Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60521

PAPER

# Intelligent Tool Design and Creative Behavior Analysis in Dance Composition Enabled by Mobile Interaction Technologies

**Yanjun Jiang** (✉)

Heze University, Heze, China

wlxjjj@126.com

## ABSTRACT

To advance the transition of dance composition from experience-driven practices to data-augmented and cognition-coordinated processes, a triadic model of mobile-interactive dance composition was proposed. Guided by this framework, a mobile intelligent composition tool integrating multimodal sensing and lightweight artificial intelligence (AI) was designed and implemented. A heterogeneous data fusion strategy combining an MPU9250 inertial measurement unit (IMU) with MediaPipe-based visual capture was adopted, achieving a three-dimensional reconstruction error below 2 cm and an interaction latency under 50 ms. A prototype implementation was developed and validated experimentally. A mixed experimental design involving 60 professional and non-professional dancers was conducted, incorporating a short single-segment creation task and a four-week longitudinal project. A cognitive load scale, a creative flow state scale, and eye-tracking measurements were employed to systematically compare the tool's performance with that of Kinect and professional motion capture systems. The findings reveal a three-stage evolutionary pattern in dance creative behavior under tool intervention, characterized by enhanced efficiency, cognitive restructuring, and expressive innovation, demonstrating the tool's comparative advantages in multi-scene adaptability and operational simplicity. This study establishes an interdisciplinary paradigm coupling theoretical modeling, tool development, experimental validation, and behavioral analysis, offering both conceptual foundations and a transparent methodological pathway for integrating mobile intelligent technologies into creative practices. The proposed tool effectively addresses critical challenges in dance education, digital heritage preservation, and real-time interaction in immersive performance, thereby providing essential support for the scalable digital transformation of dance art.

## KEYWORDS

mobile interaction technology, dance composition, intelligent tool, creative behavior analysis, triadic theoretical framework, multimodal sensing, lightweight AI

Jiang, Y. (2026). Intelligent Tool Design and Creative Behavior Analysis in Dance Composition Enabled by Mobile Interaction Technologies. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 60–74. https://doi.org/10.3991/ijim.v20i04.60521

Article submitted 2025-09-07. Revision uploaded 2025-12-30. Final acceptance 2026-01-08.

## 1 INTRODUCTION

The deep convergence of mobile computing, edge AI, and digital humanities [1–3] has been driving a fundamental transformation in the field of dance composition, shifting creative processes from choreographer-dependent experiential modes toward data-augmented and cognition-coordinated paradigms. During this transition, the demand for portable and intelligent composition tools has intensified across the industry [4, 5]. However, existing solutions continue to exhibit pronounced limitations: conventional systems are constrained by fixed creative spaces and high feedback latency, making them unsuitable for improvisational choreography and multi-scene adaptation [6, 7]; meanwhile, mainstream intelligent tools are often characterized by high equipment costs and complex operational procedures, resulting in substantial technical barriers to adoption [8]. These constraints further exacerbate persistent pain points in key application scenarios, including the difficulty of delivering personalized movement feedback in dance education, the challenge of balancing efficiency and precision in the digital preservation of traditional dance heritage, and the limited real-time interaction and dynamic adjustment capabilities in immersive performance creation.

The root of these challenges lies in the lack of a systematic theoretical foundation for integrating mobile interaction technologies with dance composition, as well as insufficient adaptation of technical implementations to the embodied cognitive characteristics of artistic creation. Existing studies predominantly focus on single technologies or isolated stages of the creative process [9–11], and an integrated analytical framework connecting technological features, cognitive processes, and artistic expression has yet to be established. Consequently, a significant research gap persists in the intersection of embodied cognition, mobile intelligence, and artistic creation. The present study is centered on the core inquiry of how mobile interaction technologies can empower dance composition. Its academic value lies in constructing a comprehensive theoretical system that supports the deep fusion of these technologies with creative practice, thereby enriching interdisciplinary perspectives on creative behavior in the digital age. Its practical significance is manifested through the development of low-threshold, high-accuracy tools designed to enhance compositional efficiency and diversify artistic output, while providing essential technological support for personalized instruction in dance education, digital transmission of traditional dance, and innovation in immersive performance.

The core research questions guiding this study may be distilled into three dimensions. First, it remains to be determined how a mobile multimodal interaction architecture can be designed to maintain both high precision and low intrusiveness while meeting technical requirements of three-dimensional reconstruction accuracy within 2 cm and interaction latency below 50 ms, and simultaneously preserving naturalness and flexibility in the creative process. Second, the extent to which the mobile intelligent tool influences creative efficiency, cognitive patterns, and artistic expression during both short-term adaptation and long-term use requires systematic examination, along with an assessment of its comparative advantages and limitations relative to Kinect and professional motion-capture systems. Third, it must be clarified how a creation ecosystem can be constructed in which technological empowerment, artistic ontology, and ethical regulation function synergistically. Such an ecosystem must avoid risks associated with movement-data privacy leakage and disputes related to AI-generated content copyright while supporting the tool's applicability for both professional and non-professional dancers across diverse cultural and choreographic contexts.

First, a closed-loop mobile intelligent composition tool architecture comprising "perception–processing–interaction–generation" was proposed. This architecture integrates a heterogeneous data fusion strategy combining an MPU9250 IMU with MediaPipe-based visual capture, and the implementation details reported in the manuscript support transparent evaluation and future implementation. Second, a triadic model of mobile-interactive dance composition was established, in which the technological adaptation layer, embodied cognition layer, and artistic expression layer are tightly coupled, enabling theoretical modeling to directly inform tool design. Third, a mixed experimental design combining short-term experiments, long-term tracking, and multi-tool comparison was implemented. The NASA Task Load Index (NASA-TLX) cognitive load scale, the creative flow-state scale, and gaze-path entropy analysis from eye-tracking data were employed to construct a quantitative analytical framework describing the dynamic evolution of creative behavior. Fourth, an ethics and universality framework was developed based on data privacy protection, copyright boundary specification, and stratified adaptation design, providing practical guidelines for the large-scale deployment of similar tools.

