Gesture and Narrative: Blending Human Performance with Visual Storytelling

Anna Savery
Creativity and Cognition Studios
University of Technology Sydney
anna.savery@uts.edu.au

Sam Ferguson
Creativity and Cognition Studios
University of Technology Sydney
samuel.ferguson@uts.edu.au

Andrew Johnston
Creativity and Cognition Studios
University of Technology Sydney
andrew.johnston@uts.edu.au

ABSTRACT

Kind Regards-for a friend is a narrative-driven audiovisual composition that examines the interplay between a human performer and a visual agent. This project was integrated with the development of a new musical interface for the violin bow, and encompassed various strategies for gesture mapping solutions and narrative development. An online audience response survey examined audience experiences of the piece, garnering insights into the effectiveness of our creative and technical processes. Reflections on the project underscored its narrative and interactive strengths, while also identifying music and visual elements that could be further refined to augment the immersive quality of the experience.

Author Keywords

Audiovisual composition, Human-Computer Interaction, Hyperinstruments, Augmented Instruments, Violin, Narrative, Storytelling, Interface Design

CCS Concepts

•**Human-centered computing** → User interface design; **Interaction techniques;** •**Applied computing** → *Sound and music computing;*

Figure 1: Violinist, summoning the bird in Storyboard 1

NIME'24, 4–6 September, Utrecht, The Netherlands.

1. INTRODUCTION

This paper details the development of *Kind Regards - for a friend* - a new audiovisual work for solo augmented violin and interactive visuals (Figure 1). *Kind Regards* evolved across the span of approximately six months, alongside the ongoing development of a new musical interface for the violin bow. Although the interface preceded the piece, its evolution was shaped by it, specifically with the addition of an accelerometer sensor which is central to the piece. The predominant intention of this audiovisual composition was to cultivate a relationship between the violinist and the visuals through gestural and audio interactions as well as positioning the visuals as having a sense of agency through independent motion and musical motifs.

This project combined physical prototyping, data mapping and an exploration of audiovisual compositional strategies, all aimed at creating a narrative-driven work intended for live performance. From this perspective, the design of the technical components was informed by the composition and the desire to connect with our audience.

In presenting our work, we discuss the physical development of our interface, our musical composition approaches within the audiovisual realm, the creation of a visual agent that interacts with the performer, and our narrative techniques, including the use of storyboarding and extensive rehearsals. Additionally, we present an outcome summary of an online audience survey and provide a thorough review of related literature on augmented instruments, audiovisual compositions, and various mapping strategies, placing our contributions within a broader research context.

2. RELATED WORK

2.1 Augmented String Instruments

Augmented musical instruments, which incorporate various sensors, broaden the potential for performers, composers, and audiences alike, revolutionizing how music is created and experienced [16]. One of the earliest significant projects in this field was Machover's *Hyperinstruments*, initiated at IRCAM in 1981 and later moved to MIT [17, 14, 24]. This project produced the *HyperCello*, designed for Yo-Yo Ma, which used sensors to analyze his gestures within a real-time computing environment, mapping movements to control sound parameters like notes and phrases [22]. Following this, the project expanded to include the *Hyperviolin* for Ani Kavafian and an updated version for Joshua Bell, utilizing similar technology to enhance violin performance [15, 21, 31].

Adding to this tradition, Frances-Marie Uitti and CNMAT researchers developed an augmented cello that integrates a mechanical tuning device and a novel rotary sensor for bow interaction, allowing dynamic control over the cello's sound in live performance [?]. Similarly, Laurel S. Pardue's *Svampolin* represents a hybrid approach, employing an electrodynamic pickup and actuated acoustic body to digitally manipulate the sound of a violin while maintaining its conventional playability, thus offering new possibilities for beginners and professionals alike [23].

