Visually-Led Design for Gestural Audiovisual Instruments

Sam Trolland
sam.trolland@monash.edu
SensiLab, Monash University
Melbourne, Victoria, Australia

Alon Ilsar
alon.ilsar@monash.edu
SensiLab, Monash University
Melbourne, Victoria, Australia

Jon McCormack
jon.mccormack@monash.edu
SensiLab, Monash University
Melbourne, Victoria, Australia

Figure 1: The Final Moment of the Final Scene from the *Cymbalism* Performance

Abstract

In this paper we present our visually-led design method for creating gestural mappings in a new audiovisual percussion work titled *Cymbalism*. Unlike most audiovisual works, *Cymbalism* was inspired by the creation of a series of interactive visual scenes that respond to the performer's real-time movements. In leading with the visual interaction, we discuss how this approach fostered a union between the physical, audio and visual elements of the work, creating a performance where the visualisation is not simply a feedback mechanism but fundamental in inspiring compositional concepts and new ways of interacting with sound. Through practice-based research, we use the insights gained through creative development and performance outcomes to guide the continued evolution of an established wearable gestural DMI.

Keywords

Gesture, Audiovisual, Instrument, Percussion, Performance

NIME '25, June 24–27, 2025, Canberra, Australia

1 Introduction

Recent advancements in the technology that enables wireless wearable gestural control over digital music making and real-time graphics have opened new avenues for audiovisual performance and expression [8, 24, 29, 33, 36, 43, 44]. Integrating movement, sound and visuals into instrument design introduces an additional dimension to the 'mapping problem' discussed in [26]. In a previous NIME paper, the second author, a DMI designer and percussionist, suggested tackling the mapping problem by leading with one of these elements; for example sound-first mapping [18], or movement-first mapping [10–12, 14, 19]. For the collaboration outlined in this paper, we decided to lean into the expertise of the first author as an engineer and interactive visual designer, acknowledging that "[t]he principles in a true collaboration represent complementary domains of expertise. As collaborators, they do not only plan, decide, and act jointly, they also think together, combining independent conceptual schemes to create original frameworks" [22]. Adopting a 'visual-first' approach to mapping design allowed the collaborators to play into the strengths of the first author, not only as an interactive visual artist [43], but also as the lead developer of the *AirSticks* gestural instrument used in this collaboration [42].

In the technical implementation of the software underpinning the *AirSticks*, visualisations were primarily employed to expose the inner workings of the instrument, with a focus on function or musical-accompaniment rather than serving as a primary means

of creative expression. The visually-led design process employed in this work allowed the first author to drive the collaboration through creative visualisation of the gestural data, which inspired new movement-to-sound mappings through the introduction of an intermediate visual-mapping layer [30]. By implementing this intermediate visual-mapping layer in a 3D game engine, complex simulations are used to transform the raw gestural data (from one or more *AirSticks*) into higher-dimensional information [30], resulting in novel mechanisms for inspiring and generating sound with movement [16, 23].

This work distinguishes the mapping design space and the resulting instrument mappings. The design space was visually-led; driven by the initial development of visual prototypes, which subsequently inspired movements and sonic mappings. This design process resulted in three audiovisual instrument mappings that translate movement to sound via an intermediate visual mapping layer. Both the visually-led design space and the resulting instrument mappings demonstrate the complex web of entangled relationships involved in integrating technology in practice; highlighting the active agency and affordances of new instruments in shaping musical expression.

In this paper, we present the practice-based research process through which these novel audiovisual mappings emerged, highlighting the improvisational and compositional affordances that arose. Our findings are presented through discussion and video-documentation of a 10-minute gestural audiovisual work titled *Cymbalism*. Quotations from a semi-structured interview between the authors, combined with the use of video documentation, serve to communicate the tacit knowledge embedded in the physical experience of playing the instrument [1, 3, 4, 25].

2 Background

2.1 Related Work

Within the NIME community, the term gesture is used broadly to describe "a motion of the body that contains information" [45]. In this paper, we discuss mappings with a gestural "open-air" instrument [35], using the term 'gestural data' to refer to the continuous stream of movement information provided by motion tracking sensors. We use this gestural data to create "audiovisual compositions", described by Harris [15] as a process of composing while giving equal weight and importance to both elements in the integrated outcome of the work.

Typical approaches to audiovisual works give primacy to sonic media, using visualisation to add extra information, expressivity and interest to musical content [5, 6, 34]. In [9], touch-screen interactions with visual information expose the internal processing of 'audio morphing'. Other works focus on audio and visual expression simultaneously, however the expression of each medium remains independent [13, 16, 34, 37]. In [37] audiovisual expression is achieved through corpus-based audio samples and still images, utilising a touch controller to navigate the independent domains. Connected audio and visual generation is achieved in [13] through interconnected feedback systems. In [36], musical gestures are linked to visual effects, approaching audiovisual composition through narrative-driven story-boarding that achieves cohesion through a unifying theme.

In [34], Perrotin discusses their music-first approach while also imagining a reverse process based on drawing visual feedback, with music added afterwards to enhance expressivity. In both of these approaches, primacy is given to one form of media with the other media seen as following afterwards. Momeni and Henry [30] outline three categories of interaction between sound and image in real-time work: (1) sound-to-image, (2) image-to-sound, or (3) concurrent generation of sound and image. Momeni describes an implementation of the third category by extending the traditional mapping with additional dynamic, independent, visual-mapping layers. These additional layers, inserted between the controller and the instrument, visually represent the internal mapping algorithms to enhance exploration and expressive control of the system. Momeni highlights that this concurrent approach enables complete synchrony between audio and visual media, creating a rich and perceivable link between the two domains.

This concept of an intermediate mapping layer is extended in [16, 23], where simulations of physical process mediate between gestural input and audiovisual output. A key distinction is that in [16] the sound produced is directly linked to the performer's gestures, whereas in [23] the performer's gestures are transformed by a fluid-simulation layer, with the higher-dimensional simulation output used for audio synthesis.

