# DanceFlow: Design and Preliminary Implementation of an End-to-End Intelligent Dance Assistance System Based on a Three-Layer Hybrid Architecture and Three-Layer Six-Dimensional Evaluation

Li Jie

2026-07-08

ChinaRxiv

## Abstract

The author is a frontline teacher who has taught dance for many years. In recent years, an increasing number of intelligent choreography tools have emerged, yet few have truly been useful in the classroom: most focus only on "generating movements," while no one tells us "whether the dance is good, and where it is good." Small intangible cultural heritage dance forms also suffer from a lack of data, making it unfeasible to rely solely on stacking models. In response to these real classroom concerns, the author attempted to develop an auxiliary tool that both the author and students could use. Since the author cannot write code, an AI programming assistant was brought in to collaborate: the author proposed requirements, the assistant wrote the code, and the author verified whether the results were correct. Technically, Labanotation was used as a "translator" to convert movement qualities into machine-readable parameters; the general skeletal format SMPL was used as an intermediary so that "evaluation" and "generation" could speak the same language. Gradually, preliminary frameworks of a "three-layer hybrid architecture" and "three-layer six-dimensional evaluation" were established. The main body of the DanceFlow v1.0 desktop version has already been developed, including modules such as a visual workflow with 11 functional nodes, three-layer six-dimensional expressiveness evaluation, multi-agent collaboration, a generation-evaluation-optimization closed-loop prototype, an intangible cultural heritage dance knowledge base, 3D motion preview, formation choreography, style locking, role consistency, and video export. The reverse mapping module for Laban effort has become basically stable after seven rounds of refinement and has been tested in daily teaching and intangible cultural heritage effort analysis. Due to the lack of a local GPU, some deep learning modules are temporarily replaced by cloud APIs. In the forward mapping, the two factors of "space" and "flow" are still not very obedient; the aesthetic layer currently only has a framework and lacks sufficient data support; the overall validation scale is relatively small, and the conclusions are more of a directional sense derived from the author' s teaching practice. Systematic experimental data are still being collected and will be supplemented and updated in subsequent versions. The path of "three-layer hybrid architecture + three-layer six-dimensional evaluation + bidirectional Laban mapping" has proven workable in the author' s own teaching. Starting from the real needs of dance teachers to build AI tools and pursuing interpretable collaboration in which "what humans can understand, and what machines cannot explain clearly, is supplemented by Laban" may be a practical route for AI to enter the dance classroom. This article is a staged record of the author' s practice, offered as a modest contribution in the hope of receiving correction and advice from colleagues.

Full Text

# DanceFlow: Design and Preliminary Implementation of an End-to-End Intelligent Dance Assistance System Based on a Three-Layer Hybrid Architecture and Three-Layer Six-Dimensional Evaluation

**Li Jie**

School of Music, Huizhou University, Huizhou, Guangdong 516007, China

ORCID: https://orcid.org/0009-0004-5869-2035

**Abstract:**

The author is a frontline dance teacher with many years of teaching experience. In recent years, intelligent choreography tools have become increasingly numerous, yet very few can truly help in the classroom—most only focus on "generating movements," while no one tells us "whether the dance is good, and where it is good." For small-scale intangible cultural heritage dance genres, data are also scarce, and relying purely on stacking models is simply not feasible. With these real classroom concerns in mind, the author attempted to work out a set of assistive tools that both the author and students could use. The author does not write code, so an AI programming assistant was brought in to work together: the author proposed requirements, it wrote the code, and the author verified whether the results were correct. In terms of the technical route, Labanotation was used as a "translator," converting movement qualities into parameters that machines can understand; a general skeleton format such as SMPL was used as an intermediary, enabling "evaluation" and "generation" to speak the same language. Gradually, a preliminary framework of a "three-layer hybrid architecture" and "three-layer six-dimensional evaluation" was built. The main body of the DanceFlow v1.0 desktop version has already been completed: it includes a visual workflow with 11 functional nodes, three-layer six-dimensional expressiveness evaluation, multi-agent collaboration, a generation–evaluation–optimization closed-loop prototype, an intangible cultural heritage dance knowledge base, 3D motion preview, formation arrangement, style locking, character consistency, video export, and other modules; the Laban effort reverse-mapping module has become basically stable after seven rounds of refinement; and it has already been trialed in daily teaching and in the analysis of effort qualities in intangible cultural heritage dance. Due to the lack of a local graphics card, some deep-learning modules are temporarily replaced by cloud APIs. In the forward mapping, the two factors of "space" and "flow" are still not very obedient; the aesthetic layer is currently only a framework and lacks sufficient data support; the overall validation scale is relatively small, and the conclusions are more of a directional sense felt out by the author in teaching; systematic experimental data are still being collected and will be gradually supplemented and updated in subsequent versions. The path of "three-layer hybrid architecture + three-layer six-dimensional evaluation + Laban bidirectional mapping" has proved workable in the author's own teaching. Developing AI tools from the real needs of dance teachers, and pursuing an explainable collaboration in which "what humans can understand, but machines cannot clearly express, is supplemented by Laban," may be a practical route for bringing AI into the dance classroom. This article is a staged record of the author's practice during this period; it is offered as a modest beginning in the hope of receiving corrections from colleagues.

**Keywords:** intelligent choreography; dance expressiveness; Laban Movement Analysis; three-layer hybrid architecture; three-layer six-dimensional evaluation; digitization of intangible cultural heritage dance; Labanotation; BESS framework

**Classification numbers:** TP391; J70-05

# DanceFlow: Design and Preliminary Implementation of a Full-Pipeline Intelligent Dance Assistance System Based on a Three-Layer Hybrid Architecture and a Three-Level Six-Dimension Evaluation

Li Jie

(School of Music, Huizhou University, Huizhou 516007, China)

**Abstract:**

In the current field of AI dance, generation tools are developing rapidly, while evaluation tools remain relatively scarce, creating a practical dilemma of "able to generate but difficult to judge quality." Meanwhile, niche dance forms such as intangible cultural heritage (ICH) dances suffer from data scarcity, making purely data-driven approaches difficult to apply directly in teaching scenarios. As a front-line dance teacher, the author has deeply felt these pain points in teaching and project practice, and therefore attempts to design a technical roadmap for an interpretable intelligent dance-assisted creation system. Starting from the current needs of dance teaching, with Laban Movement Analysis as the semantic bridge and the SMPL 3D skeleton as the motion intermediate representation, a three-layer hybrid architecture and a preliminary framework of three-level six-dimensional evaluation are tentatively built, seeking a balance between semantic interpretability and physical consistency. Adopting the approach of "few-shot + high-quality effort annotation," the paper explores feasible paths for the digitalization of ICH dances. The author, a dance teacher without programming skills, proposed the functional requirements and verified the results, while the actual coding was completed by AI programming assistants under the author' s guidance. The main development of the DanceFlow v1.0 desktop application has been preliminarily completed: including a visual workflow frontend with 11 functional nodes, a three-level six-dimension dance expressiveness evaluation framework, a multi-agent collaboration module, a generation-evaluation-optimization closed-loop prototype, a dance knowledge base (including a dedicated ICH dance section), 3D motion preview, a formation choreography editor, style locking, character consistency control, and a video export pipeline. The Laban effort inverse mapping module has basically stabilized after seven rounds of iteration. The system has been trialed in daily teaching assistance and ICH dance effort analysis. Limited by the lack of a local GPU environment, deep learning modules such as forward mapping diffusion

model training and end-to-end video generation have not yet been fully developed, and cloud API solutions are used as alternatives. The controllability of spatial and flow factors in forward mapping still requires improvement. The aesthetic layer is currently more framework-oriented and lacks sufficient benchmark data support. The overall scale of validation is small, and the conclusions are more directional impressions derived from personal practice. The technical route of a "three-layer hybrid architecture + three-level six-dimensional evaluation + Laban bidirectional mapping" is feasible in practice. Designing AI tools based on the real needs of dance teaching and pursuing an interpretable human-machine collaboration approach may be a viable direction for bringing AI into dance classrooms. This paper is more a phased summary of personal practice, offered as a modest contribution, and it welcomes criticism and corrections from peers.

