PART II
THE THEORY OF LIVE VISUALS

Chapter 10 – A Parametric Model for Audio-Visual Instrument Design, Composition and Performance

Adriana Sá
https://orcid.org/0000-0003-3919-3121

Atau Tanaka
https://orcid.org/0000-0003-2521-1296

Introduction

We may think of an audio-visual performance as a construction of experienced time, which depends on multiple, intertwining variables. While the performer may compose and perform by thinking directly or indirectly about these variables, the audience will likely prefer to focus on the experience itself, and not try to glean insight on low level details; however, creative compositional processes can benefit from analysis, and understanding how a work is perceived.

There are many ways of elaborating a performance: creating design specifications, scoring the sonic and visual materials, and planning the arrangement of the physical space. Each of those activities requires a its own process and method, and some decisions are more subject to change than others, depending or not on complementary aspects. As practitioners, we feel need of a unifying, complementary method, which can be used to set out our general approach for those different work stages at once, as well as in parallel.

We introduce the notion of the parametric visualisation model as a tool for the compositional process. By systemising a set of variables, i.e. parameters in a graphical way, a parametric visualisation model can reveal relationships and interdependencies between those parameters. Instruments and performance situations can then be represented with a set of axes, and analysed accordingly. As an example, the model created by Birnbaum et al. reveals relations between interaction, sound organisation, physical distribution in space, and semantics of sound.[1] Alternatively, the framework created by Thor Magnusson reveals how digital music devices condition interaction and sonic results,[2] while the one created by Marko Ciciliani reveals how an electronic music performance might draw the focus to the performer or the environment.[3]

In this chapter we propose a parametric model that is useful in audio-visual instrument design, composition and performance. We can draw a separation between those activities, but in practice that separation might not be so obvious: ultimately, the iterative creation process must always consider the final, global experience. We take a perceptual approach in conceiving our model, and intend it to be applicable to the broad diversity of aesthetical options and technical platforms. The model is modular: one can discard part of the parameters so as to analyse any time-based work, including recorded audio pieces and films. On the one hand, the model enables the separate analysis of performer-instrument interaction, sound, image, audio-visual relationship and physical setup. On the other, it enables the analysis of how the combination conducts the audience's experience.

A panorama of creative questions

We use the term 'spatial presence' to describe the subjective experience of being physically located in a mediated space. Studies in sound art[4] [5] and media theory[6] [7] [8] [9] indicate that spatial presence is influenced by individual predispositions, the characteristics of sound and image, and audio-visual relationships. Other discussions focus on how the psychological space of the work is informed by loudspeaker placement.[10] [11] From these theorizations, we conclude that spatial presence depends on multiple variables, whose combination informs the semantics of a work. We define semantic typologies to assess the global semantics of a composition or performance based on their relative weights. In addition, we draw upon Simon Emmerson's notion of 'performative arena' [12] to enable the definition of an additional high-level parameter, which can provide cues that other, detailed parameters may not provide.

Michel Chion observed that the combination of sound and image generates a third audio-visual element.[13] However, only a part of the sensory information reaches conscious awareness. The question is, how can that combination benefit the experience of music, be it sonic or visual? The abstract filmmaker Stan Brakhage created silent films out of a feeling that sonic music dominates the subtler rhythms of vision. In the last two decades of his life he created audio-visual movies, where he breaks any direct connection between picture and sound.[14] The converse exists in the musical domain, where the Salomé Voegelin points out how vision dominates over audition in our oculocentric culture.[15] From Pierre Schaeffer [16] to recent acousmatic composers many people have argued that sounds must be detached from their originating cause to be fully experienced. Jeff Pressing, a composer and cognitive researcher who investigated the audio-visual relationship in digital 3D environments, noted that perception operates from vision to audition whenever a direction of causation is discernible.[17] The graphic performer Meghan Stevens believes that the music remains dominant when the audio-visual relationship is partially congruent, yet she is the first to admit that her theories rely on limited evidence.[18]

Questions of sensory dominance therefore pose a challenging problem. In neuroscience, Sinnett et al. found that it depends on attention, and that attention can be manipulated so that one sense dominates over another.[19] He does not say how, but extrapolating from creative practice, audio-visual theory and the science of perception enables us to identify the variables that influence how the audience's experience can be driven through the sound, the image or the audio-visual composite. Our model considers not only the audio-visual fit, but also the sonic and the visual dynamics.

