Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 22 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i22.59057

PAPER

Real-Time Feedback and Audience Engagement Optimization in Music Performance through Interactive Mobile Technologies

Zihan Fan (✉)

Zhumadian Preschool Education College, Zhumadian, China

10158@zyz.edu.cn

ABSTRACT

To address the challenge of real-time measurement and dynamic optimization of audience engagement in music performances under mobile Internet environments, the real-time feedback and audience engagement optimization were explored through mobile interaction technologies. A multimodal real-time emotional feedback analysis model and a mobile interaction behavior evolution tracking framework were established, and a dynamic quantification approach to audience engagement was proposed, enabling continuous perception and integrated assessment of audience emotional states and interactive behaviors during performances. An adaptive optimization mechanism based on real-time engagement data was further constructed, whereby audiovisual rendering, interactive tasks, and content orchestration strategies were dynamically adjusted, forming a closed-loop engagement enhancement scheme. Experimental results demonstrated that the proposed method significantly strengthened audience participation experiences and provided both theoretical underpinnings and technical pathways for the design and implementation of intelligent performance systems. The findings overcome the limitations of conventional offline evaluation modes and advance the paradigm shift of music performance from one-way dissemination to bidirectional interaction.

KEYWORDS

music performance, audience engagement, interactive mobile technologies, real-time emotion feedback, mobile interaction evolution, multimodal fusion, adaptive optimization

1 INTRODUCTION

With the deep integration of digital media technologies [1–3] and mobile intelligent terminals [4, 5], music performance is undergoing a critical transformation from one-way dissemination to bidirectional interaction and from static spectating to dynamic participation. Interactive mobile technologies-such as smartphones [6, 7], wearable devices [8, 9], and augmented reality (AR) interfaces

Fan, Z. (2025). Real-Time Feedback and Audience Engagement Optimization in Music Performance through Interactive Mobile Technologies. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(22), pp. 69–83. https://doi.org/10.3991/ijim.v19i22.59057

Article submitted 2025-08-05. Revision uploaded 2025-09-30. Final acceptance 2025-10-02.

[10, 11]-provide audiences with unprecedented channels of engagement. Through modalities, including touch, gesture, voice, and even biofeedback [12–15], audiences are able to intervene in real time within the performance process, reshaping both the artistic content and the experiential flow. In this context, the audience is no longer a passive recipient but has become an active and integral component of the performance system. Consequently, the scientific measurement and effective optimization of audience engagement states have emerged as central issues for enhancing both artistic expressivity and interactive technological experiences.

Although the significance of audience engagement has been widely acknowledged in performance studies and interactive design, existing methodological approaches still display substantial limitations. Most conventional approaches [16, 17] rely on post-performance questionnaires or offline behavioral analysis, which are unable to capture engagement states in real time, continuously, or with fine granularity. In terms of measurement dimensions, existing methods tend to emphasize behavioral indicators while neglecting emotional responses, which constitute a deeper level of audience engagement [18]. More critically, engagement measurement and optimization have typically been treated as separate processes. The absence of closed-loop regulation mechanisms based on real-time feedback [19, 20] has hindered the dynamic adjustment of interactive strategies and artistic content during performances, thereby limiting the realization of truly adaptive audience engagement enhancement.

In response to these challenges, the present study centers on real-time feedback and audience engagement optimization in music performance based on interactive mobile technologies. The research encompasses two closely interconnected components. First, a dynamic audience engagement measurement model integrating real-time emotional feedback and mobile interaction evolution was constructed. Through the synchronous acquisition and fusion analysis of multimodal data—including visual emotional responses, physiological signals, and streams of interactive behaviors—the continuous variation of audience engagement states across the temporal dimension of performance was characterized, and a comprehensive index for quantifying engagement intensity was established. Second, an engagement optimization framework driven by real-time measurement results was proposed. Within this framework, a closed-loop control mechanism dynamically regulates audiovisual rendering, interactive task generation, and narrative content orchestration during performances, thereby enabling a continuous optimization cycle spanning perception, decision-making, intervention, and evaluation. This study is intended to advance the transformation of music performance from technological enhancement toward intelligent adaptation, thereby providing both theoretical foundations and methodological support for the development of next-generation performance systems that are genuinely centered on audience experience.

2 MEASUREMENT AND OPTIMIZATION OF AUDIENCE ENGAGEMENT IN MUSIC PERFORMANCE

A real-time emotional feedback measurement method for music performance, based on both the tendency and the rate of change of emotional states, was established in this study. In parallel, an audience mobile interaction behavior evolution measurement method grounded in a “stimulation–decay” mechanism was developed. By integrating the dynamism of emotional feedback with the evolutionary patterns of mobile interactive behaviors, a comprehensive audience engagement evaluation model for music performance was constructed. This model relies on the real-time acquisition of multimodal emotional data and mobile interaction information to quantify the intensity and persistence of audience emotional responses, as well as the depth and evolutionary trajectory of interactive behaviors. Through this integration, a dynamic measurement and optimization mechanism was achieved, enabling the transition from passive spectating to active participation across the entire performance process.

2.1 Real-time emotional feedback

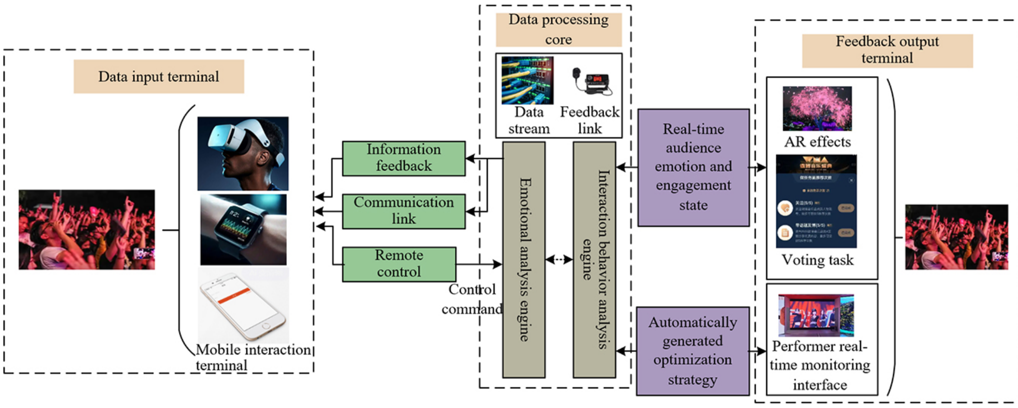


Fig. 1. Composition of the real-time emotional feedback system for audience engagement in music performance

