Sensors and Materials, Vol. 38, No. 5 (2026) 2891–2908
MYU Tokyo

S & M 4476

Design and Performance Evaluation of a Multisource-sensor-fusion-based Immersive Interactive Projection System

Dan Zhong,[1] Juan Lu,[2] and Cheng-Fu Yang[3,4]*

[1]Department of Network and New Media, Dongguan City University, Dongguan 523000, China
[2]School of Creative Design, Dongguan City University, Dongguan 523000, China
[3]Department of Chemical and Materials Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan
[4]Department of Aeronautical Engineering, Chaoyang University of Technology, Taichung 413, Taiwan

(Received May 11, 2026; accepted May 22, 2026)

Keywords: multisource sensor fusion, immersive interactive projection system, projection system, depth sensing, human–computer interaction, system performance evaluation

In this study, we propose and experimentally validate a multisource-sensor-fusion (MSF)-based immersive interactive projection system (IIPS) integrating a depth camera, a distance-sensing module, and an audio-input module. The system is designed to enhance interaction recognition accuracy, response consistency, and robustness in dynamic environments, where single-sensor approaches are often affected by occlusion, environmental noise, and event misclassification. A sensor fusion framework is employed to integrate heterogeneous sensing signals at the decision level for real-time interaction detection and event triggering. By leveraging complementary sensing modalities, the system has an improved environmental perception capability and interaction reliability, enabling stable and adaptive user–system interaction in complex exhibition scenarios. Two complementary evaluations were conducted. First, system-level performance was quantitatively assessed using interaction recognition accuracy, response latency, trigger success rate, and operational stability. The proposed system achieved 93.4% recognition accuracy, 186 ms average response latency, 91.8% trigger success, and only two abnormal events over 30 continuous test rounds, demonstrating high robustness and real-time performance. Second, user evaluation data from 384 participants were analyzed using exploratory factor analysis, confirmatory factor analysis, and structural equation modeling. Unlike conventional IIPS studies that primarily focus on visual rendering or isolated interaction performance, in this study, we emphasize the integration of multisource sensing reliability, system interaction performance, and user-experience evaluation within a unified analytical framework. The primary objective of this study is to investigate how MSF affects both technical interaction performance and experiential outcomes in immersive environments. The main contribution lies in establishing a sensor-oriented evaluation framework that directly links sensing quality with immersion, satisfaction, and revisit intention through empirical structural modeling. However, the current framework is limited to a specific experimental environment

---

*Corresponding author: e-mail: cfyang@nuk.edu.tw

https://doi.org/10.18494/SAM6417

ISSN 0914-4935
https://myukk.org/

and a fixed combination of sensing modules, which may affect the generalizability of the results. Overall, the results demonstrate that MSF is a critical design strategy for improving sensing reliability, interaction performance, and experiential quality, providing practical insights for the development of sensor-integrated IIPS.

1. Introduction

Immersive interactive projection systems (IIPSs) have become increasingly important media in digital media art, exhibition design, stage technology, and cultural–technology integration. IIPSs are interactive digital environments that combine projection technologies, environmental sensing, real-time computation, and multimedia rendering to create responsive spatial experiences. Unlike conventional projection systems that only display static or predesigned visual content, IIPSs continuously detect environmental changes and user behaviors through sensing devices and dynamically modify projected images, sound effects, or interaction feedback accordingly. In practical applications, these systems are widely utilized in digital museums, smart exhibitions, interactive installations, educational spaces, virtual art environments, and intelligent entertainment systems. The core concept is to establish a closed-loop interaction framework in which environmental sensing, signal transmission, data processing, and projection feedback are continuously synchronized in real time.(1–3)

In a typical IIPS environment, multiple sensing devices are distributed around the interaction space to detect user movement, position, gesture behavior, proximity variation, and sound-trigger events. In the proposed framework, the depth camera was installed above the interaction zone to capture spatial body movement and depth information, the distance-sensing module was positioned near the projection boundary to detect user proximity and entrance behavior, and the audio-input module was deployed within the exhibition environment to receive sound-trigger interactions and ambient acoustic signals. The sensing signals were transmitted to the central processing computer through wired USB and digital communication interfaces for real-time synchronization and event analysis. After signal acquisition, the data-processing module performed feature extraction, event classification, and interaction decision-making, and the processed results were subsequently delivered to the projection and audio-rendering system to generate responsive multimedia feedback. From the perspective of *Sensors and Materials* readers, IIPSs can be regarded as integrated sensor-based intelligent environments rather than merely visual presentation systems. The effectiveness of such systems strongly depends on sensing accuracy, signal stability, environmental adaptability, and real-time interaction reliability. Therefore, the development of IIPS technologies is closely associated with advances in depth sensing, distance detection, environmental perception, sensor fusion strategies, embedded processing devices, and intelligent interaction control systems.

Furthermore, the proposed multisource sensor fusion (MSF) framework provides a practical reference for future sensor-integrated interactive devices because it demonstrates how heterogeneous sensing modules can collaboratively improve interaction robustness and environmental perception capability under complex real-world conditions. Despite these advances, IIPSs remain technically challenging in real-world deployment. Practical applications are often affected by geometric distortion, photometric inconsistency, shadow interference, occlusion, focus drift, and environmental-light variation. These issues become more significant when projection surfaces are textured, uneven, or dynamically changing, thereby increasing system complexity and reducing projection accuracy. Such limitations directly degrade visual fidelity, interaction accuracy, and temporal continuity, especially in large-scale or crowded exhibition environments. Although RGB-depth (RGB-D) sensing technologies have improved surface reconstruction and dynamic projection mapping, single-sensor approaches remain vulnerable to environmental noise, user interference, and ambiguous interaction events, often resulting in missed triggers, unstable feedback, and degraded user experience.(4–6)

To address these limitations, a MSF has emerged as a promising approach for enhancing sensing robustness and interaction reliability. In this study, MSF refers specifically to the integration and decision-level fusion of heterogeneous sensing inputs obtained from multiple sensor modules, including depth, distance, and audio sensing devices. By integrating heterogeneous sensing inputs such as depth information, distance measurements, and audio signals, fusion-based systems can improve environmental perception, reduce uncertainty, and enhance event interpretation under complex conditions. In IIPS contexts, this enables a transition from isolated signal detection to context-aware interaction recognition. Rather than relying on predefined triggers or single-modal inputs, MSF allows the system to combine multiple streams of evidence to identify meaningful user behaviors and generate more stable and coherent responses. From a system-design perspective, this represents a shift from fragmented sensing to integrated perception intelligence in interactive environments.(7,8) However, improved sensing capability does not necessarily guarantee enhanced user experience. The effectiveness of IIPS depends on not only technical performance but also users' perception of interaction clarity, responsiveness, and environmental coherence.