Interaction – between media, between performer and material, and between performance and spectatorship – is fundamental in the exploration of perceptual experience. Different artists address interaction through a wide range of strategies and with very different creative motivations. Tod Machover and his team at MIT Media Lab developed interfaces intended to compensate for “people’s limitations;”[20] this also implies that the software prescribes which output results are desirable, and which are not. In contrast, Joel Ryan (who pioneered the digital signal processing of acoustic instruments) encourages us to make control as difficult as possible, linking musical expression to effort.[21] Indeed, one interface might allow for unpredictability so as to convey the reciprocal interaction between performer and instrument, whilst another interface prescribes the outcome beforehand. Interaction design is thus determined by different creative principles corresponding to different notions of expression. Our model should include a distinguishing parameter, applicable to any sonic, visual and audio-visual system and instrument. We can quantify effort relative to the amount of cognitive processing required in a given task. Effort might further reflect in the dynamics and the semantics of sound and image.

In the next sections of this chapter we begin by presenting each parameter independently. We illustrate their use by giving a range of artistic examples. We then explain how their combination facilitates the analysis of expression, spatial presence and sensory dominance. Finally, we demonstrate how the model can be used in creative practice and show its usefulness as a compositional tool.

Interaction effort

One of our objectives is to facilitate the analysis of how interaction designs convey expression, be it in the sonic, the visual or the audio-visual domain. The term ‘design’ is here not constrained to the activity of designers; we purposefully do not distinguish between idealising, crafting, composing and performing.

Birnbaum et al.[22] and Magnusson[23] propose parametric models to analyse interaction with digital music devices. Both their models include parameters related to the performer’s control over the device and the prior system knowledge required for interaction. Here, we summarise those variables into a single parameter: the performer’s cognitive effort, that is their mental information processing, conscious and unconscious.

In previous work on new musical instruments, we elaborated on how different levels of effort convey different notions of musical expression.[24] This understanding is equally applicable to the sonic, visual and audio-visual domains. As a parameter, effort can be characterised as follows:

- **Little effort** means one of two things: either the work does not depend much on real-time interaction, or the relationship between intention and resulting output is linear and clearly perceivable.
- **Medium effort** means that the performer needs particular skills to play the instrument, but a sense of immediacy conveys fluency and timing, and/or technical configurations rule out undesired outcomes.
- **High effort** implies particular skills and/or high cognitive demand; the interaction with the system does not feel immediate, and/ or the system does not rule out any outcomes.

In our parametric visualization model, a single axis suffices to represent the real time interaction effort, motor and/or conceptual.

An example of low interaction effort can be found in Phill Niblock's *Movements of People Working*, performed since 1973.[25] These works show repetitive movements of manual labour combined with massive drones of sound, rich in harmonics and overtones. The images are created in advance, and Niblock's graphic scores for the sound have been interpreted by many musicians. His interaction with sound and image in performance is very sparse. A different example of low effort can be found in *Music for Solo Performer* by Alvin Lucier (1965),[26] where he uses a brain interface to activate multiple percussion instruments. The interface was crafted so that the lesser is the brain wave energy, the stronger is the actuation over the instruments. This work raises an interesting issue: the interaction with an effortful interface can be effortless. Indeed, brain waves are hard to control. But in an interview Lucier explained that he didn't want to show mind control, because he preferred the discovery of how his brainwaves sounded.[27] To him, composition was about how to deploy the loudspeakers and what instruments to use. For him, brain music performance was not about making an effort to create certain brain activity, but rather enter into a meditative state of biofeedback.

