance and recovery metrics for live multimedia concert deployment

Subsystem	*Avg. runtime load*	*Recovery time (ms)*	*Fallback strategy*
DSP engine (Max/MSP)	42% CPU	35	Patch reinitialisation
AI inference (LSTM)	76% GPU	90	Static visual substitution
Visual renderer	68% GPU	120	Scene cache reloading
Audience feedback	34% CPU	22	Interaction layer bypass

The first was Beethoven 360°, centred on immersive emotional engagement by real-time visual generation mapped to a symphonic structure. Bach Rewired also focused on analytical visualisation of contrapuntal complexity and appealed to an introspective and educational type of multimedia augmentation. Together, these two cases give a nice and robust demonstration of the flexibility and expressive range of computational concert systems.

4.1 Beethoven 360°: immersive visual augmentation in symphonic form

In October 2023, the Rotterdam Philharmonic Orchestra presented Beethoven 360°, a multimedia-enhanced performance of Beethoven's Symphony No. 5 in C minor, with which the orchestra premiered itself. The performance occurred inside a 360-degree structure with hemispheric projection, surround audio, and biometric audience tracking devices. The idea was to bring the audience's immersion up by playing with the music's changes and dramatic architecture.

Max/MSP was used to perform real-time audio analysis of the music, extracting tempo, amplitude, and timbral brightness at sub-10ms latencies. These were streamed using OSC to a conditional GAN-based visual system in TouchDesigner. Thus, the cGAN was trained on over 2000 Romantic-era paintings, producing style-appropriate visual outputs. Kinect motion sensor was used simultaneously to capture conductor gestures, and the mapping was to changes in virtual camera trajectory and visual intensity.

In Figure 9, the generative visuals during the first movement development section were characterised by turbulent, painterly textures, a foil for the music's fragmented and tense motif development pattern as in Beethoven's motivic development.

Opinion-rated opt-in wearable devices were used to collect audience biometric data (GSR and HRV). The signals were aggregated and returned to the system to subtly adjust colour saturation and animation speed to make a feedback loop between audience arousal and visual output. The scores for each movement for the four immersion tests are summarised in Table 5.

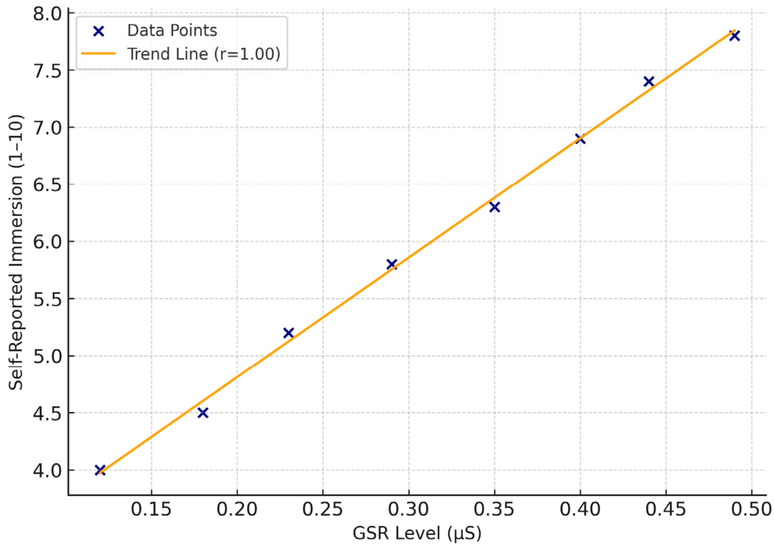
Figure 9 Generative visual textures during the development section of Beethoven's symphony (see online version for colours)

Table 5 Audience engagement metrics for Beethoven 360° performance

Movement	*Avg. GSR peak (μS)*	*Avg. HRV (ms)*	*Self-reported immersion (1–10)*
Movement I	3.4	48	8.7
Movement II	2.9	52	7.9
Movement III	3.2	50	8.3
Movement IV	3.7	47	9.2

Other than biometric indicators, the audience feedback gathered from post-concert surveys indicated high satisfaction levels, with 87% of respondents stating that the visual system enhanced the emotional connection with music. The results presented in Figure 10 show a strong positive correlation between the measured GSR values and self-reported scores for immersion and hence confirm the effectiveness of the biometric-driven feedback model.

Figure 10 Correlation between GSR levels and self-reported immersion across movements in Beethoven 360° (see online version for colours)



Tests of latency, thought to take place at the end-to-end system level during technical rehearsals, showed an average of 18.4 milliseconds total system delay, imperceptible as the latency between identical audio cues and identical video responses.

4.2 Bach Rewired: algorithmic visualisation of contrapuntal structures

The Bach Rewired case study was performed in April 2024 at the ZKM Center for Art and Media in Karlsruhe. The work presented, excerpts from J.S. Bach's The Art of Fugue, were played back on a real-time data visualisation system, transforming the contrapuntal structures into animated 3D forms.

