

Navigating a Guzheng Phrase Corpus Through Ballet-Driven Granular Synthesis

Yueshen Wu
yueshen.wu@network.rca.ac.uk
Royal College of Art
London, UK

Yuting Xue*
yuting-xue@my.cityu.edu.hk
City University of Hong Kong
Hong Kong, SAR China



Figure 1: *Fluid Body*: the performer is in the immersive environment.

Abstract

This paper presents an interactive audio-visual system that reimagines cultural movement as a public interface for accessible musical expression. It allows participants to reshape a Chinese Guzheng soundscape through ballet-inspired body movement. Based on granular synthesis, we recontextualize Chinese Guzheng phrases and performance techniques as a sonic corpus that can be navigated by ballet postural vocabulary, complemented by expressive sound descriptor shaping to enrich the sonic result and legibility. In addition, we develop a hybrid classification-regression mapping strategy in which “in-between” and deviated postures are supported by interpolating

*Both authors contributed equally to this work.



This work is licensed under a Creative Commons Attribution 4.0 International License.

NIME '26, June 23–26, 2026, London, UK

© 2026 Copyright held by the owner/author(s).

between learned examples. The contribution is an intercultural co-adaptation experiment in NIME practice by proposing a shared interaction language between communities.

Keywords

Embodied sonification, intercultural communication, supervised machine learning system, audio-visual installation art

1 Introduction

In recent years, digital and computational art practices have created new opportunities for the documentation, reinterpretation, and presentation of cultural materials in interactive forms [16]. In such practices, interactive creative systems are often designed to support broader participation, enabling the public to engage with cultural expression through embodied interaction and responsive means [2, 19]. Based on machine learning (ML), these practices provide adaptive mappings between human action and audio-visual transformation, allowing artists to work with cultural materials as dynamic, performable resources [24, 31].

In the realm of sound art and NIME-related practices, the establishment of granular synthesis has enabled a variety of time-based effects, such as time-stretching and pitch-shifting of recorded materials while retaining perceptible traces of the source, supporting rich sonic transformations grounded in the materials' original tone [25]. When combined with parameter mapping sonification, data streams such as bodily signals can be sculpted into responsive auditory structures that emphasize interaction, authorship, and aesthetic emergence [14]. Accordingly, ML-driven creative systems have been integrated into sonification interfaces and dance performances to extend how movement can be mapped to musical structures and expressive nuance [7, 21, 38].

Traditional Chinese music and Western classical dance provide a valuable lens for this exploration. The Guzheng has a rich performance tradition in which techniques such as “vibrato,” “portamento,” and “glissando” shape fluid pitch and timbral variations that are closely linked to embodied modulation and expressive nuance [32]. Ballet, on the other hand, is a codified movement practice in which spatial orientations and a formal system of arm and foot positions define body geometry, balance, and flow, and enable continuous phrasing in time [17, 35]. Although the two traditions emerge from different cultural and historical contexts, they share concerns with continuity, controlled gesture, internal rhythm, and breath-like phrasing. These shared qualities suit them for mutual computational reinterpretation as interactive systems, where movement can serve as a navigational method, and sound correspondingly remains in culturally situated material while becoming responsive and transformable.

Drawing on Laban Movement Analysis (LMA) [10] as a descriptive framework for movement, we present *Fluid Body*: an ML-mediated intercultural body instrument in which ballet-inspired postures serve as a navigational language for recomposing a corpus of Guzheng phrases via granular synthesis. The system combines supervised learning for discrete state selection (classification) with continuous estimation (regression). It can control both corpus region access and continuous textural change based on articulation variation. This design is intended for mixed communities of participation, regardless of experience; everyone can share the same interface, each bringing different embodied knowledge, curiosity, and interpretive strategies to the instrument.

More importantly, the system does not require movements to follow the canonical ballet postures strictly. By learning interpolations between demonstrated examples to support continuous correspondences beyond fixed pairings, this system allows participants to perform “in-between” or deviated postures, thereby creating intermediate versions of the Guzheng soundscape. In this setup, ML serves as an algorithmic bridge that enables culturally encoded vocabularies to remain in place while becoming executable through novel navigation and interpretation. Through *Fluid Body*, this paper discusses how sound synthesis design and intercultural interaction can be developed together in NIME, with attention to mapping strategy, accessibility across participants, and implications for interfaces that engage intercultural practices beyond a single community.

2 Background and Related Work

2.1 Movement-based musical interfaces and ML-based strategies

Practitioners in NIME and creative sonification fields have treated the body as an instrument for decades and designed systems in

which movement becomes a meaningful control surface. Prior examples include *Borrowed Gestures* [9], which frames non-musical movement vocabularies as latent musical knowledge and shows how participatory and example-based approaches can support personally and culturally situated gesture languages in performance. Interactive environments for performance, such as *Somacoustics* [3], similarly emphasize the felt body, transforming motion and social interaction into shared sonic events and positioning the performer's body as a resonant medium rather than a neutral controller. Apart from motion capture, sound-generating interfaces based on biosignals extend this embodied mapping into the physiological field. Donnarumma et al. [11] combined EMG and MMG in their project, which demonstrates how muscular activity is mapped to sound synthesis parameters. This method enables performers to articulate sonic texture layers within a single gesture. Tanaka and colleagues [33] have designed an end-to-end EMG-to-sound instrument to further highlight how machine learning can mediate between embodied signals and synthesis, shifting the design focus from fixed mappings to learned translation pipelines.

