residUUm: user mapping and performance strategies for multilayered live audiovisual generation

Ireti Olowe
Center for Digital Music
Queen Mary
University of London, UK
i.o.olowe@qmul.ac.uk

Giulio Moro
Center for Digital Music
Queen Mary
University of London, UK
g.moro@qmul.ac.uk

Mathieu Barthet
Center for Digital Music
Queen Mary
University of London, UK
m.barthet@qmul.ac.uk

ABSTRACT

We propose residUUm, an audiovisual performance tool that uses sonification to orchestrate a particle system of shapes, as an attempt to build an audiovisual user interface in which all the actions of a performer on a laptop are intended to be explicitly interpreted by the audience. We propose two approaches to performing with residUUm and discuss the methods utilized to fulfill the promise of audience-visible interaction: mapping and performance strategies applied to express audiovisual interactions with multilayered sound-image relationships. The system received positive feedback from 34 audience participants on aspects such as aesthetics and audiovisual integration, and we identified further design challenges around performance clarity and strategy. We discuss residUUm's development objectives, modes of interaction and the impact of an audience-visible interface on the performer and observer.

Author Keywords

computer-generated audiovisuals, mapping strategies, sonification, performance tool, performance analysis

ACM Classification

H.5.1 [Information Interfaces and Presentation] Multimedia Information Systems—Evaluation/methodology, H.5.2 [Information Interfaces and Presentation] User Interfaces—Interaction styles, H.5.5 [Information Interfaces and Presentation] Sound and Music Computing—Signal analysis, synthesis, and processing, J.5 [ARTS AND HUMANITIES] Performing arts (e.g., dance, music)

1. INTRODUCTION

To express is to convey feelings, thoughts or ideas "through the manipulation of a medium" such that some degree of emotional arousal is evoked within an observer [12]. The expressivity of an audiovisual tool, mediated by a performer, is revealed to an observer through its aural and/or visual aesthetic, which, over a duration of time, emerges as a consequence of a programmed interpretation that materializes the translation of sound by image and/or the converse [6][19]. The interconnection between audition and vision is negotiated by mapping — creating links that control parameterized characteristics in one domain utilizing behaviors of parameters in the other [4][13]. The actions of a performer can influence the degree of expressivity that is disclosed to a viewer. However, movements and gestures executed from behind a laptop and the unobserved direct manipulation of software interfaces or hardware devices can obscure the communication between performer and audience [7][2]. Hook et al. explore the concept of expressive interactivity through the lens of the VJ and the context of HCI for the purpose of proposing modern performance techniques "that are sensitive to the more nuanced spaces of user activity," which concerns the relationship between performer and observer [12]. Correia and Tanaka have also explored the landscape of audiovisual tools to investigate themes, which they hypothesize are fundamental to contemporary performance and composition practice: expression, usability and the relationship between the practitioner and viewer [9]. Correia formulates his audiovisual research objectives into a developing framework entitled AudioVisual User Interfaces or AVUIs to overcome the challenges of performance practice and interaction methods accessible through the use of current GUIs, which do not afford audiences visual cues from the performer that establish causal relationships between sound and image[1] [8]. In this paper, we propose residUUm and focus on its visible interface in our pursuit to build a successful AVUI.

2. THE CONNECTION BETWEEN PERFORMER AND OBSERVER

The contribution a musical performer makes while performing with a traditional instrument can easily be deduced by an observer. The actor's physical display and gestures made with the body demonstrate effort and commitment. The causal relationship is also confirmed by immediate feedback between a cause and its effect. However, when a performer performs using software while interacting with an audiovisual instrument, the causal relationship between sound and image may not be so clear [2]. The contribution of the actor (the cause) may be attributed to the system from which the effect is created or to the user of the tool [19]. Furthermore, the awareness about the provenance in such performances is dependent on the perception of the audience [7]. Auslander discusses the different views that challenge the "traditional relationship" between the performer and observer: Stuart assigns the onus of comprehension to the observer who need only rely on the act of listening, whereas, the traditionalist point of view states that the relationship between performer and audience is one based on trust and comprehension on the part of the observer. However, the adoption of new composition and performance technologies challenges these views. Schloss, declares that the act of listening should be accompanied by a mechanism that appends vision to audi-

[1] http://avuis.goldsmithsdigital.com/

NIME'16, July 11-15, 2016, Griffith University, Brisbane, Australia.

tion to "clarify the nature of causality" [2].