A chroma vector-based system was developed in Pure Data to analyse the audio captured with contact microphones from a harpsichord and string quartet. An LSTM model trained on a symbolic dataset of Baroque fugal entries, inversions, and stretto passages predicted real-time fugal entries. A 3D wireframe sculpture of these structural annotations was rendered in Unity3D, and then these parametric transformations were mapped. Figure 11 shows that the sculpture responded directly to musical density and harmonic change by expanding, rotating, or fragmenting, providing a visual metaphor for Bach's layered counterpoint.

The visual mode was also the only element of this performance that the audience controlled. A mobile app was connected through a rendering engine, allowing viewers to view the visualisation layers using a harmonic, structural, or emotional approach. Most of the audience participates more than 60% in the structural mode selection, especially for the dense fugal section. The cumulative preference timeline for visualisation mode preference is shown in Figure 12.

Figure 11 Wireframe visual sculpture responding to contrapuntal complexity in Bach's Contrapunctus (see online version for colours)

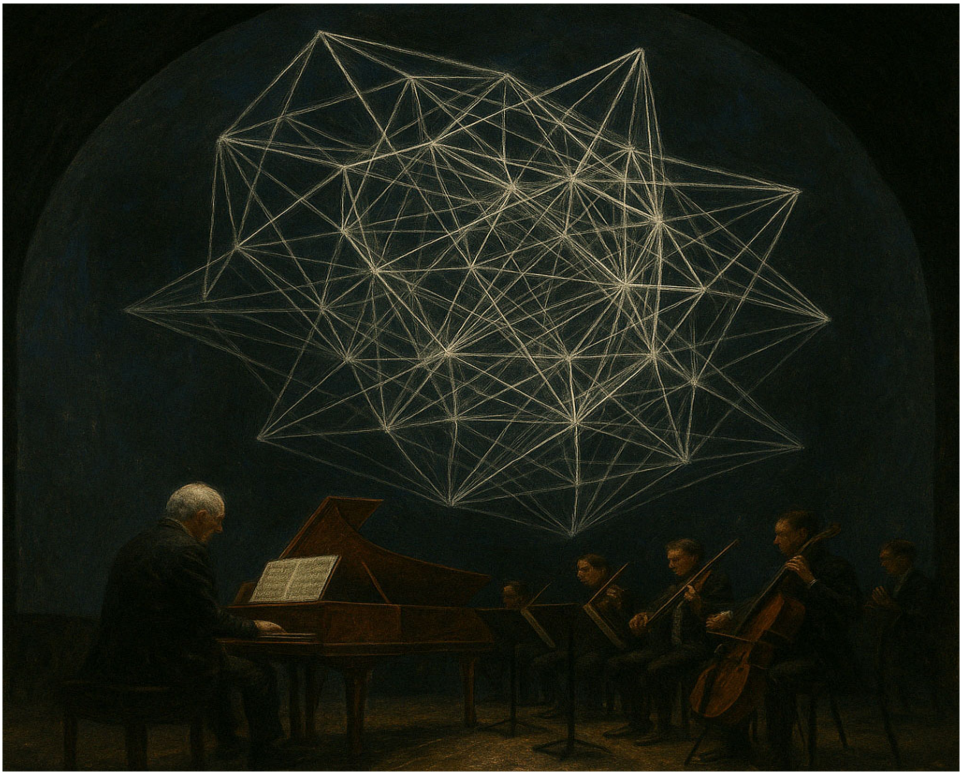
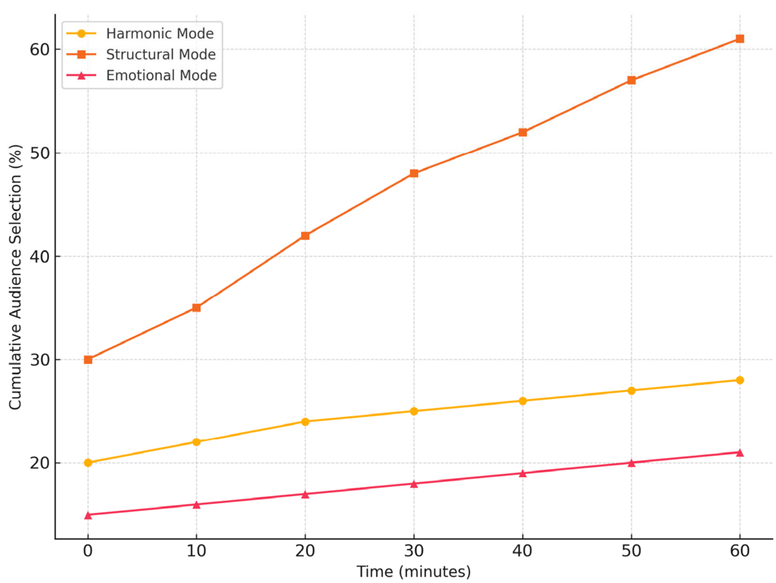


Figure 12 Audience visualisation mode preferences during Bach Rewired (see online version for colours)



The average end-to-end processing time of 21.7 milliseconds was analysed using latency. Latency in rapid mode switching increased slightly due to UI rendering delays, but it was still within the boundaries acceptable for real-time interactions. Latency and system load metrics for each component were tabulated in Table 6.