Alongside these instruments, interactive machine learning has become a key mapping paradigm in NIME, particularly through approaches that allow artists to train mappings by demonstration. Wekinator is a foundational example of this paradigm, enabling artists to record examples, modify mappings, and retrain models iteratively in real time as part of the creative process [13]. This approach has been deployed notably in NIME practice; for example, *Wekinating 000000Swan* [28], uses Wekinator to prototype both discrete and continuous models from the Kinect sensor for live audio-visual performance. More recent Wekinator-based examples include its usage as a classification approach for recognizing tactile materials from audio features, offering a tangible interaction layer to instrument design [12]. Other ML toolkits have embedded similar capabilities within musician- and designer-oriented creative environments. The *Gestural Sound Toolkit* [6] supports the design of gestural interactions with sound in Max/MSP, combining motion input, ML modules, and sound-synthesis modules within a framework for rapid prototyping. Its support for classification and regression in posture recognition and gesture-sound mapping is particularly useful in movement-based musical interface development. In parallel, the *Fluid Corpus Manipulation Toolkit* [34] expands the role of ML towards sound-material exploration by integrating signal decomposition, machine listening and machine learning into coding environments including Max and Pure Data, enabling musicians to analyze, navigate, and manipulate sound banks as part of what the authors call “programmable data mining as musicking.” These deployments show how interactive ML in music and sound practices supports both discrete-state recognition and continuous control, and how its value lies not only in model accuracy but also in how learning is integrated into the authoring and rehearsal process for playable, interpretable instruments.

2.2 Granular and corpus-based instruments for navigation and sound design

The establishment of granular synthesis has provided a foundational sound design strategy for turning recorded material into continuous morphing textures, enabling time scale transformations while maintaining perceptual traces of the source [25]. Within NIME, granular engines have been repeatedly explored as performative instruments and interactive soundscapes.

GrainProc [27], for example, designs a real-time granular interface that emphasizes intuitive, continuous manipulation of grain parameters for live performance. More recent work, such as the *Tangible Granular Device* [12], manipulates granular control through material exploration and manipulation, demonstrating how granular engines can support exploratory interaction and perceptual learning through touch and gesture.

In parallel, corpus-based and concatenative approaches treat a collection of sound units as a navigable instrument. *CataRT*, developed by Schwarz et al. [30], is a corpus-based concatenative synthesis system. It frames the “sound space” itself as a musical instrument, where interaction involves traversing and selecting from a real-time, descriptor-analyzed corpus. *Grainspace* [20] similarly combines concatenative techniques with embodied navigation using a Kinect sensor, reinforcing the idea of corpus exploration as an experiential, spatialized act in performance.

2.3 Culturally situated instruments in NIME practices

A growing body of work positions interactive systems as mediators between cultural situated practices and contemporary publics. Yu et al. introduce *VirGuzhin* and *MRGuzhin* [37] to reduce barriers to learning Guqin techniques and notation, demonstrating how motion tracking and mixed reality can support museum and classroom engagement with historically specialized practices. In NIME, Bryan Kinns and Li’s project addresses cross-cultural co-creation around Duxianqin, explicitly reflecting on how “essence” is negotiated when a traditional instrument is reimaged through digital technologies [4]. The epistemic DMI “O-” [22] further addresses interculturality by reflecting on the concept of time across Eastern and Western cultural contexts, and invites performers to improvise and discuss how cultural assumptions are embedded in instrumental metaphors.

This body of prior work has established robust foundations for embodied mapping strategy in sonification, interactive machine learning, granular and corpus-based sound synthesis, and culturally situated instrument design. However, the intersection of these fields remains underdeveloped in the following ways. First, a large amount of movement-based sonification and interactive ML instruments prioritize technical exploration, including mapping flexibility and performative expressivity. However, the underpinning meaning and cultural aesthetics of the series of interactions are less discussed. Second, granular and corpus-based concatenative synthesis often positions the samples as a generic collection, and fewer works address how the corpus is culturally meaningful and can be recontextualized as a shareable, interactive environment for the public. More importantly, relevant culturally situated musical interfaces typically focus on mono-cultural access, reflection, and preservation, sometimes with interculturality as a conceptual layer rather than as a coupled, shared, and bidirectional interaction design language between two embodied traditions. These potential gaps have motivated *Fluid Body*, which brings together Eastern and Western cultural aesthetics: Guzheng, the Chinese traditional music and ballet, the Western codified movement using granular synthesis and corpus navigation based on a hybrid ML mapping strategy. The system supports continuous, interpretable control across both canonical and in-between gestures, exploring how one embodied vocabulary provides access to another sonic tradition while keeping both legible within interaction.

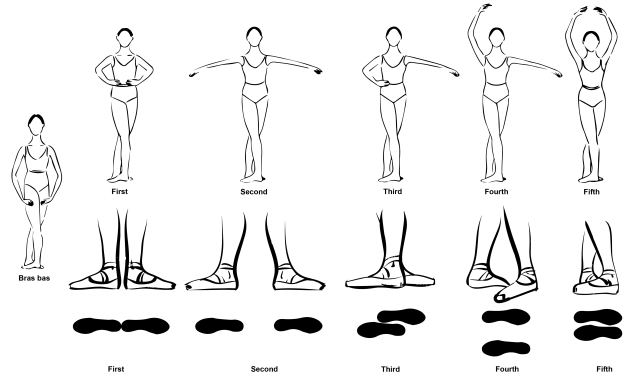


Figure 2: The illustration of ballet postures: *Bras bas*, five arm positions and five foot positions.

3 *Fluid Body*: Ballet-Navigated Guzheng Corpus Through Granular Synthesis

Fluid Body is an intercultural, embodied musical interface mediated by machine learning. It treats ballet-derived movements as a learnable control surface for recomposing Guzheng phrases and shaping timbre based on granular synthesis. It is positioned in a 180-degree projection space that supports full-body performance and can provide immediate spatial feedback. Such a setup allows participants to orient their bodies freely, react to sound changes, and verify control outcomes through immediate visual cues without being confined to a frontal screen. This spatial setup helps maintain a tight loop between bodily intention, continuously evolving sonic texture, and the feedback needed to learn the mapping in real time.