The control and generation mechanisms used in a tool such as residUUm via a laptop, has been identified as a problematic instrument by Lew. He found that the connection between audience members who were uneducated in the culture of audiovisual presentation and the performance was determined based on assumption (prerecorded video) and misinterpretation (the role of the performer). Therefore, he proposes a transparent interface and performative style to engage an audience with the performer's actions, process, and the relationship between the action and process that mediates the outcome [16][2].

2.1 Multimodal Linking and Mapping

The multimodal relationship between sound and image may be an aesthetic matter, which can be addressed with design or left to be answered by psychological empirical inquiry [5]. That "perceptual aspects [of image and sound] connect and influence each other" aligns with the traditionalist perspective. Auslander distinguishes between the sound/image relationship as learned from empirical study and its value in a cultural context. With respect to experimental psychology, both visual and audition contribute to the listening experience as the brain compensates and seeks a source for what is only heard. With respect to culture, previous experience and established norms dictate what audiences of different musical performances expect to occur. It is a competence acquired through experience and subculture [2].

Achieving some degree of "temporal simultaneity" in audiovisual performance between sound and image is a basic condition for acceptance [19]. Sound mapped to an image such that its appearance (i.e. size, shape, opacity, position) is associated with its characteristics communicates the impression that the image is linked to its structure [4]. Linking between sound and image serves as a "general perceptive function" to describe the relationship between the two. Kapuscinski's concept of linkages presents a merged modality such that sound and image are equally attended. Linking is a temporal criterion whose use is to establish relationships between audio characteristics (i.e. density, dynamics, texture, tempo) and visuals [5]. Real-time linking can be accomplished externally and manually by the performer or internally and automatically by the system. External linking relies on human effort and two independent sources of aural and visual information. Internal linking is a feature of a system whose visuals are mapped to the internal structure of an audio signal and are generated directly from an audio signal in real time [13].

Applied mapping strategies mediate the relationship between sound and image whatever the complexity of the underlying methodology may be. Callear explores the four classes of mapping, their subdivisions, and their affordance for clarity. A chosen implementation can impact an observer's experience. The causal relationships between sound and image may be apparent, however, the viewer's attention may wane due to banality that develops over time. Complexity and dynamism can resolve the perceptual tedium that unambiguous mapping strategies can cause [4]. Hunt et al. introduce the concept of mapping layers by which mappings are separated into independent controls for greater complexity [14].

Ciufo sees a lack of explicit incorporation of sound with image "on a structural level." The challenges of mapping color (and light) and music were illustrated in reviews of early live performances that involved color organs. The sound and image relationships, the influence of music on the visuals, were perceptively obscure [5]. The attribution of mapping as a "blind medium" can be directed towards current audiovisual performances during which relationships are either composed on a "superficial level of image and sound simultaneity" or are methodically designated [5][19][18][2]. The visible interface of an AVUI is meant to convey which effects are triggered by which causes.

3. OVERVIEW OF THE residUUm AVUI

residUUm,[2] is a screen-based audiovisual performance tool, which uses sonification to orchestrate a particle system of shapes. It was developed during the second of two hackathons on generative audiovisuals (GEN.AV2) in support of Correia's research into AudioVisual User Interfaces (AVUIs) at Goldsmiths, University of London, UK. Other sonification projects developed during the hackathon were DrawSynth, Esoterion Universe, and Modulant.[3] The purpose of the hackathon was to prototype easy-to-use, generative, audiovisual tools whose interaction was entirely visible for an audience to clearly interpret the causal effects between sound and image [8]. These requirements informed the interaction and interface design for residUUm whose system is shown in Fig. 1. It employs a computer screen and cursor as a visual interface which, during live performances, is mirrored to a projection screen visible to the audience. The interaction is conducted using the computer's keyboard and mouse or trackpad.