Figure 1 illustrates the composition of the real-time emotional feedback system for audience engagement in music performance. In the quantification of real-time audience emotional feedback, the initial emotional states of the audience during the performance are first obtained. Based on the real-time acquisition of multimodal physiological and behavioral data, a dynamic time-window segmentation approach is applied to calculate the emotional feedback values of audience members across different temporal intervals. Specifically, let the time window of the emotional feedback analysis be denoted as S_s^t, and let L represent the total number of physiological and behavioral data changes within this window. The emotional feedback value of audience member i_u within time window S_s^t is computed as follows:

$$SE^{S_s^t}(i_u) = \frac{\sum_{k=1}^{L} SE(z_k^{S_s^t})}{L} \tag{1}$$

For emotional state determination, a threshold of 0.5—widely applied in SnowNLP—was adopted as the classification boundary. When the emotional feedback value exceeds 0.5, the state is regarded as positive; when it is lower than 0.5, the state is regarded as negative. Accordingly, the criterion for identifying whether an emotional feedback transition occurs is expressed as

$$t^{S_s^t} = \begin{cases} 1, SE^{S_{u-1}^t}(i_u) \le 0.5 \ AND \ SE^{S_s^t}(i_u) > 0.5 \\ -1, SE^{S_{u-1}^t}(i_u) > 0.5 \ AND \ SE^{S_s^t}(i_u) \le 0.5 \\ 0, SE^{S_{u-1}^t}(i_u) > 0.5 \ AND \ SE^{S_s^t}(i_u) > 0.5 \\ 0, SE^{S_{u-1}^t}(i_u) \le 0.5 \ AND \ SE^{S_s^t}(i_u) \le 0.5 \end{cases} \tag{2}$$

That is, when $SE_{s-1}^{St}(u_i) \le 0.5$, the emotional feedback tendency is negative, and when $SE_{s-1}^{St}(u_i) > 0.5$, the tendency is positive. Based on the changes in emotional tendencies across adjacent time windows during the performance, emotional feedback transition events can be identified. A transition from negative to positive is recorded as a positive feedback event ($t_s^{St} = 1$); a transition from positive to negative is recorded as a negative feedback event ($t_s^{St} = -1$); and when no transition occurs, it is recorded as $t_s^{St} = 0$.

Because audience emotions may fluctuate multiple times during a performance, each emotional transition constitutes a critical feedback point. These points divide the performance into several continuous emotional intervals, denoted as h_r. Based on the differences in emotional values and the time spans between adjacent intervals, the emotional transition rate can be computed to reflect the intensity and sensitivity of the audience's dynamic emotional adjustments. In this study, the transition rate was calculated according to the changes in audience emotional feedback values between adjacent intervals, expressed as

$$j = \left(\frac{\sum SE_{h_{r+1}}^{S_s^t}(i_u)}{m_{h_{r+1}}} - \frac{\sum SE_{h_r}^{S_s^t}(i_u)}{m_{h_r}} \right) / \frac{\sum SE_{h_r}^{S_s^t}(i_u)}{m_{h_r}} \tag{3}$$

where, the denominator represents the difference between the emotional feedback values of audience member i_u within unit time windows of interval h_{r+1} and interval h_r, while the numerator denotes the emotional feedback value of i_u within interval h_r. The parameter m is used to represent the length of the interval. Finally, by combining the emotional transition tendency and transition rate, an emotional feedback correction coefficient was introduced to dynamically adjust the initial emotional evaluation values. This approach allows a more accurate characterization of both the depth of audience emotional engagement and the real-time response patterns during music performances. Assuming that a negative transition from interval h_r to h_{r+1} is represented by $t_{r+1}^h = -1$, and a positive transition is represented by $t_{r+1}^h = 1$, with the number of intervals denoted by v, the correction coefficient is calculated as

$$\mu_{h_r} = \begin{cases} \dfrac{1}{1 + e^{-j\left|SE_{h_r}^{S_s^t}(i_u) - 0.5\right|}}, t^{h_{r+1}} = -1 \\ 1 + e^{-\frac{1}{j}\left|SE_{h_r}^{S_s^t}(i_u) - 0.5\right|}, t^{h_{r+1}} = 1 \end{cases} \tag{4}$$

The final emotional feedback value of the audience is then obtained by

$$SE(i_u) = \frac{\left(\dfrac{\sum SE_{h_v}^{S_s^t}(i_u)}{m_{h_v}} + \sum_{r=1}^{v-1} \dfrac{\sum \mu_{h_r} \times SE_{h_v}^{S_s^t}(i_u)}{m_{h_v}} \right)}{v} \tag{5}$$

This real-time emotional feedback mechanism not only reveals the evolutionary characteristics of audience emotions as performance content progresses but also provides a data foundation for subsequent integration with mobile interaction behaviors and realization of a holistic dynamic measurement and optimization of audience engagement.

2.2 Mobile interaction evolution

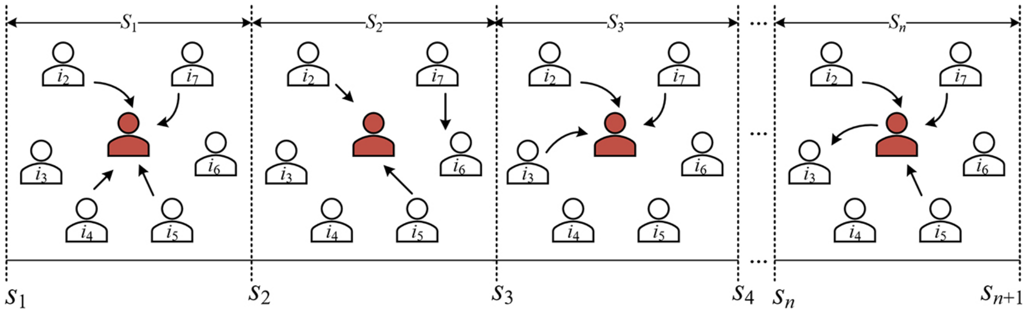


Fig. 2. Network of audience mobile interaction evolution

In music performance environments, audience members generate interactive behaviors through mobile devices such as smartphones and wearable sensors. These behaviors—including real-time comments, lighting control, rhythm interaction, and the use of virtual props—exhibit pronounced dynamism and diversity. They evolve continuously with time and with the progression of performance content, forming a dynamic participation network in which audiences are represented as nodes and mobile interactions as edges. A schematic representation of this network is provided in Figure 2. Formally, the network is expressed as $H = \{i_1, i_2, \ldots, i_v\}$, where, $I = \{i_1, i_2, \ldots, i_v\}$ denotes the set of audience members, $R = \{r_1^{TA}, r_2^{TA}, \ldots, r_v^{TA}\}$ represents the set of interaction behavior edges recorded within different time windows, and $S = \{S_1^{ACT}, S_2^{ACT}, \ldots, S_v^{ACT}\}$ denotes the set of time windows segmented according to the performance sequence. Conventional approaches often simplify continuous dynamic networks into discrete static snapshots, which neglect the spatiotemporal continuity of behavioral evolution. Consequently, an evolution measurement mechanism that better reflects the characteristics of audience behaviors is required.