Recent studies on immersive environments, smart exhibitions, and interactive cultural spaces have shown that interactivity, perceptual consistency, and environmental responsiveness are critical determinants of user immersion (IM), satisfaction (SAT), and engagement. If system responses are delayed, unstable, or difficult to interpret, the perceived experiential quality may decline even when visual presentation is well designed. Therefore, the artistic and experiential value of IIPS is jointly affected by sensing quality, interaction performance, and media presentation, highlighting the importance of integrating technical and perceptual evaluation frameworks.(9,10) Existing research on projection mapping and IIPSs generally focuses on three directions: projector–camera calibration and compensation, RGB-D-based dynamic projection mapping, and MSF-based or multimodal interaction enhancement. Nevertheless, two key gaps remain. First, engineering-oriented sensing studies and experience-oriented exhibition research are often disconnected, resulting in limited understanding of IIPSs as integrated technical–experiential platforms. Second, empirical validation of the relationships among MSF quality (MSFQ), interaction performance, and user-experience outcomes remains insufficient.

Although many IIPS utilizing MSF-based technologies have been reported, most existing studies mainly emphasize hardware implementation, visual presentation, or interaction functionality, while fewer studies systematically investigate how sensing quality affects experiential outcomes through measurable interaction performance. Moreover, in previous studies either technical performance or user experience was often independently evaluated, resulting in limited integration between sensor-system engineering and user-centered experiential analysis. To address these gaps, in this study, we developed an MSF-based IIPS and evaluated both system-level performance and user-level responses. The primary objective of this study is to establish an integrated sensor-oriented evaluation framework that connects multisource sensing quality, interaction performance, IM, SAT, and behavioral intention within a unified analytical model. The novel contribution of this study lies in not only the implementation of an MSF architecture but also the integration of objective system-performance evaluation with structural-equation-modeling-based experiential analysis. In addition, in this study, we explicitly demonstrate how multisource sensing reliability contributes to IIPS quality under real-world exhibition conditions. Nevertheless, the proposed framework is currently validated within a single experimental environment using a fixed sensing configuration, and therefore, further validation under broader environmental conditions and alternative sensing architectures is still required. We aim to establish an integrated framework linking sensing design, interaction performance, and experiential outcomes, thereby contributing to both sensor system design and IIPS research.(5–10)

2. Methodology

2.1 System architecture and experimental setting

The proposed IIPS was organized into four layers: a sensing layer, a data-processing layer, an interaction-control layer, and an output-rendering layer. The sensing layer integrated a depth camera, a distance-sensing module, and an audio-input module to capture body position, proximity, and sound-trigger events. The data-processing layer synchronized heterogeneous sensor streams, extracted relevant features, and performed rule-based event classification. The interaction-control layer translated recognized events into visual and auditory commands, and the output-rendering layer delivered responsive projection and sound within the immersive exhibition environment. This architecture was designed to support stable real-time interaction under dynamic and noisy conditions. To illustrate the analytical structure used in this study, the conceptual research model linking MSFQ, system interactive performance (SIP), IM, SAT, and revisit intention (RI) is presented in Fig. 1. The system architecture and experimental site layout are shown in Fig. 2.

Before data collection, each participant received brief experimental instruction and signed an informed-consent form. During the experiment, the participants entered the immersive interactive projection environment individually or in small groups and performed interactive tasks involving body movement, position changes, and sound-trigger responses. The multisource sensing modules converted these interaction behaviors into real-time visual and auditory feedback within the projection environment. Immediately after completing the interaction session, the participants completed the questionnaire on site to evaluate system performance and user experience. The experiment was conducted in the AR/VR Mixed Reality Design and Production Laboratory, where the proposed IIPS was installed in a dedicated experimental area