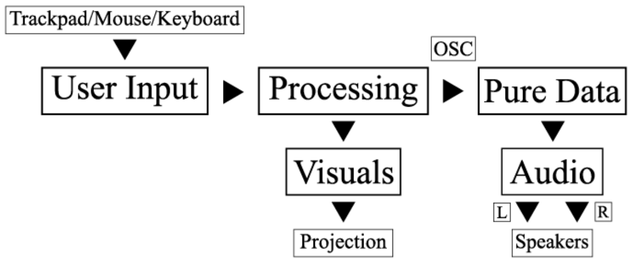

Figure 1: Schematic of the residUUm system

In residUUm, particles interact with their neighbors and respond to external influences as they spawn, move, age, and die while colliding with one another, contributing to a din of polyphonic noise. The sonic characteristics of each object are defined by its graphic attributes: shape, color fill, size and scene location. The cast bodies, which remain after colliding with one another are the sounds generated from the impact of collisions. The term residuum refers to the remains of a body following some process or event [1]. Additionally, the suffix -duum is reminiscent of the latin word "duum", neuter genitive for "duo", referring to the inherent audio-visual duality of the system.

In residUUm, the mapping between visual characteristics and sound parameters was crucial. Mapping is the last stage of the creation process. The graphics and audio were independently developed by the authors. As a general rule, we tried to avoid linear, one-to-one mappings between the visual and audio parameter spaces. Occasionally, random components were added to afford the system to react unpredictably. Some mappings are based on intuition, dictated by the laws of physics or semantic associations, while other mapping decisions are entirely arbitrary [11].

Each particle within the canvas is linked to a single voice of the sound generator. The visual characteristics of the particle affect the sound that is generated:

[2]https://github.com/AVUIs/residUUm
[3]https://github.com/AVUIs

Size. The size of the particle is mapped to the frequency of the oscillator. Smaller sizes correspond to higher frequencies. This mapping builds on the common knowledge that smaller objects emit higher pitched sounds.

Shape. Each of the particle shapes is associated with a distinct waveform for the voice's main oscillator. The concept of visual shape is directly transformed into the shape of the waveform.

Color. The color of the particle affects the frequency of the notches in the filterbank, conditioning the timbre of the voice. The timbre of a sound is also called "sound color", [*klangfarbe* in German] [20]. The vocabulary used to describe timbre is poor, therefore, we often use semantic descriptions adopted from other senses (i.e. vision) to refer to them. The choice of mapping the color to timbre is based on these considerations.

Group. Particles generated during the same mouse press are assigned to the same group. The waveform of the Low Frequency Oscillator that modulates the amplitude is unique for a given group. The group also affects the envelope of the collision sound of the particle. Therefore, particles created concurrently will retain a signature of their common characteristic even after they navigate away from one another.

Position-X. The position along the X-axis controls the panning of the particle in the stereo field, with a direct correspondence between the physical location and the position in the stereo field. The same parameter is used to control the frequency of the amplitude modulation LFO.

Position-Y. The vertical position controls the overall amplitude of the sound generated by the particle. It also affects the central frequency of the band pass filter, which only influences the signal coming from the granular oscillator which sonifies collisions.

Lifespan. Two characteristics of the sound are influenced by the lifespan of the particle. As the lifespan decreases, the overall amplitude of the sound generated from the particle decreases. As the lifespan decreases, the amount of noise in the amplitude envelope is increased. These mappings account for the idea that as the particle ages, it becomes weaker (quieter) and exhibits instability (noise).

Background. The color of the screen background is associated with audio effects applied to the master stereo bus. These effects purposely degrade the audio signal to obtain a low-fidelity result.