To quantify mobile interaction intensity within a single time window, the interaction strength of audience member i_k toward performance content o_u in window S_l^{ACT} is defined below. The interaction strength is computed as a weighted composite index that integrates multiple mobile interaction events—such as touchscreen operations, device-shaking amplitude, and message transmission frequency—derived in real time from device sensor and log data. This indicator is employed to characterize the level of audience engagement within a given period. Specifically, let v denote the number of interacting audience members in the time window S_l^{ACT} for audience member i_u. The weights corresponding to bullet-screen messages, lighting control, rhythm interaction, and the use of virtual props are denoted by η_1, η_2, η_3, and η_4, respectively. The calculation is then expressed as

$$R_{i_k o_u}^{S_l^{ACT}} = \frac{\eta_1 m_{i_k o_u}^{S_l^{ACT}} + \eta_2 z_{i_k o_u}^{S_l^{ACT}} + \eta_3 e_{i_k o_u}^{S_l^{ACT}} + \eta_4 l_{i_k o_u}^{S_l^{ACT}}}{\sum_{k=1}^{v} \left(\eta_1 m_{i_k o_u}^{S_l^{ACT}} + \eta_2 z_{i_k o_u}^{S_l^{ACT}} + \eta_3 e_{i_k o_u}^{S_l^{ACT}} + \eta_4 l_{i_k o_u}^{S_l^{ACT}} \right)} \tag{6}$$

where, the numerator represents the weighted sum of the frequencies of interactive events—real-time comments, lighting control operations, rhythm interactions, and virtual prop usage—by audience member i_k toward content o_u within time window S_l^{ACT}. The denominator represents the weighted sum of the same interactive events generated by all audience members toward content o_u within the same time window.

To capture the continuous evolutionary characteristics of audience interaction behaviors, a mobile interaction update model incorporating a stimulation–decay mechanism was introduced. The interaction intensity is updated once for each time window. When the interaction intensity in the current window S_l^{ACT} is no lower than that of the preceding window S_{l-1}^{ACT}, the audience's willingness to participate is regarded as enhanced, and a stimulation factor is applied to increase the interaction weight. Conversely, when the interaction intensity decreases, a penalty factor is applied to reduce the weight. In addition, a temporal decay factor is introduced to simulate the natural decline of audience attention as the performance progresses, reflecting the cognitive principle that recent interactions exert a stronger influence on the current engagement state than historical behaviors. Specifically, the feedback situation of the preceding time window is represented by λ_u^e and λ_u^o. When positive feedback occurs, $\lambda_u^e = \lambda_u^e + 1$; when negative feedback occurs, $\lambda_u^o = \lambda_u^o + 1$. The temporal decay parameter is denoted as z. The computation formulas for the stimulation factor d^+, the penalty factor d^-, and the temporal decay factor σ are defined as follows:

$$d^{+} = R_{i_k o_u}^{S_l^{ACT}} \times \frac{LN\left(R_{i_k i_u}^{S_l^{ACT}} - R_{i_k o_u}^{S_{l-1}^{ACT}} + 1\right)}{LN2} \times \lambda_u^e \tag{7}$$

$$d^{-} = R_{i_k i_u}^{S_l^{ACT}} \times e^{\frac{\left(R_{i_k i_u}^{S_l^{ACT}} - R_{i_k i_u}^{S_{l-1}^{ACT}}\right)}{\lambda_u^o}} \tag{8}$$

$$\sigma = e^{-z\left(S_v^{ACT} - S_l^{ACT}\right)} \tag{9}$$

By integrating d^+ and d^-, the dynamically adjusted interaction intensity of an audience member within time window S_l^{ACT} is obtained as:

$$Q_{i_k i_u}^{S_l^{ACT}} = \begin{cases} \left(R_{i_k i_u}^{S_l^{ACT}} + d^{+}\right) \times \sigma, R_{i_k i_u}^{S_l^{ACT}} \geq R_{i_k i_u}^{S_{l-1}^{ACT}} \\ \left(R_{i_k i_u}^{S_l^{ACT}} - d^{-}\right) \times \sigma, R_{i_k i_u}^{S_l^{ACT}} < R_{i_k i_u}^{S_{l-1}^{ACT}} \end{cases} \tag{10}$$

Through this mechanism, audience mobile interaction behaviors can be dynamically quantified and evolutionarily modeled. This not only reveals the trajectories of individual and collective participation levels but also provides behavioral data essential for integration with emotional feedback in the comprehensive measurement of audience engagement.

2.3 Dynamic audience engagement evaluation

Within the interactive environment of music performance, audiences demonstrate diverse participation patterns and response levels due to differences in individual characteristics, behavioral habits, and aesthetic backgrounds. To effectively characterize this heterogeneity, a quantitative baseline of initial audience engagement must first be established. An initial engagement evaluation model was constructed based on observable behaviors and interactive features of audiences at the performance venue. The model integrates three dimensions: participation willingness, interaction ability, and real-time responsiveness. Participation willingness was quantified by the average start-up delay of the interactive application, defined as the time interval between the beginning of the performance (or the appearance of an interaction prompt) and the successful activation and connection of the interactive mobile application by the audience. This indicator reflects the subjective initiative and readiness of audiences to participate in interaction; shorter delays indicate stronger willingness. Interaction ability was measured by the average task completion time, defined as the mean time required by the audience to complete each valid operation in a standard sequence of interactive tasks provided at the beginning of the performance. This indicator captures familiarity with interactive mobile technologies, operational proficiency, and efficiency of human–machine coordination; shorter times indicate higher ability. Real-time responsiveness was evaluated by rhythm synchronization deviation, defined as the average absolute time difference between audience interaction behaviors (collected via mobile device sensors) and the main musical rhythm of the performance. This indicator objectively reflects the degree of attention devoted to the ongoing performance and the precision of behavioral alignment with the musical tempo; smaller deviations indicate timelier responses and higher attentional focus. Based on these three dimensions, the formula for calculating initial engagement is expressed as