2.2 Project Background

The *AirSticks* are an ever-evolving gestural instrument that has been developed over 6 years by a team of researchers at *SensiLab*.[1] The second author is a professional drummer and inventor of the *AirSticks*, creating the original implementation of the instrument in 2012 [17, 21]. With the engineering expertise of the first author, the instrument became wireless and handheld, evolving from a musical to an audiovisual instrument [42]. This wearable instrument uses an Inertial Measurement Unit (IMU) to measure orientation and acceleration data, which is wirelessly transmitted to a nearby computer for further processing and audiovisual generation. A detailed technical overview of the *AirSticks* can be found in [42].

The instrument has been developed through an iterative process involving hundreds of public performances, many of which have involved both authors [2, 38–40]. The development of hardware, software and gestural mappings for these performances has been pivotal in advancing the *AirSticks*, with a focus on maintaining its status as an Accessible Digital Musical Instrument [20]. Through this collaboration, we leverage and extend the capabilities of the *AirSticks* to create a performance with tightly integrated audiovisual interactions.

In this work, we developed an intermediate visual simulation layer between the raw gestural data and audio generation. This layer, built in a 3D game engine (Unity),[2] takes an input stream of gestural data via the Open Sound Control protocol (OSC) [46] to influence movement parameters of virtual objects within the simulation. Interactions between these virtual objects, such as proximity and collision detection, are visually rendered and simultaneously transmitted as discrete and continuous MIDI events [28] for audio generation in Ableton Live.[3] The input stream includes 3D orientation (yaw, pitch, roll) and 3D linear-acceleration (x,y,z) gestural data from two wrist worn *AirSticks*. Movement '*Energy*' is calculated using equations from [42] by summing the magnitude of the acceleration vector over a time-window. The *Energy* value smooths spikes in the acceleration data and can be used to determine periods of stillness or sustained movement, regardless of direction.

[1] SensiLab: https://sensilab.monash.edu
[2] Unity: https://unity.com
[3] Ableton Live: https://www.ableton.com/live/

Through this collaboration we continue our performance-inspired approach, using performances to guide the evolution of the technology, and performance deadlines to avoid creative choice "paralysis" [27], which can occur when creating new musical mappings in a DAW with near-limitless mapping possibilities. The insights gained from this collaborative creative process will inform future revisions of the *AirSticks* technology and future approaches to making new work with the *AirSticks*.

2.3 Motivations

Our motivations for this collaboration arose from the intent to harness the technical and creative strengths of the first and second authors. The first author's intimate knowledge of the gestural data and their expertise in visualising this data, both functionally and creatively, led to our first key motivation;

> *How can interactive visualisations inspire new ways of improvising, composing and performing live music?*

The second author's acoustic percussion and *AirSticks* practices have begun to merge in recent years, but have often remained separate; that is, both digital and acoustic instruments may be used in performances, but not simultaneously. The recent advancements in the *AirSticks* technology that now fit in a glove have allowed for a merged instrument practice; the percussionist can play acoustic percussion at the same time as the *AirSticks*. This led to our second key motivation;

> *How can we best develop a system that allows the percussionist to perform naturally and combine digital and acoustic instruments?*

Guided by these motivations, we set out to create a unified audiovisual instrument that the second author could confidently improvise with, playing percussion naturally without explicitly focusing on also 'playing' the *AirSticks*. A visually-led design method was explored to help overcome the "creative paralysis" [27], but also as an experiment to see how it would shape the second author's movement and sonic decisions in both improvisation and composition. This approach allowed the visualisations to have more influence over the improvisation, composition and performance of the work, while still representing a clear and transparent relationship between the movement and sound.

3 The Design Process

Cymbalism emerged through an iterative cycle of development, evaluation and reflection, based on Research-through-Design principles [47]. This process included nine rehearsal sessions over six months, with artefact development periods between sessions. During rehearsals, the first author (henceforth referred to as 'the visual artist') introduced a new interactive visualisation, or 'scene', for exploration by the second author (henceforth referred to as 'the percussionist'). Three 'performance scenes' emerged, forming the backbone of the work and inspiring new sonic mappings. The performance was then composed through improvising with the interactive audiovisual system.

3.1 Visual Inspiration from Compositional and Musical Elements

The collaboration began with a preliminary meeting to outline our motivations and define the creative scope at the intersection of percussive gestural performance and audiovisual interaction design. The visual artist was motivated to personalise the interactive system to accommodate the percussionist's 'natural' style of movement when playing the drums. This approach aligns with the concept that drum-kits, and particularly cymbals, are personalised; they are set up at different heights and in different positions according to individual preferences.

Inspiration was drawn from discussions on musical elements used in solo drumming performances, including tempo, straight and swung rhythm, loudness, regular and dynamic accents, legato and rests. Phasing was highlighted as a key concept for the percussionist, stating: "as a solo drummer you are working with time, and one way of influencing time is phasing". Simultaneous gestural data and video recordings of these drumming techniques were captured without digital augmentation, allowing movement ideas to flow uninfluenced by the technology. Compositional ideas for integrating these percussive elements into the performance were discussed, along with visualisation concepts to pair with different motifs. These discussions laid the foundation for the work, inspiring the creation of the first visual scenes.

3.2 Visually-Led Development

In response to the compositional and percussive elements, the visual artist created a series of 'prototype scenes' to embody these concepts using the percussionist's recorded movements. At this stage the prototypes were visualisations only and did not produce any sound. Each scene was designed to inspire new ways of interacting with digital and acoustic instruments by either:

(1) embodying a musical concept through visual feedback mechanisms, or
(2) modelling a simulation with simple behaviours that combine to produce complex interactions.