Further diversifying the range of augmented instruments, Mari Kimura developed the *Augmented Violin Glove* at IRCAM in 2010, which integrated all necessary sensor technology within a glove [12]. Ko and Oehlberg take a more intrusive approach in their *TRAVIS II*, by integrating touch sensors into a custom-built 3-D printed fingerboard to track finger gestures and placing interactive sensors on the body of the instrument [13]. Similarly, *Overtone Violin* serves as a gestural computer music controller that requires both expert traditional and new extended techniques to master its performance potential [13, 20, 18]. A continuation of novel playing approaches through sensor enhancement is seen in *BoSSA*, which, while inspired by traditional instruments, require mastering entirely new skills [29, 19].

2.2 Approaches to Audiovisual Composition

Augmented musical instruments and performance technologies create new ways for performers to express and connect with audiences. Kapuscinski's work, for example, focuses on blending expressive elements with simpler ones, using both sound and gestures to create cohesive narratives [25, 11]. His work *Counterlines* shows how visuals can significantly affect the audience's experience [11, 3]. Harris outlines the basic components of audiovisual composition, emphasizing design principles such as balance, unity, and variety [9].

Blue Space creates a virtual environment where fluid simulations respond to live oboe music, providing the performer with opportunities to blur the lines between instrumental music performance, dance and real-time audio-visual synthesis [30].

R. Luke Dubois' *Moments of Inertia* and Alexander Schubert's *Weapon of Choice* integrate technology to alter video and sound in real-time based on the performers' actions [6, 28]. These innovations enhance the connection between the performer's movements and the resulting audiovisual effects [27].

Further extending these ideas, the theater production *Creature: Dot and the Kangaroo* uses interactive visuals that respond to actors' movements to support the narrative. This method shows how technology can blend seamlessly with traditional storytelling, enhancing the audience's experience by making the performance more engaging and multidimensional [1]. Of particular note are the authors' observations that the work's impactful 'peak' scenes tended to involve high levels of interactivity between performers and real-time visuals.

2.3 Approaches to Data Mapping

A sensor-enhanced musical interface incorporates various input controls and synthesizers, vital for design, development, composition, and performance [8, 2]. The *Video-Organ*, used in live audio-visual performances, demonstrates this integration. Bongers and Harris simplified these complex performances by deconstructing and reconstructing components, integrating interface design and mapping into the compositional process. This creates a unified system where human movements impact both audio and visual outputs [2].

Instruments like *AirSticks* embody a similar design philosophy, transforming performer gestures into sound and visuals through sophisticated mappings that range from simple direct correlations to complex configurations involving multiple inputs and outputs. This approach, as highlighted by Hunt and others, improves performance and expressivity by utilizing intricate, layered mapping strategies, which are crucial for achieving high-level expressive control and virtuosity in performance [10, 4, 5].

Technological advancements extend to systems like Rebecca Fiebrink's *Wekinator*, which applies interactive machine learning to music information retrieval in live settings, allowing performers to directly mold and refine machine learning models for audio-visual control [7]. This merges with technologies like the *Real Lisp* and *Agent Based Score Follower*, which interpret instrumental inputs to responsively engage with musical interactions, fostering innovative performances that challenge conventional boundaries of audiovisual composition [15, 27].

2.4 Motivation

The development of new musical interfaces and audiovisual works, along with various approaches to mapping sensor data, have been widely studied within the NIME community and beyond. Despite this extensive research, there is a noted gap in projects that focus on strong narrative elements and the dynamic between human interaction and technology for storytelling [1].

Kind Regards addresses this gap by emphasizing a narrative communicated through the interaction between a violinist and a simple bird animation. This project integrates the development of a new musical interface with an audiovisual composition, utilizing call and response techniques, expressive gestures, and music motifs. Designed for live performance, the project balanced overwhelming design choices by sketching strong initial ideas and maintaining flexibility to adapt, with the ultimate goal of creating a compelling and emotionally engaging audiovisual experience.

3. PHYSICAL DESIGN

In early 2022, we initiated the development of a minimally invasive wireless interface for the violin bow, aimed at enhancing the creative capabilities of the first author in live performance contexts, particularly for audiovisual compositions [26]. The design of this interface was tailored to the specific compositions it was used in, with ongoing modifications based on each project's feedback.