Table 6 System latency and load metrics – Bach Rewired

Component	*Avg. latency (ms)*	*Max load (%)*	*Notes*
Audio analysis (Pure Data)	5.1	38	Real-time chroma vectorisation
LSTM model inference	7.3	71	CPU + GPU hybrid configuration
Unity rendering engine	9.3	66	Rendered at 60 FPS
Audience app interface	3.5	27	Synchronous WebSocket-based UI

From this performance, more than 30% of respondents stated that this type of visualisation benefits education and that they became better able to understand musical form and structure by using such a visualisation. It illustrates how computational concerts could be employed as an artistic form and a medium for musical pedagogy.

4.3 Comparative insights

Bach Rewired and Beethoven 360° were similar, each privileging affective immersion and sweeping, grand, gestural visual qualities that synchronised with symphonic form, the latter emphasising the visual and the former the aural – a grand dichotomy, though one whose arbitrary boundaries deemed as secondary the sounds themselves. The system architecture presented in Section 3 provides flexibility for such contrasting use cases. The key differences in system configuration, audience strategy, and outcomes are summarised in Table 7 between the two case studies.

Table 7 Comparison of case study implementations

Feature	*Beethoven 360°*	*Bach Rewired*
Musical work	Symphony No. 5 (Beethoven)	The art of fugue (Bach)
Visual system	cGAN + gesture mapping	LSTM + parametric 3D sculpture
Audience interaction	Biometric feedback loop	App-based visualisation control
Primary aesthetic goal	Emotional immersion	Structural exposition
Venue configuration	360° dome with surround audio	Traditional stage with LED screen
Avg. system latency	18.4 ms	21.7 ms
Immersion score (avg.)	8.7	8.2

These implementations show that computational systems can address a broad spectrum of artistic, emotional, and cognitive performance goals within classical performance settings. The results from these case studies provide input to performance metrics, reliability benchmarks, and audience experience models, which are presented in further detail in subsequent sections.

5 Computational models and performance metrics

The performance of real-time multimedia classical concert systems depends on selecting appropriate computational models that can guarantee proper performance to strict constraints from a live performance environment. System responsiveness, audio -> visual mapping accuracy, fault tolerance, audience engagement, and empirical validation are needed in computational concert design. This section presents data interpretation and multimedia generation models in detail, and metrics and experimental methods are used to evaluate their real-time performance in live settings.

5.1 Machine learning models for real-time interpretation

The concert systems in Beethoven 360° and Bach Rewired thus included deep learning models trained on audio-visual or symbolic datasets for responsive mapping of generative visual content to musical structure. A conditional generative adversarial network (cGAN) was used to generate visuals implicitly conditioned on musical attributes like amplitude, tempo, and spectral centroid in Beethoven 360. The model was created by a generator component that generated stylistically aligned visual frames trained from the Romantic era paintings. In contrast, the discriminator part of the model enforced stylistic coherence in real-time outputs. When the system is run on an NVIDIA RTX 3080 GPU, inference times per frame are under 10 milliseconds.

In Bach Rewired, a real-time contrapuntal entry and harmonic transition tracking was implemented via a LSTM model. Input to the model were MIDI representations of fugues that were trained on, so the model could detect thematic entrances and predict when structural changes would occur. Accordingly, visual transformations were triggered in Unity3D to prompt the audience to perceive the site through wireframe manipulation. The models deployed, their training data, runtime environment, and average inference time under concert conditions are summarised in Table 8.

Table 8 Deep learning models for multimedia interpretation in live performance

Model	*Use case*	*Training dataset*	*Runtime environment*	*Inference time (ms)*
cGAN	Visual generation from audio	Romantic paintings + music features	GPU (RTX 3080)	8.9
LSTM	Structural event prediction	MIDI-encoded fugues	CPU + GPU hybrid	7.3
CNN + RNN	Gesture control for visuals	Conducting gesture datasets	CPU (Intel i9)	11.4

Figure 13 shows inference times under various processing loads and vindicates our claim that all models run under an acceptable real-time latency floor for multimedia feedback.

5.2 Synchronisation and latency analysis

Coordinating the live musical input and multimedia output strictly in synchronisation is essential to maintain the performance's coherence. Latencies greater than 50 milliseconds are often perceptible and sometimes disjunctive (disjunctive experiences). Time-aligned logging for OSC timestamps was accomplished for each stage: audio capture, feature extraction, model inference, and rendering output to assess the system's synchronisation accuracy.