The subsequent sections are structured according to the logical progression of "theory–technology–experiment–application." In Section 2, the core theoretical framework is constructed and the internal mechanisms of the triadic model are articulated. Section 3 presents the design principles and technical implementation details of the intelligent tool. Section 4 outlines the experimental design and data acquisition procedures employed for the analysis of creative behavior. Section 5 reports the results of tool performance validation and the observed patterns of behavioral evolution in the creative process. The study concludes with a synthesis of key findings and a discussion of future research directions, thereby establishing a complete research closed loop.

## 2 THEORETICAL FOUNDATIONS AND ORIGINAL FRAMEWORK

### 2.1 Core theoretical foundations

The theoretical system underpinning this study is supported by three major pillars: mobile human–computer interaction theory [12, 13], embodied cognition theory [14], and human–machine co-creative theory [15, 16]. Together, these frameworks provide the conceptual foundation necessary for understanding the deep integration of mobile interaction technologies with dance composition. Mobile human–computer interaction theory serves as the basis for the design of the technological architecture. Its principles of multimodal sensory fusion emphasize the spatiotemporal alignment and complementarity of visual and inertial data. Visual data ensure global accuracy in joint-pose estimation, while inertial data compensate for local motion-capture deficiencies in occlusion-prone scenarios. Kalman filtering is employed to achieve redundant data calibration between these inputs. The design principles of portability, low intrusiveness, and real-time responsiveness directly correspond to the fundamental requirements of unrestricted movement expression and immediate feedback during dance composition. Embodied cognition theory elucidates the intrinsic mechanisms of dance creation, positing that bodily experience constitutes the core medium through which affective expression is produced. Through bodily movement, choreographers perceive spatial structures, construct imagery, and transmit emotional intent. Within this process, technology is not regarded as an external and detached apparatus but as a cognitive scaffold that extends the body, enhancing perceptual sensitivity to movement detail and emotional mapping through data augmentation. Human–machine co-creative theory defines the relational boundary between technology and the creative agent, outlining an evolutionary trajectory in which technology transitions from a passive assistive tool to an active creative partner. This theory delineates the conservation mechanism of artistic ontology, indicating that while technology may optimize creative workflows and expand expressive dimensions, it must remain anchored to the emotional core, cultural symbolism, and embodied experience that characterize dance, thereby preventing creative alienation driven by technological dominance.

## 2.2 Original theoretical framework: The triadic model of mobile-interactive dance composition

Building on the aforementioned theoretical foundations, a triadic model of mobile-interactive dance composition was constructed to achieve an integrated alignment of technological characteristics, cognitive processes, and artistic expression. The model comprises three layers—the technological adaptation layer, the embodied cognition layer, and the artistic expression layer—which together form an organic structure through a dynamic closed loop. The technological adaptation layer serves as the foundation of the model and incorporates three core modules: multimodal sensing, edge computing, and lightweight AI. Multimodal sensing enables comprehensive acquisition of movement data; edge computing ensures real-time data processing capability; and lightweight AI facilitates movement feature extraction and auxiliary generative functions. The embodied cognition layer operates as the central mediating component and encompasses three progressive stages: movement perception, decision adjustment, and creative generation. This layer receives quantitative data from the technological adaptation layer and transforms them into perceptible creative cues for choreographers. The artistic expression layer constitutes the final output stage, translating cognitive-level creative intentions into concrete dance expression across three dimensions: emotional mapping, stylistic presentation, and cultural symbolism.

The core mechanism of the model is reflected in the bidirectional interaction among the three layers. Immediate data feedback from the technological adaptation layer supports creative and decision-making processes within the embodied cognition layer, while the evolving demands of the embodied cognition layer drive parameter optimization within the technological adaptation layer. Creative concepts emerging from the embodied cognition layer directly shape the expressive form of the artistic expression layer, whose expressive outcomes, in turn, feed back into the cognitive process, enabling iterative refinement. A distinctive feature of the model lies in its integration of mobile interaction characteristics. Through the synergy between immediate feedback and improvisational creation, traditional latency barriers in the "perception–adjustment" cycle are effectively removed, enabling dynamic optimization of the creative process. A standardized framework diagram delineates the core elements, logical relationships, and interaction pathways across all layers, providing a precise theoretical mapping for subsequent tool design.

## 2.3 Dialogue with existing theories

The triadic model advances targeted innovations that enable a substantive dialogue with major cognitive theories such as distributed cognition theory [17] and extended mind theory [18]. Through this dialogue, conceptual breakthroughs are achieved by addressing the unique characteristics of dance creation while remaining grounded in foundational principles of cognitive science. The model thus contributes a theoretical innovation situated at the intersection of technology, cognition, and artistic practice. The core differences and contributions of the triadic model, relative to existing theories, are summarized in Table 1.

**Table 1.** Core differences between the triadic model and existing theories

| Comparison Dimension | Distributed Cognition Theory | Extended Mind Theory | Triadic Model |
|---|---|---|---|
| Primary focus | Distributed division of cognitive tasks | Extension of cognition through external tools | Co-evolution of technology, the body, and artistic expression |
| Role of the body | Execution substrate for cognitive tasks | Passive participant in cognitive processes | Principal agent in creative generation |
| Positioning of technology | Auxiliary node for cognitive tasks | External extension of cognition | Partner-like intermediary in co-creative processes |
| Artistic dimension | Not included in the core analytical framework | Emotional and cultural attributes de-emphasized | Emotional mapping and cultural symbolism are positioned as core components |