$$\begin{aligned} U_{i_u}^0 &= \sigma_1 N_1(i_u) + \sigma_2 N_2(i_u) + \sigma_3 N_3(i_u) \\ &= \sigma_1 \frac{\tau_{MAX} - \tau(i_u)}{\tau_{MAX} - \tau_{MIN}} + \sigma_2 \frac{\delta_{MAX} - \delta(i_u)}{\delta_{MAX} - \delta_{MIN}} + \sigma_3 \frac{\phi_{MAX} - \phi(i_u)}{\phi_{MAX} - \phi_{MIN}} \end{aligned} \tag{11}$$

where, adjustment factors are represented by σ_1, σ_2, and σ_3. N_1 denotes the normalized function of application start-up delay, τ is the measured delay, and τ_{MAX} and τ_{MIN} represent the pre-set upper and lower bounds, respectively. N_2 denotes the normalized function of average task completion time, δ is the measured completion time, and δ_{MAX} and δ_{MIN} represent the pre-set upper and lower bounds. N_3 denotes the normalized function of rhythm synchronization deviation, ϕ is the measured average deviation, and ϕ_{MAX} and ϕ_{MIN} represent the pre-set upper and lower bounds.

Initial engagement reflects the static baseline of audience participation prior to the performance; however, during the actual event, engagement levels are dynamically influenced by real-time emotional feedback and the evolution of mobile interaction behaviors. To account for these influences, emotional feedback correction factors and mobile interaction adjustment coefficients were introduced to update engagement dynamically. The emotional feedback correction factor was quantified based on the tendency and rate of change in audience emotional states, capturing fluctuations in engagement caused by affective responses. The mobile interaction adjustment coefficient was derived from the stimulation–decay mechanism, through which the evolving intensity of audience interactions via mobile devices at different time intervals was modeled, thereby characterizing the continuous transformation of participation behaviors.

Specifically, assuming that $SE(i_u)$ represents the emotional feedback transformation value of audience member i_u within the time window S_l^{ACT}, the engagement of audience member i_u after j iterations in the same time window is calculated as:

$$\begin{cases} U^1_{i_k,S_l^{ACT}}(j) = U^0_{i_u,S_l^{ACT}} \times (SE(i_u))_{S_l^{ACT}} \times \\ \sum_{k=1}^{v} \left(\dfrac{1 + U^0_{i_u,S_l^{ACT}} \times Q^{S_l^{ACT}}_{i_k i_u}}{U^0_{i_u,S_l^{ACT}} + \sum_{l=1}^{n} Q^{S_l^{ACT}}_{i_k i_u}} \times U^0_{i_k,S_l^{ACT}} \right) \\ U^j_{i_u,S_l^{ACT}}(j) = U^{j-1}_{i_u,S_l^{ACT}}(j-1) \times (SE(i_u))_{S_l^{ACT}} \times \\ \sum_{k=1}^{v} \left(\dfrac{1 + U^{j-1}_{i_u,S_l^{ACT}}(j-1) \times Q^{S_l^{ACT}}_{i_k i_u}}{U^{j-1}_{i_u,S_l^{ACT}}(j-1) + \sum_{m=1}^{v} Q^{S_l^{ACT}}_{i_k i_u}} \times U^{j-1}_{i_u,S_l^{ACT}}(j-1) \right) \end{cases} \tag{12}$$

By integrating both emotional and behavioral dynamic signals, a comprehensive dynamic evaluation model of audience engagement for music performance was constructed. This model not only enables real-time reflection of variations in individual and collective engagement levels but also supports the adjustment of interactive strategies and the optimization of engagement experiences during performances, thereby providing a quantitative basis and regulatory interface for achieving highly engaging musical performances.

2.4 Dynamic optimization of audience engagement

Building upon the results of dynamic audience engagement measurement, a three-tier optimization framework was proposed to substantially enhance audience participation experiences in music performances. The implementation process is illustrated in Figure 3.

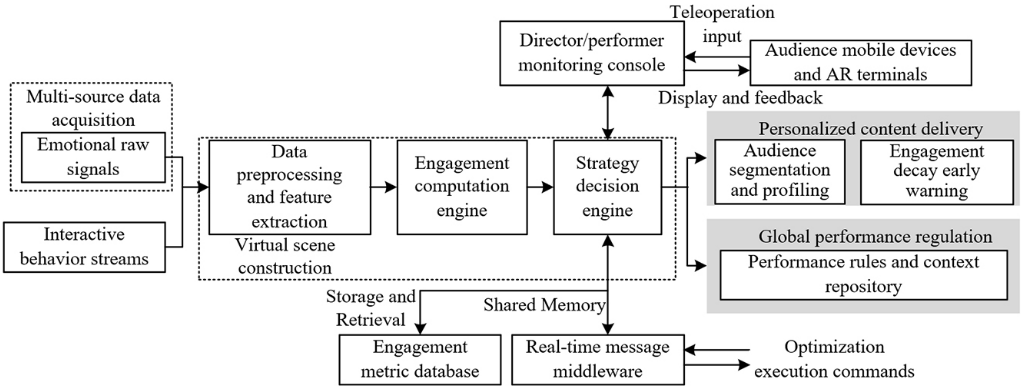


Fig. 3. Workflow of the dynamic optimization strategy for audience engagement

First, a dynamic regulation mechanism driven by real-time engagement data was established to adjust performance elements adaptively. Based on the trends of engagement indicators, the system automatically activates strategies from a predefined optimization repository. For example, when audience emotional engagement is detected to decline, audiovisual effects are dynamically adjusted: low-frequency rhythms are enhanced to stimulate physiological resonance, lighting colors and motion patterns are modified to strengthen visual attraction, and immersive virtual elements are projected through the stage background. At the interactive level, when mobile interaction intensity falls below expectations, personalized invitations are delivered to audience devices, such as initiating time-limited voting to determine the next piece, offering reward tasks with virtual props, or launching location-based collective light-and-shadow interaction games. These interventions effectively reshape behavioral participation pathways through technological means.

Second, a dual-channel optimization framework for human–machine collaboration was constructed. Automated regulation and the performance director team operate in closed-loop coordination: optimization recommendations generated by the system are displayed on a visualization dashboard at the control console, where directors or interaction designers select, modify, or reject them according to artistic intentions. Meanwhile, the system continuously learns from the effectiveness and contextual conditions of human interventions, refining the accuracy and artistic adaptability of automated strategies. This collaborative mechanism ensures both the responsiveness of real-time regulation and the preservation of artistic subjectivity, thereby preventing the mechanical impressions or aesthetic conflicts that may arise from fully automated systems.

