MUTABLE GESTURES: A NEW ANIMATED NOTATION SYSTEM FOR CONDUCTOR AND CHAMBER ENSEMBLE

Ciaran Frame, Alon Ilsar, Sam Trolland
SensiLab, Monash University
Melbourne, Australia
{ciaran.frame, alon.ilsar, sam.trolland}@monash.edu

ABSTRACT

This paper outlines the creation of a new real-time scoring work, *Mutable Gestures*, for any combination of chamber instruments with a conductor. The work translates a conductor's gestures into real-time animated notation, relayed to performers over a wireless network to generate musical material for improvisation. Drawing on recent real-time notation works, *Mutable Gestures* proposes a new form of gestural notation creation through the use of the AirStick, a new gestural musical instrument. The creation of this work contributes to the growing field of real-time animated notation, a field that reinterprets the traditional roles of score, conductor, composer and performer.

1. BACKGROUND

1.1 The role of the conductor

Conductors are often viewed as the mediator between musical work and performer, conveying the composer's intentions through a single position of leadership and direction. However, the prevalence of conductors in smaller chamber ensembles has contracted, with many ensembles such as quartets usually opting out of a conductor entirely.

Contrary to this trend, recent findings by Wyatt and Hope suggest that the conductor is in fact desired in an animated notation context within the chamber ensemble, offering greater accuracy and musical insight beyond the animated score itself [1].

With this in mind, it is interesting to posit what other roles a conductor may play in the ever-growing field of animated notation, perhaps ceding some roles like timekeeping and gaining others, like running notation systems or software debugging, particularly in an environment where so-called 'virtual conducting' is already possible through smartphones [2] or infrared technology [3].

Thus, the combination of new technologies in conducting and the changing role of the conductor suggests fertile ground for a notated work that generates musical material from the conductor themselves.

1.2 Gesture and notation

Kinetic movement, particularly gestural movement from musicians, is a productive field for real-time notation. The use of motion sensors such as IMUs (inertial measurement units) provides clear input data, offering exploration of causal relationships between action and sound and inferring meaning from gesture.

Work in both Digital Musical Instrument (DMI) and music notation communities of practice has seen an explosion of complex, disembodied musical works that use kinetic movement to produce sound or sound representations [4].

The data provided by a gestural interface in the context of music has been the subject of much related work. Of particular note is Dori's recent work *Arcos*, which captures a cellist's gesture, informing the creation of both action-based notation and sound [5].

Notably though, the source of gesture in musical works is usually a musician, dancer or unspecified performer whose movements and actions are translated into notation [5, 6] or sonic material [7, 8, 9].

It is therefore interesting to note that the conductor is rarely used in a notation or sonic material context, despite being one of the clearest examples of gesture and prescriptive actions provided to musicians.

Conductors have been the subject of gestural scrutiny in music technology, particularly in a motion capture and gesture recognition space [10, 11]. However, this work is almost exclusively focused on musical control and understanding actions and gesture within the conductor's usual scope of practice, such as in a traditional Western Classical concert.

While the use of musicians with gesture data provides intimate composition possibilities, especially in a solo context, the conductor holds a unique perspective and position of power within the chamber ensemble - a power explored in *Mutable Gestures*.

2. MUTABLE GESTURES

2.1 Overview

Mutable Gestures was written for up to five chamber players, and was created from a simple premise of a conductor creating musical content from their baton, for each respective player to play. A conductor gestures towards one or many players in an ensemble, and generates notation. Pointing at a specific player generates live notation on that player's device (such as in in Figure 1), and general gestures produce notes for all players at once.

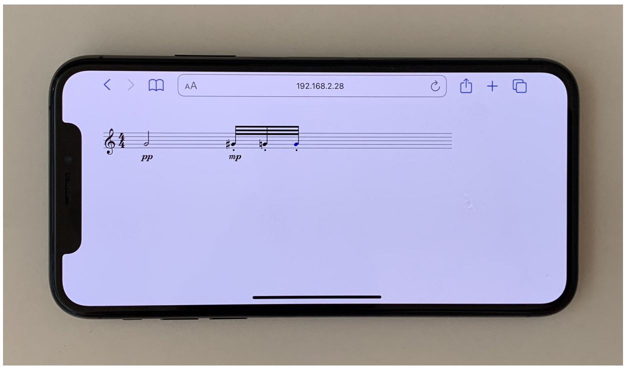

Figure 1. An example of live notation streamed to an iPhone on Safari.