3.1 Design Concept: Cultural and Rhythmic Resonance

The design concept of *Fluid Body* is based on the cultural and rhythmic resonance between Guzheng and ballet movement. Both of them are refined by history, emphasizing physical cultivation and rhythmic structure, and transforming controlled postures into aesthetic expression.

From a cultural perspective, both Guzheng and ballet are associated with traditions of cultivated refinement. The Guzheng has been connected with literati, courtly, and classical musical cultures in China, where performance is shaped by gesture, restraint, and expressive interiority [18]. Ballet, emerging from European court culture and later formalized through theatrical and academic systems, similarly carries an aesthetic of discipline, elegance, and bodily refinement [1]. This shared association with refinement provides a basis for placing them into dialogue through a system that emphasizes precision, control, and expressive transformation. A more specific point of connection lies in the relationship between gesture and restraint. In Guzheng performance, expressivity often emerges from a controlled economy of the performer’s hands. Small changes in finger, wrist, or pressure can produce significant sonic effects through timbral variation, pitch bending, vibrato, resonance, and the continuation of sound after the initial pluck [23]. Similarly, ballet is based on the disciplined regulation of posture, balance, extension, and transition [15]. In both practices, expression is produced through intensification of subtle variation within codified constraints. This makes ballet a productive counterpart to Guzheng for this

project: both traditions make small differences in bodily control aesthetically meaningful.

A further connection concerns continuity and energetic flow. The concept of *Qi* “气” offers a useful framework for understanding how form is animated from within. In Chinese aesthetic thought, *Qi* may be understood as breath, vitality, or dynamic force, shaping the continuity between body, gesture, and expression [26]. In Guzheng performance, this continuity is materially produced through timing, pressure, vibration, resonance, and the interaction between the two hands. The right hand initiates the sound through plucking, while the left hand continues, bends, or destabilizes the pitch after the initial attack [23]. Ballet also gains expressive force through continuity rather than through isolated positions alone. The dancer does not simply move from one fixed pose to another; aesthetic meaning emerges through transition, breath, torso organization, weight shift, pelvic drive, and the sustained flow between postural states [36].

3.2 Mapping Strategies

To facilitate navigation of the corpus and perception of timbral changes during public performances, we use a perception-based mapping approach. We begin with the sonic characteristics that the instrument must clearly show. Then, we select movement features that can effectively produce those sounds without confusion. Our mapping system combines structural navigation with continuous articulation. Recognizable postures serve as reference points for accessing different sections of the Guzheng phrase corpus. At the same time, continuous movement descriptors influence the details within and between those sections. This approach promotes stability and exploration. Participants can recognize a phrase while still creating a shifting soundscape through subtle movements. Such a double-stage mapping strategy echoes recent NIME work, where direct mappings combine with ML models to balance responsiveness and reliability.

Building on this mapping approach, we use Laban Movement Analysis (LMA) as a theoretical intermediary to translate between the visible movement of the ballet and the Guzheng playing articulation [10]. LMA uses “Space,” “Shape,” “Effort,” and “Body” to define a relation between spatial orientation, bodily configuration, dynamic quality, and temporal phrasing. This relational view corresponds to the Guzheng’s own performance logic, and supports our three mapping layers: spatial orientation informs register navigation, bodily configuration supports posture-based corpus access, and movement quality provides continuous control for granular rearticulation, timbral density, and sonic instability. For example, vertical arm orientation provides a perceptually legible way to navigate register, since downward reach can bias the system toward lower and darker materials, while upward reach can bias it toward brighter and higher-pitched materials. LMA thus allows the mapping to be framed as a shared grammar of form, flow, and expressive modulation between two culturally distinct practices.

3.2.1 Register Navigation. Guzheng materials in the corpus are curated into four register bands: “Sub bass,” “Bass,” “Middle,” and “High.” We map the performer’s arm spatial orientation to simplified ballet terms describing body directions, including “en bas,” “en avant,” “à la seconde,” and “en haut,” to a continuous navigation bias across these bands. In this way, downward reach increases the likelihood of lower, darker material, while upward reach biases the system toward brighter, higher-pitched phrases.

This mapping treats spatial reach as a culturally situated and learnable cue for navigating pitch spaces.

3.2.2 Corpus Access. To keep the corpus manageable, we use recognizable posture frames from ballet arm and foot positions (Fig. 2) to select phrase regions. We support “in-between” and deviated postures through interpolation between learned examples. This approach changes phrase playback from simply “triggering a correct excerpt” to navigating a flexible phrase ecology. Users can move through hybrid states and access intermediate textures, avoiding sudden jumps.

3.2.3 Timbral Articulation. Fine-grained movement nuance drives how the selected corpus material is rearticulated as granulated textures. Movement intensity and temporal characteristics are mapped to continuously varying sound controls that remain perceptually interpretable in performance, such as grain density, scan speed, and texture spread. By doing so, we recontextualize Guzheng phrases and techniques as a malleable sonic material library that can be reshaped into continuous, morphing textures by bodies with diverse movement backgrounds.

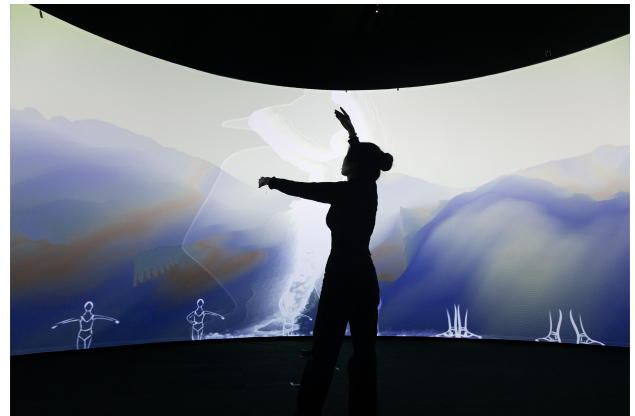


Figure 3: The silhouette of participant movements leaves traces in the ink.