In comparison with distributed cognition theory, both frameworks acknowledge the mediating role of technology within cognitive processes. However, distributed cognition theory emphasizes the distributed allocation of cognitive tasks across the "individual–tool–environment" system, treating the body merely as an execution substrate while overlooking the dominant role of bodily experience and affective motivation in dance creation. In contrast, the triadic model foregrounds the real-time coordination of "technology–body–environment," situating the embodied cognition layer as the central hub and highlighting the body's authority in data transformation and creative generation. This emphasis results in a theoretical construct that is more closely aligned with the embodied characteristics of dance composition. Relative to extended mind theory, both perspectives recognize the capacity of technology to augment cognitive abilities. Nevertheless, extended mind theory conceptualizes technology as an external cognitive tool and limits its concern to improvements in cognitive efficiency, thereby diminishing the emotional core and cultural attributes fundamental to artistic creation. The triadic model, through the explicit delineation of the artistic expression layer, incorporates emotional mapping and cultural symbolism as central components. The model asserts that technological augmentation must ultimately serve the conservation of artistic ontology, clarifying that the role of technology is to strengthen rather than replace the distinctive nature of artistic expression. This safeguards against the reduction of dance creation to a purely computational cognitive task. This theoretical dialogue and advancement provide a novel research perspective for the intersection of art and technology—one that simultaneously respects technological principles and preserves the essential characteristics of artistic practice.

# 3 INTELLIGENT TOOL DESIGN AND IMPLEMENTATION

## 3.1 Requirements analysis and design objectives

The requirements analysis for the tool was derived from a comprehensive investigation of dance composition scenarios and an extraction of core challenges, resulting in a three-dimensional requirements framework encompassing scenario adaptability, interaction experience, and functional extensibility. Scenario adaptability requirements emphasize the diversity of creative environments. The tool must support movement capture in standard indoor rehearsal studios, open outdoor spaces, and small temporary sites, while satisfying the need for unencumbered capture to avoid restricting choreographic bodily expression. Interaction experience requirements prioritize low-latency feedback, with interaction latency required to remain below 50 ms to accommodate real-time adjustments inherent to improvisational creation. Functional extensibility requirements include dual support for individual creative work and multi-user collaborative composition to address the coordination demands of different creative modes.

Based on these requirements, three core design objectives were established. First, the balance between technological precision and artistic freedom is pursued by constraining three-dimensional reconstruction error to within 2 cm and joint angle error to below 0.5°, while employing lightweight design to prevent operational burdens from disrupting the creative process. Second, personalized adaptation capability is implemented through the construction of dedicated feature libraries tailored to stylistic characteristics across dance domains—including the expressive bodily resonance of classical dance, the explosive dynamics of modern dance, and the highly stylized motion patterns of ethnic dance—while also providing tiered user interfaces for professional and non-professional dancers. Third, a scenario-based modular architecture is constructed, offering rapidly deployable modules for education, digital heritage preservation, and immersive performance. These modules respectively address the requirements of pedagogical feedback, high-fidelity recording, and real-time interactive engagement.

### 3.2 Technical architecture design

The tool adopts a four-layer closed-loop architecture comprising perception, processing, interaction, and application. These layers are integrated through standardized data interfaces to ensure efficient coordination and precise alignment between technical performance and creative requirements. The perception layer employs a multimodal sensing fusion scheme equipped with an MPU9250 nine-axis IMU and the MediaPipe v0.10.9 monocular vision system. Joint angle, center-of-mass trajectory, force application characteristics, and motion velocity data are synchronously captured at a sampling rate of 100 Hz. The IMU ensures local precision in limb movement, while the visual system provides global pose calibration. A spatiotemporal alignment algorithm is implemented to eliminate heterogeneity between the two data sources. The processing layer integrates edge computing with lightweight AI. Model deployment on the device side is implemented using TensorFlow Lite v2.15. An unscented Kalman filter is applied to mitigate data drift, with the process noise matrix $Q$ set to $[1e^{-4}, 1e^{-4}, 1e^{-4}]$ and the observation noise matrix $R$ set to $[1e^{-3}]$, ensuring stable dynamic capture. A lightweight Transformer model comprising four encoder layers and a 128-dimensional hidden layer is used to extract dance style features at 10 ms per frame, enabling real-time style adaptation.

The interaction layer is designed with a multimodal feedback mechanism. Force distribution patterns are visualized through heatmaps, haptic rhythm cues are transmitted via Bluetooth 5.3 Low Energy technology, and voice-based interaction delivers style optimization suggestions. Overall feedback latency is maintained below 38 ms. The application layer adopts a modular architecture. The composition management module supports movement library storage, version iteration, and export in BVH, FBX, and JSON formats. The collaborative creation module enables multi-terminal data synchronization using a 5G network and Redis caching, supporting concurrent creation for up to eight users. The scenario adaptation module enables rapid switching among educational, digital heritage, and performance modes, corresponding respectively to movement correction, high-fidelity archival recording, and real-time generation. Data transmission is implemented through Bluetooth 5.3 Low Energy for real-time streaming. Local data are stored using AES-256 encryption, and cloud backups are processed through anonymization protocols, balancing data security with accessibility.

Figure 1 presents the complete technical architecture of the mobile-interactive intelligent dance composition tool, encompassing the tool's client interface, the motion data relay service, and the dance interaction processing module. The modular composition and data flow logic of the four-layer architecture—perception, processing, interaction, and application—are clearly illustrated, providing an intuitive representation of the technical pathways underlying multimodal motion acquisition, lightweight style model loading, and related system functions.