The work is an improvisation system, in which musical material is provided in the form of real-time traditional notation, to be interpreted alongside the conductor's gestures. One does not replace the other — notation cannot replace the nuance in a conductor's gesture, nor can gesture provide a level of detail in musical content that notation supplements. Instead, the work provides a platform for 'comprovisation' [12], where notation and the conductor's gesture provide musical material to be used alongside the players' own inspiration and improvisation in a chamber ensemble context. The degree to which players balance these competing inspirations is up to them.

Attached to the conductor's baton is an *AirStick* (shown in Figure 2), which relays sensor data back to a computer, outlined in Section 2.2. The mapping system in turn makes decisions about what notation appears on performers' devices, outlined in Section 2.3.

Figure 2. The *AirStick* attached to a conducting baton.

2.2 Technology

2.2.1 System architecture

Notation generation begins at the *AirStick* [13, 14], an established wireless gestural Digital Music Instrument (DMI) now in its second design iteration. Attached to the baton, the *AirStick* is capable of determining absolute orientation (the direction the baton is pointing) and linear acceleration (the acceleration of the baton along the three axes local to the device's frame of reference).

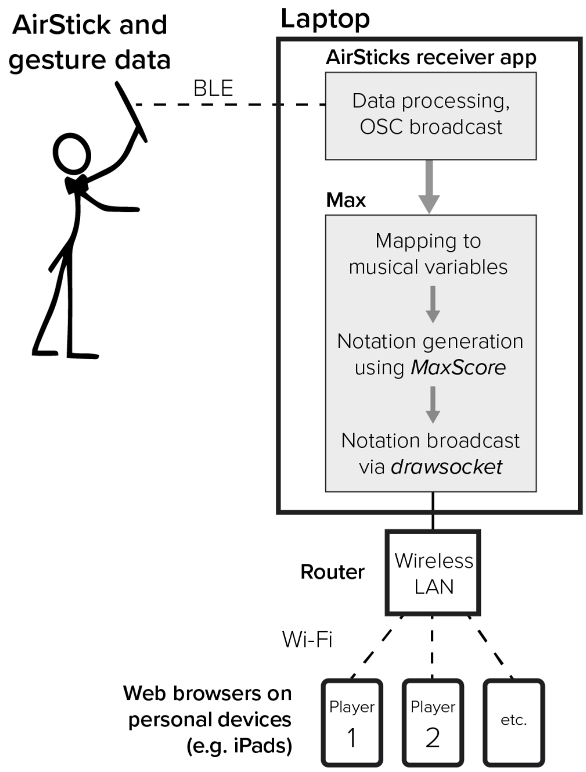

Figure 3. Technical flowchart of *Mutable Gestures*.

The *AirStick* is a gestural instrument by design, and is best placed to deliver these types of gestural data that a conductor conveys, particularly with its natural drumstick-like affordances capturing percussive gestures and orientation data.

The piece is designed to cater for up to five players, in order to distinguish between players from the conductor's point of view, set up in an arc around the conductor. Instrumentation and performer seating positions are calibrated before the performance, to ensure notation is within playing range and using the correct clef.

The *AirStick* communicates via a low-latency Bluetooth Low Energy (BLE) connection to the receiver software on a laptop. Here, sensor data is transformed into readable OSC data, broadcasting sensor data locally to an instance of Max. [1] This sensor data is separated into five key data streams, outlined in detail in Section 2.3.1).