**Keywords:** intelligent choreography; dance expressiveness; Laban movement analysis; three-layer hybrid architecture; three-level six-dimensional evaluation; ICH dance digitalization; Labanotation; BESS framework

## I. Research Background

The wave of AI has arrived with great force. From ChatGPT and DeepSeek to today' s task-performing agents such as OpenClaw that can directly get work done, development is advancing almost by the day. The impact on educators is especially strong: when knowledge can be searched everywhere and AI is readily available, the classroom question has shifted from "Will AI come?" to "Now that AI is here, how should classes be taught?" The author comes from a dance background and does not have independent programming ability, yet has long been thinking about how AI can be used in dance, how it can support teaching, how it can assist creation, and how it can improve efficiency. If one does not want to be left behind by the times, the only option is to embrace it.

Thinking alone is of no use; only by learning while doing can one find a direction. For reasons of network and data security, the author mainly used tools from major domestic companies—ByteDance Trae, Tencent WorkBuddy, Alibaba qoder, and Baidu dumate—and essentially tried them all. The pitfalls encountered are more instructive than the tools themselves: free quotas are insufficient, and complex tasks cause them to stall; paid use is too expensive, and building a teaching system can cost over one hundred yuan; stable tools have few functions, while feature-rich tools come with piles of bugs, and disconnections and lost records are common; even more troubling are agents' "correct but empty talk," "serious-sounding fabrication," and "hallucinations" —common problems for heavy users. These forced the author to cross-check with several tools, making the efficiency pitifully low.

The earliest project was ThinkFar (https://gitee.com/lihaoshenjing/ThinkFar), in which the ideas of a "data flywheel" and "cyclical iteration" were later applied directly to the dance system. After turning to dance, the hands-on work began

DanceFlow (main repository: https://gitee.com/lihaoshenjing/danceflow; demo version: https://gitee.com/lihaoshenjing/danceflow-demo) set out to build a one-stop AI choreography tool covering topic selection, scripts, storyboarding, movement, music, formation, and video. After more than a dozen rounds of functional design and over a month of polishing, 11 core nodes were ultimately implemented. Yet once it could actually be used, its flaws became obvious: movements could be generated, 3D could be previewed, and videos could be exported, but the results were not good; instead of improving efficiency, it created more disorder. This, in turn, forced the author' s working method into shape: use whatever runnable version was available, judge what was not easy to use, then feed the problems back to the AI for revision and improvement. Every step forward in the system was "used" into existence in this way, rather than fully conceived in advance. A demo version was therefore rebuilt and submitted in July 2026 as an entry to the preliminary-round area of ByteDance TRAE AI Creativity Competition' s "Learning and Work Track" [1]. In web form, it integrated modules such as intelligent topic selection, storyboarding, 3D movement generation, visual synthesis, dance analysis, AI sparring, and digitization of intangible cultural heritage. It was intended as a lightweight entry point for non-technical users and educational scenarios. (Taking it to the competition was also a way to force myself to refine the system through a real teaching scenario —the eyes of judges and peers are better at spotting problems than closed-door revision.)

To confirm that the "difficulty of evaluation" was not merely a niche illusion, the author systematically reviewed the current state of dance-AI products and research. On the product side, there are two main paths: motion capture (using professional equipment to record human movement accurately) and pose estimation (using AI to infer a motion skeleton from ordinary video) (overseas examples include Move.ai, DeepMotion, and RADiCAL; domestic examples include Dance AI, Zhiwufei Young), as well as dance generation (ByteDance's Jimeng Seedance, Kuaishou Kling, MACE-Dance developed in collaboration by Alibaba with Renmin University of China and Tsinghua University, etc.). On the research side,

1

Beijing Dance Academy, the Chinese National Academy of Arts, and Shanghai Theatre Academy have mostly focused on digital preservation, motion-capture-assisted teaching, and the datafication of intangible cultural heritage, while the broader community has largely revolved around generation and motion capture (such as SteadyDancer and MusePose, launched by Nanjing University in collaboration with Tencent PCG). In the end, it became clear that, whether in industry or academia, "evaluation of dance expressiveness," "aesthetic analysis," and "judgment of artistic quality" are almost entirely blank areas. This is not because they are unimportant, but because they are the hardest: motion capture is an engineering problem, generation is a matter of effect, whereas evaluation of expressiveness is aesthetic judgment. Defining what is "good," quantifying what is "moving," and explaining what has "flavor" easily becomes metaphysical, so most people detour around it. But the generation side has already matured enough that the blank on the evaluation side is especially conspicuous.[2]

Interestingly, several key upgrades of this system were not conceived while sitting in the study, but were forced out by the project. At first, I only wanted to build a hierarchical evaluation framework; the "three-layer, six-dimensional parametric architecture" was gradually polished during the process of building DanceFlow. The architecture was truly clarified only after work related to digital asset modeling (DAMO) and after participating in the TRAE competition. Initially, I tried "fully automatic generation," feeding Laban parameters to the model and expecting it to output

corresponding texture; as a result, when the dynamics were right, the logic of the movement was wrong, and when the movement became smooth, the texture drifted away. Later, it became clear: Laban is an "evaluative language," not a "generative language"—it is good at diagnosing "what kind of texture this is," but not at directly specifying "which action to perform." Generation still depends on kinematic models; Laban is responsible for "controlling direction" and "diagnosing problems," just as a doctor first makes a diagnosis and then prescribes medicine. Once this turn was made, SMPL could handle generation (physical correctness), Laban could handle semantic control and evaluation (whether the texture is right), and the two-way mapping and translation between them could finally make the closed loop of "generation—evaluation—optimization" truly operate.