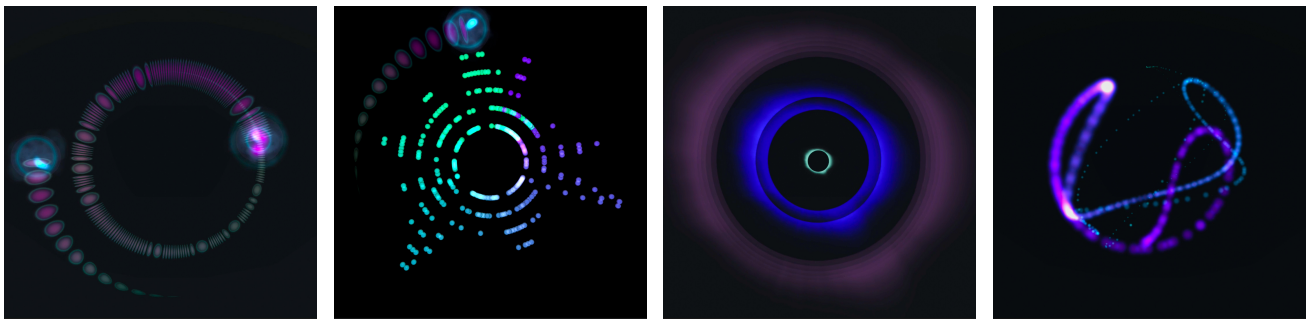
3.3 Exploration and Instrument Mapping

The rehearsal sessions provided an opportunity to explore and evaluate the prototype scenes. During these rehearsals the percussionist would start by 'playing' the visualisations, wearing the *AirSticks* and learning how the visual scenes responded to various gestures such as strikes, shakes, pokes, short and sharp, or sustained movements. After exploring the visualisation space and the movements it inspired, the percussionist would experiment with musical mappings to 'hear' what the visualisation might sound like.

The musical mappings that emerged were paired with each scene, resonating with both the visual aesthetic and the style of movement the visualisation inspired. This process of exploring the sound space of each scene often inspired new interaction and visualisation ideas to extend the system. Rehearsal sessions concluded with a discussion about the creative affordances of each scene and how they influenced compositional decisions or inspired improvisational elements.

3.4 Visual Prototype Scenes

Five prototype scenes were developed and iteratively refined throughout the collaboration, each inspiring a different set of movements and styles of play. Often, these interactive visual scenes inspired movements and compositional elements that differed from how the visual artist expected they would be used. Guided by exploration and discussion, the prototype scenes were expanded, merged, and shaped into the three performance scenes. Some prototypes were discarded after exploration, as the style of movement and composition they inspired did not align with our



(a) Orbiting Planets (b) Standing Waves (c) Emerging Eclipse (d) Orientation Heat-Map

Figure 2: Visual prototype scenes representing rhythm with (a) orbital speed, (b) orbital distance and (c) distance separation. (d) represents movements through the orientation space, glowing white in regions of increased activity.

musical and interaction aims. The insights gained from creating, exploring and discarding prototypes fed back into refining the remaining scenes.

3.4.1 Orbiting Planets. In the 'Orbiting Planets' prototype (Figure 2a), two planets revolved around the centre of the simulation, with *Energy* from each hand independently controlling the orbital speed of each planet. When *Energy* was applied through a strike or shaking gesture, the planet would accelerate and leave an elongated trail. The trail, elongating and contracting in response to phrases of higher or lower *Energy*, visually represented the rhythm of each hand in a short historical window. When exploring this prototype, the percussionist expressed:

> "It feels like a task . . . for me to figure out ways to make the visual do certain things . . . like catch up, keep them together in *Energy*."

Although the visual representation of rhythm was promising, this scene did not make it into the final performance. However, the concept of simultaneously representing the *Energy* of each hand evolved into an interaction that combined the *Energy* of both hands. This concept of combined *Energy* became a key interaction mechanism in the 'Morphing Sequencer' performance scene (outlined in subsection 4.3).

3.4.2 Standing Wave. Similarly, the 'Standing Wave' prototype (Figure 2b) also represented the *Energy* from each hand with two orbiting entities. However, in this prototype the *Energy* from each hand caused a spike in the orbital distance. While only the representation of *Energy* differed, and the interaction mechanism remained unchanged, this prototype inspired a very different style of movement. A key difference was that each orbit now had a constant period, with the percussionist noting:

> "What's really interesting about this one is . . . my instinct was to try and hit at the top every time, and then do a phrase, and then hit again at the top."

3.4.3 Emerging Eclipse. The 'Emerging Eclipse' visualisation (Figure 2c) responded to *Energy* by generating rings that emerged from the centre of the screen. These rings expanded at different rates based on the *Energy* of each strike, representing the timing of strikes through the distance separation of the rings. Three tiers of 'Strike Energy' were established: soft strikes created slowly expanding rings, while higher-energy strikes produced rings of different colours that expanded quickly and overtook slower rings.

This differed from the earlier prototypes as it reacted to discrete gestures rather than a continuous representation of *Energy*. The strike gestures were identified using a new dynamic linear acceleration triggering algorithm, extending the triggering system described in [42]. In this dynamic triggering system, the threshold for detecting a discrete note event adjusted based on the recent linear acceleration history. As shown in Figure 3, this enabled the system to adapt by lowering the threshold during periods of softer strikes and raising it during periods of harder hits and faster movements.

When exploring this prototype, the percussionist found that after a big hit, he could softly play acoustic drums below the falling dynamic threshold. The visual artist initially considered this an error to be corrected, but the percussionist saw performative potential, stating:

> "Compositionally that's great, we can do things on the drums that it doesn't pick up."

This prototype was extended in later rehearsals, becoming the 'Planes of Play' performance scene (outlined in subsection 4.2).

3.4.4 Orientation Heat-map. Two other prototypes focused on plotting the percussionist's movements through the orientation space, highlighting regions of increased activity where significant time and *Energy* were spent. The first of these prototypes, the 'Orientation Heat-map' (Figure 2d), displayed a short history of the path through the orientation space, glowing 'hotter' in regions of increased activity. This direct visualisation of gestural data was initially perceived by the percussionist as "too clear . . . too literal".