Within Max, this data is mapped to predetermined note values, articulation and dynamics, separated into different parts for respective musicians depending on where the conductor is gesturing towards (described further in Section 2.3). These musical variables are transformed instantly to music notation across multiple parts using MaxScore, [2] then broadcasted to live web pages using Drawsocket, [3] which integrates seamlessly with the MaxScore environment.

Each musician connects to their own 'part' through a unique

[1] https://cycling74.com/products/max
[2] http://www.computermusicnotation.com/
[3] https://forum.ircam.fr/projects/detail/drawsocket/

web address, meaning performers only require a personal device (such as an iPad), and a Wi-Fi connection to the local performance network. This reduces barriers to performing the piece, bypassing additional software and hardware requirements such as the installation of apps in favour of a system that will work with no prior purchases or setup time.

This process is described in Figure 3.

2.3 Gesture and mapping

When the conductor gestures, notation is generated live based on a select number of variables from the *AirStick*. Different gestures produce different notation, loosely based on the norms of traditional conducting in order to create a sense of embodiment of the notation, and therefore sound the conductor is triggering.

This can manifest itself in simple (e.g. big gesture maps to strong dynamics) and complex (e.g. long-term variables maps to piece structure) forms, an example of which can be seen in Figure 4.

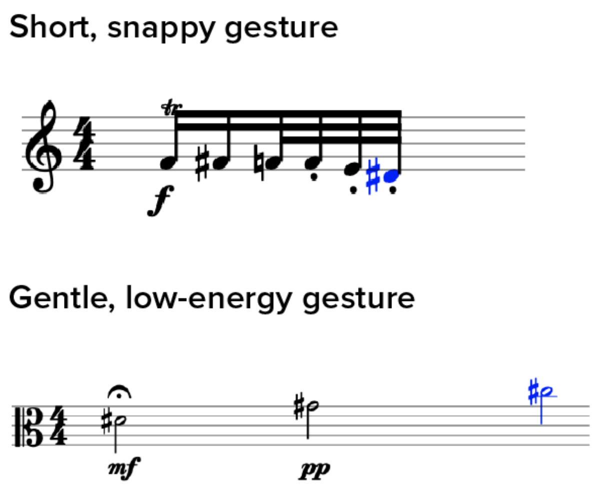

Figure 4. Two examples of generated notation based on different gestural inputs from the baton.

Importantly though, the work does not seek to categorise and learn different styles of conducting, instead choosing to directly ‘map’ kinetic variables directly to musical ones. This is partly to bypass the adjacent discipline of Machine Learning and gesture (discussed further in Section 3.3), but also to explore how this gestural data might translate to musical material. Indeed, the motivation behind mapping particular kinetic variables to specific notation elements evokes sonification of physical gestures, with the added layers of musical interpretation by conductor, system, and performer adding further lenses through which one might view bodily movement. This point becomes more relevant when we consider the fact that the conductor does not receive instructions around how to conduct with the *AirStick*, nor are they aware of the exact notation their own movements are generating.

The real-time notation is an intentionally imperfect representation of the conductor’s gesture, in order to create notation materials that inspire and provoke musical improvisation from performers in conjunction with the gestures they are receiving from the conductor.

AirStick variables are translated from their raw form (described in Section 2.3.1) to notation (described in Section 2.3.2).

2.3.1 AirStick variables

The *AirStick* produces five key data streams that are used in notation generation in the piece. They are:

- **‘Energy’ over time** — the amount of force put into the *AirStick* at a point in time. Calculated in the receiver software as *linear acceleration* mean over n samples in a given time window. This variable also used in ‘Energy variation’, the rolling Standard Deviation over n samples in time window, calculated in Max.

- **Yaw** — where the *AirStick* is pointing across a horizontal plane (i.e. left to right), relative to its position in space. This variable also used in ‘Yaw over time’, the mean over n Yaw samples in a given time window. Euler angle calculated in receiver software from quarternion values, values over time calculated in Max.

- **Pitch** — where the *AirStick* is pointing across a vertical plane (i.e. up and down), relative to its position in space. Euler angle calculated in receiver software from quarternion values.