Finally, a long-term optimization mechanism based on incremental learning and multi-session data iteration was formed. Comprehensive engagement data, regulatory operations, and feedback from each performance are structurally stored and utilized to train predictive and optimization models of engagement. By analyzing the mapping relationships between historical strategies and corresponding engagement responses, a progressively refined repository of personalized optimization strategies is established by the system, tailored to different audience groups, performance styles, and even specific musical pieces. Through this iterative process, optimization accuracy is improved across successive performances, while a long-term data-driven basis and an evaluation baseline are provided to support the continuous refinement of content design, interaction processes, and technical systems.

3 EXPERIMENTAL RESULTS AND ANALYSIS

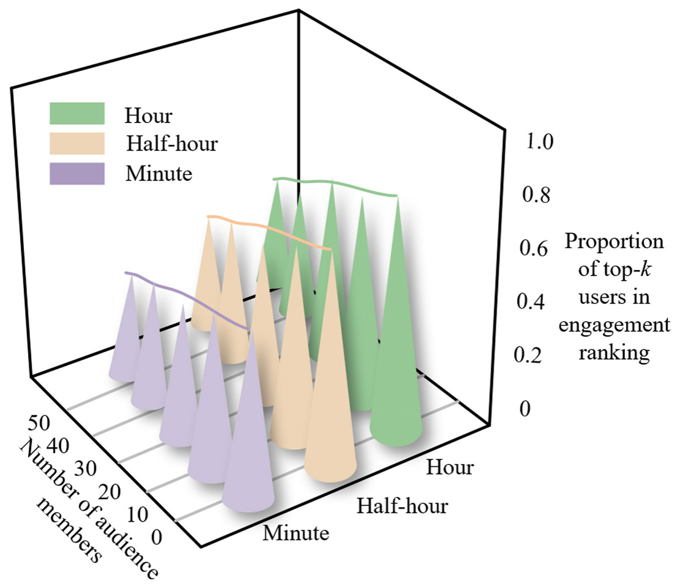


Fig. 4. Results of time-window segmentation

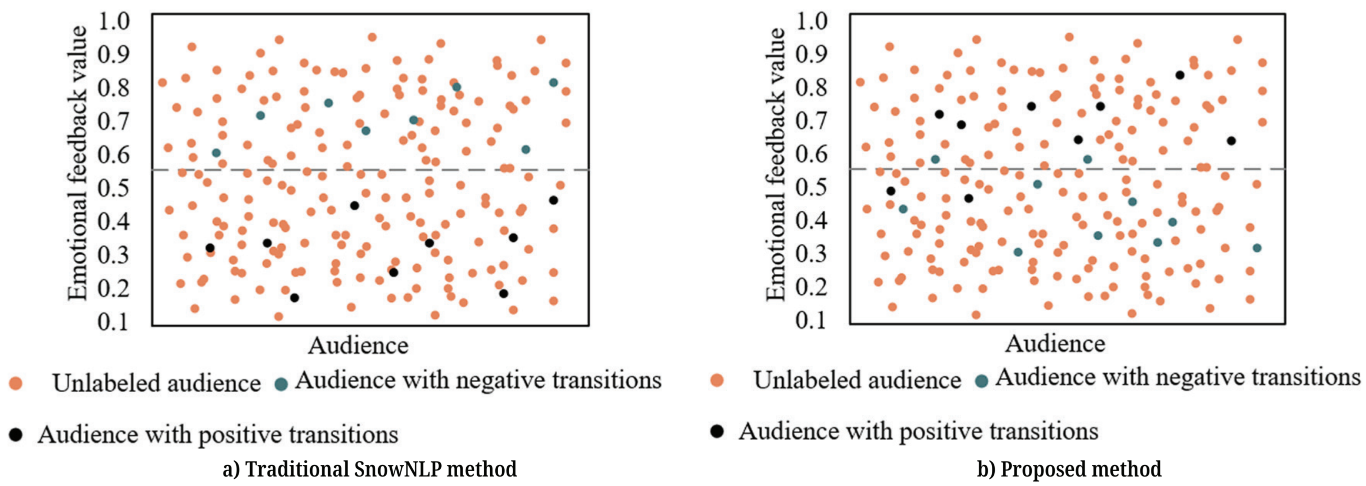
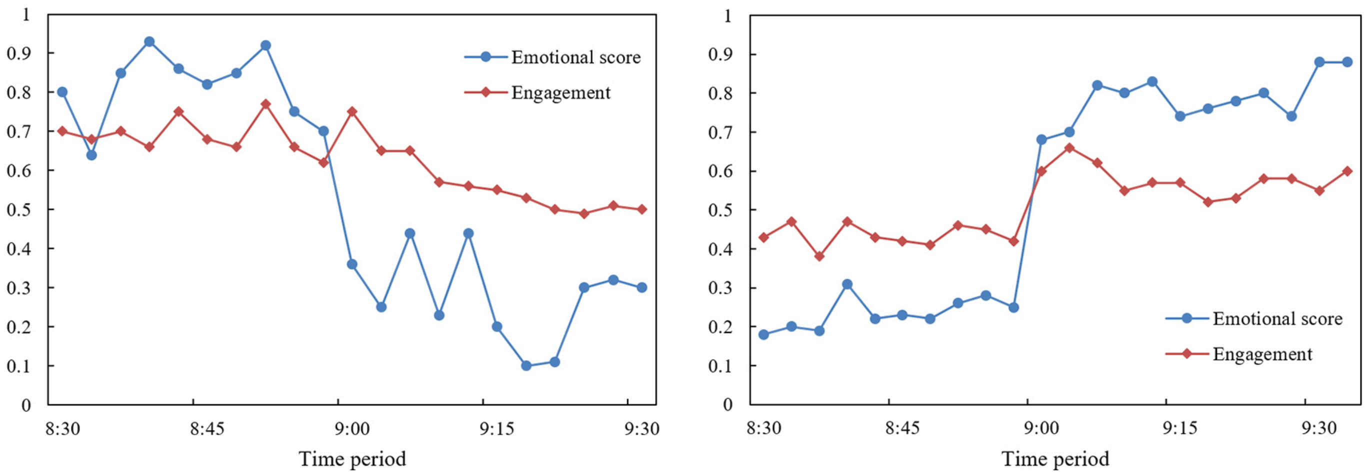


Fig. 5. Distribution of audience emotional feedback

To evaluate the influence of different time-window segmentation strategies on the accuracy and timeliness of dynamic audience engagement measurement, a set of comparative experiments was conducted. As illustrated in Figure 4, the distribution curves of the proportion of top-*k* users in engagement rankings were compared under hour-, half-hour-, and minute-level segmentation. The curve corresponding to minute-level segmentation reached a high plateau earliest and exhibited the steepest shape, indicating the highest sensitivity of resolution. In contrast, the hour-level segmentation produced the smoothest and most delayed response curve. These experimental results demonstrate that finer-grained time-window segmentation significantly enhances both the real-time responsiveness and the precision of engagement measurement. This finding provides critical validation for the proposed dynamic time-series–based engagement measurement model, ensuring that subsequent optimization strategies can be executed in a timely manner on the basis of the most up-to-date and accurate audience states.