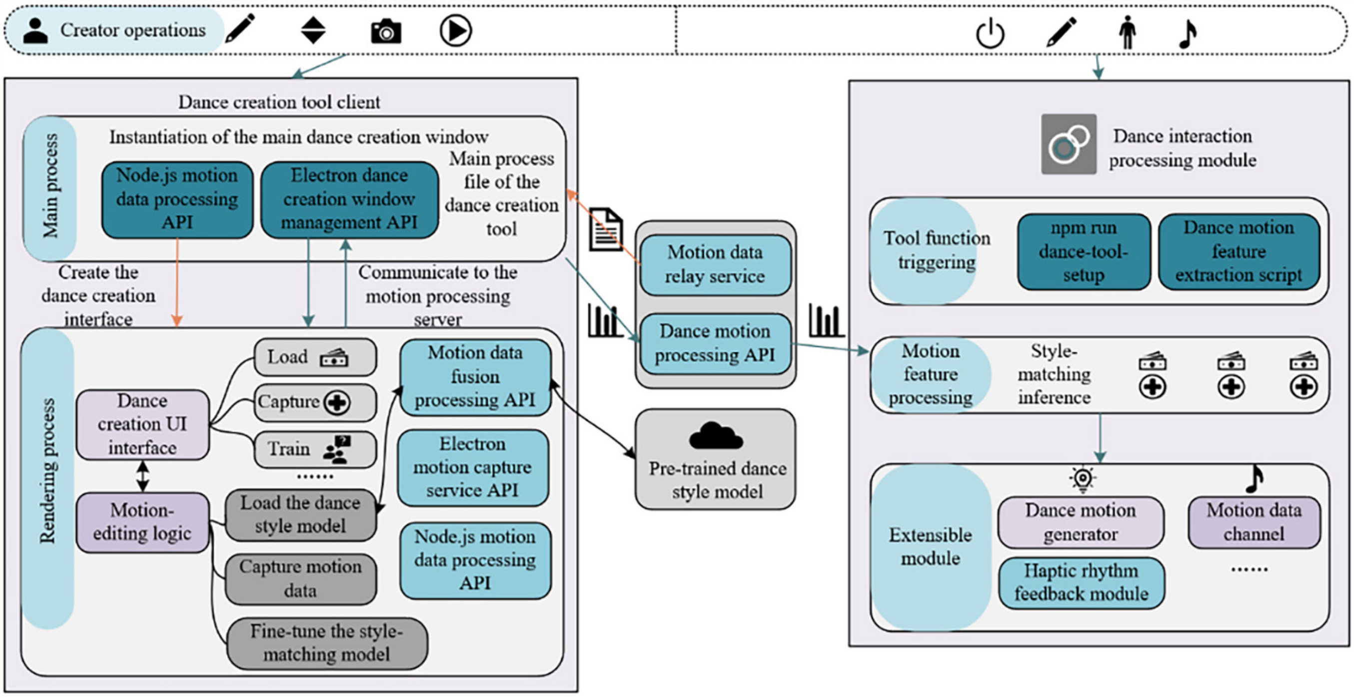


**Fig. 1.** Overview of the system architecture for the mobile-interactive intelligent dance composition tool

## 3.3 Implementation of core functional modules

Based on the four-layer technical architecture, four core functional modules were implemented. These modules operate through deeply integrated data pathways to support the full creative workflow of capture–generation–feedback–adaptation. Their key technologies, performance indicators, and application scenarios are summarized in Table 2. The mobile motion capture module functions as the primary data entry component and adopts a fusion scheme combining MediaPipe-based visual pose estimation with IMU data. The visual system outputs 2D joint keypoints, while the

IMU provides 3D motion pose data. A Perspective-*n*-Point (PnP) algorithm is applied to solve three-dimensional pose estimations. To address abnormal data resulting from motion blur or occlusion, a Random Sample Consensus (RANSAC) algorithm is introduced to remove outliers. The final reconstruction accuracy is maintained within 2 cm, satisfying precision requirements across multiple dance genres. The intelligent-assisted generation module is constructed using a lightweight Generative Adversarial Network (GAN). The encoder extracts semantic features from the input movements, and the decoder, combined with pre-trained style libraries for five major dance genres—classical, modern, ethnic, jazz, and contemporary—generates motion variants. A semantic consistency loss function is applied to ensure that generated variants remain aligned with the core semantics of the original movement, preventing unintended deviations in creative intent during style transfer.

The real-time feedback optimization module integrates biomechanical modeling with a dance style database. The biomechanical model computes joint loads and force efficiency based on the human skeletal–muscular structure, while the style database stores characteristic motion parameters representative of various dance genres. Their combined output yields three quantitative indicators—accuracy, coordination, and style congruence—along with corresponding visualized optimization suggestions. The scenario-adaptive module enables configuration-based switching across functional combinations. In the educational mode, emphasis is placed on real-time motion correction and instructional feedback, including numerical indicators of joint angle deviation. In the digital heritage mode, high sampling rate capture and multi-format archival functionality are activated to ensure long-term preservation value of motion data. In the performance mode, multi-user collaboration interfaces and real-time motion generation are optimized to support improvisational interactive creation in stage environments.

**Table 2.** Technical parameters and application scenarios of core functional modules

| Module Name | Core Technologies | Performance Indicators | Primary Application Scenarios |
|---|---|---|---|
| Mobile motion capture module | MediaPipe–IMU fusion and RANSAC | 3D reconstruction error ≤ 2 cm | Movement data acquisition across diverse environments |
| Intelligent-assisted generation module | Lightweight GAN and semantic consistency loss | Supports five dance style categories and generation latency ≤ 50 ms | Style transfer and motion variant generation |
| Real-time feedback optimization module | Biomechanical modeling and a dance style database | Three quantitative indicators and feedback latency ≤ 38 ms | Precision enhancement and style optimization of movements |
| Scenario-adaptive module | Modular functional configuration | Mode switching time ≤ 1 s | Education, heritage preservation, and immersive performance |

## 3.4 Prototype validation and reproducibility

A three-tier validation framework was designed to assess the tool's technical effectiveness and methodological transparency, covering technical performance, data consistency, and user usability, while transparency was further supported through detailed reporting of the implementation and experimental procedures.