A second orientation prototype was developed, incorporating a delay that allowed the visual path to build and then retract (Figure 6). The percussionist observed that this interaction began to "feel a bit more like call and response", further elaborating that it "reminds me of a tape measure that you pull out then it snaps back in . . . the shorter the phrase, the quicker it repeats". This concept later inspired the 'Gestural Looper' performance scene (outlined in subsection 4.2), featuring a looper where "gestures are recorded when performed, then looped when still". This novel style of interacting with music was highlighted by the percussionist, who noted:

> "I've never dealt with it in that way, of changing delay time based on the phrase length."

This was a pivotal moment in the development of the work, with the percussionist highlighting "[this] is what we wanted this

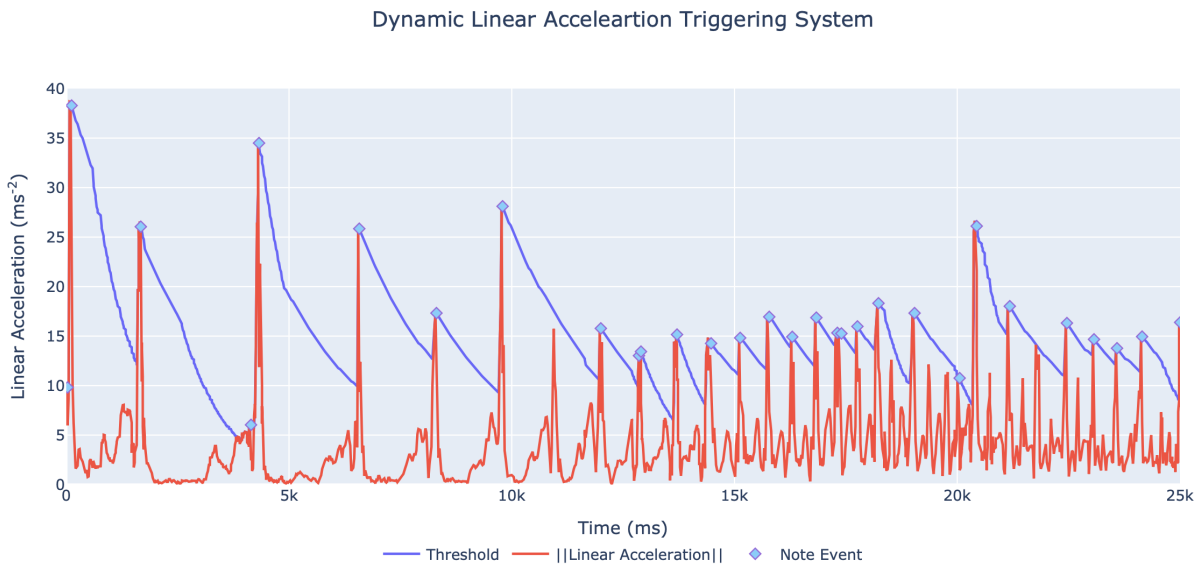


Figure 3: Dynamic Linear Acceleration Triggering System identifying discrete note events in drumming gestures with varied tempo and strength.

project to be about: what music is inspired by the visuals that *you* [the visual artist] make?"

3.5 Cohesion through Creative Themes

After the open exploration of movements and musical mappings inspired by each visual scene, the next pivotal moment in the development involved narrowing the possibility space to focus on elements we wanted to take forwards into the performance. This turning point was achieved by deciding on a creative theme for the work: a set of four cymbals. Establishing this unifying theme was led 'completely visually', inspired by the idea to incorporate the visual textures of the cymbals into the existing prototype scenes.

The cymbals were sampled as both photographic and sonic textures, creating a tight cohesion between the audio and visual media depicted in Figure 4. Concentrating the audio, visual and acoustic interaction to just four cymbals inspired the decision to project the visualisations down onto the cymbals from above. As the percussionist describes, "the limitation was liberating", elaborating:

> "We had so many options before that, but then as soon as we sat down and went, no, this is now on the floor, two *AirSticks*, four cymbals, projection onto the floor, which was great because it meant I could relate to [the digital instruments] and the acoustic instruments at the same time."

4 The Performance

Cymbalism is an augmented reality percussion performance that weaves interactive visualisations, live percussion and digital echoes of sampled instruments. The performance comprises three distinct scenes, each introducing a new mapping of movement to sound through an intermediate visual-mapping layer. Throughout the performance the percussionist is seated on the floor, surrounded by four cymbals arranged in a semi-circular configuration. Projections are cast down from above onto the stage, the cymbals and the percussionist. The percussionist is wearing two *AirSticks* on the top side of each hand in fingerless gloves. Video documentation[4] of the performance is available online [41].

4.1 First Scene: 'Planes of Play'

In the first scene an invisible horizontal plane delineates the boundary between the acoustic and digital realms of the instruments. When the percussionist interacts with the cymbals below this plane he engages them acoustically, with each strike triggering visualisations that emanate from the cymbal's centre (Figure 5a). These visualisations consist of photographic textures of each cymbal, visually sampled and layered with each strike gesture, creating a dynamic and immersive visual experience.

Additionally, the audio textures consist of prerecorded samples from the same cymbals in the same room. When the percussionist raises their wrist above the horizontal plane, they modulate a recording of the cymbal's resonating sound. Discrete strike gestures in the air above each cymbal further augment the resonance by triggering additional samples that reveal the initial attack of the cymbal. The auditory spatialisation corresponds to the physical arrangement of the cymbals; the audio textures of the cymbals on the percussionist's left and right are panned hard left and right, while the two central cymbals are panned soft left and right.