To validate the superiority of the proposed real-time emotional feedback model compared with traditional text-based sentiment analysis approaches, particularly in capturing dynamic emotional transitions, a comparative experiment was conducted against the SnowNLP-based method. As shown in Figure 5, the emotional value distribution generated by the traditional SnowNLP method was highly concentrated around the neutral value of 0.5, making it difficult to effectively distinguish audience members with different emotional states and incapable of recognizing dynamic transitions. By contrast, the results of the proposed method exhibited a markedly dispersed distribution, in which three distinct audience groups were clearly identified: those with stable emotional states, those undergoing positive transitions, and those experiencing negative transitions. These findings provide strong evidence that the proposed real-time feedback measurement method—based on transition tendency and transition rate—enables a more fine-grained and sensitive capture of micro-level fluctuations and dynamic evolution in audience emotions during music performances. This capability ensures that reliable emotional data are supplied as a critical dimension for the subsequent dynamic measurement and optimization of audience engagement.



a) Synchronous-response scenario **b) Emotion-driven scenario**

Fig. 6. Relationship between audience emotional scores and engagement in different scenarios

To investigate the dynamic relationship between audience emotion and engagement under different performance contexts, two representative scenarios were analyzed. Figure 6a (synchronous-response scenario) corresponded to climactic segments of the performance, where intense emotional stimuli—such as powerful choruses or striking visual effects—triggered emotional scores and engagement curves to rise synchronously and sharply to their peaks. This pattern indicates that artistic content itself can directly drive immersive participation. By contrast, Figure 6b (emotion-driven scenario) corresponded to interactive guidance segments of the performance. In this case, emotional scores were observed to increase first, while engagement curves rose slightly later but then grew significantly, suggesting that positive emotional arousal can effectively prime and be converted into subsequent interactive behaviors. The experimental findings demonstrated that audience emotion functions as a key leading indicator and driving factor of engagement, while the mode of influence is modulated by specific performance contexts. In highly impactful segments, emotion and engagement were tightly synchronized, whereas during transitional interactive phases, increases in emotional scores served as precursors and predictors of subsequent engagement behaviors. This discovery strongly validated the necessity of incorporating real-time emotional feedback into the engagement measurement model and provided essential evidence for the design of context-sensitive optimization strategies.