The turning point came from Labanotation: the four effort factors quantify "movement texture"—whether it is heavy depends on the weight factor, whether it is crisp depends on the time factor, whether it is floating depends on the space factor, and whether it is contained depends on the flow factor. This led to the idea: when data are insufficient, supplement them with theory; when AI does not understand texture, first translate texture into parameters it can understand. By making Labanotation into an "intermediate semantic layer" for AI, when a choreographer says "make it heavier," this can be mapped directly to increasing the weight factor by 0.2; when they say "make it crisper," this corresponds to moving the time factor toward the sudden end, giving the

2

system something to grasp. First, movements generated purely by technology may be "accurate in rhythm and free of joint jitter," yet professional students still say, "the movements are all correct, but they don' t look like dancing" — the quality of dance is a question of "whether it has enough" (texture, spirit, flavor), rather than "whether it is correct," while AI has happened to learn only the layer of "correct or incorrect." Second, when generating according to choreographic language, expressions such as "heavier" and "crisper" cannot be mapped to specific parameters; the control language of AI (joint angles, trajectories) does not align with the choreographer' s language (heaviness/crispness, hardness/softness, looseness/tightness). Third, the digitization of intangible cultural heritage dance is the most difficult: publicly available professional motion-capture data are almost nonexistent; self-collecting a professional-grade motion-capture segment with annotation costs several thousand yuan, and dozens of segments for one dance genre would amount to hundreds of thousands. A purely data-driven approach simply does not work here.

The core issue is no longer "whether there are tools," but "how tools enter real scenarios and truly solve problems." Most existing tools are made by technical people for technical people and are disconnected from the workflows of dance practitioners; their usage costs (hundreds of thousands of yuan for high-precision motion capture, annual fees for large-model APIs [interfaces for calling large models]) and learning thresholds (parameter tuning, writing prompts, handling errors) keep most people out. Even more fatal is the reliability of large models in professional domains: hallucinations (misapplying terminology), contextual amnesia (contradictions before and after analyzing long works), and reduced efficiency ( "one-click generation" in fact requiring every node to be waited for, adjusted, and revised, making it slower than a human). Fundamentally, this stems from a mistaken positioning of AI—in the field of dance, AI cannot replace the choreographer' s aesthetics, the teacher' s experience, or the dancer' s bodily perception. Its proper role is assistance rather than replacement: AI helps with data organization, pattern recognition, and large-scale preliminary screening, while humans take charge of aesthetic judgment, creative decision-making, and quality control. This evaluation framework is therefore also the "access-

… "interface," and it must be low-cost—layered computation, going deeper only as needed, and avoiding unnecessary full-scale computation. This is both more reasonable and faster, and it can run on ordinary devices.

## II. Research Status

Intelligent dance generation and evaluation has been one of the fastest-developing directions in artificial intelligence in recent years. In simple terms, it means enabling computers to "understand" human movements and then "choreograph" new dance movements. Its development has roughly gone through three stages: first, motion capture was used to record real dancers and store the dances in a general-purpose "digital skeleton" format; later, rules were used to score movements; now, deep learning directly "learns" to generate them, with steadily improving quality. This "generation" line has already become quite mature—from foundational work on large-scale motion-language datasets and text-driven generation [3][4][5], to the emergence in recent years of music-driven generation and beat-alignment methods.

On the "generation" side, recent methods each show their own strengths: some focus on the objective quality of "whether it looks like a real person dancing," some pursue motion diversity and controllable style, and others make music-beat alignment more natural [6][7][8]. It can be said that the question of "whether a plausible-looking dance segment can be generated" has basically been solved. [9][10][11] But "generation" and "evaluation" have long remained two separate tracks. Motion quality has objective indicators that can be measured—for example, how human-like the motion is and how diverse it is; some researchers have also specifically used vision-language models to score motion quality [12][13][14]. Yet once the question becomes "whether this dance is expressive enough, whether the emotion is right, and whether it is artistically good," existing methods can hardly answer. The systematic review by Wirdnam et al. points out that evaluating dance expressiveness and artistic quality remains an open challenge to this day [15]. Another highly enlightening route is to use Laban Movement Analysis (a symbolic system for recording "movement texture" ) to quantify such hard-to-verbalize dance sensations as "force, time, space, and flow" [16][17][18]. It has already been applied to folk-dance evaluation [17], emotion recognition [19], and dance-genre recognition [20], and dance academies in China have also produced systematic reviews [21]. However, most work remains at the level of "recognition and classification," failing to connect "texture" back to "generation" and "evaluation" so as to form a closed loop.

For intangible-cultural-heritage dance, digitization has mainly focused on "storing movements, building archives, and generating pose variants" [22][23], with little attention to quantifying the "inner texture" within the movements. Many intangible-heritage movements have had their "form" preserved, but not their "spirit."

| Research Direction | Representative Work | Main Authors and Institutions | Main Progress | Unresolved Issues / Entry Point of This Paper |
|---|---|---|---|---|
| Intelligent choreography generation | Li and Huang [12], Qu et al. [13], vision-language model evaluation [14], SoulDance, EDGE, MACE-Dance, etc. | Li H & Huang X (Sensors); Qu J (Sci. Reports); Monte Freitas et al. (CVPR Workshop); Yang K et al. (arXiv) | From rule-based scoring to learning-based generation; movement quality continues to improve | Most work remains at the objective distance metric of "looks like / does not look like," lacking a framework for expressiveness, style, and aesthetics |
| Evaluation of dance expressiveness | Systematic review by Wirdnam et al. [15]; objective metrics such as FID/BAS/DIV | Wirdnam M et al. (JDMS) | Objective distance metrics are relatively mature | Judgments of "expressiveness, aesthetics, and artistic quality" are almost entirely absent |
| Laban movement analysis | Laban [16], Aristidou et al. [17], Tsachor and Shafir [18], Turab et al. [19], Zhou Yu et al. [21] | Laban R (monograph); Aristidou A et al. (ACM JCCH); Tsachor R P & Shafir T (Front. Psychol.); Turab M et al. (arXiv); Zhou Yu et al. (Journal of Beijing Dance Academy) | Quantifies "movement texture" for folk-dance evaluation and emotion recognition | Mostly remains at recognition and classification, and has not yet formed a "semantics—generation—evaluation" closed loop |

| Research Direction | Representative Work | Main Authors and Institutions | Main Progress | Unresolved Issues / Entry Point of This Paper |
|---|---|---|---|---|
| Digitization of intangible cultural-heritage dance | Pataranutaporn and Klunchun [22], Guo Hong et al. [23] | Pataranutaporn P & Klunchun P (ACM MOCO); Guo Hong et al. (Sports Science) | Collection, archiving, and generation of pose variations | Rarely touches the quantification of "inner texture" ; mostly remains at the level of form |

Putting these three lines together, the problem becomes very clear: the generation side is already quite capable, and objective evaluation also has available methods; only the level of "expressiveness, aesthetics, and artistic quality" is almost a shared blank. Meanwhile, although the Laban system can quantify texture, it has not been connected into a closed loop of "semantics—generation—evaluation." This is precisely the gap this paper seeks to fill—DanceFlow does not build yet another generation model, but instead stands on the shoulders of predecessors, using Laban Effort as a "translator" to turn the dance teacher' s inner sense that "something feels wrong" into something the system can revise and the teacher can verify.

## III. Architecture Design and Methods

First, let us clarify the author' s role: the author is a dance teacher who does not know how to write code. In this system, the author does only three things—raise requirements, set standards, and test the loop; the actual coding and model tuning are handed over to the AI programming assistant. Below, in plain language, I explain the author' s design logic: matters of semantics are handed to Laban (which understands the "texture" of movement), matters of motion are handed to the SMPL digital skeleton, and evaluation is divided into three layers, connecting them with a closed loop so that the results improve over time. The author did not use the kind of end-to-end approach that "does everything in one step" —it is like a black box, and what happens in the middle is completely unclear;

What dance creation values most is explainability, controllability, and small-data operation. The shift to a hybrid architecture is precisely to regain these three qualities. Following a three-layer division of labor— "semantics—motion—presentation" —the author proposed the design requirements, which the AI then implements accordingly (see Figure 1): the Laban semantic layer is the "brain," responsible for unifying understanding and control language; the SMPL generation layer is the "body" [24], transforming Laban parameters into physically consistent 3D motion and using bidirectional mapping to understand Laban instructions; and the rendering layer is the "display window."