4.2 Second Scene: Gestural Looper

The percussionist flips the cymbals over and positions them closer to his body and the centre of the projections to commence the second scene. This scene introduces a *Gestural Looper* and an arpeggiator with varying pitch and velocity. A gesture that raises *Energy* above a threshold initiates the recording of a gestural loop. When *Energy* falls below the threshold, the loop ceases

[4] *Cymbalism*: https://bridges.monash.edu/articles/media/Cymbalism/28792673

(a) Planes of Play (b) Gestural Looper (c) Morphing Sequencer

Figure 4: Performance scene visualisations incorporating photographic textures of cymbals.

(a) Planes of Play (b) Gestural Looper (c) Morphing Sequencer

Figure 5: The three performance scenes from *Cymbalism*.

recording and immediately commences playback on a separate MIDI channel, allowing the percussionist to improvise over the loop acoustically. This process operates independently for both hands, creating a total of four virtual instruments: two live and two looped.

The initiation of each gestural recording and the playback of each loop are not synchronised to any clock, occurring whenever the *AirSticks* cross the *Energy* thresholds. This results in a section characterised by chaotic and unpredictable rhythms that progressively build in energy and phase against the quantised rhythm of the arpeggiated samples. The arpeggiator processing each cymbal's sample can be further manipulated by the percussionist's movements. The stereo panning is preserved, with the live and looped rhythms from the left hand panned left, and those from the right hand panned right.

The visualisation directly maps the percussionist's movements through orientation space, generating a trail that follows the movement of each hand and each gestural loop (shown in Figure 5b). The scale of the trail is mapped to the *Energy* of the *AirStick*, with faster movements generating a more substantial visual impact. Thus, the visual path of each channel represents a short historical window of orientation and *Energy*.

4.3 Third Scene: Morphing Sequencer

The final scene explores the concept of a morphing rhythm controlled by a visual sequencer. A 'playhead' orbits the percussionist, triggering audio and visual samples of the cymbals as it passes each beat in the sequencer (Figure 5c). The sequencer has three beats per loop: one fixed at the top of the bar, while the second and third beats morph based on the orientation of the percussionist's hands. As the hands move, the positions of the second and third beats shift accordingly. The high fidelity of the *AirSticks* orientation tracking facilitates 'getting out of the grid' by allowing the second and third beats to occupy any position on the circle with a precision of 1/100th of a degree.

This is the first mapping where the authors combined the *Energy* of both hands to collectively control parameters. This combined *Energy* influences the pitch of the triggered samples and the overall tempo of the playhead, with increased *Energy* accelerating the playhead and raising the pitch.

When exploring the affordances of the *Morphing Sequencer*, the percussionist discovered that quick staccato movements would cause a temporary spike in the playhead speed and pitch of the cymbal samples. This created a unique musical interaction; as the timbre of the cymbals morphed they began to sound like 'little bells'. The percussionist appreciated the musical affordances of the combined *Energy*, stating:

> "It was great to have overall energy like that. I felt I could just play and not worry about one hand leading more than the other. That discovery of being able to have one hand spike and the other hand keep it down was really nice as well."

5 Discussion

This section revisits the motivations established at the commencement of this collaboration, reflecting on how the performance, creative process and technology evolved to meet these motivations.

5.1 Influence of the Visually-Led Design

Our first key motivation was to investigate how interactive visualisations can inspire new ways of improvising, composing and performing live music. Our approach was to 'compose through improvisation' [7]; beginning the development of each scene by 'playing' the visualisations first, allowing the gestures to inspire musical mappings, and then improvising with these mappings until a composition took shape. This method enabled the visual artist to influence the entire process by providing expressive visual representations of the percussionist's movements.

5.1.1 Creative Data Visualisation. Each visualisation inspired the percussionist to move differently, generating new musical interactions that were iteratively refined. The percussionist noted the significant impact of the "relationship of the visualisation influencing the mapping of the sound of the *AirSticks*", particularly the "extension of the data in ways that are not physically possible in the natural world". Through the intermediate visual mapping layer, "*Energy* is being represented visually in these ways that are physically impossible, or not in the reality of playing the percussion instrument acoustically". The percussionist describes that "the visual content inspired the sound", noting the compositional decision to use:

> "[the] energy of the two sticks together to change the pitch and speed of the tape head ... [was] a really strong point, one moment where that conversation between the visuals and the sound was really clear and it really changed the way the instrument worked."

5.1.2 Visualising Energy vs. Orientation. During the exploration of each prototype scene, it was observed that some interactions felt more like tasks rather than instruments, while others allowed compositional elements to emerge. Mappings that responded to *Energy* were found to inspire movements akin to shaking or small strikes in the air or on the knees. In contrast, mappings that responded to orientation inspired more dynamic movements, with both hands traversing the entire orientation space.

The percussionist found that the three prototypes responding to *Energy*, but not orientation, were "best in intensifying the rhythms in an interesting way". The effect is that "it inspires to hit different rhythms with different velocity, [however] because there is no tracking, it doesn't inspire to move in [the yaw axis]". We found that these prototypes "work well as a shaker", displaying a short history of the rhythmic gestures that highlight the timing and periods of increased and decreased activity, and the pauses between them.

The visual artist observed that the orientation mappings inspired the percussionist to explore the entire orientation space, whereas in the *Energy* mappings, the percussionist only played in a condensed region in front of them. Even though the data from each hand was independent, we found that the orientation mappings inspired two-handed gestures, posing with each hand at different heights and angles to 'fill up' distinct regions in orientation space as shown in Figure 6.

5.1.3 Visually Inspired Creative Theme. The unifying theme of cymbals was another element that emerged from the visual development to influence the musical mapping. The percussionist described the influence of portraying visual textures of the cymbals, stating:

> "The fact that we took photos of the cymbals inspired us to then sample the cymbals and keep everything contained within that [audiovisual] world."

This decision significantly influenced the direction of the work by providing a unifying theme that brought cohesion to the audiovisual composition (in similar approach as [36]), creating a "rich and perceivable link" [30] between the generated audio and visual media.