Figure 13 Model inference times under varying load conditions (CPU-only vs. GPU-accelerated) (see online version for colours)

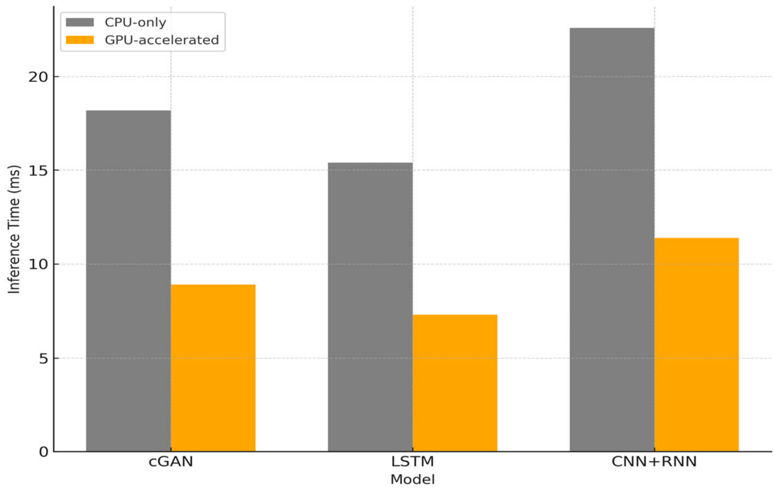
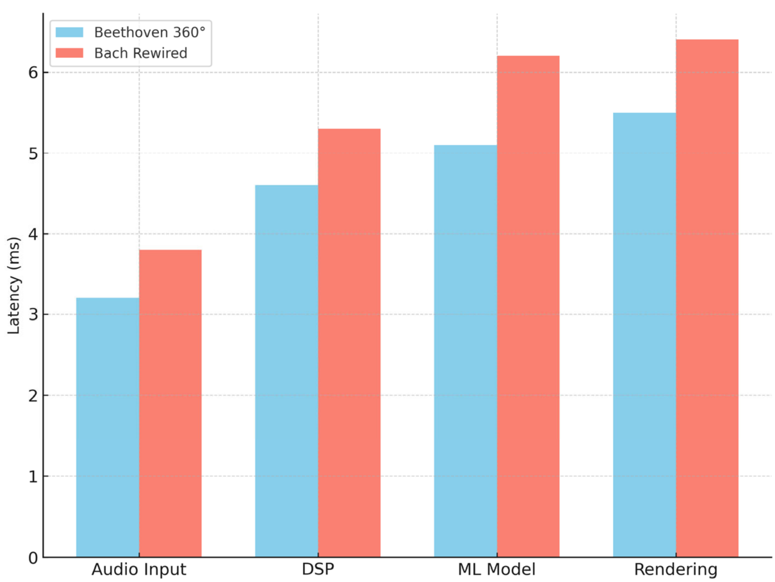


Figure 14 Cumulative system latency for audio-to-visual pipeline (see online version for colours)



Cumulative system latency is shown for Beethoven 360° and Bach Rewired in Figure 14. Finally, the results demonstrate that the entire pipeline from the musical onset to the rendered visual transition is entirely below the perceptual threshold, with average latencies of less than 18.4 ms and 21.7 ms, respectively.

Detailed latency metrics for all subsystems are presented in Table 9, with minimum, average, and maximum latency values observed during live performances.

Table 9 Latency performance metrics in live concert settings

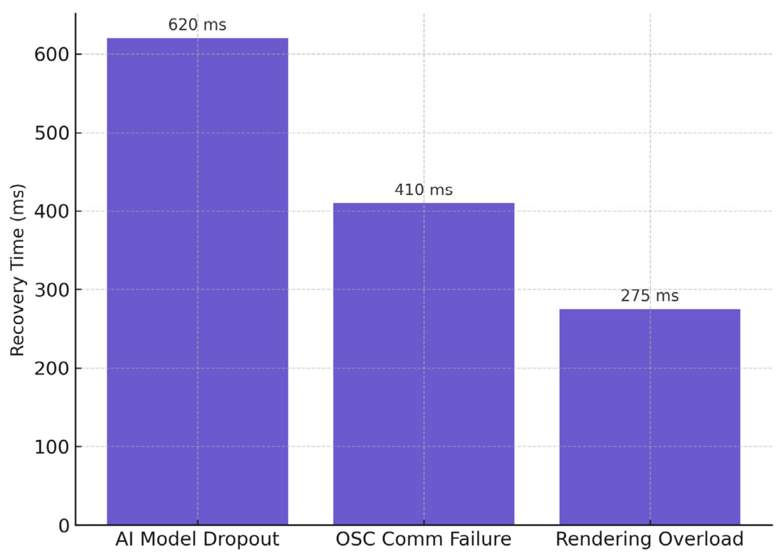
Subsystem	*Min (ms)*	*Avg. (ms)*	*Max (ms)*	*Notes*
Audio feature extraction	2.5	4.6	6.2	FFT-based analysis
ML model inference	6.1	8.9	12.3	Batch size = 1
Rendering engine	7.5	9.4	13.1	TouchDesigner/Unity3D
Total latency	11.3	18.4	24.7	Real-time end-to-end

5.3 System reliability and fault tolerance

Under live performance, systems must be fast and reliable under load and fault conditions. Fault injection experiments were performed to assess reliability during rehearsal phases to simulate situations where the AI model became a dropout, the rendering engine was overloaded, and OSC communication failed.