Medium effort can be manifested in a range of behaviours over time, an adaptation to unpredictable conditions, a monitoring of results in relation to a reference source, or an anticipation of changes in oneself or the environment. Jeff Pressing coined the term 'dynamic complexity'[28] to describe this in music. We can say that medium effort implies behavioural deviations, and reactions to those deviations.

A musical example of medium effort is in a performance by Joel Ryan (electronics) and Evan Parker (soprano saxophone).[29] Ryan's instrument is a digital signal processing of the saxophone where processing parameters are performatively manipulated. The two instruments are therefore in an interdependent relationship. The sounds of the saxophone and the electronics converge when their loudness and tone are the same, then they cause attention to focus on subtle tonal shifts, diverging progressively as one timbre emerges from the other, so as to converge again. Each performer plays their instrument with its normal cognitive load, but also must pay attention to the consequences that their play has on the material provided to the other performer. Another example of medium effort is found in the *Violin Power* series of audio-visual performances by Steina Vasulka[30], where she controls laser-disc video from an electric MIDI violin[31] (Figure 10.1). Vasulka creates audio-visual tension whenever certain musical gestures on her instrument cause the video player to seek to different segments of video footage. These transitions are abrupt at the visual level, although not at the sonic level. The subsequent return to audio-visual synchrony creates a convergence, which causes a sensation of release.

Figure 10.1 - Steina Vasulka playing her audio-visual instrument From the STEIM "Waisvisz archive." Photographer unknown. Used by permission of STEIM.

A high level of effort conveys yet another type of expression. A paradigmatic example can be found in the work of Martin Howse, where he uses the conductivity of earth as part of a system to perform noise and electronic music.[32] He investigates the links between geophysical phenomena, software and the human psyche, proposing a return to animism within a critical misuse of scientific technology. In performance, his interfaces combine a diversity of chemical substances, earth materials and computers. Similar to Lucier, Howse emphasises discovery, as opposed to control. But very differently, the output is highly dependent on his real-time decisions and actions; on direct intervention with his materials – literally digging into the ground (Figure 10.2).