Throughout each iteration of this musical interface, an XBee module was used to send data directly to an XBee receiver, using the ZigBee wireless protocol[1] (Figure 4).

After several iterations (Figure 2 3), we developed the current version of the interface, which uses stretchy cotton material to hold the hardware components.

This design allows for easy sensor interchange and relocation, helping to determine the optimal sensor placement. Key enhancements include the addition of a Grove ADXL345 accelerometer for real-time gesture tracking, attached to the back of the Arduino Fio[2], and two neopixels for live visual feedback (Figure 5). The technological setup of our interface allows for real-time sensor manipulation using right-hand fingers, as well as gesture tracking and analysis. Arduino data is sent directly into Max/MSP, which then communicates with Processing software (Figure **??**).

[1]https://csa-iot.org/all-solutions/zigbee/
[2]https://wiki.seeedstudio.com/Grove-3-Axis_Digital_Accelerometer-1.5g/

Figure 2: First interface prototype, without an accelerometer

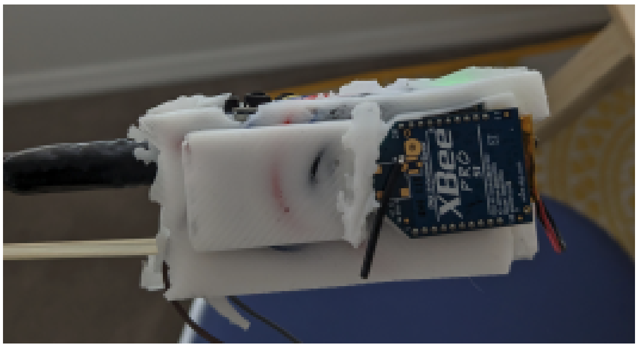

Figure 3: More recent prototype using a silicone encasement with the addition of an acceleromter and neopixels

4. COMPOSITION PROCESS

Narratively, the piece progresses through four stages, beginning with a symbolic call and response between the violinist and a bird-like animation, representing an inner monologue that escalates into a dramatic interaction with flocks of birds, symbolizing a crescendo of anxious thoughts and memories. This narrative arc peaks with the bird's overwhelming feeling of suffocation from the smothering surrounding flocks, resulting in its explosion and disappearance, marking the performance's end.

Developing this piece presented several challenges due to the complexity of integrating multiple components such as creative coding in two separate software environments (Max for Live and Processing), musical composition, narrative development, and mapping sensor data. To manage these complex elements, we used storyboards as a means to facilitate the organization of all moving parts, including the music score. The composition was structured to last approximately 6 minutes, with the initial segment allowing some flexibility in timing during the call and response interaction between the violinist and a tentative bird-like creature. The subsequent three sections were timed according to the duration of pre-recorded samples.

4.1 Tonality, Rhythm and Improvisation

The composition emerged from ongoing dialogue between the violinist (composer/first author) and the bird, shaping a narrative that guided the musical elements, which include free improvisation and pre-recorded violin samples.

The piece is set in D minor, allowing the open strings of the violin to resonate against the muted tones of the Fs and Cs. A majority of the improvised performance is spent in the higher positions of the G and D strings, conveying a sense of vulnerability through the darker, raw tones of that register.

Rhythmically, the piece alternates between long expressive notes and fast lyrical passages. This approach evolved