5.2 Affordances for Natural Interactions

Our second key motivation was to create a system enabling the percussionist to perform naturally while integrating digital and acoustic instruments.

5.2.1 Digital and Acoustic Affordances. This was achieved in various ways across the three performance scenes:

(1) In *Planes of Play*, an invisible horizontal plane divided the digital and acoustic realms, allowing the percussionist to alternate between them by varying the angle of his wrists.
(2) In *Gestural Looper*, the division was based on *Energy* rather than angle, with higher *Energy* gestures triggering digital loops that could then be improvised upon acoustically.
(3) In *Morphing Sequencer*, combined orientation and *Energy* produced a morphing three-beat sequence, laying a digital rhythmic foundation for further acoustic improvisation.

5.2.2 Physical Affordances. The physical affordances of the wearable instrument further contributed to produce an instrument the percussionist could perform with naturally, without explicitly focusing on the movement of the *AirSticks* themselves. We observed that the placement of the *AirSticks* introduced different affordances or limitations when worn on the hands or attached to a drumstick.

When attached to a drumstick the percussionist describes enhanced movement capabilities, allowing him to "to twirl it, move it more away from the palm," and subtly adjust the orientation by varying how he holds the stick. Additionally, situating the *AirSticks* on the end of a drumstick provides more nuanced linear acceleration data as the sensor is positioned further from the pivot point in the hands and wrists. However, this configuration also imposed restrictions, with the percussionist describing:

> "Anytime I try to hit something while there's an *AirStick* attached to a stick, it feels like I can't play it as well, and I don't want to hit as hard ... Having them on sticks and trying to hit things doesn't really work."

To accommodate the significant acoustic percussion throughout the performance, a compromise was made to wear the *AirSticks* inside gloves on the back of the hand. The percussionist describes that this placement was "as far down the body that we can put the *AirSticks* without putting them on the stick", maximising the quality of the available gestural data while maintaining acoustic playability.

5.2.3 Spatial Affordances. Projecting the visualisations onto the percussionist, cymbals, and stage ensured that the percussionist was immersed in the visual world of the instrument. This direct projection mapping aligned the orientation space of the real and virtual worlds, providing crucial visual feedback that allowed the percussionist to adjust for yaw-drift and confidently strike the intended virtual cymbals.

It is important to note that in the first and third scenes, the musical mappings were heavily influenced by the yaw orientation of the *AirSticks*, which can be prone to 'drift' [31, 32]. The visualisations in these scenes could be used to identify and adjust for yaw-drift, with the percussionist noting "I would think about the *AirSticks* in these moments where I needed to hit a certain cymbal".

The spatial affordances of the immersive projections mitigated the *AirSticks* lack of positional information by placing the percussionist in the centre of the virtual world. The percussionist's

Figure 6: Two-handed gestures inspired by the 'Orientation Mapping with Delay' in a rehearsal session.

seated placement limited positional variation, ensuring orientation alignment was not significantly altered by changes in position. These combined affordances allowed the percussionist to focus exclusively on the music during the performance, without needing to consider what the visualisations or the *AirSticks* are doing.

5.2.4 The Duality of Adapting Instruments. Although we set out to create an interactive system that was adapted to the percussionist's 'natural' style of movement, throughout the creative process we observed that percussionist's movements also adapted to the creative themes and compositional elements that arose. Reflecting on this, the percussionist described the process as personalising the instrument to his style of movement in this particular performance:

> "I think we adapted the system to make it work better for the piece ... [to] my style of movement within the context of sitting down on the floor and playing cymbals."

This highlights an interesting tension point between adapting the instrument to the performer, and the performer adapting to the instrument. The percussionist describes this point of tension:

> "You either play it as it lands, you go, okay, this is the visualisation - What do I do within this context? Or you start to question it and change it and tweak it and bring different things in."

The percussionist describes that "we found a nice balance" by alternating between these two approaches throughout the collaborate development of the work.

6 Conclusion

Reflecting on the motivations behind this practice-based research, we aimed to develop a performance that showcases the current capabilities of the *AirSticks*, while allowing the insights gained through practice to inform development and further extend the technology. We approached this aim through the development of a series of interactive visualisations, exploring how different styles of interaction could influence the composition through improvisation. As outlined, this was an iterative process of exploring, refining, and sometimes discarding prototypes before finalising the three performance scenes. Through creative exploration led by 'playing' the visualisations, novel movement movement-to-sound mappings emerged, including a new dynamic note triggering system, a *Gestural Looper*, and a *Morphing Sequencer*.

The visually-led design method discussed in this paper has inspired new ways of improvising, composing and performing live music through creative data visualisation. By developing a system that enabled the percussionist to perform naturally while integrating digital and acoustic instruments, we have explored the physical and spatial affordances of our audiovisual instrument, and addressed the dual nature of adaptation between the performer and the instrument. These affordances, which shape and are shaped by practice, illustrate the multilayered contexts in designing new instruments that cohesively entangle audio and visual elements with digital and acoustic expression.

7 Future Work

Through this collaborative development, the authors combined their professional skills and creative practices to produce a new gestural audiovisual performance. We aim to continue developing this work, using this 10-minute version as a 'Proof of Concept' for further expansion and refinement.

In the next stage of development, we plan to extend both the technology and the creative concepts behind the work. Technical improvements will focus on fine tuning interaction parameters and exploring additional technologies to integrate alongside the *AirSticks*, addressing yaw-drift and adding positional tracking to complement the existing orientation tracking. Creatively, we are motivated to delve deeper into the theme of cymbals, exploring their history, the process of making them, and their incredibly rich and complex sound properties. Future development will also expand the set design for a larger stage, considering additional scenes with both standing and seated arrangements, and the inclusion of multiple percussionists.

8 Ethical Standards

This paper complies with the NIME ethical standards and all research activities were conducted in accordance with the ethical standards of Monash University and the national code for the responsible conduct of research.