For technical performance evaluation, three representative creative environments were selected: an indoor rehearsal studio, an outdoor plaza, and a confined 10 $m^2$ space. Standardized movement sequences performed by 10 professional dancers were used as test samples, and the OptiTrack Prime 13 high-precision optical motion capture system served as the ground-truth reference. The results indicate that the average three-dimensional reconstruction errors across the three environments were 1.5 cm, 1.8 cm, and 1.7 cm, respectively; interaction latency remained stable between 32 and 38 ms; and the system sustained continuous operation for up to eight hours on a single charge, meeting the core performance requirements for mobile creative work. Reproducibility testing employed the publicly available Human3.6M and AIST++ dance motion datasets for cross-validation. The average capture error for the 15 action categories in Human3.6M was 1.6 cm, while the corresponding error for the five dance genres in AIST++ was 1.8 cm. Deviations from the benchmark optical system were consistently maintained within 0.3 cm, demonstrating performance stability across standardized datasets.

An initial usability assessment was conducted with 10 professional dancers and 10 non-professional dancers over a two-week pilot trial. Task completion time, operational error rate, and satisfaction scales were used as evaluation metrics. Professional users reported a satisfaction score of 8.2/10 for motion capture accuracy and style generation, while non-professional users rated the satisfaction of the operation workflow at 7.8/10. Based on this feedback, refinements were made to the interface's icon layout and feedback prompt mechanisms. The study provides detailed descriptions of the core methodological components, including the multimodal sensing configuration, data fusion workflow, model architecture, and experimental procedures. Technical aspects such as hardware configuration, software environment, parameter settings, and data preprocessing steps are explicitly reported to support methodological transparency. These descriptions establish a practical foundation for future implementation, comparison, and extension of the proposed approach by other researchers.

## 4 CREATIVE BEHAVIOR ANALYSIS: RESEARCH DESIGN

### 4.1 Research hypotheses

Based on the core logic of the triadic model of mobile-interactive dance composition and the technical characteristics of the tool, four progressive research hypotheses were formulated, establishing a systematic validation framework spanning tool effectiveness, behavioral evolution, group adaptability, and cognitive mechanisms. The four hypotheses are interrelated: Hypothesis 1 anchors the comparative advantages of the tool; Hypothesis 2 focuses on the dynamic behavior evolution; Hypothesis 3 examines the adaptability across groups; and Hypothesis 4 investigates the underlying cognitive mechanisms. Together, they form a complete validation chain encompassing tool performance–behavioral outcomes–cognitive characteristics. The formulation of the hypotheses is grounded in theoretical foundations and empirical evidence. Hypothesis 1 is derived from the tool's mobile adaptability and lightweight design, positing superior efficiency, flexibility, and lower cognitive load relative to Kinect and professional systems. Hypothesis 2 draws from embodied cognition theory regarding the reshaping effects of technology on cognition, emphasizing the co-evolution of data-driven and experiential elements. Hypothesis 3 is motivated by the group differences identified in the requirements analysis, highlighting the universal value of modular design. Hypothesis 4 incorporates cognitive psychology mechanisms linking feedback and immersion, introducing gaze-path entropy as a quantitative indicator of cognitive complexity.

### 4.2 Experimental design

To systematically validate the hypotheses, a mixed experimental design was implemented, combining controlled variables with repeated measures to ensure the reliability and validity of the results. Sixty creators participated in the study, divided into professional and non-professional groups with thirty participants each. The professional group included creators from classical, modern, and ethnic dance backgrounds, with ten participants from each style and more than five years of choreographic experience. The non-professional group consisted of individuals with over one year of dance-learning experience but no choreographic background. Age and gender ratios were matched across groups to control for extraneous variables. A $2 \times 3 \times 2$ mixed design was adopted. Tool type served as a between-subjects factor with three levels: the intelligent tool developed in this study, Kinect v2, and the OptiTrack professional system.

Each tool group contained 20 participants with matched proportions of professional and non-professional creators. Creative duration served as a within-subjects factor with two levels (short-term and long-term). Creator type was a between-subjects factor, forming an interaction structure with tool type. The experimental procedure consisted of short-term and long-term stages. The short-term experiment required participants to complete a three-minute dance-segment composition task, focusing on assessing immediate tool effectiveness. The long-term experiment spanned four weeks, during which participants engaged in weekly sessions to develop a sequence of four progressively advanced works, enabling the tracking of behavioral evolution. A unified thematic cue—natural imagery—was used to ensure comparability across groups. The short-term task required the creation of a single thematic segment, whereas the long-term task involved a series of creations that gradually deepened across stages of imagery extraction, movement design, and stylistic integration. This experimental design allowed between-subjects comparisons to capture differences across tool types, while within-subjects measurements revealed longitudinal evolution. When combined with analyses of group-level differences, the design fully addressed the validation needs of the four hypotheses. Task standardization and theme consistency further controlled for confounding variables such as task difficulty and creative motivation.

## 5 EXPERIMENTAL RESULTS AND ANALYSIS

Figure 2 illustrates the accuracy differences among the three motion capture tools across an indoor rehearsal studio, an outdoor plaza, and a confined 10 $m^2$ space. The intelligent tool developed in this study achieved accuracy errors between 1.5 and 1.8 cm across all environments, with a fluctuation range of only 0.3 cm, indicating stable performance under mobile conditions. Kinect v2 exhibited errors between 3.2 and 3.5 cm, maintaining consistent variability but demonstrating substantially lower precision overall. The OptiTrack professional system delivered the highest accuracy, with errors between 0.7 and 0.9 cm; however, its physical footprint and deployment requirements prevented operation in outdoor settings and small confined spaces. In terms of functional responsiveness, the proposed intelligent tool maintained an average core function latency of 32–38 ms, while synchronization latency during eight-user collaborative creation remained below 50 ms, meeting the requirements of real-time interactive composition. Kinect v2 exhibited latency in the range of 72–80 ms, and OptiTrack maintained latency between 18 and 22 ms. Performance validation results further demonstrated the technical stability of the proposed intelligent tool. Across the public datasets Human3.6M and AIST++, the average capture errors were 1.5 cm and 1.8 cm, respectively, with deviations from reference data remaining within 0.3 cm.