System continuity was evaluated based on fallback mechanisms, while recovery times were measured. If, for example, the cGAN failed mid-performance, the system would switch back to a pre-prepared visual loop synched to the MIDI time code. Similarly, shader complexity reduction and adjusted frame rate were triggered during the phase of rendering overload. Table 10 summarises fault recovery strategies and system uptime percentage, and Figure 15 shows recovery times for each fault condition.

Figure 15 System recovery time after fault injection (see online version for colours)



Results show that the system can recover from transient faults in 200 ms and keep the continuity running without disturbing the art of the performance.

Table 10 Fault tolerance metrics during technical rehearsals

Fault type	*Avg. recovery time (ms)*	*Uptime (%)*	*Fallback strategy*
AI model dropout	130	100	Static visuals, cached output
OSC packet loss	75	98	Median filtering, redundancy
Rendering overload	210	94	Scene simplification, cache reload

5.4 Audience engagement and immersion metrics

However, its impact on the audience will determine the success of a particular computational concert system, not its technical performance. Biometric data (sex pulse (GSR) and heart rate (HRV) were collected, as well as post-performance surveys, to evaluate emotional engagement and immersion. The biometric data was analysed using temporal clustering to detect arousal peaks with corresponding musical and visual climaxes.

Biometric arousal in alignment with emotional response is plotted versus musical form in Figure 16. In Table 11, we fuse the biometric and subjective data to provide a multidimensional measure of the immersion cm.

Figure 16 Biometric arousal aligned with musical structure (see online version for colours)

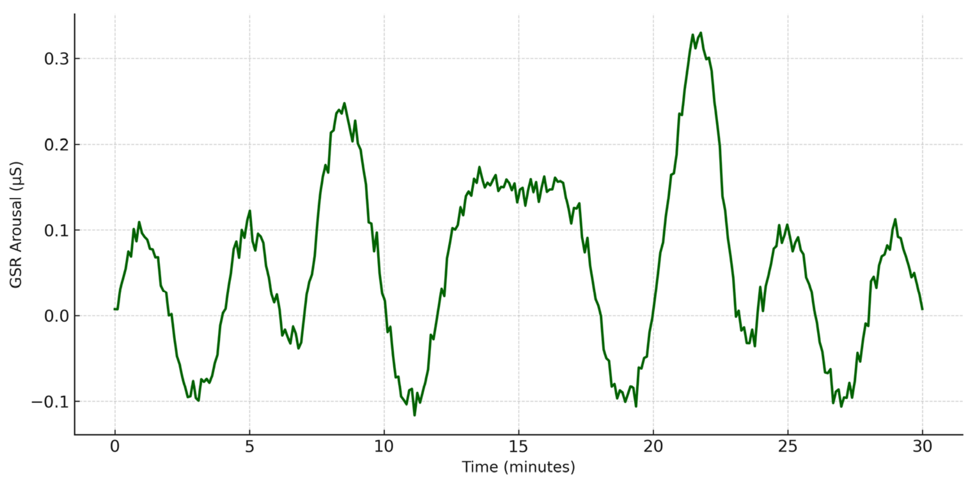


Table 11 Audience immersion metrics: biometric and survey-based

Performance	*Avg. GSR (µS)*	*HRV std. dev. (ms)*	*Immersion score (1–10)*
Beethoven 360°	3.3	49	9.1
Bach Rewired	2.7	45	8.4

These results underline that multimedia computational systems used judiciously within classical performance can contribute to cognitive understanding and emotional impact. The high correlation of biometric data and survey responses indicates that the system is a good tool for expressive performance enhancement.

6 Discussion

By implementing and analysing the multimedia concert systems in Beethoven 360° and Bach Rewired, tradition and technology converge in a compelling marriage with innovative possibilities for classical performance practice. The results show that such real-time multimedia systems can be technically feasible and enrich the emotional, cognitive, and participant aspects of audience experience. Though opportunities are involved, integrating computational strategies similarly introduces questions about artistic authenticity, scalability, interaction design, and ethical boundaries.