Acknowledgments

We acknowledge the team at *SensiLab* that have contributed to the *AirSticks* project, including Alon Ilsar, Sam Trolland, Elliott Wilson, Ciaran Frame, Steph O'Hara, Jonathan Carter and Jon McCormack.

References

[1] Michael Biggs. 2000. the foundations of practice-based research. *Working papers in art & design* (2000).
[2] BoilOver, Ciaran Frame, and Sam Trolland. 2022. Buoy. Theatre Performance at Bowery Theatre, St Albans, 2022-06-03, 60min, Performed By: BoilOver Theatre Ensamble.
[3] Hendrik Anne Henk Borgdorff. 2012. *The conflict of the faculties: Perspectives on artistic research and academia.* Leiden University Press.
[4] Linda Candy. 2006. *Practice Based Research: A Guide.* Technical Report 1. Creativity and Cognition Studios. 1–19 pages.
[5] Krishnan Chandran, Lars Engeln, Daniel Zeidler, and Matthew McGinity. 2024. Sonic Structures: 4E Visual Sound in Multi-User Mixed Reality. In *Proceedings of the International Conference on New Interfaces for Musical Expression.* 419–423. https://doi.org/10.5281/zenodo.13904897 ISSN: 2220-4806.
[6] Daniel Chin, Gus Xia, Yian Zhang, Tianyu Zhang, and Junbo Zhao. 2020. Interactive Rainbow Score: A Visual-centered Multimodal Flute Tutoring System. In *Proceedings of the International Conference on New Interfaces for*

Musical Expression. 208–213. https://doi.org/10.5281/zenodo.4813324 ISSN: 2220-4806.

[7] Katya Davisson. 2022. Improvisation as a method of composition: Reconciling the dichotomy. *British Journal of Aesthetics* 62, 3 (2022), 373–386. Publisher: Oxford University Press UK.

[8] Kristian Eicke, Stefano Fasciani, and Cagri Erdem. 2024. Xyborg: A Wearable Hand-based Instrument for Musical Expression. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 474–479. https://doi.org/10.5281/zenodo.13904921 ISSN: 2220-4806.

[9] Lars Engeln, Dietrich Kammer, Leon Brandt, and Rainer Groh. 2018. Multi-Touch Enhanced Visual Audio-Morphing. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 152–155. https://doi.org/10.5281/zenodo.1302711 ISSN: 2220-4806.

[10] Sidney Fels, Ashley Gadd, and Axel Mulder. 2002. Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound* 7, 2 (Aug. 2002), 109–126. https://doi.org/10.1017/S1355771802002042

[11] Rebecca Fiebrink, Perry R. Cook, and Dan Trueman. 2009. Play-along mapping of musical controllers. In *Proceedings of the 2009 International Computer Music Conference*. International Computer Music Association, Montreal, Quebec, Canada, 61–64.

[12] Jules Françoise and Frederic Bevilacqua. 2018. Motion-sound mapping through interaction: An approach to user-centered design of auditory feedback using machine learning. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 8, 2 (2018), 1–30. Publisher: ACM New York, NY, USA.

[13] Costa K. Colachis Glass and Fabio Morreale. 2023. TAILSPIN: An Inquiry into the Audio-Visual Feedback Equilibrium. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Mexico City, Mexico, 542–548. ISSN: 2220-4806.

[14] José Carlos Godinho. 2016. Miming to recorded music: Multimodality and education. *Psychomusicology: Music, Mind, and Brain* 26, 2 (2016), 189. Publisher: Educational Publishing Foundation.

[15] Louise Harris. 2021. *Composing Audiovisually: Perspectives on audiovisual practices and relationships*. Focal Press, London. https://doi.org/10.4324/9780429327179

[16] William Hsu. 2011. On Movement , Structure and Abstraction in Generative Audiovisual Improvisation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Oslo, Norway, 417–420. https://doi.org/10.5281/zenodo.1178047 ISSN: 2220-4806.

[17] Alon Ilsar, Mark Havryliv, and Andrew Johnston. 2013. The AirSticks: a new interface for electronic percussion. In *Proceedings of the Sound and Music Computing Conference*. KTH Publishing, Stockholm, Sweden, 220–226. http://www.speech.kth.se/smac-smc-2013/

[18] Alon Ilsar, Matthew Hughes, and Andrew Johnston. 2020. Nime or mime: A sound-first approach to developing an audio-visual gestural instrument. In *International conference on new interfaces for musical expression 2020*. New Interfaces for Musical Expression, 315–320.

[19] Alon Ilsar and Andrew Johnston. 2015. Choreography in the mapping of new instruments. In *Proceedings of the 2015 ACM SIGCHI conference on creativity and cognition*. 161–164.

[20] Alon Ilsar, Gail Kenning, Sam Trolland, and Ciaran Frame. 2022. Inclusive improvisation: Exploring the line between listening and playing music. *ACM transactions on accessible computing* 15, 2 (2022), 1–21. https://doi.org/10.1145/3506856 Type: Journal Article.

[21] Alon Alexander Ilsar. 2018. *The AirSticks: a new instrument for live electronic percussion within an ensemble*. phd.

[22] Vera John-Steiner, Robert J Weber, and Michele Minnis. 1998. The challenge of studying collaboration. *American educational research journal* 35, 4 (1998), 773–783. Publisher: Sage Publications.

[23] Andrew Johnston. 2013. Fluid Simulation as Full Body Audio-Visual Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Daejeon, Republic of Korea, 132–135. https://doi.org/10.5281/zenodo.1178572 ISSN: 2220-4806.

[24] A. Kapur, E.L. Yang, A.R. Tindale, and P.F. Driessen. 2005. Wearable sensors for real-time musical signal processing. In *PACRIM. 2005 IEEE pacific rim conference on communications, computers and signal processing, 2005.* 424–427. https://doi.org/10.1109/PACRIM.2005.1517316

[25] Susanne K Langer. 2009. *Philosophy in a new key: A study in the symbolism of reason, rite, and art*. Harvard University Press.