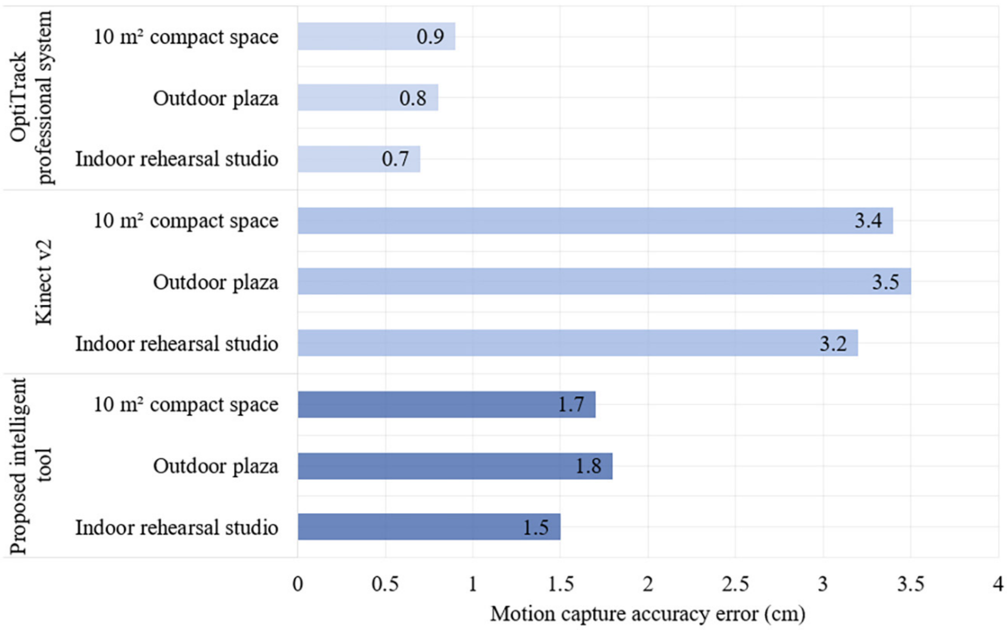


**Fig. 2.** Comparative accuracy error of motion capture tools across creative environments

Figure 3 demonstrates the dynamic evolution of creative behavior. For professional creators using the intelligent tool, the movement generation rate increased from 12 movements per minute in the short-term phase to 18 movements per minute by Week 4 of the long-term phase, while cognitive load (NASA–TLX score) decreased from 15 to 8. Non-professional creators exhibited a similar pattern, with movement generation rates increasing from 8 to 14 movements per minute and cognitive load decreasing from 18 to 12. In comparison, the Kinect v2 group showed only moderate improvements: professional creators' movement generation rates increased from 10 to 14 movements per minute, and cognitive load decreased from 17 to 14. These findings indicate that the intelligent tool more effectively facilitates a shift in creative cognition from experience dependence to a data–experience co-evolution model. Differences across creator types and dance genres were also observed. Professional creators demonstrated an 85% utilization rate of the style-matching module, compared with 60% among non-professional creators. Among dance genres, creators working in ethnic dance exhibited a 90% utilization rate of the cultural symbol extraction function, higher than that of classical dance (80%) and modern dance (75%). These patterns indicate that the tool's modular design effectively accommodates diverse functional needs across creator groups.