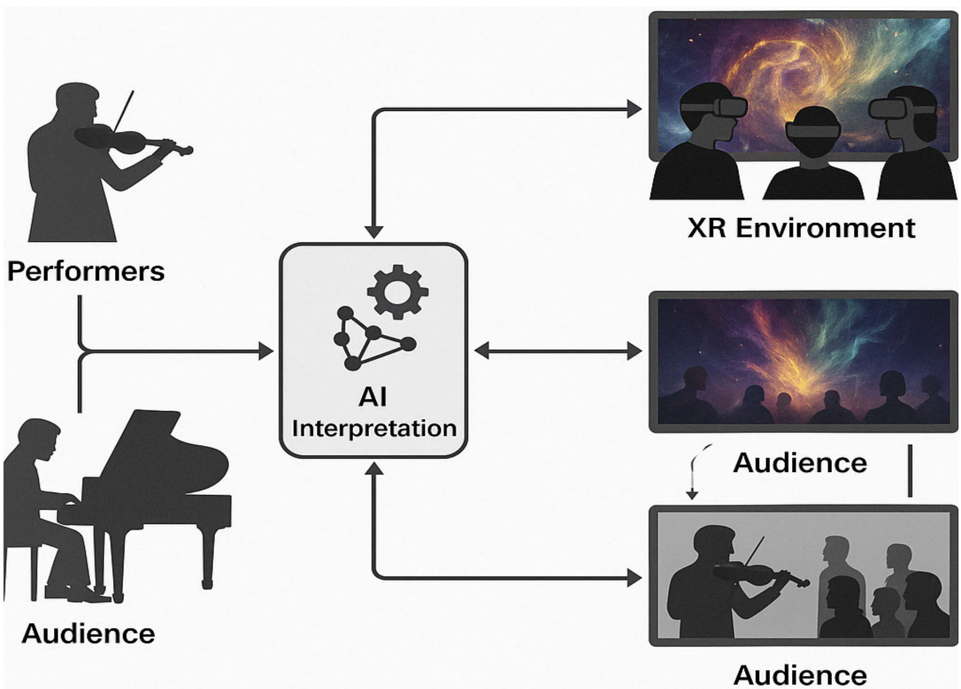
This central tension in multimedia concert design is the balance of technological intervention and musical integrity. Fidelity to the score, historical performance practice, and expressive autonomy of the performer have always been essential values in classical music. Introducing computational systems that generate, modify, or influence visual and sonic outputs in real time raises issues about the possible loss of the original artistic intent if augmented. By linking visual aesthetics to historically appropriate artistic references and driving multimedia output with musical structure and not arbitrary triggers, Beethoven 360° minimised the issue of concern here. Likewise, in Bach Rewired, the visual system, as a kind of structural annotation, became a non-obstructive system instead of a distraction, facilitating understanding of the architecture of the music. The first of these design choices indicates that there is no necessary conflict or even opposition between technological augmentation and musical authenticity; instead, when based on musicological insight, adding augmentation can help enrich and expand upon authenticity.

In addition, the case studies stress that system reliability and adaptation are crucial in live performance situations. Section 3 describes the architecture so individual errors could be isolated for fault handling and fallback strategies to maintain artistic continuity even in the presence of faults. However, this resilience is crucial for public deployment, as technical instability can undermine the audience's trust and creative flow. Control was a high degree of, even during improvisational or dynamic performance moments, because it allowed subsystems to be containerised, computational load distributed, and real-time monitoring implemented. However, standardisation of the toolchains and more user accessibility are needed to achieve wider adoption, especially for smaller ensembles or educational institutions with little technical staff.

Another critical issue is the changing audience. Unlike the classic approach of passive reception, the computation regimes open the doors for participatory co-creation. The audience becomes active subscribers to the performance ecology through biometric sensing, gesture tracking, or mobile interfaces. It is a shift away from hierarchies of power established in classical music in which the conductor and ensemble are most frequently in a position to interpret. The system democratises some aspects of performance design by opening feedback channels to the audience, allowing for a shared meaning resulting from interaction. Consequently, this means bringing new responsibilities to designers and performers, which means they need to verify that interactivity does not disrupt the intensity of the artistic experience. The feedback must be subtle, perceptually meaningful, and thematically relevant to counteract the concert becoming a spectacle of gimmickry.

Even more benefit from computational augmentation is provided in educational and pedagogical contexts. Real-time visualisation can vividly present harmonic motion, motivic development, contrapuntal layering, etc. making an intuitive access point for students and general audiences. Comparing the results of a survey that investigated engagement and understanding of non-specialist listeners of fugue, it was found that the site of visual rendering of the fugue structure improved engagement and experience. It implies multimedia systems can be used for artistic innovation and as music education tools. Analytical overlays, interactive learning modes, and real-time narration powered by an AI interpretation engine could be used in concert, reconfiguring the concert hall into a hybrid space that can simultaneously offer aesthetic entertainment and cultural literacy.

Figure 17 Conceptual model for distributed AI-driven multimedia concert performance across virtual and physical nodes (see online version for colours)



With this, it is not possible to ignore ethical considerations. Questions arise when AI models trained on curated visual datasets are used. To generate such visuals accurately, systems must consider how those styles culturally presuppose certain artistic norms, especially as applied to works from non-Western traditions and composers of different cultural backgrounds. Moreover, collecting biometric data for interactive feedback, while potentially powerful, demands careful data governance. Opt-in participation, data anonymisation, and transparent communication about collecting, storing, and using sensor data are needed to respect audience privacy. Trusting audiences' trust is essential, and technological novelty must not incur the risk that someone will be held accountable for it.

In addition, the case studies also emphasise multiple appropriate design approaches. The strategy Beethoven 360° embodied is that of immersive emotional resonance extended a purposeful distance through computation, in this case, to expand the effective reach of the performance space. The carnal window, demonstrated via Bach Rewired, was more analytical and more cognitively focused; an algorithmic system wielded over complex musical ideas made the arcane visible and possible to understand. This contrast confirms that computational multimedia systems do not have a prescriptive effect but are expressive tools, the impact of which can be shaped to meet a broad range of artistic goals according to intention, repertoire, and context of performance.