[26] Pieter-Jan Maes, Marc Leman, Micheline Lesaffre, Michiel Demey, and Dirk Moelants. 2010. From expressive gesture to sound. *Journal on Multimodal User Interfaces* 3, 1 (March 2010), 67–78. https://doi.org/10.1007/s12193-009-0027-3

[27] Thor Magnusson. 2010. Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal* 34, 4 (2010), 62–73. Publisher: MIT Press.

[28] Peter Manning. 2004. The Development of the MIDI Communications Protocol. In *Electronic and Computer Music*, Peter Manning (Ed.). Oxford University Press, 263–278. https://doi.org/10.1093/acprof:oso/9780195144840.003.0014

[29] Eduardo Reck Miranda, Marcelo M. Wanderley, and Ross Kirk. 2006. *New Digital Musical Instruments: Control and Interactions Beyond the Keyboard* (1st ed. ed.). A-R Editions, Inc., Middleton, Wisconsin.

[30] Ali Momeni and Cyrille Henry. 2006. Dynamic Independent Mapping Layers for Concurrent Control of Audio and Video Synthesis. *Computer Music Journal* 30, 1 (2006), 49–66. https://www.jstor.org/stable/3682026 Publisher: The MIT Press.

[31] Mundla Narasimhappa, Arun D Mahindrakar, Vitor Campagnolo Guizilini, Marco Henrique Terra, and Samrat L Sabat. 2019. MEMS-based IMU drift minimization: Sage Husa adaptive robust Kalman filtering. *IEEE Sensors Journal* 20, 1 (2019), 250–260. Publisher: IEEE.

[32] Lakshmi Prabha Nattamai Sekar, Alexander Santos, and Olga Beltramello. 2015. IMU drift reduction for augmented reality applications. In *Augmented and virtual reality*, Lucio Tommaso De Paolis and Antonio Mongelli (Eds.). Springer, Cham, 188–196.

[33] Luciana Perc. 2024. Sonic Wings: A Wearable Live Electronics Device for Performing Mixed Music. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Utrecht, Netherlands, 178–180. https://doi.org/10.5281/zenodo.13904822 ISSN: 2220-4806.

[34] Olivier Perrotin and Christophe d'Alessandro. 2014. Visualizing Gestures in the Control of a Digital Musical Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. London, United Kingdom, 605–608. https://doi.org/10.5281/zenodo.1178901 ISSN: 2220-4806.

[35] Joseph Rovan and Vincent Hayward. 2000. Typology of tactile sounds and their synthesis in gesture-driven computer music performance. *Trends in gestural control of music* (2000), 297–320. Publisher: Editions IRCAM: Paris.

[36] Anna Savery, Sam Ferguson, and Andrew Johnston. 2024. Gesture and Narrative: Blending Human Performance with Visual Storytelling. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Utrecht, Netherlands, 188–193. https://doi.org/10.5281/zenodo.13904828 ISSN: 2220-4806.

[37] Diemo Schwarz. 2023. Touch Interaction for Corpus-based Audio–Visual Synthesis. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Mexico City, Mexico, 394–401. https://doi.org/10.5281/zenodo.11189220 ISSN: 2220-4806.

[38] SensiLab, Alon Ilsar, Melinda Smith, Ciaran Frame, Sam Trolland, Alessio Dilettoso, Amanda Betlehem, Robbie Avenaim, and Lucija Ivsic. 2022. AirSticks 2.0. Music Performances for Melbourne Design Week at Jolted Theatre, Northcote, 2022-03-25, 60min..

[39] Melinda Smith, Alon Ilsar, Anthea Skinner, Sam Trolland, Anna Cordingley, and Matt Scott. 2023. Conduit Bodies. New-Media Dance Performance at Victorian College of the Arts, University of Melbourne, 2023-02-25, 60min, Performed By: Melinda Smith, Alon Ilsar.

[40] SpeakPercussion, Alon Ilsar, Eugene Ughetti, Kaylie Melville, Ciaran Frame, and Sam Trolland. 2021. Before nightfall #14. Music Performance at Grainger Museum, University of Melbourne, 2021-12-04, 60min, Performed By: Alon Ilsar, Eugene Ughetti, Kaylie Melville.

[41] Sam Trolland and Alon Ilsar. 2025. Cymbalism. https://bridges.monash.edu/articles/media/Cymbalism/28792673 Augmented Reality Percussion Performance at SensiLab, Monash University, 2024-03-27, 10min, Performed By: Alon Ilsar.

[42] Sam Trolland, Alon Ilsar, Ciaran Frame, Jon McCormack, and Elliott Wilson. 2022. AirSticks 2.0: Instrument Design for Expressive Gestural Interaction. In *International Conference on New Interfaces for Musical Expression*. https://doi.org/10.21428/92fbeb44.c400bdc2

[43] Sam Trolland, Melinda Smith, Alon Ilsar, and Jon McCormack. 2024. Visual instrument co-design embracing the unique movement capabilities of a dancer with physical disability. In *Proceedings of the 9th International Conference on Movement and Computing (MOCO '24)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3658852.3659075

[44] Kyriakos Tsoukalas, Joseph Kubalak, and Ivica Ico Bukvic. 2018. L2OrkMote: Reimagining a Low-Cost Wearable Controller for a Live Gesture-Centric Music Performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Blacksburg, Virginia, USA, 275–280. https://doi.org/10.5281/zenodo.1302581 ISSN: 2220-4806.

[45] Alan Wexelblat. 1995. An approach to natural gesture in virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)* 2, 3 (1995), 179–200. Type: Journal Article.

[46] Matthew Wright and Adrian Freed. 1997. Open SoundControl: A New Protocol for Communicating with Sound Synthesizers. In *International Computer Music Conference (ICMC)*. 101–104.

[47] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 493–502.