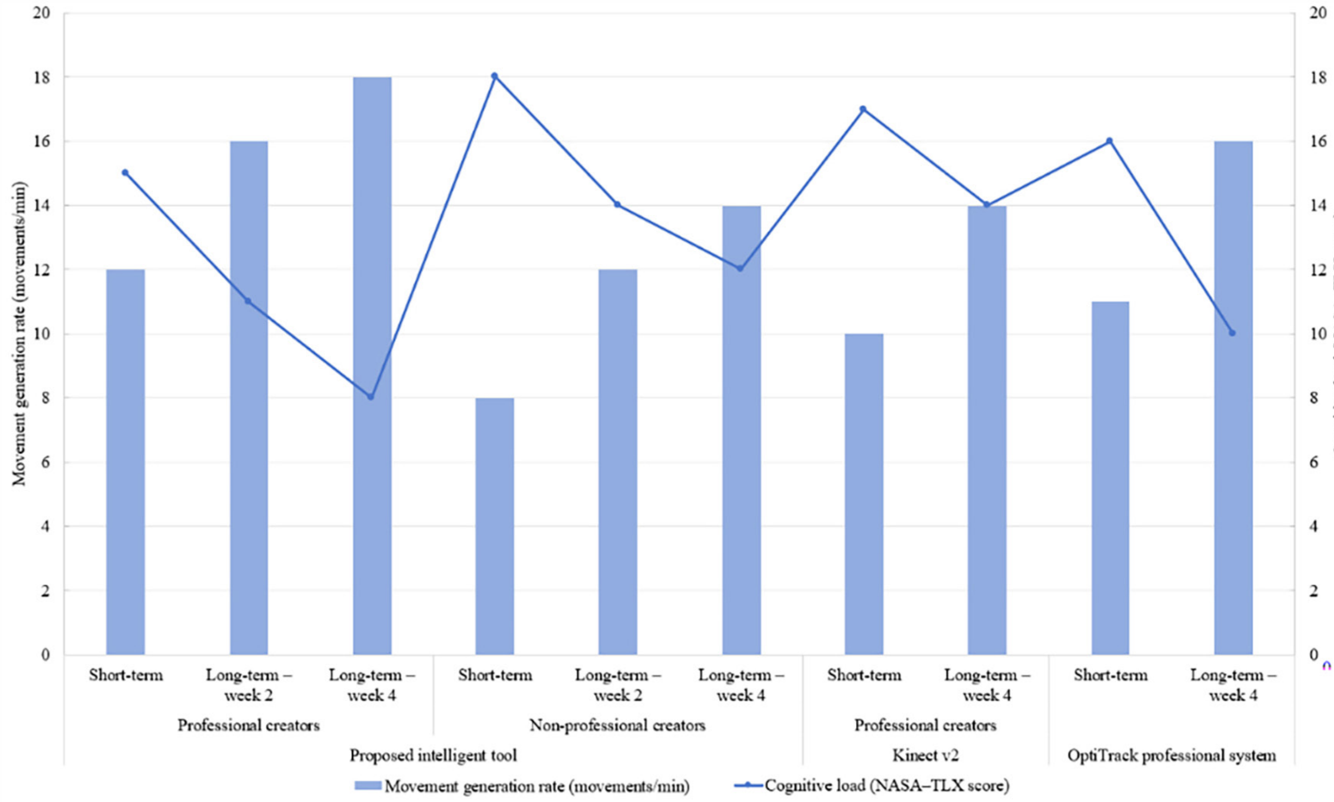


**Fig. 3.** Evolution of movement generation rate and cognitive load for professional and non-professional creators across tools

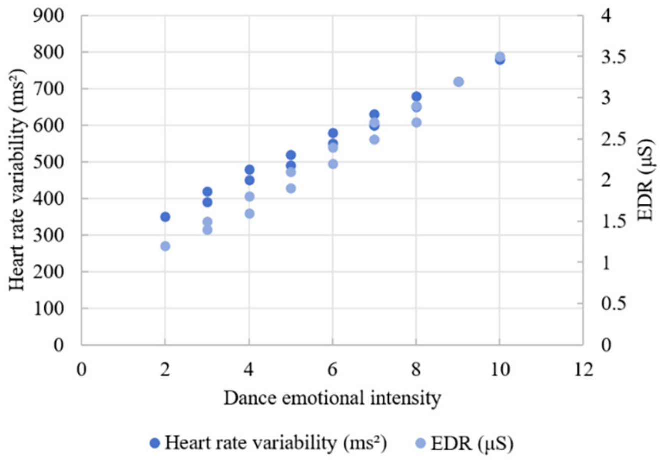


**Fig. 4.** Correlation distribution between dance emotional intensity and physiological indicators

As shown in Figure 4, a significant positive correlation was observed between emotion intensity and heart rate variability ($r = 0.68$, $p < 0.01$): as emotion intensity increased from 2 to 10, heart rate variability rose from 350 $ms^2$ to 780 $ms^2$. Electrodermal response (EDR) was also positively correlated with movement force ($r = 0.75$, $p < 0.01$), increasing from 1.2 μS to 3.5 μS as force ratings increased from 3 to 10. These results demonstrate the capacity of physiological indicators to serve as quantitative proxies for artistic expressive processes. Analysis of creative flow states showed that the intelligent tool group achieved a significantly higher flow score (8.2) than the Kinect v2 group (6.8), with no significant difference from the OptiTrack group (8.0). Creative flow scores were negatively correlated with cognitive load ($r = -0.72$, $p < 0.01$). Qualitative feedback further revealed that 90% of professional creators perceived no reduction in the uniqueness of artistic expression when using the tool, and 85% of non-professional creators reported that motion-assisted generation lowered the entry barrier to creative composition. These findings validate the tool's ability to balance technological augmentation with the preservation of artistic ontology.

**Table 3.** User experience and usability evaluation results

| Evaluation Dimension/ Tool Type | Proposed Intelligent Tool (Professional Creators) | Proposed Intelligent Tool (Non-Professional Creators) | Kinect v2 (Professional Creators) | Kinect v2 (Non-Professional Creators) | OptiTrack Professional System (Professional Creators) |
|---|---|---|---|---|---|
| Tool learning time (min) | 8 | 12 | 15 | 20 | 25 |
| Operational satisfaction (1–5) | 4.8 | 4.5 | 3.2 | 2.8 | 4.0 |
| Task error rate (%) | 5 | 8 | 12 | 18 | 6 |
| Core function usage rate (%) | 92 | 80 | 75 | 60 | 85 |
| Operational smoothness (1–5) | 4.7 | 4.3 | 3.0 | 2.5 | 4.2 |

To evaluate usability and creator acceptance under practical operational conditions, a user experience and usability assessment was conducted. As shown in Table 3, professional creators required only eight minutes to learn the intelligent tool, substantially shorter than the 15 minutes required for Kinect v2 and the 25 minutes required for the OptiTrack system. Non-professional users demonstrated similar advantages, requiring 12 minutes compared with 20 minutes for Kinect v2. Operational satisfaction scores for the intelligent tool reached 4.8 for professional creators and 4.5 for non-professional creators—significantly higher than the scores for Kinect v2 (3.2 and 2.8) and comparable to the OptiTrack system (4.0). Core function usage rate and operational smoothness exhibited similar superiority. These findings indicate that the tool's low-threshold design markedly enhanced acceptance across diverse creator groups. The usability advantages demonstrated relative to mainstream comparison tools provide experiential support for the tool's broader practical deployment.

## 6 CONCLUSION

This study addressed the central inquiry of integrating mobile interaction technologies with dance composition by constructing a triadic theoretical framework linking technology, the body, and cognition and by designing and implementing a mobile intelligent dance composition tool that integrates multimodal sensing with lightweight AI. Multidimensional experimental validation demonstrated the tool's technical effectiveness, creative value, and adaptability across diverse scenarios. As shown by the results, stable motion capture accuracy and low latency responsiveness were maintained across multiple creative environments; creative efficiency was substantially improved for both professional and non-professional creators;

and creative cognition was shifted from experience dependence toward a data–experience co-evolution model. Furthermore, the system achieved a balance between technological augmentation and the preservation of artistic expressiveness. Its practical effectiveness in dance education and digital heritage preservation further confirmed its potential to address core challenges within the field. This study fills a theoretical gap at the intersection of embodied cognition, mobile intelligence, and artistic creation, while also providing a transparent methodological framework for future implementation and comparison that may serve as a reference for subsequent interdisciplinary research on mobile-intelligent creative practices.