ResidUUm was developed using Daniel Koehler's punktiert[4] particle engine library in Processing[5] and communicates over OSC[6] to send parametric data to a polyphonic synthesizer bank in Pure Data[7]. Processing handles the input from the performer, generates the particles, computes how they interact, move, and evolve, renders the graphics, and is responsible for voice allocation. Pure Data generates audio from mapped visual parameters sent over OSC. Each voice of the synthesizer in the Pure Data patch corresponds to one particle in the Processing sketch. The maximum number of voices is limited to prevent system overload and all the resources for the synthesizer voices are allocated upon startup.

The visible interface in residUUm is controlled by a cursor that exposes the actions of the performer as its position moves around the canvas. The position of the cursor, the size, position, shape, color and interactions of the particles are sonified to indicate causal relationships. However, the concept of the visible interface is challenged by the laptop and the system due to the constraints of the software and parameter programming [11][17]. Keyboard controls add extra functionality, but their use interferes with the connection between the performer's actions and the audience.

4. residUUm — A PERFORMANCE TOOL

Makela categorized five spaces that define elements of live cinema, a form of audiovisual performance. Normally, these spaces are the digital space, the laptop's storage capacity for performance assets; the desktop space, the landscape enclosed within the laptop screen and the interface where the performer organizes, operates and directs the components of the performance; the performance space, the stage; the projection space, the surface upon which the performance is projected; and, the physical space, the setting and location of the performance. When performing with an AVUI, the desktop, performance and projection spaces are compressed into one. The luxury of working with separate interfaces or multiple windows to access encapsulated functions within the software or laptop screen does not exist. The addition of peripheral interfaces, which extend the control of performance parameters is prohibited. And, the ability to work beyond the viewers' observation, the split screen between the projection surface and laptop screen, is eliminated. Here, the three spaces are shared by the performer and the viewer. The challenge of the AVUI, and residUUm's design, is to exhibit the expressivity of the tool using a transparent interface. All actions of the performer are observed by an audience, who should be able to interpret causal cues displayed through audible and visible interactions [17][19].

The following sections discuss two strategies employed to address the impact of an AVUI on residUUm as a audiovisual performance tool.

4.1 Performance Strategy 1 — a gaming approach to sound composition

residUUm was presented and performed, by the second author, to an audience on July 30, 2015 in an auditorium on the campus of Goldsmiths College, University of London.[8]

For this performance, the author aimed to use residUUm as a traditional musical instrument. As a performer's visual demonstration is an essential part of the music creation process, he physically gestured in full view of the audience while employing residUUm as a sound composition tool, primarily utilizing the graphic interface as a means to interact with the audio engine. Throughout the performance, the author planned to challenge the randomness in residUUm and layer quiet drones in the background while visually composing alternate sections with a number of spawned, short-lived sounds and large groups of loud particles in the foreground as shown in Fig. 2.

Randomness is a significant component of the residUUm engine. The visual attributes (i.e. shape, color and size) assigned to the particles are completely random as are the attraction forces and characteristics of the generated sound: pitch and timbre. Randomness, in conjunction with the

[4]http://www.lab-eds.org/punktiert
[5]https://processing.org/
[6]http://opensoundcontrol.org/
[7]http://puredata.info/
[8]https://youtu.be/eidlWCgSoJY