**Figure 1. Overall architecture of DanceFlow (input → three-layer hybrid architecture → 11-node workflow → output → evaluation closed loop)**

**DanceFlow Overall Architecture**

1. The teacher states requirements in plain language: "sink a little more," "make it crisper," "choreograph a Yingge dance segment."

**Three-layer hybrid architecture**

- **Semantic layer · Laban ( "brain" )**
  Understands the "quality" of movement, translates the teacher' s language into machine-understandable instructions, and unifies the command language.
- **Generation layer · SMPL ( "body" )**
  Converts instructions into "physically correct" 3D motion; understands Laban through bidirectional mapping.
- **Rendering layer ( "display window" )**
  Turns motion into viewable images: lightweight preview and high-definition final output.

**11 functional nodes** (like an assembly line; the teacher can simply drag and drop)

Topic selection · Script · Storyboard · Music · Rhythm synchronization · Motion recognition · Formation choreography · Knowledge retrieval · Preview · Video export · Viral-hit analysis

2. **Outputs:** 3D preview · video · effort curve
3. The three-layer, six-dimensional evaluation reports back to the system "what is wrong and how to revise it," forming a "generation → evaluation → revision" closed loop (see Figure 2).

The horizontal layered architecture combines with a vertical quality scale to form an initial closed-loop pattern in which "generation has direction, evaluation has hierarchy, and optimization has evidence." A prototype of the choreographer-oriented generation–evaluation–optimization loop has been implemented, while the other two closed loops remain at the preliminary framework stage. Three core principles ensure that the system is interpretable: embedded interpretability (the semantics of every layer and every node are clear); progressive semantic layering (abstraction increases while determinacy decreases); and bidirectional mapping throughout (Laban–SMPL bidirectional translation serves as the technical pivot). The output of any layer is interpretable and traceable. DanceFlow is based on the complete BESS framework [21], with Effort as the semantic pivot. Laban Effort—the four dimensions used in Laban theory to describe the "quality" of movement—includes the four factors of weight, time, space, and flow [16]. Each factor has polar attributes, whose combinations produce highly varied movement qualities; the polar combinations of the first three factors form the eight basic effort actions, while flow affects overall tension, release, and containment. (See Table 1)

**Table 1. Computational Definitions of the Four Elements of Laban Effort**

| Element name | Meaning | Bipolar dimension | Corresponding movement characteristics |
|---|---|---|---|
| Weight (gravity) | Magnitude of movement force and sense of weight | Light Heavy | Light: high center of gravity, elastic exertion of force; Heavy: lowered center of gravity, solid and substantial |
| Time (time) | Tempo and temporal quality | Sudden Sustained | Sudden: rapid increase in acceleration, with a sense of impact; Sustained: smooth change, continuous and stable |
| Space (space) | Spatial path and directional intention | Direct Indirect | Direct: short path, clear target; Indirect: winding path, rich use of space |
| Flow (flow) | Sense of control and degree of release in the flow of energy | Bound Free | Bound: strong control, compact rhythm, movements with marked pauses; Free: smooth and continuous, unconstrained, energy projected outward |

The author proposes the need for reverse mapping—automatically "reading" the effort qualities of any performed movement. The AI implements this algorithmically by extracting movement data from joint coordinates and mapping the factors of effort quality: weight corresponds to the vertical motion of the center of gravity, time corresponds to the rate of change of acceleration, space corresponds to the amplitude of bodily movement, and flow corresponds to the smoothness of motion. On this basis, effort-quality types can then be identified. The significance is that the author can independently analyze any SMPL movement without having to rely on others. Forward mapping (Laban→SMPL) is the other half of the requirement proposed by the author: given curves of effort-quality parameters, corresponding stylistic movements are generated, allowing the AI, during generation, to continuously regulate the texture parameters. This is currently being advanced and implemented by the AI, with reverse mapping used during training as a closed loop for self-verification. Evaluation is not a simple weighted sum, but a dynamic process that advances layer by layer, with lower layers supporting higher layers. (See Figure 2 and Table 2.) The three-layer, six-dimensional system refers to a three-layer structure—objective quality layer L1–L2, Laban semantic layer L3–L4, and artistic-aesthetic layer L5–L6—with each layer containing two dimensions, for a total of six evaluation dimensions.

**Figure 2. Three-layer, six-dimensional evaluation system (objective quality→Laban semantics→artistic aesthetics + closed-loop improvement)**

**Three-Layer, Six-Dimensional Evaluation System**

**First layer · Objective quality**

Is the technique good?
Joints / speed / rhythm
Like a physical-examination report

↓

**Second layer · Laban semantics**

Is the texture right?
Four effort elements + five-dimensional expressiveness
Like a teacher's felt impression

↓

**Third layer · Artistic aesthetics**

Does the artistry work?
Structure / climax / *qi-yun*
Like a judge's eye

The higher up, the more it emphasizes feeling.

**Closed loop (self-improvement)**

Generation → three-layer, six-dimensional evaluation → locate problems and provide revision methods → regeneration

The core logic progresses layer by layer from the bottom upward: the conclusion at each layer can be traced downward to concrete data; the emphasis of the output is "problem localization + direction for improvement" —for every deduction, it makes clear where the deduction occurs, why it occurs, and how to revise it.

Note: objective indicators are the "raw materials," while the three-layer, six-dimensional system is the "cooking method" —turning isolated numbers into meaningful diagnostic evidence.

**Table 2 Structure of the Three-Layer, Six-Dimensional Evaluation System**

| Layer | Dimension | Core Question | Calculation Method | Output Form |
|---|---|---|---|---|
| Third layer: Artistic aesthetics | L5 Formal structure (rhythm/climax/contrast)L6 Aesthetic concepts (Gestalt/form-spirit/*qi-yun*) | Does the artistry work? | Structural analysis + aesthetic judgment | Structured analysis + directional suggestions |
| Second layer: Laban semantics | L3 Laban semantic analysis (BESS framework + effort types)L4 Expressiveness scoring (5 dimensions) | Is the texture right? | SMPL → Laban mapping + effort → expressiveness mapping | Itemized scores + effort diagnosis + problem localization |
| First layer: Objective quality | L1 Kinematic/geometric featuresL2 Standardized indicators (FID/DIV/BAS) | Is the technique good? | Purely objective computation, distance measurement | Numerical indicators + quality-detection results |

The core logic proceeds from the bottom upward, becoming abstract layer by layer, and the conclusion of each layer can be traced downward to concrete data; the emphasis of the output is "problem localization + direction for improvement" —for every deduction, it makes clear where the deduction occurs, why it occurs, and how to revise it. It follows "small samples + high-quality effort annotation" : dual-channel collection (professional motion capture + pose estimation from public videos), four-layer annotation (effort curves → expressiveness scores → cultural labels), and human–machine collaboration (automatic pre-annotation + professional verification).