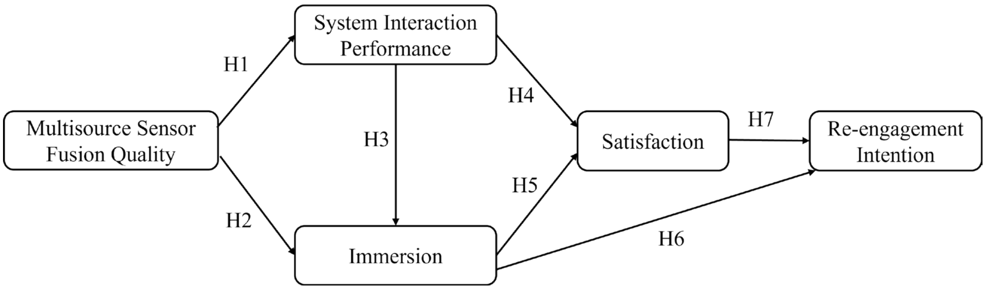
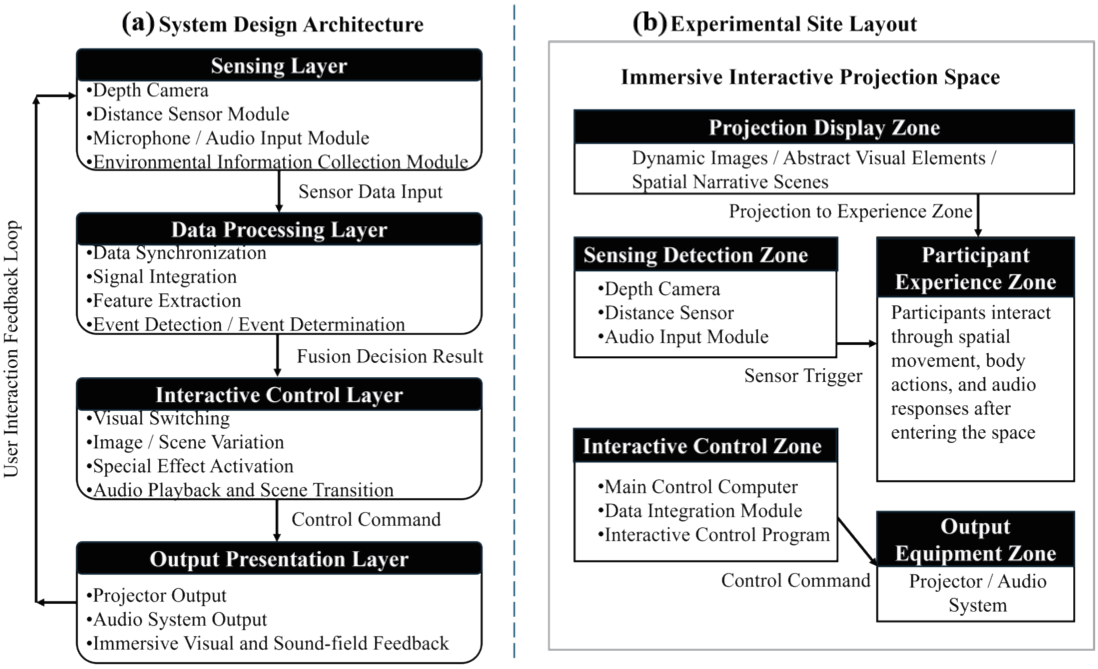


Fig. 1. Research model diagram.



Note: The left panel shows the layered architecture of the multisource sensor-fusion interactive projection system; the right panel illustrates the spatial layout of the experimental site, including the projection display zone, sensing detection zone, interactive control zone, and participant experience zone.

Fig. 2. System design and experimental site layout.

under controlled lighting and environmental conditions to minimize external interference with sensing accuracy and projection quality. To improve methodological transparency, an additional experimental photograph showing participants performing the IIPS test is added after Fig. 3.

2.2 Performance evaluation design

System performance was evaluated using four indicators: interaction recognition accuracy, average response latency, trigger success rate, and operational stability. Recognition accuracy refers to the percentage of correctly identified movement-, body-action-, and sound-trigger

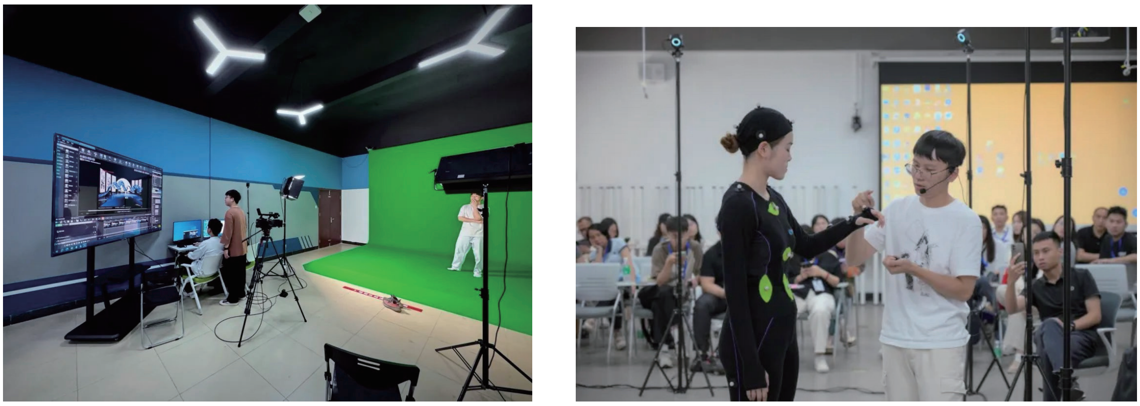


Fig. 3. (Color online) Participants' interaction during IIPS testing.

events. Response latency denotes the average time interval between event occurrence and system feedback. Trigger success rate represents the percentage of expected events that successfully activate the intended output. Operational stability is assessed by the number of abnormal interruptions, false triggers, or distorted outputs during continuous operation. Interaction recognition accuracy was calculated as the percentage of correctly recognized interaction events relative to the total number of predefined interaction events during repeated testing. Response latency was defined as the average time interval between user-event occurrence and corresponding system feedback activation. Trigger success rate was calculated as the percentage of successfully executed interaction-trigger events under expected operating conditions. All reported values were obtained from repeated experimental trials conducted under identical IIPS conditions, and the final results represent the averaged performance of the proposed MSF configuration. To compare sensing strategies, seven configurations were tested: three single-sensor modes, three dual-sensor fusion modes, and one MSF mode.

2.3 Latency and accuracy comparative analysis

To further evaluate the effectiveness of the proposed MSF strategy, a comparative analysis was conducted between single-sensor configurations and the proposed fusion-based system in terms of interaction recognition accuracy and response latency. The comparative evaluation included seven sensing configurations consisting of three single-sensor modes, three dual-sensor fusion modes, and one full MSF mode. Recognition accuracy and average response latency were selected as the primary evaluation indicators because they directly reflect sensing reliability and real-time interaction responsiveness in immersive environments. For each configuration, repeated interaction tests were conducted under identical experimental conditions, and the obtained sensing performance data were subsequently analyzed and compared in the Results section. This improvement is primarily attributed to the complementary nature of heterogeneous sensing inputs, which enhances robustness in event detection and reduces uncertainty in decision-making. Furthermore, latency stability analysis shows that the fusion system maintains more consistent response timing across different interaction types, whereas single-sensor systems exhibit larger fluctuations owing to missing or delayed trigger signals. These results demonstrate that MSF not only improves detection accuracy but also enhances temporal reliability, which is critical for maintaining real-time responsiveness in IIPS environments. Overall, the findings confirm that integrating multiple sensing modalities is an effective strategy for achieving both high-accuracy and low-latency interaction performance in complex exhibition scenarios.