- **Roll** — where the *AirStick* is rotating along the axis of the baton (i.e. twisted right or left), relative to its position in space. Euler angle calculated in receiver software from quarternion values.

- **‘Poke’ gesture** — a poke in the forwards direction, sent as a single trigger. Calculated in the receiver software as a threshold of directional *linear acceleration* in a forwards direction.

2.3.2 Notation variables

In mapping the above variables, different structural levels are applied to map notation to structural levels of the piece:

- **Note-level structure** — decisions on a note-to-note basis (e.g. the pitch and duration of each note)

- **Phrase-level structure** — decisions across a small passage of up to three bars (e.g. the influence of modes of pitches)

- **Piece-level structure** — decisions across the work (e.g. duration of piece, structural movements or sections)

AirStick and notation variables are applied according to the mapping criteria outlined in Table 1, illustrated in the example in Figure 5.

Whilst in theory a replicated gesture will produce the same notation result, in reality, the subtle differences in motor movements and the high sensitivity of the AirStick

Musical variable	AirStick variable	Structural level	Example
Player part	Yaw	Note	Notation generated only on Player 1's part
Pitch values	Pitch and Yaw	Note	An E♭
Rhythmic values	Yaw	Note	A dotted quaver note value
Articulation	Energy variation over 5 sec	Note	A tenuto marking on a note
Phrase length	Energy	Phrase	The end of a phrase in a player's part, marked by an instruction to augment it for 1 minute
Dynamics	Energy over 10 sec window	Phrase	A *pp* dynamic marking on a phrase
Tempo	Energy over 30 sec window	Phrase	A *Largo* tempo marking on a phrase
Pitch influence	Pitch and Yaw over 30 seconds	Phrase	Pitch values limited to natural minor note set
Extended techniques	Roll	Phrase	An *air tone* marking on a flute phrase
Structure	'Poke' gesture	Piece	Change the structural direction of the piece, resetting variables

Table 1. Musical variables and their respective AirStick mappings.

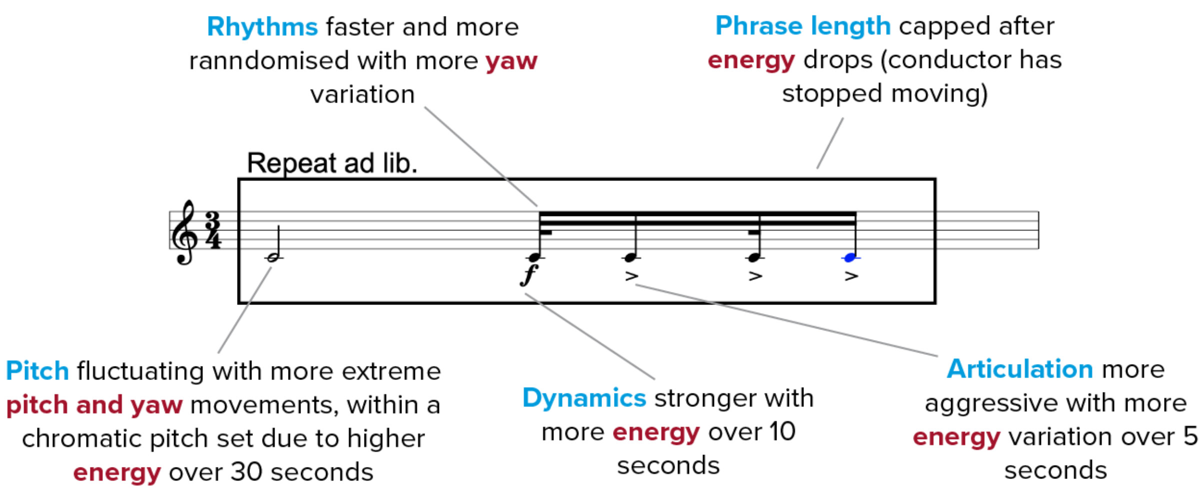

Figure 5. An an example of mapping variables applied to a single part's phrase described in Table 1.

sensor gives rise to varying notation results – an interesting proposition for improvisation. Additionally, the presence of phrase- and piece-level structures in mappings means that 'local' structures and patterns in notation tend to form, such as when the conductor remains at a particular energy, pitch and yaw over an extended period of time, shaping both the pitch mode and dynamics of the notation.