3.3 Sonic-visual Interaction

The visual presentation is organized into two synchronized layers. A persistent background layer provides a stable spatial reference. It reflects the slow-changing qualities of the sound, including overall energy, spectral brightness, and density, enabling participants to perceive long-term shifts in the soundscape. Meanwhile, another silhouette layer mirrors the participant’s current movement and is situated on top as a simplified luminous form. Its blur, persistence, and trace length help display the motion continuity and help users recognize when they are producing gradual transitions versus discrete switches. At the lower edge of the projection, a compact set of ballet arm and foot position illustrations serves as mapping cues, echoing the current posture classification state and suggesting adjacent anchor postures that the participant can explore next. Together, these visual elements serve not only as an artistic narrative but also as practical feedback that supports learnability (Fig. 3).

In this setup, visual feedback helps participants understand the mapping while they move. The silhouette layer shows the continuity of bodily motion, while the lower posture cues indicate the current or nearby posture states. These visual elements make

the system easier to learn because participants can compare their movement with the sound changes. The audio-visual relation therefore supports the intercultural mapping at a practical level, where ballet-derived posture cues guide bodily navigation, while the evolving Guzheng soundscape gives immediate feedback on corpus selection and granular transformation.

4 System Implementation

The installation is conceived as an interactive audio-visual instrument in which ballet-inspired movements navigate and reshape a live soundscape featuring Guzheng textures. We aim to achieve intercultural coupling using interaction grammar rather than quantifying movement as generic data for sound control. Interculturality is therefore approached as a mapping mechanism: ballet-derived postures provide an interaction basis through which participants approach and move between regions of the Guzheng corpus. The Guzheng corpus is organized according to register, phrase contour, and instrumental techniques such as glissando based on the meaning of the posture, so that the sound material is not treated as neutral. Classification identifies the posture anchors, regression supports deviations and in-between states, and granular synthesis reshapes the selected material through buffer position, grain behavior, descriptor manipulation, and effects processing. This architecture supports two complementary actions: navigation through a culturally coded phrase corpus and continuous shaping of texture through granular synthesis, so that both expert and novice participants can engage the same interface with different levels of intention and interpretation (Fig. 4).

The technical workflow follows three sections: motion feature extraction, movement-sound correlation learning, and audio-visual rendering. A Kinect sensor first captures participants' body movements and transmits skeletal raw data to TouchDesigner, which acts as our system's central information-processing hub. Processed skeletal features are then sent via OSC to Wekinator, where models are trained to decode posture patterns using supervised ML. These outputs are returned to TouchDesigner and forwarded to Ableton Live, where they drive Granulator 3, a built-in granular synthesis engine and an effects chain. The resulting audio is finally routed back into TouchDesigner through a virtual audio interface, allowing sound and image to evolve together. Although the pipeline spans multiple applications, OSC messaging and audio routing are configured to behave as a single responsive instrument, with immediate sonic and visual feedback for each movement.

4.1 Motion Sensing and Feature Extraction

Motion sensing is treated as a means of extracting a simplified choreographic description that can remain legible in a public setting. We work with a reduced ballet-inspired vocabulary comprising six arm positions and five foot positions. To avoid overfitting and reduce the cognitive load of training, we track a compact set of joints and derive a limited number of channels from their three-dimensional coordinates, including indexes of three shoulders ("shoulder_left_tx/ty," "shoulder_right_tx/ty," "shoulder_centre_tx/ty"); two elbows ("elbow_left_tx/ty," "elbow_right_tx/ty"); the two hands ("hand_left_tx/ty," "hand_right_tx/ty"); and two feet ("foot_left_tx/tz," "foot_right_tx/tz"). Upper-body features provide the primary input to the learned mapping, while foot-related features are retained for direct parameter mapping. In the current prototype, each posture is recorded as a short held demonstration

lasting approximately 5–8 seconds. At TouchDesigner's default frame rate of 60 fps, this produces roughly 300–480 training samples per class. The relatively brief recording window was chosen to keep the training process fast and iterative, while also reducing the likelihood of posture drift during sustained holds. In parallel, silhouette masks are extracted from the same Kinect stream and integrated into the visual rendering pipeline, so the participant's body serves as both the interaction source and a visible part of the output.

Table 1: Mapping information of ballet postures and sound events.