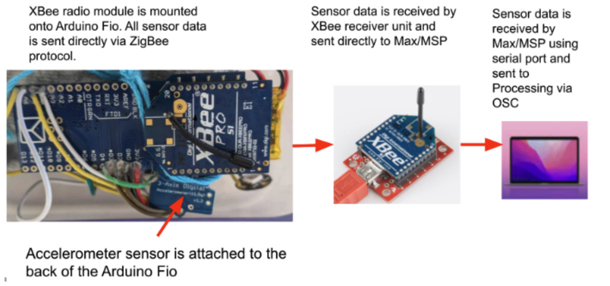

Figure 4: XBee data flow

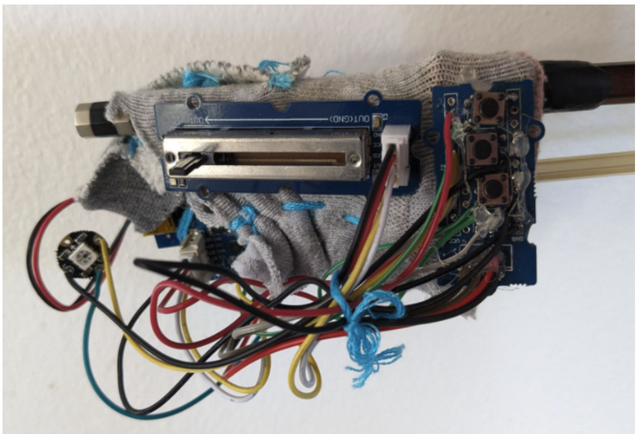

Figure 5: Current interface design

from the interaction between bow gestures and accelerometer data mapped to the bird's flight and wing movements. The pre-recorded audio is intended to sonify the bird, allowing the violinist to interact with both components as one entity, structuring the improvisation around the interaction with the visuals and the musical development of the bird's accompanying sounds.

4.2 Storyboard 1

Storyboard 1 is a call and response interaction between the violinist and the bird. This section is the opening minute or so of the piece and is focused on creating a sense of listening and connection between the human and the virtual being. This interaction is driven by an audio feed from the violin, analyzed in Max using the external Sigmund~[3]. If the violin sound is sustained for at least 180 milliseconds, the bird gradually appears, echoing a slightly distorted violin line (Figure 10). The triggered audio sample activates the bird's

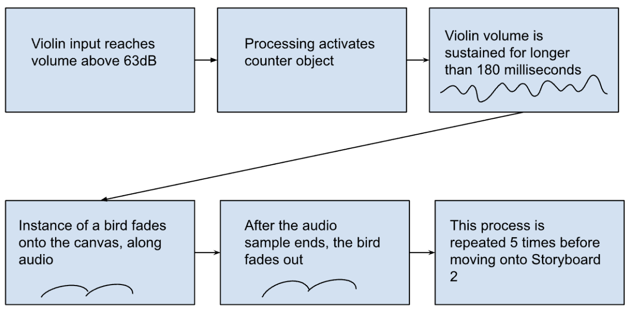

Figure 6: Development of Storyboard 1

[3] Sigmund 64-bit in GitHub

flight path, making it seek a target that is a random value within the canvas window. Once the sample audio is played out (approximately 10-15 seconds in length), and volume drops to 0, the bird's stroke color shifts from white to black, deactivating the flight path. This process is repeated five times, until the final call and response audio loop is heard, triggering the start of Storyboard 2. (Figure 6).

4.3 Storyboard 2

Here, the solo bird takes on a more assertive stance. It's accompanying prerecorded sample is two minutes in length, gradually building in intensity and volume. The bird's flight trajectory is shaped by the real-time bow gestures recorded alongside the audio.

At the half way point of this section, a flowing line made up of a tiny ellipses with a long trail is triggered, following the movement of the violinist's bow across the canvas window. The line's path is slightly delayed through easing, which obscures its direct relationship with the violinist's bow movements, making it more of a playful counterpart to the bird, rather than a visual translation of the bow arm.

4.4 Storyboard 3

The conclusion of the audio sample from Storyboard 2 triggers a two-minute audio sample accompanying the solo bird as well as a sequential appearance of flocks of birds, entering the canvas window one array at a time (Figure 7). Each flock array is timed to appear sequentially between 10 - 20 seconds apart, with 9 flocks in total, retrieving with it a processed audio snippet recorded during Storyboards 1 and 2, with flight paths mapped to bow movement data. As the flocks appear, their layering creates a gradual volume and textural build in audio and visuals. The flocks gradually gravitate towards the solo bird, reaching the climactic point of the piece.