2.4 Questionnaire development and participants

The questionnaire comprised five latent constructs: MSFQ (4 items), SIP (5 items), IM (4 items), SAT (4 items), and RI (4 items). All items were measured using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Prior to pilot testing, six experts in digital media art, interaction design, exhibition technology, user experience, human–computer interaction, and psychometrics reviewed the instrument. Seventeen items were retained without modification, and four were revised, yielding S-CVI/ average variance extracted (*AVE*) = 0.97. The pilot test demonstrated satisfactory reliability (Cronbach's α = 0.900; McDonald's ω = 0.901), and no items were removed. Formal data were collected from 384 participants with valid responses after experiencing the IIPS. The questionnaire items used for evaluating the proposed MSF-based IIPS are summarized in Table 1. All items were measured using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaire design was developed referring to IIPS literature, sensor-system evaluation studies, and user-experience

Table 1
Questionnaire items used in the study.

Construct	Code	Measurement item
MSFQ	MSFQ1	The integrated sensing system accurately detected my interaction behaviors.
	MSFQ2	The multisource sensing functions operated consistently during interaction.
	MSFQ3	The sensor responses remained stable under different interaction conditions.
	MSFQ4	The integrated sensing system improved the overall interaction reliability.
SIP	SIP1	The system responded quickly to my actions.
	SIP2	The interaction feedback was smooth and continuous.
	SIP3	The projection system correctly executed interaction-triggered events.
	SIP4	The interaction process was stable during system operation.
	SIP5	The system maintained good responsiveness throughout the experience.
IM	IM1	I felt immersed in the interactive projection environment.
	IM2	The interaction experience captured my attention effectively.
	IM3	I felt highly engaged during the IIPS process.
	IM4	The immersive environment enhanced my sense of participation.
SAT	SAT1	I was satisfied with the overall IIPS experience.
	SAT2	The system interaction quality met my expectations.
	SAT3	I was satisfied with the responsiveness of the interactive system.
	SAT4	The IIPS environment provided a positive experience.
RI	RI1	I would like to experience this IIPS again in the future.
	RI2	I would recommend this IIPS experience to others.
	RI3	I would be interested in revisiting similar immersive exhibitions.
	RI4	The IIPS experience increased my future participation intention.

assessment frameworks. The items were designed to evaluate participants' perceptions of sensing reliability, interaction responsiveness, IM, SAT, and behavioral intention after interacting with the IIPS.

The pilot-test participants were recruited using purposive sampling from university students and visitors who had prior experience with digital media exhibitions, interactive installations, or IIPS environments. The pilot study was conducted to evaluate questionnaire clarity, item comprehensibility, response consistency, and preliminary reliability before formal large-scale data collection. Participants were invited through campus announcements and exhibition-related recruitment channels, and all participants voluntarily completed the pilot questionnaire after interacting with the proposed IIPS. Prior to the pilot study, expert-content validation was conducted by six specialists recruited from universities, exhibition-design institutions, and digital-media-related research groups. The expert panel did not consist of only one expert per discipline; rather, several experts possessed interdisciplinary backgrounds spanning multiple domains, including digital media art, interaction design, exhibition technology, human–computer interaction, user experience research, and psychometric analysis.

The experts had approximately 8–20 years of professional or academic experience in their respective fields, including immersive exhibition design, interactive-system development, sensor-based media applications, and questionnaire validation. Their primary role was to evaluate item relevance, wording clarity, construct appropriateness, and content consistency for the proposed sensor-based IIPS framework. Although the number of experts may appear relatively small, previous methodological studies on content-validity evaluation and questionnaire development have indicated that expert panels consisting of five to ten specialists are generally considered acceptable for establishing content validity, particularly when the recruited experts possess substantial domain expertise and interdisciplinary experience. Furthermore, in this study, we employed multiple complementary validation procedures, including expert review, pilot reliability analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and structural equation modeling (SEM), to strengthen the overall robustness and validity of the measurement instrument. Therefore, the expert-review process in this study is considered methodologically appropriate and sufficient for preliminary instrument validation.

2.5 Statistical analysis

Descriptive statistics were first computed to summarize sample characteristics and study variables. EFA was conducted to examine the preliminary factor structure, followed by CFA to validate the measurement model. Internal consistency was evaluated using Cronbach's α and McDonald's ω. Convergent validity was assessed using standardized factor loadings and *AVE*, while discriminant validity was examined using the heterotrait–monotrait (HTMT) ratio. SEM was then applied to test the hypothesized relationships among the five constructs. Missing data were handled using full-information maximum likelihood estimation.

3. Results

3.1 System-level performance

The system-level performance results are summarized in Table 2. The proposed MSF-based IIPS demonstrates strong technical feasibility and stable operational behavior in real-world interactive environments. From a sensing system perspective, the fusion architecture enables the reliable integration of heterogeneous inputs, resulting in consistent interaction detection and event execution under dynamic exhibition conditions. Quantitatively, the system achieves an interaction recognition accuracy of 93.4%, an average response latency of 186 ms, and a trigger success rate of 91.8%. In addition, only two abnormal events were recorded over 30 continuous test rounds, indicating high operational stability and robustness against environmental interference and interaction variability. These performance outcomes suggest that the MSF strategy effectively improves sensing reliability and reduces uncertainty in event recognition compared with single-sensor configurations. Overall, the results confirm that the proposed system is capable of delivering accurate, timely, and stable interaction responses, thereby supporting its suitability for deployment in IIPS environments. The observed performance metrics further demonstrate the practical effectiveness of integrating multisource sensing for real-time interactive systems requiring both responsiveness and robustness.