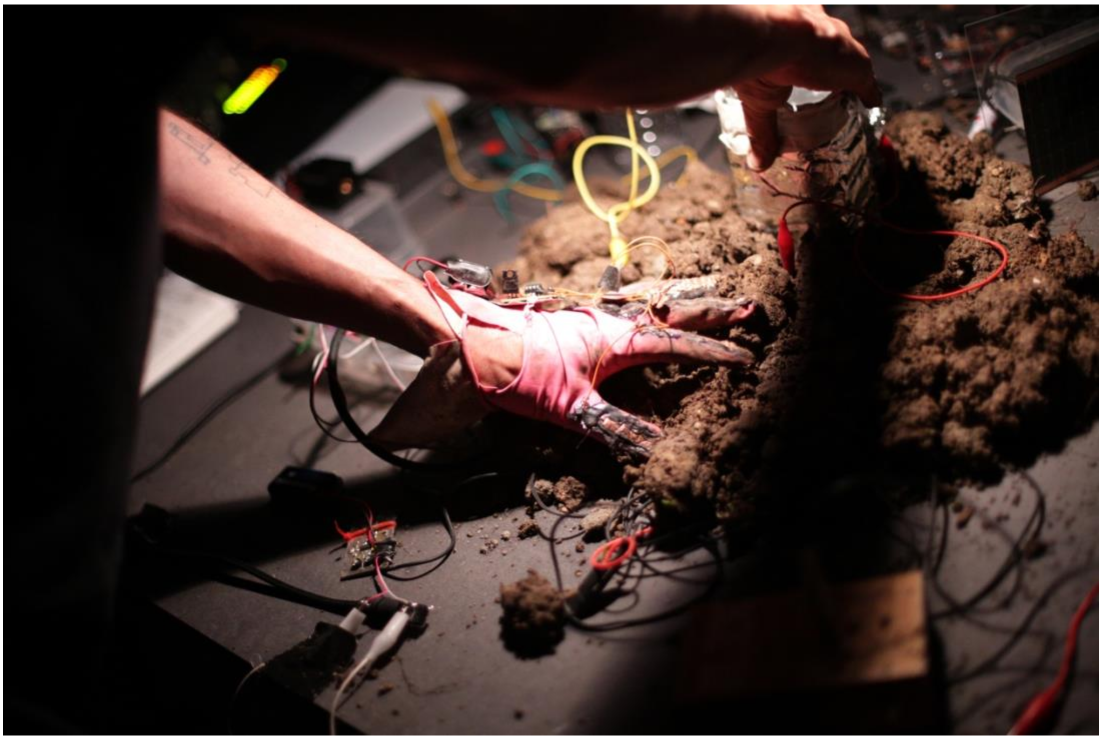

Figure 10.2 - Martin Howse, Earth Voice, Media Mediums, 2014, Paris. Used by permission.

Continuity

Turning to the audience's experience, one important concern is how the dynamics of sound and image drive attention. Drawing from neuroscience and psychology, we created a taxonomy of continuities and discontinuities related with intensity and attention.[33] We defined intensity as the psychophysical impact of any change in the chain of stimuli causing an increase in neural activity. That neural activity is a measure of attention, which means that we can quantify intensity based on how attention works.

Attention is automatic when driven by salient events, such as the sudden appearance or disappearance of a stimulus. Such events counteract biophysical expectations, causing a great increase in neural activity; that increase is consistent with primary survival instincts. Conversely, attention is under individual control when expectations are fulfilled; it evokes less neural activity, then, because there are no significant changes in sensory information. It is important to note that expectations depend greatly on the panorama – previous and simultaneous events, as well as the time length of experience. Meanwhile, the threshold between deliberate and automatic attention can be fuzzy, as attention causes us to optimize perceptual resolution, so as to better process information related to the attention target.[34] As deliberate attention makes detail changes more intense, we also become more susceptible to automatic attention.

Intensity is therefore proportional to perceived discontinuity. It depends on the event itself, on the panorama, and a person's perceptual resolution. We propose the following taxonomy:

- **Steady continuity** has no intrinsic motion; it is of lowest intensity, dispensing with attention. Attention is likely to deviate and focus upon any simultaneous stimuli, or upon internal states.
- **Progressive continuity** occurs when successive, non-abrupt events display a similar interval of motion. It fulfils the expectation that once something begins to move in a certain direction, it will continue to move in that direction (Gestalt principle of *good continuation*).
- **Ambivalent discontinuity** refers to the threshold between continuity and discontinuity. At low perceptual resolution a predictable logic is shifted without disruption. At high resolution, discontinuities become more intense. Higher intensity implies greater attention, and lower intensity implies less attention.
- **Radical discontinuity** violates psychophysical expectations, prompting automatic attention. It is of highest intensity, implying greater neural activity. Radical discontinuities are always prioritised in the stimuli competition to reach conscious awareness.
- **Endogenous continuity** corresponds to the mental representation of perceptual motion; it can embrace all the other types of continuities and discontinuities. We use the term endogenous to stress that perceiving a coherent relationship between them depends greatly on the individual.

Our parametric visualization model employs two axes to represent the sonic and visual dynamics, as shown in Figure 10.8, ahead in this chapter; “SC” means steady continuities, “PC” means progressive continuities, “AD” means ambivalent discontinuities and “RD” means radical discontinuities. Endogenous spans all the dynamics. Other types of continuities and discontinuities can be illustrated with paradigmatic examples.

A musical example of steady continuity is Elaine Radigue’s *Triologie de la Mort* (1998),[35] a three-hour drone piece where we hardly perceive any overtones; the work relates strongly to Tibetan Buddhism. An audio-visual example is La Monte Young and Marian Zazeela’s *Dream House*,[36] an ongoing installation (since 1962) that defines a vibratory space through the combination of continuous sound and light frequencies, experimenting on how people are drawn to inhabited it.