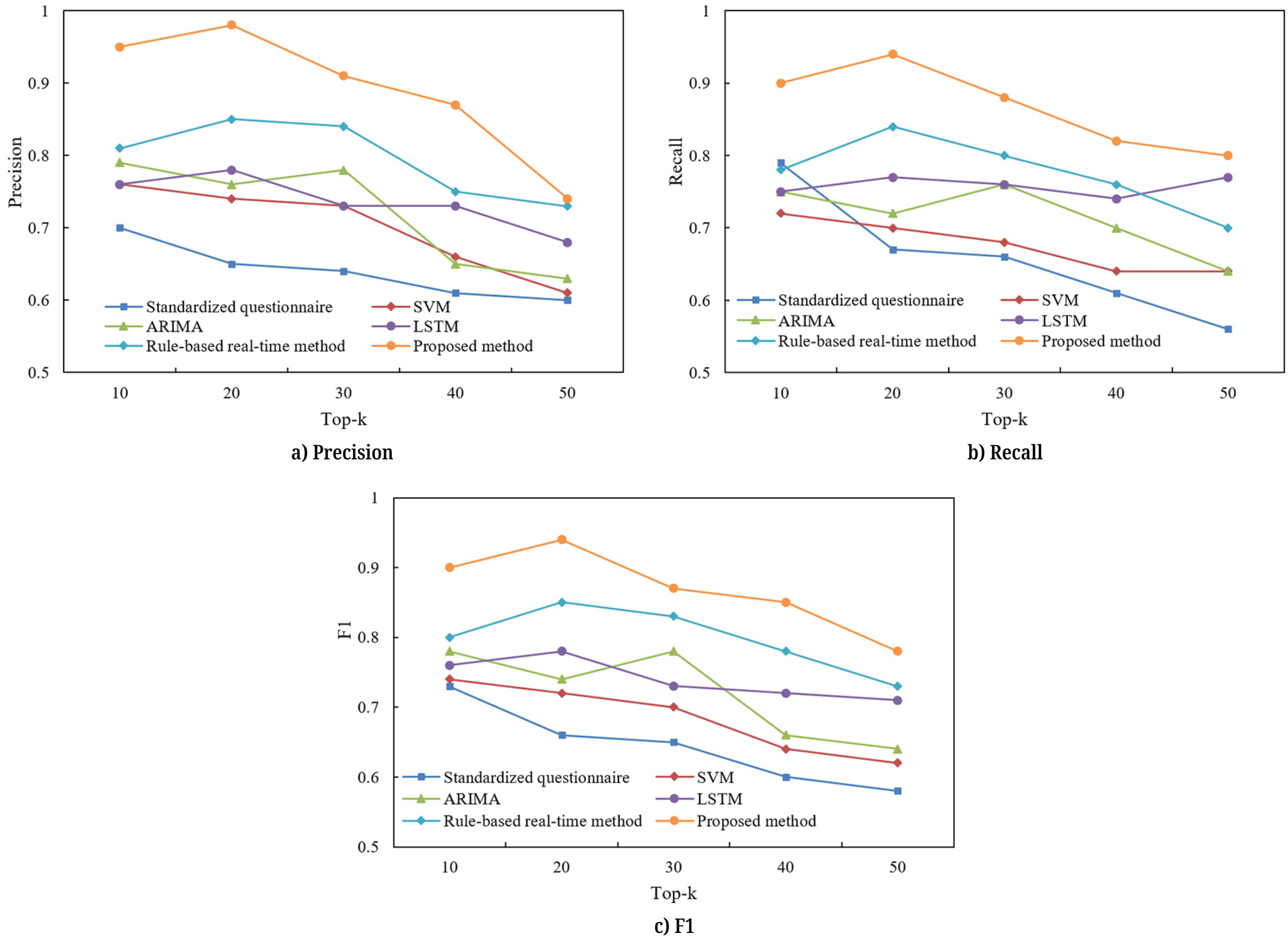


Fig. 7. Comparative experimental results

To systematically evaluate the overall performance of the engagement measurement model that integrates real-time emotional feedback with mobile interaction evolution, comparative experiments against multiple baseline approaches were conducted. As shown in Figure 7, in terms of the accuracy of identifying top-*k* highly engaged audience members, the traditional standardized questionnaire method exhibited only moderate effectiveness due to its subjective and delayed nature. The autoregressive integrated moving average (ARIMA) model failed to capture nonlinear interaction features. Although the rule-based real-time method responded quickly, its precision was limited. Support vector machine (SVM) and long short-term memory (LSTM) models demonstrated partial effectiveness but did not adequately account for the dynamic coupling of emotion and behavior. In contrast, the proposed multimodal dynamic fusion model significantly outperformed all baselines across evaluation metrics. Its performance curves rose rapidly to higher levels and consistently maintained stable superiority. These results demonstrated that the proposed approach achieved notable advantages in measurement accuracy, real-time responsiveness, and robustness. The ability to more precisely identify and predict audience engagement states provided a reliable and powerful data-sensing foundation for subsequent real-time optimization and intervention.

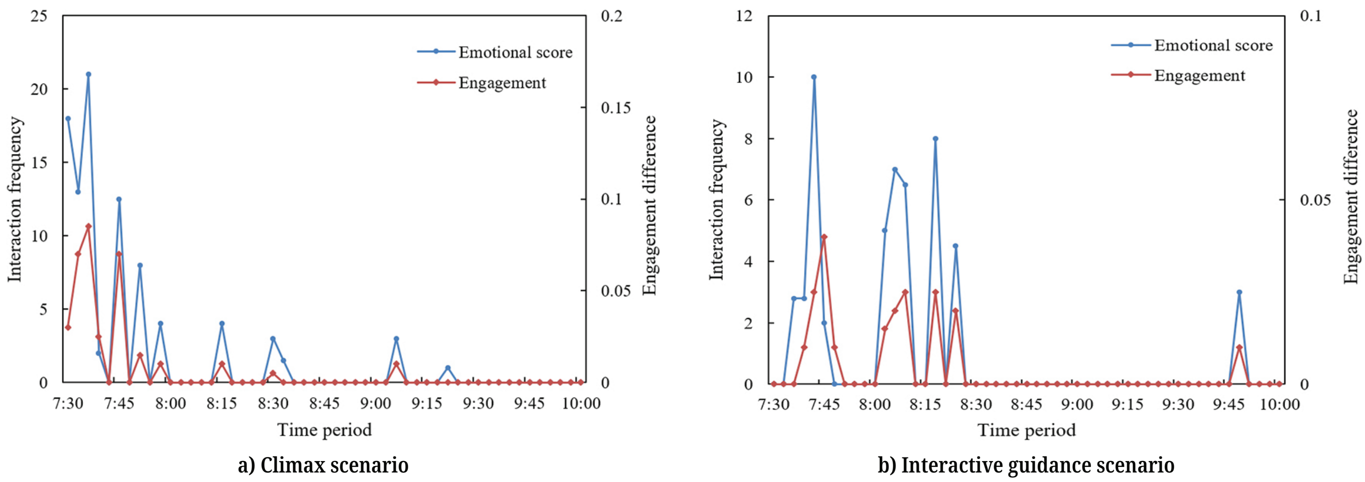


Fig. 8. Relationship between audience mobile interaction evolution and engagement across different scenarios

To further examine the dynamic relationship between audience mobile interaction behavior and engagement across different performance contexts, time-series curve characteristics were analyzed in two representative scenarios. In Figure 8a (climax scenario, 8:00–8:15 and 9:15–9:30), the interaction frequency curve exhibited sharp peaks, while both emotional scores and engagement difference values dropped to minimal levels. This indicated that during highly impactful segments, strong external stimuli simultaneously activated audience interaction behavior and emotional investment, producing a state of high synchrony and immersion. In contrast, Figure 8b (interactive guidance scenario, 8:45–9:00) showed a different pattern. The interaction frequency curve increased first and stabilized at a plateau, whereas emotional scores and engagement differences initially widened before narrowing. This suggested that after the issuance of interactive prompts, audiences responded behaviorally through mobile devices before their emotional states and deeper engagement were gradually activated and ultimately synchronized with behavior.

The experimental findings demonstrated that mobile interaction behavior serves as an explicit manifestation of audience participation, while its coupling with implicit emotional states is modulated by performance content. In the climax scenario, emotional response, behavioral interaction, and engagement increased synchronously, whereas in the interactive guidance scenario, behavioral responses led and emotional as well as deeper engagement followed. This evidence confirms the effectiveness of mobile interaction data as a critical indicator for real-time engagement measurement and underscores the necessity of tailoring optimization strategies to performance segment characteristics so that interaction content can foster the eventual alignment of behavior and emotion.

4 CONCLUSION

A dynamic approach for measuring and optimizing audience engagement in musical performances was proposed through the integration of real-time emotional feedback and mobile interaction evolution. A multimodal data acquisition framework was established to enable the synchronous perception and quantification of audience emotional states and interaction behaviors. On this basis, a dynamic emotional feedback adjustment model grounded in "transition tendency and rate" and an interaction evolution tracking model based on a "stimulation–decay" mechanism were designed, leading to the construction of a comprehensive engagement index. Experimental evidence demonstrated that this approach outperformed traditional questionnaires, static machine learning models, and classical time-series methods in terms of accuracy, timeliness, and interpretability. Audience engagement trends and critical transition points were effectively identified, thereby providing a reliable foundation for real-time interventions during performances. At the optimization level, a closed-loop regulation strategy driven by real-time engagement data was implemented, through which interactive content and audiovisual elements were dynamically adjusted, resulting in a marked enhancement of audience experience and immersion. This study thus advances beyond the traditional paradigm of offline, subjective engagement assessment by constructing an integrated "perception–measurement–optimization" framework, offering both theoretical foundations and practical pathways for the realization of intelligent performance systems.