(See Table 3)

**Table 3. Framework of the Effort Annotation System for Intangible-Cultural-Heritage Dance**

| Annotation level | Annotation content | Output form | Annotation method |
|---|---|---|---|
| L1 Basic data | Human keypoint coordinates; SMPL parameters | Motion-data files | Automatically obtained through motion capture / pose estimation |
| L2 Effort annotation | Time curves of the four Effort elements; BESS multidimensional analysis; segmentation of Effort types | Effort time series + multidimensional analysis report | Automatic pre-annotation + manual verification |
| L3 Expressiveness annotation | Five-dimensional expressiveness scores; itemized comments | Score sheet + text | Annotation by professionals |
| L4 Cultural annotation | Dance-genre style; cultural connotations; movement functions | Metadata tags | Joint annotation by cultural scholars and dancers |

The intangible-cultural-heritage dance genres currently being built include representative Guangdong projects such as Qilin Dance, Yingge Dance, Xizi Lion, and Fire-Dog Dance.

"Evaluation first, gradual integration, and implementation in three stages": first, build a solid evaluation module—the computation of Laban Effort plus a three-layer, six-dimensional evaluation system (L1–L6); then carry out system integration, including a visual workflow, multi-agent collaboration, a knowledge base, and 3D preview; finally, move into deep generation, temporarily using cloud APIs because of the lack of GPU resources. The principle of "evaluation first" enables the system to enter the classroom and provide assistance even when the generation side is not yet mature.

(1) Laban Effort inverse-mapping module (core; the author and AI iterated through seven rounds): the author defined what it should do—automatically convert SMPL pose sequences into time curves of the four Effort elements, and extract accompanying BESS features; AI implemented it accordingly (777 lines of core code, `evaluator/core/laban_engine.py`). It is now used for intangible-cultural-heritage Effort analysis and for evaluating the quality of instructional movements; whether it is useful is verified by the author.

(2) Three-layer, six-dimensional evaluation framework (completed): the author designed its output to follow a three-part structure of "score + diagnosis + recommendation," with every point deduction clearly justified. Based on inverse mapping, AI built and implemented the three-layer, six-dimensional pipeline (L1–L6); the accompanying parameter tuner, HTML visual reports, and batch video-comparison tool were also implemented by AI.

(3) Backend-service infrastructure (completed): the author required the system to provide services externally, and AI implemented a unified external API that includes modules such as knowledge retrieval, dance analysis, and video generation.

The author planned an 11-node workflow: topic selection, script, storyboard, music creation, rhythm synchronization, movement recognition, formation choreography, knowledge retrieval, preview, video export, and viral-hit analysis. Using a flowchart-style tool, AI constructed it as a node-based visual workflow (see Architecture Figure 1). The accompanying project management, 3D preview (MMD virtual avatar), multi-agent collaboration, dance knowledge base (including topics on intangible cultural heritage), and formation-choreography editing

···tool, style locking, character-consistency control, and the video-export pipeline were also specified by the author and implemented by AI; the main body has been completed. The desktop application has completed basic packaging, with single-file distribution and one-click launch. The full front end has already been deployed on the local-area network of a teaching institution for trial use, and students can access it with ordinary laptops.

**Forward mapping (in progress):** The author proposed that it should be able to perform "texture-based control," so that, during AI generation, texture parameters are used to regulate motion data such as speed and center-of-gravity trajectories. During training, inverse mapping is used to form a closed loop for self-checking by the author. Because of the lack of GPU resources, the system mainly relies on cloud computing power. End-to-end video generation adopts a cascaded-expert scheme. For deeper development of the aesthetic layer, more aesthetic research and large-scale annotation are planned.

DanceFlow is an open-source project: main repository (complete system) https://gitee.com/lihaoshenjing/danceflow, demo-version repository

https://gitee.com/lihaoshenjing/danceflow-demo, public trial address (web page)

https://danceflow-demo-1421531034.cos-website.ap-guangzhou.myqcloud.com/. The demo version has been submitted as a competition entry to the preliminary-round section of the ByteDance TRAE AI Creativity Competition[1]. Secondary development and collaborative research based on this architecture are welcome.

Verified as usable: Laban Effort inverse mapping (7 rounds of iteration), the basic version of the three-layer, six-dimensional evaluation framework, the visual workflow front end (11 nodes), the basic version of multi-agent collaboration (multiple AI assistants each responsible for its own area), the generation–evaluation–optimization closed-loop prototype, the basic version of the dance knowledge base (including an intangible-cultural-heritage topic), 3D preview, formation arrangement, style locking and character consistency, and basic desktop-application packaging. Under development and awaiting full validation: the forward-mapping diffusion model, end-to-end video generation, deepening of the aesthetic layer, and high-quality video export. Framework designs not yet fully implemented include: large-scale multi-dance-type validation, multi-person/group-dance generation, multimodal comprehensive evaluation, and fully automated high-quality optimization. The maturity of the modules varies considerably. This paper mainly releases the complete architecture and technical roadmap, as well as the results of Phases I and II; the Phase III models are still being iterated and have not been sufficiently validated.

**Version note:** This paper corresponds to DanceFlow v1.0, with emphasis on releasing the system-architecture design, technical roadmap, and preliminary implementation of the core modules—Laban Effort inverse mapping, the three-layer, six-dimensional evaluation framework, and the visual workflow. Work such as large-scale controlled experiments, validation data for multiple dance types, and quantitative research on the aesthetic layer is still in progress and will be supplemented successively in the form of version updates. The preprint version number and update log will be synchronized on the project homepage.

## IV. Verification and Pilots

This study' s verification follows the idea of "small-step validation and learning through use." It does not seek one-time completeness; rather, once each module can run, it is put into trial use in real teaching and intangible-cultural-heritage scenarios, with feedback driving iteration. The completed verification is summarized below from three aspects—core modules, teaching trials, and intangible-cultural-heritage pilots (see Table 4).

**Table 4. Summary of Verification Status**

| Verification item | Verification content | Method | Sample / scope | Result |
|---|---|---|---|---|
| Inverse mapping of Laban effort | Automatically computes time curves of the four effort elements from SMPL motion | Extract motion data from joint coordinates and map effort factors; iteratively refine with AI | Several segments from teaching videos and personal recordings | After seven rounds of iteration, it is basically stable and has been used for intangible-cultural-heritage effort analysis and teaching evaluation |
| Three-layer, six-dimensional evaluation framework | A three-part structure of scoring + diagnosis + suggestions, explaining the basis for deductions item by item | Build a three-layer, six-dimensional pipeline based on inverse mapping | Teaching trial | Students most recognized that "the reasons for score deductions can be clearly explained"; feedback changed from vague impressions into trainable points |
| Generation–evaluation–optimization closed loop | Prototype | Multi-agent collaboration and visualized workflow | Internal trials and trials by teaching units | Usable; the iteration cycle is still long, and there remains a gap from real-time interaction |

| Verification item | Verification content | Method | Sample / scope | Result |
|---|---|---|---|---|
| Intangible-cultural-heritage effort analysis | Representative Guangdong dance genres such as Qilin Dance, Yingge Dance, Mat Lion Dance, and Fire Dog Dance | Dual-channel acquisition + four-level annotation: effort curves → expressiveness scores → cultural tags | Pilots on four dance genres | An initial effort archive has been established, verifying the feasibility of the "small sample + high-quality annotation" route |
| Teaching deployment | LAN deployment, accessible from ordinary laptops | Front-end packaging + single-file distribution | Trials by teaching units | Student feedback has been positive, and the tool has entered the classroom |

The scale of the above verification is limited and the conclusions are still preliminary, but all of them come from real teaching and intangible-cultural-heritage scenarios, confirming the direction. In follow-up work, the weights of each layer and the efficiency of the closed loop will be systematically calibrated on larger samples.