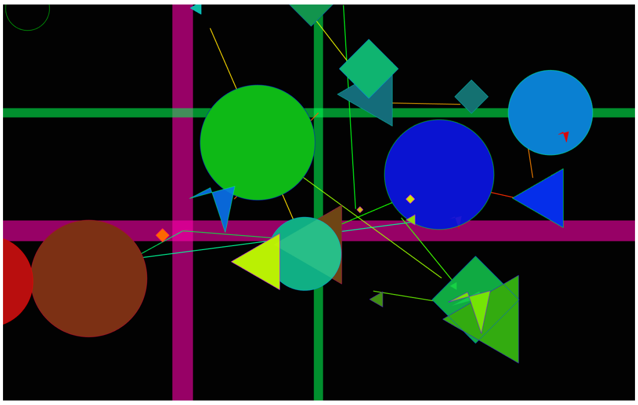

Figure 2: Scene from residUUm

limited set of controls available, continuously challenges the performer to control a system that is not designed to be entirely controlled. For example, the performer can use the gravitational attraction force of the cursor to move particles throughout the screen, however, it is difficult to interact with an individual particle. Separating two or more colliding particles or moving only the intended particle(s) while leaving others is a challenge. Keyboard controls are available to the performer to set the initial lifespan of newly generated particles and the rate at which particles on screen fade out. The user can control the lifespan of particles in advance of their creation. After a group of particles is created, the performer has no control over the lifespan of individual particles, but can adjust the global fade rate of all the particles on screen. The performer must plan deliberate actions that disrupt the function of the system in order to control individual particles.

To obtain a background drone, a particle has to be selected and its lifespan sustained infinitely. However, obtaining quiet drones to keep in the background is not a straightforward process. ResidUUm is not programmed to generate quiet particles. Particles have to be created and their voice and appearance faded to a point at which they almost disappear. A strategy was devised to select a single particle. First, the fading rate of the lifespan was set to an average value and, using the mouse, one particle was generated at a time. If the particle did not meet the required characteristics of pitch and timbre, it was discarded by allowing the lifespan to immediately fade out. If the particle's characteristics were selected and retained, its voice was dampened before its lifespan was sustained. A technique used to smoothly transition from a loud section with many particles to a quiet section was to allow most of the active particles to fade out until only a few remained barely visible and then set them to sustain for the duration. These particles became the background drones for the next section.

Overall, the interaction with the system shifted the intended presentation from an arranged musical performance towards that of a video game predicated on chance. The particles generated by the computer were seen as random challenges, which the author could discard — continuing to curate particles that aligned with his plan — or use as a creative stimulus to spontaneously compose based on the characteristics presented.

The keyboard controls were used to create multiple layers of interaction between different sets of particles. The use of layering in this performance reestablished the landscape of Mekela's desktop space in the elemental hierarchy of audio-visual performance by using depth to expand the volume of the screen. Although the condition of an entirely visible interaction was not satisfied due to the lack of explicit feedback communicated to the audience when the keys are engaged, implicit feedback is indicated by the relationship between groups of particles. Some sustain their lifespan while others fade over varying periods of time. Some project their sonic characteristics while some remained quiet. This use of the canvas created visual and musical relations between separate groups and individual particles that sought to extend the limits of residUUm's visible interface.

4.2 Evaluation of Performance 1

4.2.1 Quantitative Results

The evaluation of the first public performance of residUUm was conducted using an audience survey created and collected by Correia. Demographic information about the age, gender, occupation and experience with sound and audio-visuals was queried from each participant. For each of the three two-part questions, the audience was asked to rate the performance on a Likert scale, from 1 to 5, based on the variety and diversity of the presented visuals, the perceptual relationship between the audio and visuals and perceptual impact of the performers actions upon the visual presentation. The second part of each question was open-answer. Participants were asked to complete one sentence about the audio and visuals, the relationship between them and the overall performance.

The results of the survey are shown in Fig. 3. Most participants, who answered, felt that the visuals in residUUm were diverse and varied. residUUm's particles are diverse in shape and varied in hue and opacity. 27 viewers felt the relationship between the sound and visuals were good. And, most felt there was a fair to clear causal relationship between the performer's actions and visual feedback. The low responses for both questions may account for the events triggered by keyboard controls, which were hard to discern, or the lack of visual feedback for sonic events.

4.2.2 Qualitative Results

We performed thematic analysis on the results of the open-answer questions following the step-by-step guide with six phases of analysis provided by Braun and Clarke. We started by getting familiar with the feedback data, generating initial codes, then searching and reviewing themes, and finally defining and naming themes[3]. Five themes emerged from the thematic analysis of qualitative data collected from the audience:

Theme 1. Viewing Experience: "Fun" as a motif was mentioned in response to all three open-answer questions: the variation of visuals, the perceived audio visual relationship and as a critique of the overall performance. The presentation was described as "evocative" and "engaging," the viewing experience as "palatable" and "tangible."