2.4 Performance system

Performers begin the work connected to the same server, each with different parts, as seen in Figure 6.

Figure 6. Multiple devices connected to real-time notation generation using *drawsocket*.

The conductor activates the notation system by picking up the *AirStick*, and bringing performers to attention. The conductor also receives a static paper score — a work composed by the author, created for the performance, that contains music that will never be 'heard' in a traditional sense. Instead, the score provides material for the conductor to follow.

The score has been deliberately designed to accentuate different gestures, and to focus on players in both a discrete (i.e. solo) and general (i.e. unison tutti) manner. In some senses, the conductor's score is a caricature of a traditional score, designed to be interpreted as a normal score, but producing accentuated gestures for the performance system. For instance, an particular section of the score may be written with a *fortississimo* dynamic marking and hocketed rhythms, encouraging strong and discrete gestures from the conductor. The static score may also be changed from performance to performance, as the notation system is independent of the score.

Both the static score and real-time notation use Western Classical notation to reduce the cognitive strain on performers and conductors (as discussed in [5]), allowing for a holistic approach to performance that allows space to improvise and listen.

While the notation received on devices is near real-time (with roughly 6ms latency between the *AirStick* and a note appearing on a device), the presence of different time scales in notation generation described in Section 2.3.2 provides an interesting context for improvising performers. That is, there is a theoretical minimum time from a gestural event to music notation (outlined in Figure 7), but this notation is not occurring at every conceivable sample of the sensor, which would produce an endless stream of notes. Instead, the presence of thresholds means that the performer is met with musical instructions on multiple time scales – from the immediate gesture of the conductor themselves, to a phrase-level dynamic marking based on 10 seconds of gestural data, to a piece-level event that might occur every 30 seconds based on the conductor using the poke gesture. The result is a series of rolling notation snapshots from a performer's perspective, that advance forward but do not edit previously generated notation.

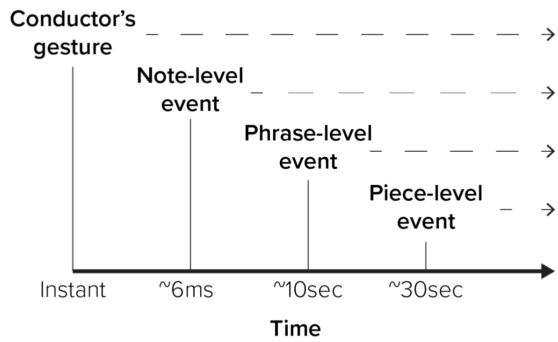

Figure 7. Theoretical minimum times from gesture to performer.

These time scales allow for a notation system that dynamically reacts to gestures, producing both clear immediate notation and longer-term structural instructions, and also ensuring that the conductor retains a certain level of control over the performance. For example, a neutral rest position provided by the conductor can still indicate silence amongst players, maintaining an intricate dialogue between the score, notation system, conductor and performers.

3. DISCUSSION

3.1 Interpretation and evaluation

Musical interpretation, a form of translation, is reshaped in *Mutable Gestures*.

In traditional chamber ensemble performance, the cause and effect relationship between conductor and performer, and performer and sound, is relatively clear, and this causal link provides a context and foundation for listening [15].

From the audience's perspective, *Mutable Gestures* could very well be a traditional notated work, with a conductor fulfilling their normal role and all players playing off a score, as illustrated in Figure 8.

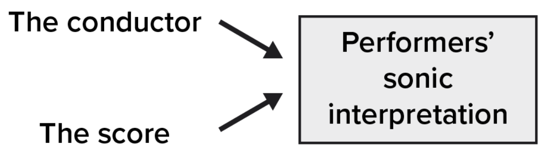

Figure 8. A possible audience perspective of interpretation in *Mutable Gestures*.