In addition, dance videos involve performers' portraits and privacy. In the use of materials, the author follows the principles of informed consent and minimum necessity, and uses them only for methodological verification. More systematic experimental data—including public dance datasets and multi-genre samples—are still being collected and checked, and will be gradually supplemented in subsequent versions.

## V. Application Scenarios

DanceFlow is positioned as an interpretable, controllable, teachable, and learnable intelligent dance-tool system. In different scenarios, it can serve as a teaching assistant, creative partner, or research tool.

the most urgently needed scenarios. The system can support classroom assessment (students record videos; a three-layer, six-dimensional evaluation automatically produces diagnoses of standardization and expressiveness, freeing teachers from "correcting movements one by one" ), personalized learning paths (recommending exercises according to weak dimensions), process-based assessment portfolios (dynamically recording effort features—i.e., effort curves, lines that plot changes in movement intensity—and five-dimensional curves), and Laban teaching support (effort visualization makes abstract theory "visible" ).

In frontline trials, what students recognized most was that "deductions can be explained clearly" : in the past, when teachers said something was "not flavorful enough," students had no concrete way to improve; now the system can locate specific effort factors, turning feedback from vague impressions into trainable points.

Choreographers can use Laban parameters to quickly verify ideas, systematically explore the effort space to break through the limits of experience, examine the overall rhythm and climaxes of a work through formal-structural analysis, and collaborate with multiple people through a node-based workflow.

A direction with distinctive features. Based on the four elements of effort, the system quantitatively records movement quality, conducts quantitative stylistic comparison within a unified coordinate system, and, through "few samples (learning from only a very small number of examples) + high-quality annotation + theory-driven methods,"enables small-sample intangible-cultural-heritage data to realize its value while providing a new quantitative perspective for cultural research. Guangdong intangible cultural heritage currently being advanced includes Qilin Dance, Yingge Dance, Mat Lion Dance, and Fire Dog Dance, though the samples remain small. What is especially crucial is that effort annotation transforms the "embodied memory of inheritors" into searchable and comparable structured data, laying an interface for subsequent "theory-driven small-sample generation."

Potential directions: objectively monitoring the quality of rehabilitation movements, designing personalized plans according to weak dimensions, and providing quantitative tools for research on the "correlation between movement and psychological states." It should be noted that dance therapy involves medical and health issues; DanceFlow serves only as a research aid and does not replace professional diagnosis or treatment.

Other possibilities include AI dance tutoring for the general public, entertaining "imitation challenge" interactions, and assistance for primary and secondary schools where dance-aesthetic-education teachers are insufficient. Current priorities are: first tier (already implemented), teaching assistance and effort annotation and analysis for intangible cultural heritage; second tier (under development), choreographic creativity assistance and structural analysis of works; third tier (under exploration), therapeutic research tools and public aesthetic education. The core logic is to first build a solid evaluation end, then expand toward the generation end, and finally broaden to wider scenarios.

## VI. Limitations and Future Directions

From the perspective of a dance teacher, three aspects of this system give the author the greatest confidence. Technically, the three-layer hybrid architecture plus the three-layer, six-dimensional evaluation breaks the large leap from "physics → semantics → aesthetics" into something that can be approached step by step.

As a step upward, compared with a framework that only stares at objective indicators, it adds two further steps: "semantic diagnosis" and "initial aesthetic exploration" (the aesthetic layer is still only a scaffold at present). In terms of architecture, semantics and motion are separated yet connected— "semantics are mapped to Laban, motion to SMPL," and the two are linked by bidirectional mapping, avoiding both the opacity of a black box and the unnaturalness of purely symbolic formulas. Culturally, it adds a route of "Effort annotation" for the digitization of intangible cultural heritage, moving a small step from "preserving the form" toward "understanding the texture." Of course, the basis for these statements mainly comes from the author' s own classroom and practice; the sample size and control experiments are both limited, so they should be taken only as directional impressions, not as conclusions.

DanceFlow stands on the shoulders of predecessors: Laban theory has existed for nearly a century, and SMPL (a digital skeleton format) is a domain standard. Using AI to generate movement has also become a mainstream approach—in recent years, the diffusion generation of SoulDance[6], Park et al.[7], the editable generation of EDGE[8], the robotic dance of Wallace et al.[25], and the large-model choreography of DanceChat[26] have all made continuous breakthroughs in generation quality and controllability; there have also been many pioneering studies on quality evaluation[12][13][14]. This paper does not invent an entirely new algorithm; rather, it reorganizes existing modules according to the needs of dance teaching, forming a relatively complete, interpretable prototype oriented toward the full creative process. The following table gives a rough comparison across multiple dimensions:

Table 5 Rough comparison of DanceFlow and typical related work

| Dimension | Pure data-driven generation(EDGE/MACE-Dance, etc.) | Laban movement analysis research(Aristidou/Tsachor, etc.) | Dance quality evaluation(FID/BAS/DIV, etc.) | Intangible-heritage digitization projects(Cyber Subin, etc.) | DanceFlow |
|---|---|---|---|---|---|
| Role of Laban | Not involved / weakly related | Core research object | Partially involved | Not involved | Serves as the semantic hub(attempts to connect generation + evaluation + control; preliminary implementation) |
| Evaluation level | Objective distance metrics(single layer) | Semantic analysis(single layer) | Objective indicators(single layer) | No systematic evaluation | Three-layer progression(objective → semantic → aesthetic; the latter two layers are still being improved) |
| Generation-evaluation relationship | Separated; evaluation is used for acceptance testing | No generation end | Separated from generation | Generation first, weak evaluation | Preliminary closed loop under a unified semantic space |

| Dimension | Pure data-driven generation(EDGE/MACE-Dance, etc.) | Laban movement analysis research(Aristidou/Fisher, etc.) | Dance quality evaluation(FID/BAS/DIV, etc.) | Intangible-heritage digitization projects(Cyber etc.) | DanceFlow |
|---|---|---|---|---|---|
| Interpretability | Low(black box) | High(theory-driven) | Medium(interpretable indicators) | Low | Relatively high(each layer is traceable; diagnosis can be localized) |
| Data dependence | High(big-data driven) | Low(theory-driven) | Medium | High(requires large-scale collection) | Medium-low(small samples + theoretical framework; validation scale is limited) |
| Intangible-heritage application | Not applicable(lack of data) | Has application potential but has not formed a system | Not involved | Mainly morphological archiving | Effort-level texture analysis and annotation(small-sample exploratory stage) |
| Target users | AIGC researchers | Movement-analysis researchers | Algorithm researchers | Cultural preservation personnel | Dance teachers, [[unclear: text begins with "学" but is cut off]] |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | Students and choreographers (an attempt to orient toward practitioners) |
| Core objective | Generation quality | Understanding of movement semantics | Evaluation accuracy | Cultural archiving | Human–AI collaborative creation and teaching assistance |
| Disciplinary foundation | Mainly computer science | Dance studies / kinesiology | Computer science / statistics | Anthropology / cultural studies | An attempted interdisciplinary integration of computer science + dance studies + motor anatomy + psychology + aesthetics |