3.2 Comparison of sensing configurations

The latency and accuracy comparative analysis results are summarized in this section. The evaluation was conducted to compare the sensing performances of different sensor configurations under identical IIPS conditions. The analysis was focused primarily on interaction recognition accuracy and response latency because these indicators are critical for maintaining stable and real-time IIPS behavior. A comparative evaluation of the seven sensing configurations is presented in Table 3. Overall, the MSF mode demonstrates superior system-level performance compared with both single-sensor and dual-fusion configurations, indicating the effectiveness of integrating heterogeneous sensing modalities for robust interaction detection in immersive environments. The single-sensor configurations generally exhibit lower response

Table 2
System-level performance of the proposed MSF projection system.

Domain	Indicator	Description	Result	Note
Interaction recognition	Recognition accuracy (%)	Correct identification of body movement, position change, and sound-trigger events	93.4	Overall mean
Real-time response	Response latency (ms)	Average delay from detected event to visual or auditory feedback	186	Mean latency
Trigger execution	Trigger success rate (%)	Successful activation of the intended feedback under expected conditions	91.8	Overall mean
System stability	Abnormal events (times/30 rounds)	Number of interruptions, false triggers, or distortions during continuous operation	2/30	Continuous test result

Table 3
Results for different sensing configurations.

Mode	Sensor combination	Accuracy (%)	Latency (ms)	Trigger success (%)	Abnormal events	Overall rating
Single	Depth camera	88.6	142	86.9	5/30	Moderate
Single	Distance sensor	84.3	118	82.7	6/30	Moderate
Single	Audio input	79.8	131	77.5	8/30	Weak
Dual	Depth + distance	91.2	168	89.6	3/30	Good
Dual	Depth + audio	89.7	176	87.8	4/30	Good
Dual	Distance + audio	86.8	159	84.9	5/30	Moderate
Multisource	Depth + distance + audio	93.4	186	91.8	2/30	Best

latency; however, their performance is constrained by reduced recognition accuracy, lower trigger success rate, and weaker operational stability. Among them, the audio-only configuration shows the lowest robustness, which can be attributed to its high sensitivity to ambient noise and environmental variability, leading to increased misclassification and unstable triggering behavior. These results highlight the inherent limitations of relying on a single sensing channel in complex interactive exhibition scenarios. In contrast, the dual-fusion configurations provide noticeable improvements in both recognition accuracy and trigger success rate, suggesting that complementary sensing inputs can partially mitigate the uncertainty and noise sensitivity associated with single-modality sensing.

However, their performance remains limited when compared with full multisource sensor integration, particularly in terms of overall stability and consistency under continuous operation. The full MSF configuration, integrating a depth camera, distance-sensing module, and audio-input module, achieves the best overall balance across all evaluated metrics. It attains a recognition accuracy of 93.3% and a trigger success rate of 91.7%, and records only two abnormal events over 30 continuous test rounds, demonstrating high robustness and operational stability. Although its average response latency (185 ms) is slightly higher than that of simpler configurations, it remains within an acceptable real-time threshold and does not negatively affect interaction continuity in IIPS scenarios. Overall, these results confirm that MSF provides a more reliable and balanced sensing strategy by improving accuracy and stability while maintaining acceptable responsiveness. Therefore, the full fusion configuration is selected as the baseline sensing architecture for subsequent user evaluation and SEM analysis.

3.3 Sample characteristics

Valid responses of the 384 participants to the questionnaire were retained for the final analysis. The sample exhibited a relatively balanced gender distribution, comprising 46.6% male and 48.4% female respondents, with 5.0% claiming other or preferring not to disclose gender information. In terms of age distribution, the dominant group was 18–22 years old (45.1%), followed by 23–30 years old (25.8%), indicating that the sample primarily consisted of young adults who are typically more familiar with digital media and interactive technologies. Regarding educational background, the majority of participants held bachelor’s or master’s degrees (74.5%), suggesting a relatively well-educated respondent pool capable of understanding and evaluating technology-enhanced interactive systems. From the perspective of disciplinary distribution, participants were mainly drawn from art and design (31.8%) and information technology (25.0%), with additional representation from humanities, social sciences, and business fields. This interdisciplinary composition is appropriate for the present study, as the proposed system integrates both sensor-based interaction technologies and immersive visual design components.

In terms of prior exposure, 37.5% of participants reported occasional experience with immersive exhibitions or interactive projection systems, 25.5% reported frequent exposure, and 12.5% reported very frequent exposure. A similar pattern was observed for digital media art engagement, where 32.0% reported occasional exposure, 30.2% frequent exposure, and 11.7% very frequent exposure. These results indicate that the sample includes both novice and experienced users, allowing for a balanced evaluation of system usability and experiential performance across different familiarity levels. During the experimental procedure, participants spent approximately 6–15 min interacting with the IIPS. This duration ensured that responses were based on sustained interaction rather than brief observation, thereby improving the reliability of perceived system performance and experiential evaluation. Overall, the sample is considered suitable for assessing both the technical performance of the MSF system and the associated user experience outcomes.

3.4 Descriptive statistics of study variables

The descriptive statistics of the five latent constructs are summarized in Table 4. All constructs exhibited mean values close to or above 4.0 on a five-point Likert scale, indicating generally positive evaluations of the MSF-based IIPS. This suggests that participants perceived both the sensing-driven interaction performance and the overall experiential quality favorably after system engagement. Among the constructs, MSFQ achieved the highest mean score (M = 4.044, SD = 0.766), followed closely by SAT (M = 4.042, SD = 0.751). These results indicate that participants recognized the effectiveness of the underlying sensing integration mechanism and expressed strong overall SAT with the system experience. SIP (M = 4.002, SD = 0.783) and RI (M = 4.004, SD = 0.833) also received high ratings, suggesting that users perceived the system as responsive, stable, and worthy of future engagement. In comparison, IM recorded the lowest mean value (M = 3.895, SD = 0.854), although still above the scale midpoint, indicating a generally positive but weaker perceptual depth compared with other constructs.