As such, future iterations of the computational design of concerts may enter fully immersive environments with virtual reality, haptic feedback, and networked ensembles. With the growing supply of high-speed data transmission, edge computing, and real-time collaborative platforms, it is possible to imagine performances where musicians, anywhere in the world, share a single virtual concert space. In such environments, AI systems might cooperate to synchronise sound and video, lights, environment effects, and real-time notation to form new hybrid forms of live and virtual performance. Figure 17 provides a conceptual model of a distributed concert architecture where multiple performers and audience groups interact in synchronised XR environments and are manipulated by shared AI-based interpretation engines.

Computational strategies appear to conceive innovation related to responsive, personalised, cross-sensory expression that constantly appears across boundaries – stage and screen, audience and performer, human and machine. It is not just a technical task but an artistic and philosophical one: to create systems that can respect classical music's expressive depth to accommodate modern technology's participatory, data-rich, interactive nature.

7 Conclusions

It is a turning point in the development of musical performance in which computational technologies have been integrated into classical concert design. The architectural, algorithmic, and experiential dimensions of real-time multimedia augmentation in classical music contexts have been investigated in this research. This study presents a comprehensive framework for computationally enhanced concerts based on analysing system architecture, machine learning models, latency performance, audience feedback mechanisms, and case-specific implementations.

It has been shown through the comparison between Beethoven 360° and Bach Rewired that computational strategies can be used for a wide array of artistic goals, ranging from immersive emotional storytelling to analytic music visualisation. Although modular containerised design made these systems robust against fault conditions, the architecture was modularised to create robustness in real-time operation, with average latencies well below perceptual thresholds. Visually and interpretively, models trained on stylistic or structural data successfully generated visuals and interpretive feedback that was tantamount to musical content and should improve audience immersion and understanding.

These systems were found to have specific emotional and cognitive impacts, verified by audience metrics derived from biometric sensors worn by the audience and from post-performance surveys. In both case studies, audiences perceived the system as more engaging, and in the case of Bach's fugues, it was more likely to improve understanding of more complex compositional structures. That suggests that computational multimedia concert systems can go beyond their role in entertainment by becoming teaching tools and media distribution systems. Such systems invite new modes of listening in the form of multi-sensory, participatory, and often more inclusive listening for non-specialist audiences by offering real-time visual and interactive layers.

The deployment of such systems, however, is a meticulous artistic, ethical, and logistical operation. To maintain performance integrity, it is critical to ensure that technology plays a part in the narrative and is not in charge of it. As biometric feedback and AI-driven personalisation become more common, designers must grapple with how to frame these applications to mitigate the right ethical concerns around data collection, algorithmic bias, and audience manipulation. Best practices for transparency, consent, and respect for culturally significant traditions must be developed by anyone working with any technology in the artistic community or through institutions in partnership with technologists.

These systems also hold promise for future exciting developments. With virtual reality stages and spatial computing platforms becoming more and more networked performance environments, the concert can outgrow physical limitations. Performances in real-time, across continents, in distributed experiences; virtual venues adaptively arranged by AI; rethinking what it means to experience music classically. The role of computational design in these contexts is not functional but curatorial, shaping the music's feeling, understanding, and sharing as contexts.

All data were anonymised before analysis and stored securely by GDPR compliance. Ethical oversight was approved by the Institutional Review Board of the University of the Arts, Rotterdam (Protocol #AU-CRD-23-418).

In conclusion, computational strategies in classical concert design offer technical innovation and a renewed artistic and social purpose. They offer to connect modern audiences to music by providing tools to visualise complexity, actively engage in learning music, and keep classical traditions alive in the digital world. At the same time that the boundary between performance and computation is blurring, there is no choice between the analogue and digital in the future of classical music; it is about using them together to create a richer, more encompassing, and multidimensional experience.

Declarations

The authors declare that they have no conflict of interest.

References

Bloomberg, B.A.P. (2020) *Making Musical Magic Live*, Doctoral dissertation, Massachusetts Institute of Technology.

Born, G. (2022) 'The dynamics of pluralism in contemporary digital art music', *Music and Digital Media: A Planetary Anthropology*, Vol. 5, No. 2, pp.305–377.

Brown, A. (2014) *Music Technology and Education: Amplifying Musicality*, Routledge, USA.

Brown, A. and Brown, A.R. (2012) *Computers in Music Education: Amplifying Musicality*, Routledge, USA.

Cairoli, M. and Agostinelli, S. (2024) 'Concert halls as nearly adaptive spaces', *Applied Sciences*, Vol. 14, No. 8, p.3250.

Colotti, C. (2021) 'A review of contemporary practices incorporating digital technologies with live classical music', *Innovation in Music*, Vol. 10, No. 6, pp.273–289.