The differences from existing work stem from the author' s perspective as a frontline teacher, working backward from teaching needs: the starting point is to solve real problems in my own teaching and projects, rather than to look for applications for a technology. Laban is used not only for analysis, but also as a "translator" between human and machine—a common language for generation, evaluation, and adjustment. Evaluation does not merely assign a score; it identifies, for example, "which passage has become rhythmically flat" or "which movement lacks sufficient force." Intangible cultural heritage is not only preserved in its forms; the texture of its effort qualities is also analyzed. These points may not count as innovation; they are simply a frontline teacher' s reorganization of existing technologies and theories according to his own needs, written down for exploration and exchange with colleagues.

MACE-Dance[^27] and similar work represent the academic frontier. DanceFlow stands on the shoulders of giants, and the two are not in competition: the former addresses "whether the generated result looks alike," follows a purely data-driven path, and is oriented toward AIGC researchers; the latter addresses "whether it can be used in teaching," follows an explainable human–machine collaborative path, and is oriented toward teachers and creators. The former' s objective evaluation and SMPL intermediate representation have provided many inspirations for this paper.

First, the controllability of forward mapping: gravity and time factors can be regulated relatively well, while the gaps in spatial and flow factors remain large, and the coupling among factors is difficult to disentangle. Second, the difficulty of quantifying the aesthetic layer: subjectivity is plural. The strategy of this paper is not to pursue an "objective aesthetic score," but to provide evidence-based analytical perspectives and return the final judgment to people. Third, the limitation of data scale: professional motion capture for intangible cultural heritage is costly, and samples are difficult to expand; this is also why the approach follows "small samples + high-quality annotation + theory-driven" methods. Fourth, closed-loop efficiency: the iteration cycle of generation–evaluation–optimization is still long, and there remains a gap from real-time interaction. Fifth, multimodal integration is insufficient: at present, the work mainly processes the single modality of movement, whereas real expressiveness is a synthesis of music, movement, facial expression, and costume. These problems are interconnected: insufficient multimodality amplifies evaluation bias, while data limitations in turn slow multimodal expansion. Subsequent work should therefore proceed through "small-step validation" rather than "one-time completeness."

Third, considerations at the level of data and ethics. Dance videos involve performers' portraits and personal images, and materials from classrooms and intangible-cultural-heritage fieldwork are also connected to privacy and the rights and interests of inheriting communities. In the use of materials, the author follows informed consent and the most

minimum-necessity principle, and are used only for methodological validation and teaching improvement; the persons involved are all anonymized or handled with authorization. More systematic plans for data compliance, portrait-right authorization, and privacy protection will be refined in subsequent versions and dataset construction.

Deepen the bidirectional mapping to improve factor controllability and precision; expand the scale of validation and systematically calibrate the weights of each layer; refine the aesthetic layer and explore the boundaries of quantification; build an intangible-cultural-heritage dataset and broaden coverage of dance genres to validate the "few-shot samples + effort annotation" route; and introduce music and facial expressions in multimodal expansion. It must be emphasized that all of the above directions are premised on "not departing from the real needs of teaching," so as to avoid deviating from the original aim of interpretability and low entry barriers merely in pursuit of technical metrics.

## VII. Conclusion

The contribution of this paper lies not in an algorithmic breakthrough, but in three reusable pathways. First, architectural innovation: embedding a Laban semantic layer into a "generation-evaluation-optimization" closed loop, using Laban Effort as a "translator" to connect semantics, generation, and evaluation. Second, a practical roadmap: proposing a feasible path by which non-technical personnel can collaborate with AI to develop professional tools—raising requirements, having AI implement them, and then having the author verify them. Third, directional validation: preliminarily verifying a route for the digitalization of intangible-heritage dance through "few-shot samples + theory-driven" methods. Over these past few months of working on dance AI, my deepest feeling has been that it "looks easy, but is hard to do." At first, I assumed that the larger the model and the more functions it had, the better; only in the process of doing it did I discover that many things cannot be solved simply by piling up models and prompts. After stepping into many pitfalls, I became more pragmatic: I no longer pursue full automation, but instead consider how humans and AI can cooperate with the greatest efficiency; I no longer pursue disruptive innovation, but instead ask which aspects of existing technology can genuinely help teachers and students. Its positioning is very simple: an assistive tool made by a dance teacher for the teacher' s own use. It contains no breakthrough technology; it merely strings together existing theories, models, and tools according to teaching needs. I have written it up as a paper not because I believe I have done it well, but because, on the path of "starting from the dance discipline and pursuing interpretable human-machine collaboration," there may be many colleagues traveling the same road; I offer this as a brick in the hope of drawing out jade.

AI lowers the threshold for content creation, but its penetration into dance remains slow—because the core of dance is not only whether the movements are correct, but also texture, flavor, and expressiveness: things that are difficult to articulate clearly and difficult to quantify. Precisely because it is difficult, it may have value: after practice, dance students can have a tool that shows where problems lie; young choreographers can quickly test ideas when creating works; inheritors of intangible heritage can not only record dances they have performed all their lives, but also systematically document their texture. DanceFlow has taken only a very small step in this direction, and many aspects remain rough; but at least it is already usable in teaching, and student feedback has been acceptable. I write this for exchange, hoping for more criticism and more fellow travelers. As an ordinary frontline dance teacher, I publish this preprint in order to give more

It also aims to offer dance professionals and technical experts in the field of AI a useful point of reference; we likewise hope that, in today' s era of rapid AI iteration, more tools that are truly applicable to the dance field will emerge, jointly promoting the better development of this domain.

This paper records the design and preliminary implementation of the first phase of DanceFlow. In the spirit of academic rigor and integrity, more comprehensive experimental validation, data-compliance development, and module refinement will be continuously updated in subsequent versions. We welcome attention and exchanges from colleagues.

**References:**

[1] Li Jie. DanceFlow Demo—An AI-Assisted Dance Creation and Dance Digitalization Platform [EB/OL]. (2026-07-01) [2026-07-06]. https://forum.trae.cn/t/topic/52745.

[2] Beijing Dance Academy. Research on Beijing Intangible Cultural Heritage Dance Based on Laban Theory [EB/OL]. (2025-09-19) [2026-07-07]. https://keyanchu.bda.edu.cn/kycg/fe5aa2a617b64c9fa8d30142039b9efa.htm

[3] Guo C, Zou S, Zuo X, et al. Generating Diverse and Natural 3D Human Motions From Text[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 5152–5161.

[4] Petrovich M, Black M J, Varol G. TEMOS: Generating Diverse Human Motions from Textual Descriptions[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2022: 480–497.

[5] Siyao L, Yu W, Gu T, et al. Bailando: 3D Dance Generation by Actor-Critic GPT with Choreographic Memory[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 7295–7304.