Theme 2. Interaction: Audience members described the interaction in terms of the interface and the relationships between the particles. The cursor was a "focal point," making the production easy to follow along with the performer's actions. One participant suggested that residUUm would easily adapt to a gestural interface using the Kinect or Leap Motion. The interaction was noted for promoting "interesting ideas" for creating sound.

Theme 3. Audio-Visual Aesthetics: The aesthetics of residUUm was likened to cartoons and recalled Norman McLaren's animations as well as video games; the colors "beautiful," "hypnotic," "bold" and, the performance "handsome" and "visually nice."

Theme 4. Integration: The relationship between the audio and visuals was perceived as seamless by most, "like

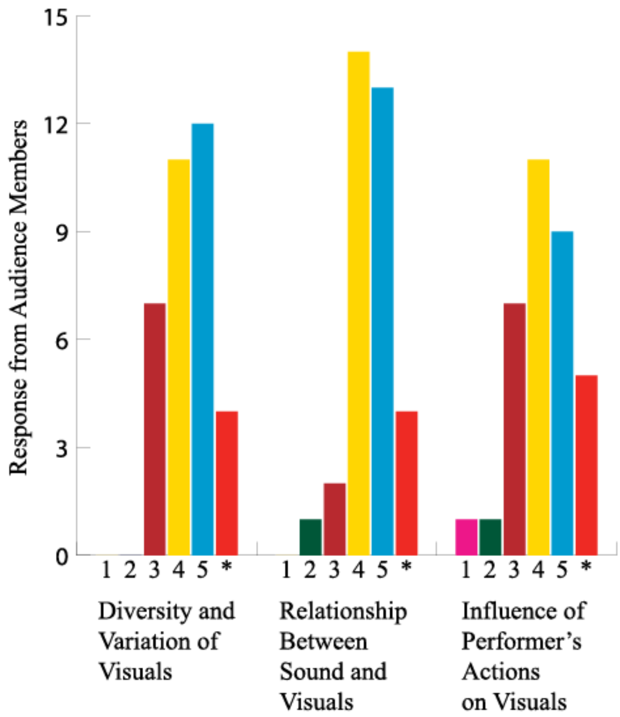

Figure 3: The questions were rated from the least to the most diversified visuals, from poor to good audio/visual relationship, and the last rated the performer's actions from obtuse to clear. The asterisk(*) denotes the number of participants who did not provide an answer.

graphic and sound were giving birth to each other." The integration between the two modalities was "inseparable," "traceable," and "onomatopoeic."

Theme 5. Performance Clarity: The relationship between the audio and visuals was described as clear enough to follow and "well established," although the function of relationship was not discernible for the entire duration of the performance. The ability of viewers to detect the causality of effects beyond the main mechanic of drawing shapes to generate sounds was "obscure." Two participants mentioned that causal actions beyond the particle generation (i.e. the keyboard effects) were hard to depict as they were "unclear what other changes the parameters did" until after they were explained.

4.3 Performance Strategy 2 — a learned approach to audiovisual composition

residUUm was presented and performed, by the first author, to an audience at the Sound/Image Colloquium on November 7, 2015 at the campus of the University of Greenwich, London.[9]

This performance sought to address the obscurity of the causal effects and visibility of the interface. The first step was to decrease the upper limit of sizes at which particles are generated. Bigger shapes commandeered large percentages of the canvas eclipsing smaller neighbors, which forced viewers to speculate about what audible, but visually occluded interactions were taking place between particles beneath and around them. Secondly, the visuals were modified to introduce contrast in the canvas. Thresholds were set to trigger changes in Saturation and Brightness of the background. This change added monochromatic moments to disrupt the monotony of the color permutations within the background and shape transitions. Thirdly, two graphic styles were created to visually indicate collisions as shown in Fig. 4, which were only perceptible through sonic events during the first performance. The line thickness was variable (controlled through the keyboard) for one and static for the other. Lastly, a parameter to adjust the framerate was added to afford the performer the ability to vary speeds throughout the audio and visual composition. At a continuous 24 frames per second, no opportunity permitted the observer's gaze to rest, or the performer to build suspense, tension or anticipation.