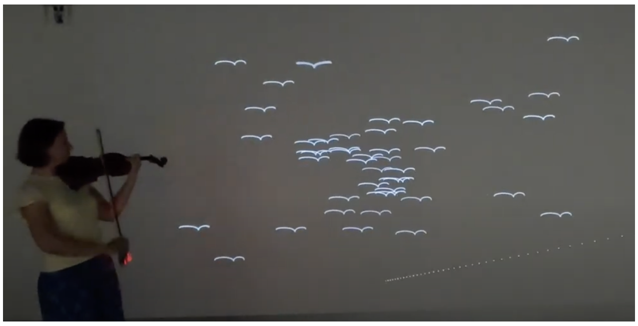

Figure 7: Flocks of birds during Storyboard 3

4.5 Storyboard 4

In the final 30 seconds of the piece, the solo bird feels overwhelmed and suffocated by the flocks. An event is triggered and the bird grows in size, then disappears(Figure 8). This is timed with the ending of the two-minute audio sample.

As the solo bird disappears, the rest of the birds start to disseminate, one flock and audio loop at a time, until the violinist is left alone with the flowing line, ending the piece.

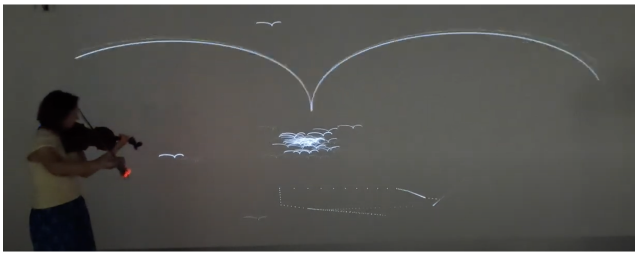

Figure 8: Exploding bird in Storyboard 4

5. VISUAL DESIGN AND BOW GESTURE MAPPING STRATEGIES

The visual component of *Kind Regards* aimed to create an identifiable yet understated bird-like animation, allowing audiences to attach personal meanings to the themes of immigration and belonging explored in the first author's previous work. Technically, the design consists of two bezier curves forming the bird's body, with wings and the center of the bird capable of moving vertically and horizontally to mimic flight.

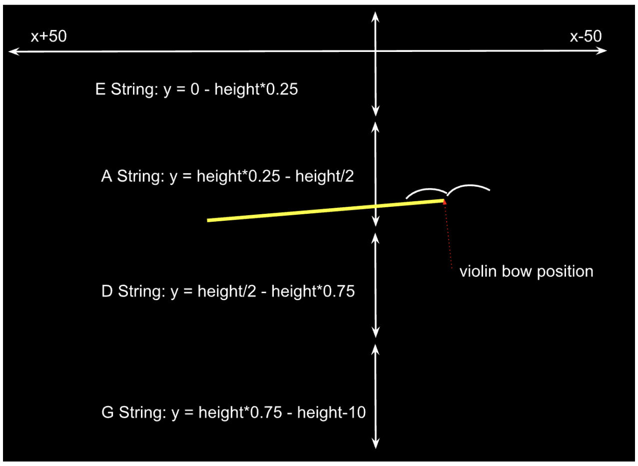

Figure 9: x and y coordinate values mapped to bow movement

Sensor data from the violin bow's interface is fed into a Max/MSP patch, where accelerometer data is converted into coordinates and acceleration values, then sent to Processing via the oscP5 library. In Processing, the bow's horizontal and vertical movements map onto the screen and different string positions—E, A, D, and G strings each in a specific quarter. This setup visualizes the bow's movement but does not translate pitch variations. For example, playing in the upper register of the D or G strings, the bow remains in the lower quadrants, contradicting the bird's flight path. Additionally, bow acceleration alters the bird's wing flapping speed, linking the dynamics of musical gestures to visual effects. Recorded bow movements and audio further enhance this interaction, triggering synchronized visual and auditory elements during the performance, deepening the performer's connection with the visual output. (Figure 9).

In order to create a relationship between the human performer and the bird visualisation, bow movement data and 2 extended audio samples were recorded into the system for Storyboards 1 and 2 and used to both, sonify and animate the visuals(Figure 10).