[6] Li X, Li R, Fang S, et al. SoulDance: Music-Aligned Holistic 3D Dance Generation via Hierarchical Motion Modeling[J]. arXiv preprint arXiv:2507.14915, 2025.

[7] Park S, et al. Not Like Transformers: Drop the Beat Representation for Dance Generation with Mamba-Based Diffusion Model[C]//WACV, 2026.

[8] Tevet G, et al. EDGE: Editable Dance Generation from Music[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 1222–1232.

[9] Li R, Wang R, Zhao S, et al. AIST++: A New AIST Dance Video Database for 3D Dance Motion Synthesis and Dance Genre Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.

[10] Punnakkal B, Chandrasekaran A, Agrawal N, et al. BABEL: Bodies, Action and Behavior with English Labels[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 722–731.

[11] Li B, Zhao Y, Zhelun S, et al. Dance Revolution: Recommendation and Choreography for Dancing to Music[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1219–1228.

[12] Li H, Huang X. Intelligent Dance Motion Evaluation: An Evaluation Method Based on Keyframe Acquisition According to Musical Beat Features[J]. Sensors, 2024, 24(19): 6278.

[13] Qu J. A dance movement quality evaluation model using transformer encoder and convolutional neural network[J]. Scientific Reports, 2024, 14: 32058.

[14] Monte e Freitas, et al. Can Vision Language Models Judge Action Quality? An Empirical Evaluation[C]//CVPR 2026 Workshop, 2026.

[15] Wirdnam M, et al. Performance Quality Assessment in Ballet, Modern and Contemporary Dance: A Two-Step Systematic Review[J]. Journal of Dance Medicine & Science, 2024.

[16] Laban R. The Mastery of Movement[M]. 4th ed. Plymouth: Northcote House, 2011.

[17] Aristidou A, Chrysanthou Y, Shamir A. Folk Dance Evaluation Using Laban Movement Analysis[J]. ACM Journal on Computing and Cultural Heritage, 2015, 8(4): 1–19.

[18] Tsachor R P, Shafir T. How shall I count the ways? A method for quantifying the qualitative aspects of unscripted movement with Laban Movement Analysis[J]. Frontiers in Psychology, 2019, 10: 572.

[19] Turab M, Colantoni P, Muselet D, et al. Emotion recognition in contemporary dance performances using Laban Movement Analysis[EB/OL]. arXiv:2504.21154, 2025.

[20] Turab M, Colantoni P, Muselet D, et al. Dance Style Recognition Using Laban Movement Analysis[J]. arXiv preprint arXiv:2504.21166, 2025.

[21] Zhou Yu, Karen Bradley, Liao Binbin. A Review of the Laban–Bartenieff Movement Studies System[J]. Journal of Beijing Dance Academy, 2017(3): 21–28.

[22] Pataranutaporn P, Klunchun P. Human-AI Co-Dancing: Evolving Cultural Heritage through Collaborative Choreography with Generative Virtual Characters[C]//MOCO ’ 24. ACM, 2024.

[23] Guo Hong, Xu Yilin, Zou Nachen, et al. Quantitative Study of Students’ Movement Performance in Online Aerobic Dance Courses Based on Nonlinear Dimensionality-Reduction Clustering[J]. China Sport Science, 2022, 42(6).

[24] Loper M, Mahmood N, Romero J, et al. SMPL: A Skinned Multi-Person Linear Model[J]. ACM Transactions on Graphics, 2015, 34(6): 1–16.

[25] Wallace B, Glette K, Szorkovszky A. How can we make robot dance expressive and responsive? A survey of methods and future directions[J]. Frontiers in Computer Science, 2025, 7: 1575667. doi: 10.3389/fcomp.2025.1575667.

[26] Wang Q, Yang X, Dong Y, et al. DanceChat: Large Language Model-Guided Music-to-Dance Generation[J]. arXiv preprint arXiv:2506.10574, 2025.

[27] Yang K, Zhu J, Tang X, et al. MACE-Dance: Motion-Appearance Cascaded Experts for Music-Driven Dance Video Generation[J]. arXiv preprint arXiv:2512.18181, 2025.

(Corresponding author: Li Jie, E-mail: 9079985@qq.com)

## Author Contributions Statement:

Li Jie: proposed the research ideas and questions; designed the overall system architecture and technical route; defined the functional requirements and evaluation criteria at the level of the dance discipline; completed the design of semantic mapping for Laban movement analysis; wrote the full manuscript and was responsible for the final revisions. (Note: all program code for the system was completed by AI programming assistants under the author' s guidance on requirements; the author does not know how to program and was responsible only for defining requirements, accepting results, and controlling the direction of iteration. The technical implementation was not carried out by the author personally.)

## Code Implementation Statement:

All program code in this paper—including the front-end interface, back-end services, Laban effort calculation module, video export pipeline, and related components—was completed by AI programming assistants Trae, WorkBuddy, qoder, and dumate under the author' s requirements guidance and functional acceptance. The author has a professional background in dance and does not know how to program; therefore, the system' s "technical implementation" is explicitly attributed to AI-assisted completion. The author' s working mode was as follows: the author put forward requirements, judged whether the tools were useful, indicated directions for improvement, and then asked AI to revise and refine the system. All code was completed through this human–machine collaboration process of "non-technical personnel proposing requirements—AI implementing them—the author accepting the results." This collaborative mode is itself also one of the practical contents addressed by this study: "how non-computer professionals use AI tools to develop professional applications."

## AI Tool Use Statement:

AI tools were used extensively in the process of this research and the writing of the paper. The specific circumstances are truthfully described as follows:

(1) Tools used: the DeepSeek-R1 large language model, ChatGPT, Trae WorkBuddy (AI programming assistant), and the Kouzi (Coze) AI Agent platform (an AI assistant capable of automatically performing tasks).

(2) Scope and role of use:

Literature search assistance: AI tools assisted in finding leads to relevant research literature and provided suggestions for search keywords; the final selection, appraisal, and reading of the literature were completed by the author.

Text polishing and format organization: AI assisted in checking sentence fluency, optimizing expression, and formatting; the core viewpoints and argumentative logic were completed independently by the author.

Code implementation: the writing, debugging, and repair of all system code were mainly completed by the Trae WorkBuddy AI programming assistant; the author was responsible for proposing functional requirements, accepting output results, and guiding the direction of iteration.

Discussion of technical schemes: issues such as architecture selection, algorithmic ideas, and technical feasibility were discussed and verified through dialogue with AI; the final decisions were made by the author.

English abstract translation: the English abstract was translated with the assistance of AI tools, and the author reviewed and revised the professional content.

In the writing of the paper, AI tools were used only for auxiliary steps such as language polishing, format organization, and English translation, and did not replace the author in completing the substantive writing of the paper (the core viewpoints, argumentative logic, and professional judgments were all completed independently by the author); in system development

In this respect, all code was implemented by an AI programming assistant under the guidance of the authors' requirements; the authors were responsible for defining the requirements, evaluating the results, and controlling the overall direction. This paper falls within the compliant category of "author-written work polished with AI assistance," and its technical implementation is likewise a form of human–machine collaboration in which "the authors provide the requirements and AI carries out the implementation," rather than "AI writing or doing the work on behalf of the authors."

*Note: Figure translations are in progress. See original paper for figures.*