Index	Sound event	Posture	Possible phrase index	Source track	Selection	Corpus size	Combo sample playback position
0	Sub bass	BIRS BAS	0-1	<i>Autumn Moon Over Han Palace</i>	00:32:17 - 00:45:17	13s	0%
			0-2	<i>Battle with the Typhoon</i>	01:13:01 - 01:17:21	4s20	17%
			0-3	<i>Lin Chong Flees by Night</i>	01:20:20 - 01:31:05	10s85	19%
1	Bass	FIRST	1-1	<i>Lin Chong Flees by Night</i>	00:29:92 - 00:42:92	13s	24.5%
			1-2	<i>Battle with the Typhoon</i>	01:25:23 - 01:30:01	4s78	32.6%
			1-3	<i>Autumn Moon Over Han Palace</i>	01:09:17 - 01:16:13	6s96	34.8%
2	Mid	SECOND	2-1	<i>Fisherman's Song at Dusk</i>	01:40:09 - 01:54:24	14s15	38.3%
			2-2	<i>High Mountains and Flowing Water</i>	01:14:03 - 01:28:71	14s68	43.3%
			2-3	<i>Lin Chong Flees by Night</i>	04:56:08 - 05:00:20	4s12	51.3%
3	Glissando	THIRD	3-1	<i>Lin Chong Flees by Night</i>	00:02:03 - 00:14:66	12s63	53.6%
			3-2	<i>Battle with the Typhoon</i>	00:39:14 - 00:51:21	12s07	59.8%
			3-3	<i>High Mountains and Flowing Water</i>	02:56:18 - 03:04:22	8s04	64.4%
4	Portamento	FOURTH	4-1	<i>Autumn Moon Over Han Palace</i>	00:00:00 - 00:14:94	14s94	68.8%
			4-2	<i>Lin Chong Flees by Night</i>	04:19:00 - 04:29:21	10s21	77.3%
			4-3	<i>Fisherman's Song at Dusk</i>	00:25:06 - 00:37:19	12s13	82.8%
5	High	FIFTH	5-1	<i>High Mountains and Flowing Water</i>	02:24:59 - 02:37:09	12s5	86.8%
			5-2	<i>Lin Chong Flees by Night</i>	02:27:02 - 02:37:00	9s98	93.5%
			5-3	<i>Battle with the Typhoon</i>	01:46:07 - 01:51:24	5s17	97.8%

4.2 Learning Movement-Sound Correlations

Movement-sound correlation learning is implemented through mapping by demonstration, aligning with established interactive ML practice in NIME [13]. Because ballet footwork can be complicated for first-time participants to reproduce, the learned models mainly focus on arm postural states. At the same time, foot movement remains available for additional continuous control to further shape the sound texture. The mapping is implemented as a hybrid strategy with two supervised models running in parallel. A classification model identifies the current postural state, providing clear boundaries for selecting phrase regions. At the same time, a regression model continuously controls navigation within the Guzheng phrase corpus, allowing intermediate movements to produce intermediate sonic outcomes. Although Wekinator provides several algorithms for classification and regression, we use a minimal configuration consisting of the default K-Nearest Neighbor Classifier and Neural Network Regressor.

The Guzheng phrases are organized as a corpus designed to preserve culturally situated sonic identity while enabling recombination through granular synthesis. We draw from five classical Guzheng pieces to provide a diverse corpus of register ranges and rhythmic profiles under different creative contexts with emotional expression.

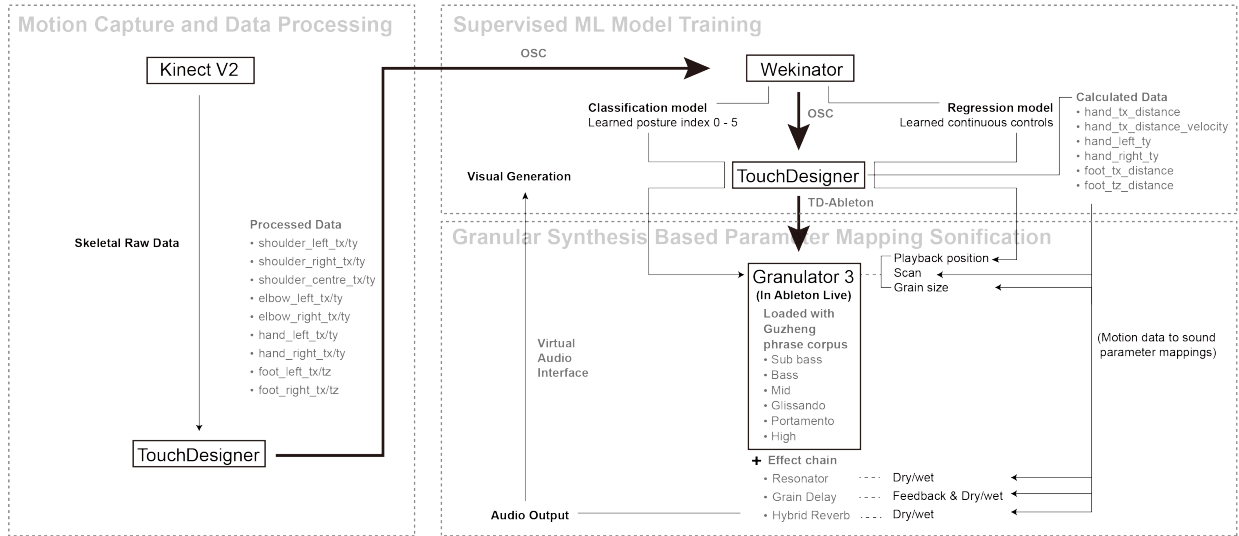


Figure 4: System overview.

- *Autumn Moon Over the Han Palace* (汉宫秋月) embodies the refined, melancholic character of traditional court music.
- *Lin Chong Flees by Night* (林冲夜奔) carries an urgent, dramatic intensity, evoking a hero's flight in the darkness and the themes of loyalty and tragedy.
- *High Mountains and Flowing Water* (高山流水) is often referred to as a symbol of harmony between human presence and the natural world, featuring the images of peaks and flowing streams.
- *Battle with the Typhoon* (战台风) surges with bold, unrelenting momentum, capturing the fierce struggle between the human and the overwhelming power of nature.
- *Fisherman's Song at Dusk* (渔歌唱晚) unfolds with a gentle, lyrical serenity, evoking the tranquil beauty of twilight over calm waters, near which people are working with lanterns flickering and lingering evening light.