Yet beneath this flowchart lies a more complicated reality, based on the mapping of gesture to sound.

The conductor uses a score which is never 'heard', but is communicated to performers, who use a translated form of these gestures (using the notation generated live and their own interpretation) to create sound.

The work becomes a translation of a musical work that is never sonically realised from its score form in a conventional musical relationship, instead interpreted through gesture and a generative notation system, with three steps of interpretation or evaluation of musical information along the chain, as seen in Figure 9.

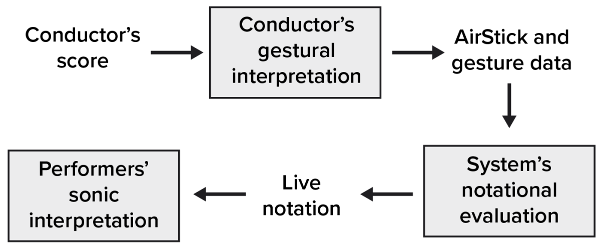

Figure 9. Interpretation and evaluation in *Mutable Gestures*.

Evidently, the link between gesture and sound making is less clear. Beneath the deceptively simple correlation between gesture and sound, the role of conductor is redefined as someone who is not just controlling, but creating musical material through gesture.

The *AirStick* is capable of generating its own sound using the receiver application, and this dimension has been explored in other work with the instrument. However, this project 'mutes' that layer, adding back in another layer of human interpretation in the process of mapping gesture to sound.

Yet, if the sound of the players were to be disembodied from the performance space, like a traditional DMI, *Mutable Gestures* could very easily be interpreted as a form of instrument - what Maestri and Antoniadis describe as "liquidizing the limit between notation and instrument" [16].

Hierarchy of interpretation and evaluation creates a complex network of relationships to technology, reminiscent of works within the DMI community.

3.2 Democratising roles

The breaking down of traditional roles in *Mutable Gestures* alludes to a form of democratisation in that the fixed structures of the relatively autocratic role of the composer are replaced by a system that distributes creative control and inverts the power structures within a traditional chamber music context. This is despite the appearance on the surface that the work is functioning as a traditional chamber performance, as discussed in Section 3.1.

Of course, many chamber works change where creative control lies. Where *Mutable Gestures* differs is in the way it takes advantage of the the natural affordances and illusions provided by roles in a chamber music context, taking fixed structures such as gesturing by a conductor, and reinterpreting and relocating within a generative system.

In a traditional sense:

- conductor becomes performer (playing a DMI, almost entertaining the audience through movement) and composer (generating notation from gesture);
- composer becomes curator (setting the rules of engagement and designing the system to the point that the piece 'works'); and
- performer adds the role of composer (improvising musical material)

One musical system with similar characteristics to *Mutable Gestures* is the concept of Soundpainting, where a 'Soundpainter' controls one or many musicians through a gestural language. Notably, Marc Duby suggests that Soundpainting provides a balance between two extremes in power relations, between an "...orchestral performance, in which the conductor plays a pivotal role, and those of freely improvised music in which, ostensibly at least, there is no leader" [17]. In contrast to Soundpainting, I suggest that *Mutable Gestures* once again reinterprets power relations, this time between improvisation and Soundpainting, retaining the leadership position of a conductor, but distributing the control of information and notation generation amongst multiple blurred roles.

3.3 Future work

Due to COVID restrictions, *Mutable Gestures* performances have been postponed. The piece will naturally evolve through workshops and performances and is created in such a manner that will allows for adaptation.

Aside from performance, the work could also benefit from expanding the number of players and instrumentation possible, perhaps in a larger ensemble or orchestral context. More players would allow for additional textural, and move the *AirStick* closer to a DMI in the sense that it would have greater sounds at its disposal, almost like a software instrument.

Additionally, machine learning in the gesture recognition field is a logical next step for the work, particularly growth of ML and gesture within other DMIs [18, 19].

4. REFERENCES

[1] A. Wyatt and C. Hope, “Conducting animated notation: Is it necessary?” in *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR’20/21*, 2020, pp. 169–174.