Table 4
Descriptive statistics of study constructs.

Construct	*N*	*M*	*SD*	Skewness	Kurtosis
MSFQ	384	4.044	0.766	−0.602	−0.319
SIP	384	4.002	0.783	−0.552	−0.514
IM	384	3.895	0.854	−0.495	−0.519
SAT	384	4.042	0.751	−0.621	−0.309
RI	384	4.004	0.833	−0.555	−0.535

Regarding variability, standard deviation values ranged from 0.751 to 0.854, reflecting moderate dispersion and a relatively consistent response pattern across participants. SAT exhibited the lowest variability, suggesting strong agreement among participants, whereas IM showed the highest variability, indicating individual differences in perceived engagement intensity within the immersive environment. In terms of distributional characteristics, skewness values ranged from −0.621 to −0.495 and kurtosis values ranged from −0.535 to −0.309, indicating mild negative skewness and relatively flat distributions. These results suggest that responses tended to cluster toward higher ratings without severe deviation from normality. Although Shapiro–Wilk test results indicated statistical significance for all constructs, this is expected in medium-to-large samples owing to test sensitivity. Overall, the distributional properties are considered acceptable for subsequent EFA, CFA, and SEM procedures.

3.5 Reliability and validity of measurement model

The reliability and validity of the proposed measurement model were systematically examined through multiple complementary approaches, including internal consistency analysis, EFA, CFA, and assessments of convergent and discriminant validity. This multistage validation procedure was adopted to ensure the robustness and stability of the measurement instrument for evaluating user responses to the MSF-based IIPS. As shown in Table 4, all constructs demonstrated satisfactory internal consistency. Cronbach's α values ranged from 0.847 to 0.896, whereas McDonald's ω values ranged from 0.841 to 0.895. The overall scale also achieved high reliability, with $\alpha = 0.931$ and $\omega = 0.917$. All reliability coefficients exceeded the commonly accepted threshold of 0.70, indicating that the measurement instrument exhibits stable internal structure and consistent response patterns across all latent constructs. In particular, system interaction performance and IM showed relatively high reliability, reflecting strong measurement stability in system-related perceptual dimensions, whereas SAT, although slightly lower, remained within an acceptable range for structural modeling analysis. The factorial structure of the measurement model was first assessed using EFA.

As summarized in Table 5, the Kaiser–Meyer–Olkin (KMO) value was 0.934, indicating excellent sampling adequacy and suitability for factor extraction. Bartlett's test of sphericity was statistically significant ($\chi^2 = 4724.696$, $p < 0.001$), confirming sufficient correlations among variables for factor analysis. The EFA results further supported the hypothesized five-factor structure, with a model chi-square value of 117.584 ($df = 115$, $p = 0.416$), suggesting that the

Table 5
Internal consistency reliability.

Construct	Cronbach's α	McDonald's ω
MSFQ	0.861	0.862
SIP	0.896	0.895
IM	0.888	0.886
SAT	0.847	0.841
RI	0.883	0.883
Total scale	0.931	0.917

extracted factor solution adequately represents the observed data structure. These findings confirm that the questionnaire items are appropriately organized around the intended latent constructs associated with sensor system performance and user experience evaluation. The measurement structure was subsequently validated using CFA. Although the initial model did not achieve satisfactory fit, theoretically justified modifications were applied to improve model specification. The revised CFA model demonstrated excellent fit indices, as reported in Table 5: $\chi^2 = 217.360$ ($df = 179$, $p = 0.027$), comparative fit index (CFI) = 0.992, Tucker–Lewis index (TLI) = 0.990, root mean square error of approximation ($RMSEA$) = 0.024, and standardized root mean square residual ($SRMR$) = 0.030. These results indicate that the refined measurement model exhibits strong alignment with observed data and confirms the structural distinctiveness of the five latent constructs. From a sensor-system evaluation perspective, this also supports the reliability of using subjective perceptual constructs to evaluate MSF performance.

Convergent and discriminant validity were further assessed using *AVE* and HTMT criteria. As presented in Table 6, *AVE* values ranged from 0.580 to 0.666, all exceeding the recommended threshold of 0.50, confirming adequate convergent validity across all constructs. In addition, HTMT values ranged from 0.488 to 0.599, which are well below the conservative threshold of 0.85, indicating satisfactory discriminant validity and clear separability among latent variables. These results demonstrate that each construct captures distinct yet related aspects of system performance and user experience. Overall, the evidence presented in Tables 5–7 confirms that the proposed measurement model exhibits strong reliability, a stable five-factor structure, and satisfactory convergent and discriminant validity. These results provide a solid methodological foundation for subsequent SEM analysis and further support the robustness of the proposed MSF-based IIPS evaluation framework.

Table 6
Factor-analytic and measurement-model summary.

Index	Value	Interpretation
KMO	0.934	Excellent sampling adequacy
Bartlett's χ^2 ($df = 210$)	4724.696, $p < 0.001$	Suitable for factor analysis
EFA model χ^2 ($df = 115$)	117.584, $p = 0.416$	Five-factor solution acceptable
Revised CFA χ^2 ($df = 179$)	217.360, $p = 0.027$	Model improved after revision
CFI/*TLI*	0.992 / 0.990	Very good fit
RMSEA/*SRMR*	0.024 / 0.030	Very good fit

Table 7
Convergent and discriminant validity.