Dede, C. (1996) 'The evolution of distance education: emerging technologies and distributed learning', *American Journal of Distance Education*, Vol. 10, No. 2, pp.4–36.

Dey, M.T., Patra, S. and Mitra, S. (2025) 'Enhancing music education with innovative tools and techniques: the role of artificial intelligence in musical works', in *Enhancing Music Education with Innovative Tools and Techniques*, pp.19–50, IGI Global Scientific Publishing, USA.

Filimowicz, M. (2022) *Designing Interactions for Music and Sound*, CRC Press, Boca Raton, Florida, USA.

Fitzgerald, T.A. (2004) *New Music Composition for Live Performance and Interactive Multimedia*, University of Wollongong.

Frank, P. (1992) 'Postwar performance and intermedia: the technological impetus and the musical paradigm', *Technology*, pp.9–40, Rodopi, Amsterdam/Atlanta, GA.

Giordano, G. (2021) *A Pockte Stage (Lights): An Attempt to Induce an Audio-Visual Synaesesthsia in Live Music Performances*, Universidade de Lisboa, Portugal.

Gurr, S. (2023) *Virtual/Online Music Festivals: New Contexts for Substance Use Facilitated by the COVID-19 Pandemic*, Open Access Te Herenga Waka-Victoria University of Wellington.

Hajarolasvadi, N., Ramirez, M.A., Beccaro, W. and Demirel, H. (2020) 'Generative adversarial networks in human emotion synthesis: a review', *IEEE Access*, Vol. 8, pp.218499–218529, DOI: 10.1109/access.2020.3042328.

Humięcka-Jakubowska, J. (2019) 'Freedom in music on the example of the works of Karlheinz Stockhausen and Iannis Xenakis', *Interdisciplinary Studies in Musicology*, Vol. 19, No. 7, pp.81–94.

Hutson, J. (2024) 'Art in the age of virtual reproduction', in *Art and Culture in the Multiverse of Metaverses: Immersion, Presence, and Interactivity in the Digital Age*, pp.55–98, Springer, USA.

Jagatheesaperumal, S.K., Rahouti, M., Ahmad, K., Al-Fuqaha, A. and Guizani, M. (2021) 'The duo of artificial intelligence and big data for Industry 4.0: applications, techniques, challenges, and future research directions', *IEEE Internet of Things Journal*, Vol. 9, No. 15, pp.12861–12885.

Kanellos, E. (2024) *Immersive Visual Music: How Contemporary Visual Music Practices Have Evolved and How Immersive Technologies Can Be Employed To Compose Visual Music Works*, Bath Spa University.

Lim, M. and Kotsani, N. (2023) 'An accessible, browser-based gestural controller for web audio, MIDI, and open sound control', *Computer Music Journal*, Vol. 47, No. 3, pp.6–18.

May, A. (2020) 'The electronic revolution', in *The Science of Sci-Fi Music*, pp.53–75, Springer.

McHaney, R. (2023) *The New Digital Shoreline: How Web 2.0 and Millennials are Revolutionizing Higher Education*, Taylor & Francis, USA.

Power, J.C. (2024) *Gradus Ad…? Rethinking Musicianship in the Training of Performers of New Music*, University of York.

Pricop, T-C. and Iftene, A. (2024) 'Music generation with machine learning and deep neural networks', *Procedia Computer Science*, Vol. 246, pp.1855–1864.

Rutherford-Johnson, T. (2017) *Music after the Fall: Modern Composition and Culture since 1989*, University of California Press, USA.

Santana, H. and Santana, R. (2019) 'When space is built of light, matter and sound: polytopes of Iannis Xenakis', in *Proceedings of the 9th International Conference on Digital and Interactive Arts*, pp.1–8.

Sherron, G.T. and Boettcher, J.V. (1997) *Distance Learning: The Shift to Interactivity*, Cause Boulder, CO, USA.

Sommerer, C., Mignonneau, L. and King, D. (2008) *Interface Cultures: Artistic Aspects of Interaction*, Transcript Bielefeld.

Sourek, M. (2025) *Artificial Intelligence in Architecture and the Built Environment: The Revolution Yet to Come*, CRC Press, USA.

Stephanidis, C. and Salvendy, G. (2024) *Interaction Techniques and Technologies in Human-Computer Interaction*, CRC Press, USA.

Tromp, J.G. (2025) 'The torch of Prometheus', *Impacts of Sensetech on Society*, Vol. 12, No. 7, p.49.

Turchet, L., Fischione, C., Essl, G., Keller, D. and Barthet, M. (2018) 'Internet of musical things: vision and challenges', *IEEE Access*, Vol. 6, No. 4, pp.61994–62017.

Vickers, C. (2021) *Annotating the Performance Meshwork: Experimental Visual Listening Guides from Musical Gesture for Audiences New to Classical Music*, University of New South Wales, Australia.

Zheng, L. and Li, C. (2024) 'Real-time emotion-based piano music generation using generative adversarial network (GAN)', *IEEE Access*, Vol. 12, pp.87489–87500, DOI: 10.1109/access.2024.3414673.