In these pieces, we select short excerpts of 4-15 seconds, chosen for how their musical gestures resonate with the selected postural states. They are divided into six representative categories inspired by the six basic ballet postures: "Sub bass," "Bass," "Mid," "Glissando," "Portamento," and "High." Each posture is associated with multiple excerpts to expand the variety within a state. All excerpts are concatenated into a single combined buffer, the corpus. Mapping information, corpus size, and other technical details are summarized in Table 1. Unlike traditional corpus-based concatenative synthesis that treats the corpus as a descriptor-driven retrieval database, our instrument views navigation as a continuous traversal across a culturally meaningful sonic landscape. Playback is controlled by setting the read position within the corpus, while a simple selection rule introduces variation by randomly choosing among the excerpts associated with a detected posture. This corpus design is intended to support the sound contribution described earlier: Guzheng phrases and techniques are recontextualized as a reorganizable sonic library

that can be shaped publicly through granular synthesis, shifting from performance reproduction toward a continuously evolving sound ecology. The specific channel-to-control assignments used during training for each posture are summarized in Table 2.

Table 2: Mapping of hand movement data and sound effect parameters in Ableton.

Movement / spatial descriptor	Arm openness	Arm-opening speed	Left-arm vertical movement	Right-arm vertical movement	Horizontal foot separation	Front-back foot separation
Description	Measures how widely the participant opens their arms	Measures the speed at which the participant opens their arms	Measures the vertical displacement of the participant's left arm	Measures the vertical displacement of the participant's right arm	Measures the lateral distance between the participant's feet	Measures the front-back distance between the participant's feet
Sample range from the movement data	hand_tx_distance $\in (0, 1)$	hand_tx_distance_velocity $\in (5, 0)$	hand_left_ty $\in (-0.32, 0.8)$	hand_right_ty $\in (-0.32, 0.8)$	foot_tx_distance $\in (-0.2, 0.5)$	foot_tx_distance $\in (0, 0.2)$
Mapping sent to Ableton	Hybrid reverb dry/wet $\in (0, 0.95)$	Granulator 3 grain size $\in (500, 2000)$	Grain delay feedback $\in (0, 0.95)$	Grain delay dry/wet $\in (0, 0.95)$	Resonator dry/wet $\in (0, 0.95)$	Granulator 3 spread $\in (0, 12)$

4.3 Audio Engine and Visual Rendering

Sound generation is realized through granular synthesis, which assembles complex textures from many short grains while retaining perceptual traces of the source [25]. In Ableton Live, we use Granulator 3 as the primary engine and select several parameters that balance structural navigation and expressive shaping. The learned regression controls buffer playback position and additional timbral dimensions, while participants also shape texture through grain-level descriptors such as scan behavior, grain size, and spread. These parameters are chosen because they offer clear perceptual affordances. They can make the sound feel denser or more transparent, more stable or more volatile in pitch, and more continuous or more punctuated, without requiring musical training. To extend the palette and connect the sonic result to the spatial language of movement, we place a small post-effects chain after the granular engine, using resonance, delay, and reverberation to modulate harmonic focus, echo behavior, and perceived

distance. In combination, corpus navigation and descriptor shaping produce a sound field in which changes of orientation and reach can translate into changes in density, register, harmonic emphasis, and spatial impression.

The visual system is created in TouchDesigner to encourage participants to explore and improve legibility. Using built-in TouchDesigner functions, the performer’s silhouette is processed through edge extraction, feedback accumulation, noise fields, and displacement to produce an ink-wash-like aesthetic in which traces of motion accumulate and dissolve. The other base layer is driven by the live audio output stream, using band energies and amplitudes to animate the geometries that interact with the sound in real time. The resulting audio-visual scene provides the participant with a responsive ink-wash painting field that visualizes bodily action and the evolving soundscape, reinforcing the instrument’s premise of intercultural co-adaptation through shared interaction and perception.

5 Reflection and Preliminary Observations

At this stage, *Fluid Body* is positioned more as an artistic research system. Following practice-based approaches, the artifact itself, the making process, and the situated interaction with the artwork are considered as research materials through which embodied knowledge is generated, documented, and critically interpreted and reflected upon [5]. This also aligns with reflective practice, in which knowledge emerges through actions, adjustments, and reflection on the consequences of design decisions [29]. In the context of NIME, this framing is particularly relevant since recent work has argued that creative practice can function as a qualitative mode of evaluating movement-based musical instruments, especially when the research concerns learning, agency, and the relations between participants and the instrument, rather than task efficiency alone [8]. The observations below therefore clarify how the current system legibly behaves as a performable intercultural instrument, for example, how participants begin to infer the relation between posture and sound, and where the current mapping constrains further exploration. These observations were generated through iterative author-performer testing and participatory demonstrations in which visitors were invited to enter the installation and explore the instrument through movement.

5.1 Behavior Over Time

We conducted participatory demonstrations in which visitors were invited to enter the installation and explore the instrument through movement. When a participant holds a posture, the classified state typically remains stable, and the associated sound world settles into a consistent configuration. Minor variations remain in the granular parameters because of sensor jitter and micro-movement, but these are more likely to be perceived as subtle changes in texture than as disturbing variations. This leads to a two-level temporal dynamic. Small-scale movement within a posture creates detailed timbral variation, while more decisive postural changes reframe the entire sound field by transitioning to a different region of the Guzheng phrase corpus. This separation of scales allows for the maintenance of Guzheng identity traces while still allowing participants to dynamically influence the “air” surrounding each phrase through continuous control. At the same time, several mappings are strongly dependent on posture frames, so longer interactions could discourage participants’ willingness to explore once they grasp the system’s feedback, potentially decreasing movement diversity. This finding inspires

a more flexible mapping that maintains the system’s stability for public use while enabling a wider variety of movement qualities to have a meaningful impact on the sonic result.