Nonetheless, several limitations remain. First, data acquisition relies on sensors and mobile devices deployed on-site, which imposes requirements on performance environments and device compatibility, potentially restricting large-scale applications. Second, although multimodal fusion has been incorporated, deeper exploration of inter-modal coupling is still required, as nonlinear or cross-modal delayed responses may exist between emotional and behavioral signals. Third, although the effectiveness of optimization strategies has been preliminarily validated, their long-term impact and generalizability across different artistic domains and audience groups remain to be further examined. Future research may advance along several directions: (i) the development of less intrusive and more natural data acquisition methods; (ii) the adoption of advanced deep learning models to better capture complex interdependencies among multimodal data; (iii) the exploration of personalized optimization strategies through audience profiling for adaptive interaction content recommendation; and (iv) the extension of the proposed framework to broader performance art scenarios to validate its applicability and effectiveness across diverse cultural contexts and artistic forms.

5 REFERENCES

[1] J. Mackenbrock, A. Gawlik, F. Pels, and J. Kleinert, "Improving students' motivation for physical activity using digital media: A quasi-experimental study in physical education using smartphones and tablets," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 4, pp. 4–25, 2025. https://doi.org/10.3991/ijim.v19i04.51969

[2] N. Terentieva, V. Karpenko, N. Yarova, N. Shkvyria, and M. Pasko, "Technological innovation in digital brand management: Leveraging artificial intelligence and immersive experiences," *Journal of Research, Innovation and Technology*, vol. 4, no. 2, pp. 201–223, 2025. https://doi.org/10.57017/jorit.v4.2(8).06

[3] C. H. Tai and M. A. Alivi, "Reinventing trust: Traditional media credibility and audience engagement in the mobile-first era," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 15, pp. 81–96, 2025. https://doi.org/10.3991/ijim.v19i15.55685

[4] B. Yang, J. L. Zhai, and S. Rho, "Basketball image trajectory analysis based on intelligent acquisition of mobile terminal," *Mobile Networks and Applications*, vol. 27, no. 6, pp. 2534–2542, 2022. https://doi.org/10.1007/s11036-022-02071-w

[5] Y. Yue and Y. Yang, "Mobile intelligent terminal speaker identification for real-time monitoring system of sports training," *Evolutionary Intelligence*, vol. 16, no. 6, pp. 1801–1812, 2023. https://doi.org/10.1007/s12065-020-00452-2

[6] M. Vallejos-Flores, K. Talledo-Sánchez, D. Carlos-Ventura, A. Caycho-Caja, J. Sullcahuaman Amesquita, and D. Rime Huamanyauri, "Technostress and smartphone addiction mediated by smartphone distraction in college students," *Revista Digital de Investigación en Docencia Universitaria-RIDU*, vol. 18, no. 2, p. e1957, 2024. https://doi.org/10.19083/ridu.2024.1957

[7] H. Lee, J. W. Kim, and T. Y. Choi, "Risk factors for smartphone addiction in Korean adolescents: Smartphone use patterns," *Journal of Korean Medical Science*, vol. 32, no. 10, pp. 1674–1679, 2017. https://doi.org/10.3346/jkms.2017.32.10.1674

[8] K. Takei, W. Honda, S. Harada, T. Arie, and S. Akita, "Toward flexible and wearable human-interactive health-monitoring devices," *Advanced Healthcare Materials*, vol. 4, no. 4, pp. 487–500, 2015. https://doi.org/10.1002/adhm.201400546

[9] T. Desplenter and A. L. Trejos, "A control software framework for wearable mechatronic devices," *Journal of Intelligent & Robotic Systems*, vol. 99, no. 3, pp. 757–771, 2020. https://doi.org/10.1007/s10846-019-01144-5

[10] H. Herwin, L. D. Prasojo, B. Saptono, and S. C. Dahalan, "Analyzing the impact of augmented reality on student motivation: A time series study in elementary education," *Ingénierie des Systèmes d'Information*, vol. 28, no. 5, pp. 1197–1203, 2023. https://doi.org/10.18280/isi.280507

[11] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic, "Augmented reality technologies, systems and applications," *Multimedia Tools and Applications*, vol. 51, no. 1, pp. 341–377, 2011. https://doi.org/10.1007/s11042-010-0660-6

[12] J. Fernando and T. Oktavia, "Evaluation of the utilization of augmented reality technology to increase interaction satisfaction in remote assistance," *Journal Européen des Systèmes Automatisés*, vol. 58, no. 4, pp. 755–761, 2025. https://doi.org/10.18280/jesa.580409

[13] U. Rauschenbach, "Interactive TV: A new application for mobile computing," *Computers & Graphics*, vol. 30, no. 5, pp. 727–736, 2006. https://doi.org/10.1016/j.cag.2006.07.005

[14] B. J. Jo, S. K. Kim, and S. Kim, "Enhancing virtual and augmented reality interactions with a MediaPipe-based hand gesture recognition user interface," *Ingénierie des Systèmes d'Information*, vol. 28, no. 3, pp. 633–638, 2023. https://doi.org/10.18280/isi.280311

[15] J. Hwang, S. Yoon, and M. J. Lee, "The impact of interactive mobile applications on guests' intention to use in hotel: The moderating role of self-construal," *International Journal of Hospitality & Tourism Administration*, vol. 24, no. 5, pp. 790–817, 2023. https://doi.org/10.1080/15256480.2022.2055694

[16] M. Medina, I. Portilla, and T. Pereira, "Exploring what audience engagement means for media companies," *Revista de Comunicación*, vol. 22, no. 2, pp. 339–352, 2023. https://doi.org/10.26441/RC22.2-2023-3159

[17] D. Rubin, I. Mohr, and V. Kumar, "Beyond the box office: A conceptual framework for the drivers of audience engagement," *Journal of Business Research*, vol. 151, pp. 473–488, 2022. https://doi.org/10.1016/j.jbusres.2022.07.021

[18] H. Moe, T. Poell, and J. Van Dijck, "Rearticulating audience engagement: Social media and television," *Television & New Media*, vol. 17, no. 2, pp. 99–107, 2016. https://doi.org/10.1177/1527476415616194

[19] J. Kim, "The value of a shared experience: Relationships between co-experience and identification with other audiences and audience engagement behaviors on social media," *Computers in Human Behavior*, vol. 152, p. 108050, 2024. https://doi.org/10.1016/j.chb.2023.108050

[20] L. Camaj, E. Çela, and G. Rexha, "The audience logic in election news reporting on Facebook: What drives audience engagement in transitional democracies of Albania and Kosovo?" *Journal of Information Technology & Politics*, vol. 21, no. 1, pp. 38–53, 2024. https://doi.org/10.1080/19331681.2023.2227847

6 AUTHOR

Zihan Fan graduated from Henan University in 2013. Work at Zhumadian Preschool Education College, Zhumadian 463000, China And is engaged in the research of music education, history of Chinese and western music (E-mail: 10158@zyz.edu.cn).