The notion of progressive continuity can be illustrated with any gradual increase or decrease in loudness, tonality, brightness, colour, density, rhythm or time length. An example can be found in Gary Hill’s film *Black and White Text* (1980),[37] which explores a relationship between geometric black and white figures and human voice. As the work unfolds, the intervals between the words and the visual shifts become progressively shorter, while sound layers accumulate and rectangles multiply on screen. Importantly, progressive continuity implies perceived motion. If the progression happens so slowly that we cannot apprehend any change (as happens in Radigue’s *Triologie*), it would instead be considered a case of steady continuity.

While progressive continuity entails motion in a clearly perceivable direction, ambivalent discontinuity entails multi-directional motion. An example is seen in Thomas Wilfred’s *Lumia* from the 1930s.[38] The Lumia are dynamic light-paintings created with a visual instrument that combines projectors, reflectors and coloured slides, so as to produce polymorphous streams of colour, which invite attention to focus on subtle detail changes. We can revisit Phil Niblock’s *Movements of People Working*[39] to see how perception focuses on the nuances of sound and image - the repetitive movements of manual labour and the continuous mass of sound. As we increase perceptual resolution, the detail variations become more intense. The fuzzy threshold between ambivalent discontinuity and radical discontinuity is particularly evident in *Stellar* (1993), one of Stan Brakhage’s silent abstract films drawn directly on the film strip. The film has elements of continuity; all frames are made with the same technique and the same colours. And yet, as we focus on the visual detail changes proliferating in the fast-changing frames, we also become very sensitive to how the interplay of discontinuities grounds the construction of time.

Radical discontinuities can be used so as to create rhythmic patterns, as happens in Vasulka's *Heraldic View* (1974)[40], where visual patterns are created by an audio synthesizer. When the duration of the experience is short, each abrupt event prompts automatic attention, causing a sudden increase in neural activity. Yet after a while, the sequence of elements fulfils expectations, as happens with any pattern; in this way, radical discontinuity becomes steady continuity. Alternatively, radical discontinuities can be explored so as to tease and counterpoint expectations. For example, Ryoji Ikeda's performance *Superposition*[41] creates radical discontinuities with sudden blackouts; the contrast with moments of progressive and steady continuity makes those discontinuities more intense (Figure 10.3).