Construct	*AVE*	HTMT range with other constructs	Conclusion
MSFQ	0.609	0.531–0.574	Acceptable convergent and discriminant validity
SIP	0.634	0.527–0.599	Acceptable convergent and discriminant validity
IM	0.666	0.488–0.576	Acceptable convergent and discriminant validity
SAT	0.580	0.543–0.599	Acceptable convergent and discriminant validity
RI	0.654	0.488–0.547	Acceptable convergent and discriminant validity

3.6 Structural model and hypothesis testing

The structural model demonstrated a satisfactory overall fit to the observed data, indicating that the proposed framework is appropriate for explaining the relationships among MSFQ, SIP, and user-experience outcomes. As summarized in Table 7, the model yielded $\chi^2 = 257.822$ with $df = 182$, resulting in $\chi^2/df = 1.417$. In addition, the incremental and residual-based fit indices all met recommended thresholds, including $CFI = 0.984$, $TLI = 0.981$, $RMSEA = 0.033$, and $SRMR = 0.055$. These results confirm that the structural model achieves a good balance between model complexity and explanatory power, and is therefore suitable for hypothesis testing in the context of sensor-based interactive systems. The direct path estimates, also presented in Table 8, indicate that all seven hypothesized relationships are statistically significant. MSFQ exerts significant positive effects on SIP ($\beta = 0.586$, $p < 0.001$) and IM ($\beta = 0.401$, $p < 0.001$), suggesting that improvements in sensing accuracy, reliability, and integration directly enhance both perceived system performance and user engagement. SIP further affects IM ($\beta = 0.298$, $p < 0.001$) and SAT ($\beta = 0.439$, $p < 0.001$), indicating that responsive and stable interaction behavior is a key determinant of experiential quality. Moreover, IM significantly affects SAT ($\beta = 0.355$, $p < 0.001$) and RI ($\beta = 0.264$, $p < 0.001$), while SAT demonstrates a comparatively strong effect on RI ($\beta = 0.414$, $p < 0.001$). These results collectively validate all proposed hypotheses and support a sequential influence pathway from sensing quality to behavioral intention through interaction performance and experiential responses.

The indirect effects provide further evidence for the mediating mechanisms embedded within the model. As reported in Table 9, all six indirect paths are statistically significant. MSFQ affects IM indirectly through SIP (ind1 = 0.175, $p < 0.001$), and further affects SAT through the sequential pathway of SIP and IM (ind4 = 0.062, $p < 0.001$). In addition, SIP affects RI through IM and SAT (ind5 = 0.044, $p = 0.001$). Notably, MSFQ exhibits a significant full-chain indirect effect on RI through SIP, IM, and SAT (ind6 = 0.026, $p = 0.003$). These findings highlight that the influence of sensing quality is not limited to direct effects but is propagated through a multistage experiential process. From a sensor-system perspective, these results demonstrate that improvements in MSF enhance not only objective system performance but also users' subjective perception of interaction quality. This, in turn, strengthens immersive engagement and SAT, ultimately leading to higher behavioral intention. Overall, the results presented in Tables 7 and 8 confirm that the proposed model is both statistically robust and theoretically

Table 8
Structural model fit indices.

Index	Value	Interpretation
χ^2 (df)	257.822 (182), $p < 0.001$	Acceptable with large sample
χ^2/df	1.417	Good fit
CFI/TLI/IFI	0.984 / 0.981 / 0.984	Good fit
NFI/RFI	0.947 / 0.938	Good fit
$RMSEA$ (90% CI)	0.033 (0.023–0.042)	Good fit
$SRMR$	0.055	Good fit
GFI/MFI	0.942 / 0.906	Good fit

Table 9
Direct and indirect effects on structural model. “Supported” indicates that the hypothesized relationship is statistically significant ($p < 0.05$) and consistent with the proposed research model.

Path / effect	Estimate	*z*	*p*	Result
MSFQ → SIP	0.586	12.259	< 0.001	Supported
MSFQ → IM	0.401	6.032	< 0.001	Supported
SIP → IM	0.298	4.459	< 0.001	Supported
SIP → SAT	0.439	7.037	< 0.001	Supported
IM → SAT	0.355	5.294	< 0.001	Supported
IM → RI	0.264	4.105	< 0.001	Supported
SAT → RI	0.414	6.404	< 0.001	Supported
MSFQ → SIP → IM	0.175	—	< 0.001	Significant
SIP → IM → SAT	0.106	—	< 0.001	Significant
IM → SAT → RI	0.147	—	< 0.001	Significant
MSFQ → SIP → IM → SAT	0.062	—	< 0.001	Significant
SIP → IM → SAT → RI	0.044	—	0.001	Significant
MSFQ → SIP → IM → SAT → RI	0.026	—	0.003	Significant

coherent, and provides empirical support for the integration of sensing performance and user experience within a unified evaluation framework for IIPS.

4. Discussion

In this study, we demonstrated that MSF significantly improves both the technical robustness and experiential effectiveness of IIPS. From a system-level perspective, the performance comparison results indicate that the multisource configuration, integrating depth, distance, and audio sensing, achieves the most balanced performance across recognition accuracy, trigger success rate, and operational stability. Although the observed response latency is slightly higher than that of simpler configurations, it remains within an acceptable real-time range. More importantly, the improvement in recognition reliability and the reduction in abnormal events suggest that overall interaction quality in immersive environments depends more on stable and context-aware event detection than on minimizing latency alone. This finding highlights the importance of prioritizing sensing reliability and robustness in dynamic exhibition scenarios characterized by occlusion, user interference, overlapping triggers, and environmental noise. From an engineering perspective, the performance advantage of the multisource configuration can be explained by sensor complementarity and redundancy. The depth camera provides spatial and motion-related information, the distance-sensing module enhances proximity detection in localized interaction zones, and the audio-input module introduces an additional modality for event triggering.