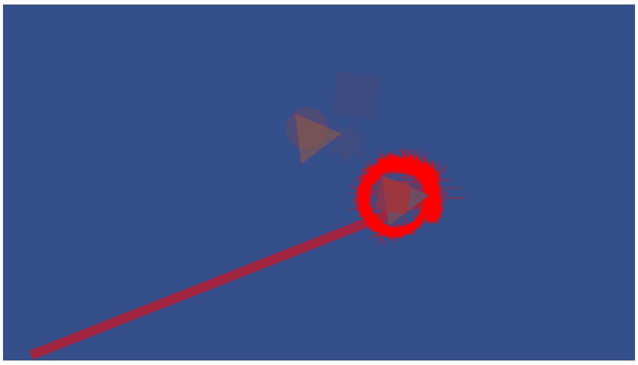

Figure 4: Visual collision between particles

The aim of this performance was to create a learned causal relationship between the performer's actions and the projected result as observed by the audience. The particle system was introduced to the audience by slowly increasing the number of shapes and their accompanying sounds to the canvas. The author interacted with individual and groups of shapes to familiarize the observer with the causal effects of cursor as larger collections of shapes were generated to coexist on the same canvas. The lifespans of the particles were varied over differing periods of time while colors evolved in the background to influence the perception of their hues [15]. Collisions were triggered to show how they were represented by a graphic style. The performer's actions responded to the spontaneous contrasted visual events by tempering movement within the canvas. The framerate was decreased at irregular, increasing intervals until a speed at which visual patterns appeared. The patterns, which interrupted the operation of the system, created multilayered interactions visually and audibly. A visual impression, which acted as an underlying rhythmic structure helped to support the audible composition like a visual beat. The framerate was slowly increased with periods of rest between slower speeds to emphasize the relationship between the discrete visual beating and the continuous audio composition. Upon reaching the normal 24 frames per second, the lifespan of currently spawned shapes were decreased until the din transitioned to silence.

4.4 Discussion of Performance 2

The author was able to follow the planned strategy for the performance, but discovered that using keyboard controls imposed new obstacles on usability. The cursor was used to indicate the shape and color of the particles and new keyboard controls were added to make adjustments to the graphic style of the collisions and framerate. The keyboard controls which regulate the lifespan of the particles and the graphic drawing style offer audible and visual feedback to

[9]https://youtu.be/AAQIPeOyCpk

the performer. However, other than the visual feedback from the canvas about the descending or ascending framerate, the performer is unaware of the precise framerate. A solution is needed to offer usability and causal cues for the user and the observer, respectively. The visible interaction of the performer is the primary tenet of a successful AVUI [8]. Finding solutions to expand interaction without relying on the keyboard as a performance interface needs to be addressed.

There are two ways this performance can be evaluated. Using an audience study to compare the new version of residUUm with the original could investigate how the changes made to the system effect the quality of the visuals, the relationship between the sound and visuals, the perception of causal actions of the performer, the overall performance and interface visibility. However, there are new questions to pose that examine only this performance. A visible indication has been added to the collision events, which leads to inquiries about the visual expectation for sonic events that occur in nature, the visual representation of a sonic event in which bodies with various sizes and sound properties adopt characteristics from the physical world, life experience and memory. In addition, defining the minimal visual characteristics necessary for an audience to distinguish subtle differences between similar events deserves investigation. The visual contrasts, pacing and variation of timing using the framerate, their effect on the composition of vision and audition individually and as an integrated couplet may be another line of inquiry. The ability to generate patterns, which is currently described as implied audio, using visuals was a consequence of varying residUUm's operating speed. At slower framerates the audio continues to provide continuous feedback from the mapped particles, however, the response from the visuals are divorced from their united coupling. Examining visual beats that occur without a paired audible response in an audience setting may reveal findings about linkages between audition and visual counterpoint, consonance and dissonance and, the impact of layered visual interactions on the performer and observer experience.