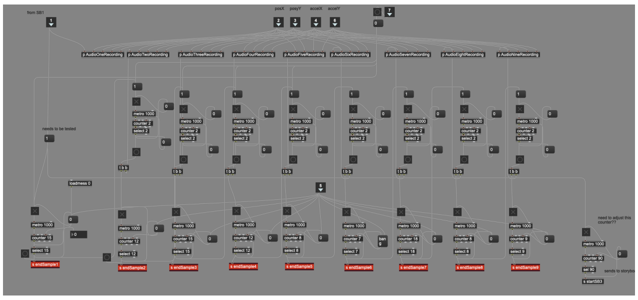

Figure 10: Accelerometer data and audio snippets recorded into Max and Live

6. AUDIENCE RESPONSE SURVEY OUTCOMES

In order to gauge an audience perspective on our work, we conducted an anonymous online audience response survey using Qualtrics. A total of 35 participants were involved, all of whom were gathered by the first author. Participants were recruited from the first author's contact list. Their familiarity with audiovisual composition and augmented instruments ranged between not familiar and very familiar.

The survey asked the following 8 questions:

- How would you describe the relationship between the performer and the visuals?
- Did you find the use of electronics/sensors attached to the violin distracting?
- Did you find the use of electronics/sensors attached to the violin distracting?
- Did you find the use of electronics/sensors attached to the violin distracting?
- Did you perceive a structure or narrative in the piece? If so, could you describe it?
- Do you have any suggestions for how the work could be improved musically?
- Do you have any suggestions for how the work could be improved visually?
- Did you want to watch the video until the end?

Audience reactions were mixed: while some clearly recognized a dialogue between the violinist and the visuals, describing it as a call-and-response interaction, others saw no distinct connection or only a partial one. Those who perceived a relationship often lacked clarity on how the bow's movements influenced the visuals, highlighting a gap in genre familiarity that could be addressed in future projects.

The most positive feedback came from questions about the narrative structure of the piece. A majority of participants identified a narrative where the music summoned a bird-like creature that interacted increasingly with the performer, culminating in a visual display of unity with a flock. This successful narrative engagement contrasted with suggestions for improvement, such as increasing rhythmic variety in the music and enhancing the visuals with more color and intricacy to make future multimedia compositions more engaging and visually compelling. Despite some critical feedback, the audience's willingness to watch the piece to the end indicated a strong overall engagement.

7. CONCLUSION

In this paper, we detailed our process of creating a new audiovisual piece while concurrently developing a musical interface for the violin bow. Our method of using storyboards as a core tool has been effective, facilitating detailed narrative, musical, and technical development.

Looking ahead, we aim to experiment with more diverse visual designs incorporating color and techniques like genetic algorithms to enhance viewer engagement. Advancing our gesture mapping through the integration of machine learning for bow gesture classification could greatly expand narrative possibilities. Furthermore, our compositional approach will evolve to focus more on rhythmic and timbral elements rather than relying on pre-recorded loops and improvisation.

From our work with *Kind Regards*, employing storyboarding has been essential for managing the complexities of the audiovisual creation process. While initially our use of this technique was basic, the success it has brought inspires us to delve deeper into its potential in future projects, aiming to fully leverage its capabilities in enhancing our audiovisual compositions.

8. ACKNOWLEDGMENTS

This research is supported by an Australian Government Research Training Program Scholarship.

9. REFERENCES

[1] A. Bluff and A. Johnston. Storytelling with Interactive Physical Theatre: A Case Study of Dot and the Kangaroo. In *International Conference on Movement Computing*, pages 1–8, 2017.

[2] B. Bongers and Y. Harris. A structured instrument design approach: The video-organ. In *New Interfaces for Musical Expression*, pages 86–91, 2002.

[3] R. B. Dannenberg. Interactive visual music: a personal perspective. *Computer Music Journal*, 29(4):25–35, 2005.

[4] C. Dobrian. Aesthetic considerations in the use of "virtual" music instruments. In *UC Irvine: Integrated Composition, Improvisation and Technology*, 2003.