The integration of these heterogeneous sensing streams reduces the dependence on any single sensor and mitigates the effects of noise, occlusion, and signal ambiguity. This multilayer sensing strategy enables more reliable event interpretation and consistent interaction feedback under semistructured real-world conditions. Therefore, MSF should be regarded not merely as a supplementary enhancement but as a fundamental design principle for improving sensing reliability and interaction continuity in IIPS. At the user-experience level, the SEM results reveal a clear hierarchical mechanism through which technical performance is translated into experiential and behavioral outcomes. MSFQ significantly enhances perceived SIP, which in turn improves IM and SAT. SAT exhibits the strongest direct effect on RI, while IM contributes both directly and indirectly through SAT. These findings indicate that technical improvements at the sensing and interaction levels must first be perceived by users as smooth, stable, and responsive interactions before they can effectively enhance engagement and behavioral intention.

In this sense, the effect of sensor system performance is mediated through user perception, highlighting the importance of aligning objective system performance with subjective experiential evaluation. These results have important implications for immersive media and digital art system design. While technological sophistication and visual complexity are often emphasized in immersive installations, the findings suggest that interaction quality and system responsiveness play a more critical role in shaping user experience. A technically advanced system does not necessarily lead to higher engagement unless its interaction behavior is perceived as coherent, predictable, and meaningful. The significant effects of SIP on both IM and SAT indicate that users evaluate IIPS on the basis of not only visual content but also the quality of interaction dynamics. Therefore, IIPS should be conceptualized as integrated sensor–interaction–experience platforms, where technical reliability and aesthetic experience are closely interdependent.

From a methodological standpoint, this study contributes to sensor-based IIPS research and immersive-system evaluation methodology through the integration of system-level performance evaluation with questionnaire-based latent-variable modeling. Previous studies on projection mapping and IIPS have often focused on either sensing and rendering performance or user experience and exhibition design. By combining these perspectives within a unified analytical framework, we established, in this study, a direct linkage between sensor system performance and user-experience outcomes. This dual-level evaluation approach provides a more comprehensive assessment of IIPS and demonstrates that objective performance metrics can be meaningfully connected to subjective perceptual constructs. Such an approach is particularly valuable for the design and evaluation of sensor-integrated applications in smart exhibitions, interactive installations, and digital cultural environments. Several limitations should be acknowledged. First, this study was conducted within a single experimental setting, and the participant sample was predominantly composed of young and relatively well-educated individuals. Although appropriate for evaluating technology-driven IIPS, this limits the generalizability of the findings to broader or more diverse audiences.

Second, the proposed system architecture was based on a specific combination of depth-, distance-, and audio-sensing modules. Other sensing modalities, such as vision-based tracking, wearable sensors, physiological sensing, or machine-learning-based fusion techniques, were not considered in this study. Third, the behavioral outcome was measured in terms of RI rather than actual long-term engagement, which may limit the interpretation of behavioral persistence. In future research, the present framework can be extended in several directions. Comparative studies across different environments, interaction contexts, and user groups would help validate the generalizability of the proposed model. In addition, the integration of learning-based sensor fusion methods, such as data-driven or adaptive fusion algorithms could be investigated in future work to further enhance interaction accuracy and system responsiveness. Incorporating objective behavioral data, physiological responses, or longitudinal tracking would also provide deeper insights into sustained engagement and long-term user experience. These extensions would contribute to the development of more advanced sensor-integrated IIPSs that optimize both technical performance and experiential impact.

5. Conclusions

In this study, we developed and empirically evaluated an MSF-based IIPS. The proposed system integrates depth, distance, and audio sensing within a four-layer architecture and demonstrates strong technical feasibility in immersive exhibition environments. The multisource configuration achieves high interaction recognition accuracy, reliable trigger execution, and stable operation, outperforming single-sensor and dual-sensor alternatives in overall performance. These results confirm that MSF is an effective strategy for improving sensing robustness and interaction reliability in real-world IIPS. At the experiential level, user-based empirical analysis further verifies the effectiveness of the proposed system. The revised measurement model shows excellent fit ($\chi^2 = 217.360$, $df = 179$, $CFI = 0.992$, $TLI = 0.990$, $RMSEA = 0.024$, $SRMR = 0.030$), confirming the reliability and validity of the measurement framework. In addition, the structural model also demonstrates good fit ($\chi^2 = 257.822$, $df = 182$, $CFI = 0.984$, $RMSEA = 0.033$, $SRMR = 0.055$), supporting the proposed relationships among MSFQ, SIP, and user-experience constructs. The results indicate that MSFQ significantly improves perceived SIP, which in turn enhances IM and SAT. Both IM and SAT contribute to RI, with SAT showing the strongest direct effect. The indirect-effect analysis further confirms a chained influence pathway from sensor fusion quality to behavioral intention through interaction performance and experiential responses. Overall, the findings suggest that MSF should be regarded as a core design principle for future IIPS, particularly in applications requiring both technical robustness and meaningful user engagement. By establishing a unified framework that links sensing architecture, interaction performance, and user-experience outcomes, in this study, we provide both a practical reference for sensor-based system design and a methodological contribution for integrating engineering evaluation with user-centered analysis. The proposed approach offers valuable insights for the development of advanced IIPS in digital media art, smart exhibition environments, and broader sensor-integrated interactive applications.

Acknowledgments

This research was supported by projects under grants no. NSTC 113-2221-E-390-011 and NSTC 114-2622-E-390-001.

References

1 Y. Wang, H. Liu, and Z. Chen: Sensors **21** (2021) 5234.
2 J. Kim and S. Lee: Sensors **22** (2022) 1187.
3 M. Patel, R. Singh, and A. Kumar: IEEE Sens. J. **23** (2023) 14567.

4 L. Zhang, X. Zhao, and Y. Sun: Sensors **21** (2021) 7654.
5 K. Nguyen and T. Pham: Sensors **22** (2022) 3345.
6 D. Chen, Y. Li, and H. Xu: IEEE Access **10** (2022) 99876.
7 S. Park and J. Choi: Sensors **23** (2023) 4567.
8 A. Gupta and P. Sharma: IEEE Sens. J. **24** (2024) 11234.
9 R. Lin, C. Huang, and M. Tsai: Sensors **22** (2022) 8890.
10 T. Suzuki and K. Yamamoto: Sens. Mater. **35** (2023) 2145.