5.2 Constraints and Prospects

The current system combines the basic ballet-inspired vocabulary with a certain number of Guzheng excerpts and a few synthesis and effects parameters. This is a deliberate limitation that helps with clarity, replicability, and artistic refinement and is in line with our goal of recontextualizing Guzheng technique and phrasing as a navigable sound corpus. However, it also reveals some challenges. First, although the regression model supports intermediate movement states, ambiguous or hybrid postures do not always become musically consequential enough. This means that the system sometimes privileges recognizable posture over movement quality. Second, the current multi-application system, including TouchDesigner, Wekinator, Ableton Live, and virtual audio routing, creates a practical overhead that may limit adoption by other artists or workshop facilitators. Third, the current observations remain grounded in artistic testing and public demonstration rather than in structured participant research. This is appropriate for the current phase, but it also limits what can be claimed about learnability, perceived agency, and intercultural legibility across different participant groups.

Future versions will tackle these problems in three ways. First, we will increase the corpus size and incorporate temporal models, such as dynamic time warping, so that movement sequences, not just static poses, can be identified and mapped to more extended musical patterns. This would enable the instrument to react to phrasing and transitions as compositional elements, further enhancing the experience of narrative co-creation. Second, we plan to lower the technical barrier to entry by integrating components where possible and offering preset templates that will make the system easier to replicate and modify for other applications. Third, we aim to develop an evaluation strategy that combines other qualitative approaches, such as structured listening tasks, interviews, and think-aloud protocols, to examine how participants experience agency, clarity, and expressive range differently as they navigate the corpus and shape texture. This could provide a more grounded form of validation, moving the project from a practice-based system towards a more design-oriented research artifact.

Overall, we consider the installation to be a form of artistic research rather than a finalized instrument. It explores the potential of interactive machine learning and sound synthesis to do more than automate control, instead facilitating a common interaction language between bodily technique, culturally situated sonic vocabularies, and computational media. In the longer term, we envision this approach as a transferable paradigm that could be applied to other movement grammars and culturally embedded sound corpora, enabling new forms of community-oriented, AI-assisted performances.

6 Conclusion

This work demonstrates how culturally situated movement and sound practices can be brought into real-time dialogue that prioritizes recomposition and public playability over reproduction. *Fluid Body* recontextualizes Guzheng phrase and technique vocabularies as a navigable sonic corpus for granular synthesis, and uses a ballet-inspired movement vocabulary as an embodied interface through which that corpus is explored, traversed, and

reshaped. By treating short Guzheng phrases as acoustic materials, the system supports an evolving soundscape that can be collectively interpreted as a shared sonic ecology.

Our main contribution is an intercultural interaction design embedded in sound synthesis in which corpus navigation and timbral articulation are coupled to movement in ways that remain legible to non-expert participants while offering expressive range to experienced movers and musicians. The hybrid classification-regression mapping approach provides stable posture recognition and facilitates continuous interpolation across intermediate postures, preventing participation from being reduced to a small set of correct poses. Recognizable poses construct a structural layer for accessing phrasing regions, while micro-gestures modulate granularity and processing descriptors that shape textural and spatial impressions. The soundscape preserves perceptible traces of the Guzheng's character while maintaining fluidity, responsiveness, and openness, allowing interpretation by diverse communities.

By emphasizing intercultural co-adaptation through interaction grammar and sound synthesis, *Fluid Body* offers implications for NIME design that address culturally situated practices across communities. It suggests that computational tools can facilitate forms of cultural dialogue in which distinct traditional aesthetics remain perceptible while being transformed through shared participation, positioning interculturality as a practice of co-presence and embodied translation. Future work will extend the corpus and explore how different participant groups negotiate agency and understanding over extended periods of participation, as well as how the system can be developed to be more straightforward for reuse and adaptation in other community-facing applications.

7 Ethical Standards

All development in this project is in accordance with NIME ethical standards. All participants involved in the participatory demonstrations were fully informed about the use of interactive technologies, the interaction procedure, and the purpose of the study. No body data were collected during the demonstrations, and we did not make any video or audio recordings. Our observations and subsequent discussion did not include any personally identifying information. Participants' consent was obtained prior to participation.