## 7 ACKNOWLEDGEMENTS

This paper was funded by the Key Project of the Ministry of Education in the "14th Five-Year Plan" for National Educational Research: Innovative Research on Ethnic Dance Teaching in Colleges and Universities (Grant No.: JYKY234208); and the 2025 Heze Social Science Planning Project "Research on the Ecological Aesthetic Imagery of the Shang Yang Dance" (Grant No.: 2025-ZZ-35).

## 8 REFERENCES

[1] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-Edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Network*, vol. 33, no. 5, pp. 156–165, 2019. https://doi.org/10.1109/MNET.2019.1800286

[2] L. Liu, "The impact of mobile applications on personalized learning paths in dance education," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 5, pp. 128–143, 2025. https://doi.org/10.3991/ijim.v19i05.54525

[3] M. del Carmen Rodríguez-Hernández and S. Ilarri, "AI-based mobile context-aware recommender systems from an information management perspective: Progress and directions," *Knowledge-Based Systems*, vol. 215, p. 106740, 2021. https://doi.org/10.1016/j.knosys.2021.106740

[4] K. Chen, Y. Zu, and D. Wang, "Design and implementation of intelligent creation platform based on artificial intelligence technology," *Journal of Computational Methods in Science and Engineering*, vol. 20, no. 4, pp. 1109–1126, 2020. https://doi.org/10.3233/JCM-204240

[5] S. K. Jha, B. H. Kumari, K. Anand, J. Kanjalkar, R. B. Gaddam, and S. Kumar, "Intelligent task prediction and partial computation offloading in mobile edge cloud computing," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 18, pp. 106–117, 2025. https://doi.org/10.3991/ijim.v19i18.57233

[6] S. M. C. Palpán and L. G. Castro, "Autonomy and dialogue during creation: A qualitative study on the experiences of creation in dance of a group of students," *Artseduca*, vol. 27, pp. 72–87, 2020.

[7] M. G. Valladares Gonzalez, "Artistic dance creation: One experience of interdisciplinary research," *Tercio Creciente*, no. 14, pp. 37–48, 2018.

[8] A. E. Yankovskaya and V. B. Obukhovskaya, "Foundations of creation of a complex of applied intelligent systems for diagnostics of psychological safety and cognitive sphere of patients with a neurological pathology," *Pattern Recognition and Image Analysis*, vol. 30, no. 4, pp. 741–747, 2020. https://doi.org/10.1134/S1054661820040252

[9] M. R. Nogueira, J. B. Simões, J. M. de Carvalho, and P. Menezes, "'Move in Tempo': Involving the audience through their movement in installation art," *International Journal of Arts and Technology*, vol. 15, no. 5, pp. 1–22, 2025. https://doi.org/10.1504/IJART.2025.146787

[10] L. Vuarnesson *et al.*, “Shared diminished reality: A new VR framework for the study of embodied intersubjectivity,” *Frontiers in Virtual Reality*, vol. 2, p. 646930, 2021. https://doi.org/10.3389/frvir.2021.646930
[11] K. M. Darda and E. S. Cross, “The computer, A choreographer? Aesthetic responses to randomly-generated dance choreography by a computer,” *Heliyon*, vol. 9, no. 1, p. e12750, 2023. https://doi.org/10.1016/j.heliyon.2022.e12750
[12] B. Paulchamy, A. Yahya, N. Chinnasamy, and K. Kasilingam, “Facial expression recognition through transfer learning: Integration of VGG16, ResNet, and AlexNet with a multiclass classifier,” *Acadlore Transactions on AI and Machine Learning*, vol. 4, no. 1, pp. 25–39, 2025. https://doi.org/10.56578/ataiml040103
[13] R. Namane, E. Boutellaa, S. E. Salem, and Y. Babaci, “A residual temporal convolutional with attention neural network for electromyogram-based hand gesture recognition,” *International Journal of Computational Methods and Experimental Measurements*, vol. 13, no. 3, pp. 739–748, 2025. https://doi.org/10.56578/ijcmem130320
[14] A. O. Shabalina, “Two meanings of ‘Cognition’ in the theory of embodied cognition,” *Tomsk State University Journal*, vol. 463, pp. 69–72, 2021. https://doi.org/10.17223/15617793/463/9
[15] M. Yang, J. Amankwah-Amoah, and H. Chunjia, “Exploring human-machine collaboration in metaverse communities: Product vs. service focus in enhancing user immersion,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 204, p. 104423, 2025. https://doi.org/10.1016/j.tre.2025.104423
[16] G. Baudoux, “The benefits and challenges of artificial intelligence image generators for architectural ideation: Study of an alternative human-machine co-creation exchange based on sketch recognition,” *International Journal of Architectural Computing*, vol. 22, no. 2, pp. 201–215, 2024. https://doi.org/10.1177/14780771241253438
[17] Z. Liu, N. Nersessian, and J. Stasko, “Distributed cognition as a theoretical framework for information visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1173–1180, 2008. https://doi.org/10.1109/TVCG.2008.121
[18] M. Bernini, “Supersizing narrative theory: On intention, material agency, and extended mind-workers,” *Style*, vol. 48, no. 3, pp. 349–366, 2014. https://doi.org/10.5325/style.48.3.0349

## 9 AUTHOR

**Yanjun Jiang** graduated from Jiangsu Normal University in 2016 and currently serves as a faculty member at the School of Music and Dance, Heze University, Heze 274015, China, specializing in research on Chinese ethnic folk dances and the appreciation of Chinese and foreign dance works (E-mail: wlxjjj@126.com).