5. CONCLUSIONS AND FUTURE WORK

This paper examines mapping and performance strategies used in residUUm, an AudioVisual User Interface for audiovisual performance. Two performers utilized different approaches that layered interactions as a method to express causal audiovisual relationships while performing with residUUm's shared interface. This multilayered approach to visible interaction aimed to examine the challenges of performing with and observing an AVUI.

residUUm, as an audiovisual tool, is limited by the keyboard interface of the laptop [11]. We propose two methods to create a visible interface through which all actions of the performer are clearly projected to the observer. Yielding part of the interaction to the system to negotiate dynamic relationships between particles may afford the user to focus attention on higher level interactions and contribute to a collaborative role with the machine as an external influence on internal audiovisual interactions [10] [11]. Currently, many of the local and global attraction forces programmed in residUUm are hardcoded. Dynamically mapping the forces to represent spatial, sonic and color relationships may aid both the observer and the performer with creating improved, visible interactions.

6. ACKNOWLEDGMENTS

This work is partly supported by EPSRC grants EP/L019981/1 and EP/K009559/1. We wish to thank Correia and Tanaka who inspired the initial development of this research.

7. REFERENCES

[1] *Merriam-Webster's Collegiate® Dictionary, Eleventh Edition*. Merriam-Webster, Inc., 2016. residuum.

[2] P. Auslander. *Sound and Vision: The Audio/Visual Economy of Musical Performance*. Oxford University Press, 2013.

[3] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

[4] S. Callear. *Audiovisual particles: parameter mapping as a framework for audiovisual composition*. PhD thesis, Bath Spa University, 2012.

[5] T. Ciufo. Real-time sound/image manipulation and mapping in a performance setting. In *Proc. MAXIS Festival of Sound and Experimental Music*, 2002.

[6] G. Cooke. Liveness and the machine.

[7] G. Cooke. Udu: A live audio-visual performance. In *Proceedings of GA2009 12th Generaive Art Conference, Politecnico di Milano, Milan, Italy*, pages 14–17, 2009.

[8] N. N. Correia and A. Tanaka. Prototyping audiovisual performance tools: A hackathon approach.

[9] N. N. Correia, A. Tanaka, et al. User-centered design of a tool for interactive computer-generated audiovisuals. In *Proc. of 2nd Int. Conf. on Live Interfaces*, 2014.

[10] T. Dillon and H. Daanen. Designing self generative systems for audio-visual composition and performance. 2005.

[11] A. Eldridge. Cyborg dancing: generative systems for man machine musical improvisation. *Proceedings of Third Iteration*, 2005.

[12] J. Hook, D. Green, J. McCarthy, S. Taylor, P. Wright, and P. Olivier. A vj centered exploration of expressive interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1265–1274. ACM, 2011.

[13] P. Hovey. Exploring audiovisual linkages. Technical report, Accessed 26/03/2012. url: http://www. create. ucsb. edu, 2009.

[14] A. Hunt, M. Wanderley, and R. Kirk. Towards a model for instrumental mapping in expert musical interaction. In *Proceedings of the 2000 International Computer Music Conference*, pages 209–212, 2000.

[15] J. Itten and F. Birren. *The elements of color*. John Wiley & Sons, 1970.

[16] M. Lew. Live cinema: designing an instrument for cinema editing as a live performance. In *Proceedings of the 2004 conference on New interfaces for musical expression*, pages 144–149. National University of Singapore, 2004.

[17] M. Makela. The practice of live cinema. *ARTECH 2008*, page 83, 2008.

[18] E. Merz. *The Signal Culture Cookbook*. Signal Culture, 2015.

[19] J. Schacher. Live audiovisual performance as a cinematic practice. *The Cinematic Experience, Sonic Acts XII*, 2008.

[20] H. Von Helmholtz. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Longmans, Green, 1885.