References

- [1] Seda Ayvazoğlu and Kerem Özcan. 2021. From court to theater in the 18th century: birth of the ballet d'action (dramatic ballet). *Journal for the Interdisciplinary Art and Education* 2, 1 (2021), 1–8.
- [2] Nicola Barbuti. 2018. From digital cultural heritage to digital culture: Evolution in digital humanities. In *Proceedings of the 1st International Conference on Digital Tools & Uses Congress*. 1–3.
- [3] Maros Suran Bomba and Palle Dahlstedt. 2019. Somacoustics: Interactive body-as-instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 95–100.
- [4] Nick Bryan-Kinns, Zijin Li, et al. 2020. Reimagining: cross-cultural co-creation of a Chinese traditional musical instrument with digital technologies. (2020).
- [5] Linda Candy and Ernest Edmonds. 2018. Practice-based research in the creative arts: Foundations and futures from the front line. *Leonardo* 51, 1 (2018), 63–69.
- [6] Baptiste Caramiaux, Alessandro Altavilla, Jules Françoise, and Frédéric Bevilacqua. 2022. Gestural sound toolkit: Reflections on an interactive design project. (2022).
- [7] Emily Carlson, Pasi Saari, Birgitta Burger, and Petri Toiviainen. 2020. Dance to your own drum: Identification of musical genre and individual dancer from motion capture using machine learning. *Journal of New Music Research* 49, 2 (2020), 162–177.
- [8] Doga Cavdir. 2024. Creative Practice as an Evaluation Method: A Case Study with a Movement-based Musical Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, SM Astrid Bin and Courtney N. Reed (Eds.). Utrecht, Netherlands. 456–464.
- [9] Doga Cavdir and Ge Wang. 2021. Borrowed gestures: the body as an extension of the musical instrument. *Computer Music Journal* 45, 3 (2021), 58–80.
- [10] Angela Loureiro de Souza. 2015. Laban Movement Analysis—scaffolding human movement to multiply possibilities and choices. In *Dance Notations and Robot Motion*. Springer, 283–297.
- [11] Rebecca Fiebrink, Daniel Trueman, Perry R Cook, et al. 2013. Muscular Interactions Combining EMG and MMG sensing for musical practice. (2013).
- [12] Vicente Espinoza and Javier Jaimovich. 2024. Resurfacing an Enactive Approach for Instrument Design: The case of the Tangible Granular Device. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 110–116.
- [13] Rebecca Fiebrink, Daniel Trueman, Perry R Cook, et al. 2009. A meta-instrument for interactive, on-the-fly machine learning. (2009).
- [14] Thomas Hermann, Andy Hunt, John G Neuhoff, et al. 2011. *The sonification handbook*. Vol. 1. Logos Verlag Berlin.
- [15] Luke S Hopper, Andries L Weidemann, and Janet Karin. 2018. The inherent movement variability underlying classical ballet technique and the expertise of a dancer. *Research in Dance Education* 19, 3 (2018), 229–239.
- [16] James Hutson. 2024. Digital Cultural Heritage Preservation. In *Art and Culture in the Multiverse of Metaverses: Immersion, Presence, and Interactivity in the Digital Age*. Springer, 99–141.
- [17] Carol Lee. 2002. *Ballet in western culture: a history of its origins and evolution*. Psychology Press.
- [18] Yanjun Li, Shazlin Amir Hamzah, and Nur Atiqah Tang Abdullah. 2025. Exploration of the Origin and Development of the Guzheng. *Journal of Ethnic and Diversity Studies (JOEDS)* 3, 1 (2025).
- [19] Polyxeni Mantzou, Xenofon Bitsikas, and Anastasios Floros. 2023. Enriching cultural heritage through the integration of art and digital technologies. *Social Sciences* 12, 11 (2023), 594.
- [20] Max Neupert. 2012. Navigating audio-visual Grainspace. In *Mensch Computer 2012 – Workshopband: interaktiv informiert – allgegenwärtig und allumfassend!?* Oldenbourg Verlag, München, 319–322.
- [21] Maria Rita Nogueira, Paulo Menezes, and José Maçãs de Carvalho. 2024. Exploring the impact of machine learning on dance performance: a systematic review. *International Journal of Performance Arts and Digital Media* 20, 1 (2024), 60–109.
- [22] Hanyu Qu, Francesco Dal Ri, Hao Zou, Hanqing Zhou, and Raul Masu. 2025. O-: An Epistemic DMI for Cross-Cultural Reflection on Time and Music. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 481–488.
- [23] Mingming Qu. 2017. The Development and Innovation of Chinese Guzheng Playing Techniques [J]. (2017).
- [24] MR Reshma, Balakrishnan Kannan, VP Jagathy Raj, and S Shailesh. 2023. Cultural heritage preservation through dance digitization: A review. *Digital Applications in Archaeology and Cultural Heritage* 28 (2023), e00257.
- [25] Curtis Roads. 2002. *Microsound*. The MIT Press.
- [26] Jana S Rošker. 2020. Classical Chinese Philosophy and the Concept of qi. *NEARCO-Revista Eletrônica de Antiguidade e Medievo* 12, 2 (2020), 116–134.
- [27] Mayank Sanganeria and Kurt Werner. 2013. GrainProc: a real-time granular synthesis interface for live performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 223–226.
- [28] Margaret Schedel, Phoenix Perry, and Rebecca Fiebrink. 2011. Wekinating 000000swan: Using machine learning to create and control complex artistic systems. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 453–456.
- [29] Donald A Schön. 2017. *The reflective practitioner: How professionals think in action*. Routledge.
- [30] Diemo Schwarz. 2012. The sound space as musical instrument: Playing corpus-based concatenative synthesis. In *New Interfaces for Musical Expression (NIME)*. 250–253.
- [31] Hugo Scurto, Baptiste Caramiaux, and Frederic Bevilacqua. 2021. Prototyping machine learning through diffractive art practice. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. 2013–2025.
- [32] Yuhua Shi. 2024. *An Exploration of Seven Tunes Heard in China by Bright Sheng for Solo Cello*. Ph. D. Dissertation.
- [33] Atau Tanaka, Federico Visi, Balandino Di Donato, Martin Klang, and Michael Zbyszynski. 2023. An end-to-end musical instrument system that translates electromyogram biosignals to synthesized sound. *Computer Music Journal* 47, 1 (2023), 64–84.
- [34] Pierre Alexandre Tremblay, Gerard Roma, and Owen Green. 2021. Enabling programmatic data mining as musicking: The fluid corpus manipulation toolkit. *Computer Music Journal* 45, 2 (2021), 9–23.
- [35] Agrippina Vaganova. 2012. *Basic principles of classical ballet: Russian ballet technique*. Courier Corporation.
- [36] Cadence Whittier. 2006. Laban movement analysis approach to classical ballet pedagogy. *Journal of dance education* 6, 4 (2006), 124–132.
- [37] Minjing Yu, Meng Zhang, Chun Yu, Xiaoguang Ma, Xing-Dong Yang, and Jiawan Zhang. 2021. We can do more to save guqin: Design and evaluate interactive systems to make guqin more accessible to the general public. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [38] Yan Zhang, Muquan Li, and Shuang Pan. 2024. Deep Learning-based Emotion Recognition Algorithms in Music Performance. *Scalable Computing: Practice and Experience* 25, 6 (2024), 4712–4719.