[5] C. Dobrian and D. Koppelman. The 'e' in nime: musical expression with new computer interfaces. In *New Interfaces for Musical Expression*, 2006.

[6] R. L. Dubois. Moments of inertia for violin, electronics, and interactive video. https://www.lukedubois.com/projects-2/moments, 2010.

[7] R. Fiebrink and L. Sonami. Reflections on eight years of instrument creation with machine learning. In *New Interfaces for Musical Expression*, 2020.

[8] C. Goudeseune. Interpolated mappings for musical instruments. *Organised Sound*, 7(2):85–96, 2002.

[9] L. Harris. *Composing Audiovisually: Perspectives on audiovisual practices and relationships.* 2021.

[10] A. Hunt and R. Kirk. Mapping strategies for musical performance. *Trends in Gestural Control of Music*, 21(2000):231–258, 2000.

[11] J. Kapuscinski and J. Sanchez. Counterlines: Studies in interfacing graphic and melodic lines. In *International Computer Music Conference*, 2009.

[12] M. Kimura, N. Rasamimanana, F. Bevilacqua, B. Zamborlin, N. Schnell, and E. Fléty. Extracting human expression for interactive composition with the augmented violin. In *New Interfaces for Musical*

Expression, pages 1–1, 2012.

[13] C. Ko and L. Oehlberg. Touch responsive augmented violin interface system ii: Integrating sensors into a 3d printed fingerboard. In *New Interfaces for Musical Expression*, 2020.

[14] T. Machover. Hyper-instruments: Musically intelligent and interactive performance and creativity systems. In *International Computer Music Conference*, pages 186–190, 1989.

[15] T. Machover. Hyperinstruments: a progress report. *MIT Media Laboratory*, 1992.

[16] E. R. Miranda and M. M. Wanderley. *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*, volume 21. AR Editions, Inc., 2006.

[17] C. C. Monteverdi. Sound gesture and rhetoric. hyper-cello as an algorithmic composer nicola baroni carlo benzi.

[18] D. Overholt. The overtone violin. In *New Interfaces for Musical Expression*, pages 34–37, 2005.

[19] D. Overholt. The musical interface technology design space. *Organised Sound*, 14(2):217–226, 2009.

[20] D. Overholt. Violin-related HCI: A taxonomy elicited by the musical interface technology design space. In *International Conference on Arts and Technology*, pages 80–89. Springer, 2011.

[21] D. Overholt. Advancements in violin-related human-computer interaction. *International Journal of Arts and Technology 2*, 7(2-3):185–206, 2014.

[22] J. A. Paradiso and N. Gershenfeld. Musical applications of electric field sensing. *Computer music journal*, 21(2):69–89, 1997.

[23] L. S. Pardue, K. Buys, M. Edinger, D. Overholt, and A. McPherson. Separating sound from source: Sonic transformation of the violin through electrodynamic pickups and acoustic actuation. In *New Interfaces for Musical Expression*, pages 278–283, 2019.

[24] N. Rasamimanana. Gesture analysis of bow strokes using an augmented violin. *IRCAM*, 2004.

[25] A. Savery. *Intermedia Storytelling*. PhD thesis, UC Irvine, 2016.

[26] A. Savery. Enhancing violin performance through real-time interaction: Design and evaluation of a wireless audio-visual interface. *Chroma: Journal of the Australasian Computer Music Association*, 39(2), 2023.

[27] A. Schubert. Weapon of choice. https://www.alexanderschubert.net/works/weapon.php, 2009.

[28] A. Schubert. *Switching worlds*. Wolke Verlag Hofheim am Taunus, 2021.

[29] D. Trueman and P. Cook. Bossa: The deconstructed violin reconstructed. *Journal of New Music Research*, 29(2):121–130, 2000.

[30] L. Walsh, A. Bluff, and A. Johnston. Water, image, gesture and sound: composing and performing an interactive audiovisual work. *Digital Creativity*, 28(3):177–195, 2017.

[31] D. Young. The hyperbow controller: Real-time dynamics measurement of violin performance. In *New Interfaces for Musical Expression*, pages 1–6, 2002.