[2] Y. K. Lim and W. S. Yeo, “Smartphone-based music conducting,” in *Proceedings of the 14th International Conference on New Interfaces for Musical Expression*, 2014.

[3] T. M. Nakra, Y. A. Ivanov, P. Smaragdis, and C. Ault, “The ubs virtual maestro: an interactive conducting system,” in *Proceedings of the 9th International Conference on New Interfaces for Musical Expression*, 2009.

[4] I. Jarvis and D. V. Nort, “Posthuman gesture,” in *ACM International Conference Proceedings*. Association for Computing Machinery, June 2018.

[5] G. Dori, “Using Gesture Data to Generate Real-Time Graphic Notation: a Case Study,” in *Proceedings of the International Conference on Technologies for Music Notation and Representation*. Hamburg University for Music and Theater, October 2020, pp. 68–74. [Online]. Available: https://doi.org/10.5281/zenodo.4105932

[6] R. Hoadley, “Making people move: Dynamic musical notations,” in *Proceedings of ICMC-SMC 2014 : 40th International Computer Music Conference*, September 2014.

[7] C. Erdem, K. H. Schia, and A. R. Jensenius, “Vrengt: A shared body-machine instrument for music-dance performance,” in *Music Proceedings of the International Conference on New Interfaces for Musical Expression*, F. Visi, Ed. Porto Alegre, Brazil: UFRGS, June 2019, pp. 63–65. [Online]. Available: http://www.nime.org/proceedings/2019/nime2019_music017.pdf

[8] J. Jaimovich, “Emovere: Designing sound interactions for biosignals and dancers,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Brisbane, Australia: Queensland Conservatorium Griffith University, 2016, pp. 316–320. [Online]. Available: http://www.nime.org/proceedings/2016/nime2016_paper0062.pdf

[9] L. Andersson López, T. Svenns, and A. Holzapfel, “Sensitiv – designing a sonic co-play tool for interactive dance,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Shanghai, China, June 2021. [Online]. Available: https://nime.pubpub.org/pub/y1y5jolp

[10] A. Sarasua, “Context-aware gesture recognition in classical music conducting,” in *MM ’13: Proceedings of the 21st ACM international conference on Multimedia*. Association for Computing Machinery (ACM), October 2013, pp. 1059–1062.

[11] H. Je, J. Kim, and D. Kim, “Hand gesture recognition to understand musical conducting action,” in *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, 2007, pp. 163–168.

[12] S. Bhagwati, “Notational perspective and comprovisation,” *Sound & Score. Essays on Sound, Score and Notation*, pp. 165–177, 2013.

[13] A. Ilsar, “The airsticks: a new instrument for live electronic percussion within an ensemble,” Ph.D. dissertation, University of Technology Sydney, 2018.

[14] A. Ilsar and M. Hughes, “A new audio-visual gestural instrument for unlocking creativity,” in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–4.

[15] G. Emerson and H. Egermann, “Gesture-sound causality from the audience’s perspective: : investigating the aesthetic experience of performances with digital musical instruments.” *Psychology of Aesthetics, Creativity, and the Arts*, vol. 12, p. 96, 2017.

[16] E. Maestri and P. Antoniadis, “Notation as instrument: From representation to enaction,” in *International Conference on Technologies for Music Notation and Representation – TENOR’15*, 2015. [Online]. Available: http://vimeo.com/80558397.

[17] M. Duby, “Soundpainting as a system for the collaborative creation of music in performance,” Ph.D. dissertation, University of Pretoria, January 2006.

[18] M. E. Benalcázar, A. G. Jaramillo, Jonathan, A. Zea, A. Páez, and V. H. Andaluz, “Hand gesture recognition using machine learning and the myo armband,” in *2017 25th European Signal Processing Conference (EUSIPCO)*, 2017, pp. 1040–1044.

[19] B. Caramiaux and A. Tanaka, “Machine learning of musical gestures: Principles and review,” in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. Graduate School of Culture Technology, KAIST, 2013, pp. 513–518.