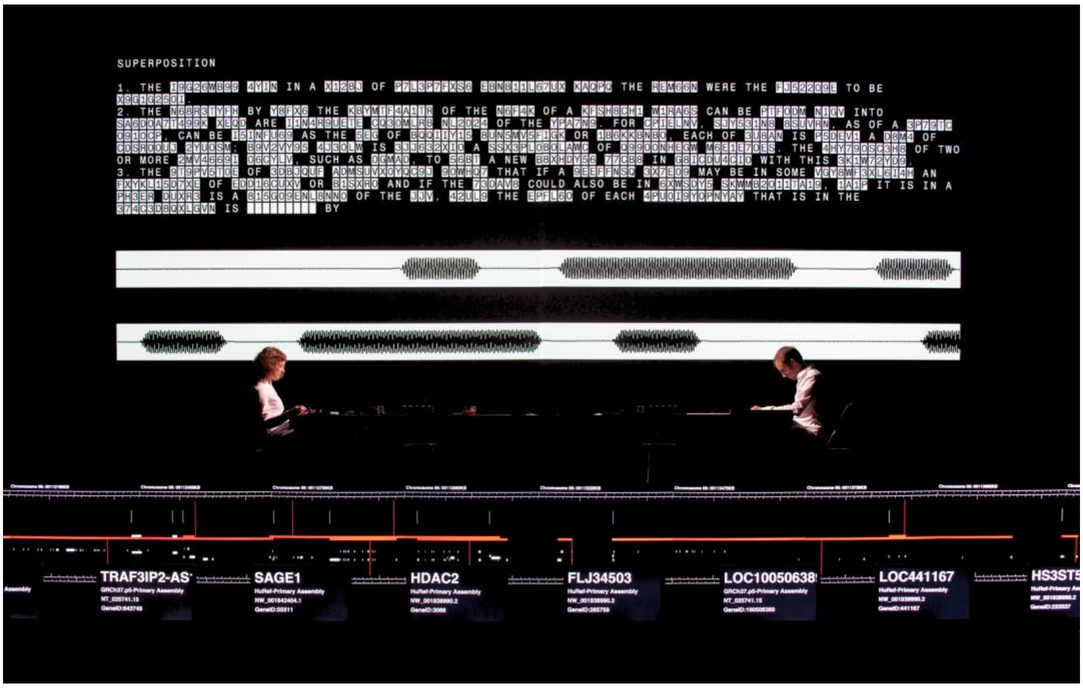

Figure 10.3 - Ryoji Ikeda, superimposition, 2012 performance. Kyoto Experiment, Kyoto Art Theatre Shunjuza, Kyoto, 2013. Photo by: Kazuo Fukunaga. Courtesy: Kyoto Experiment. © Ryoji Ikeda. Used by permission.

Audio-visual fit

The way perception prioritises sensory information is influenced by the dynamics of sound and image, but the audio-visual relationship is equally important. In audio-visual theory, Chion coined the term *added value* to describe the surplus of synchronization.[42] It is crucial not to misinterpret the term, because the meaning of the audio-visual composite is not really added to the meanings of the sound and the image. On the contrary, it tends to override those meanings. In experimental psychology, Kubovy & Schutz showed that the aural discounts the visual and the visual discounts the aural based on concepts of causation.[43] [44] They coined the term 'ecological fit' to describe how automatic interactions between the senses are governed by those concepts.

The greater is the ecological fit, the more we ignore any diverging sensory information. We can explain this in terms of cognitive 'efficiency.' A high level of fit leads to integrated perceptual encodings and representations, which require less neural activity than separate ones.[45] Drawing from the science of perception and audio-visual theory, we investigated perceived audio-visual relationships,[46] and identified three levels of ecological fit:

- **High fit** means that the audio-visual relationship conveys conclusive information about causes and effects. Our perceptual mechanisms prioritise information that comply with those conclusions, producing integrated mental representations. High fit is of low intensity, because it does not require much cognitive processing.
- **Medium fit** means that one senses causation without understanding the base cause and effect relationships. It is of medium intensity and requires a medium level of cognitive processing. It conveys perceptual chunking, but the process of audio-visual binding remains ambivalent: one can form integrated as well as separated representations of the sounds and the images.
- **Low fit** means that the pairing of sound and image does not activate prior memories of causation. Perceptual binding is weak, requiring perception to create new chunks of memory, with a large amount of cognitive processing. That means high intensity.

Our parametric visualisation model uses two axes to represent the extent to which what we see fits with what we hear: one for the fit between sound and image, and the other for the fit between physical gesture and system output.

High ecological fit can be illustrated with Norman McLaren's abstract animation film *Dots* (1940),[47] where sound and image are synchronised one-to-one. The visual elements consist of dots, which McLaren painted directly onto clear frames of film. The sounds were created in the same way, with dots painted directly into the area on the filmstrip usually reserved for the soundtrack. Another example is *Noise Fields* (1974),[48] a fully synchronised video by Steina and Woody Vasulka. Made with analogue video synthesis processors, this work visualises and sonifies the energy of the electronic signal. Beyond the film and video art, many systems and instruments were designed to emphasise the union of audition and sight through one-to-one synchronisation. The *Ocular Harpsichord* created by Louis Castel (1730) is an early example. It consisted of a harpsichord with coloured glasses and curtains; when a key was struck, a corresponding curtain would lift briefly to show a flash of corresponding colour.[49] As a contemporary example, 3D positional audio-effects used in video games are intended to create a high audio-visual fit, and creative works such as Tarik Barri's *Versum*[50] explore this as a means of composition.

Figure 10.4 - Top: *Performance by Adriana Sá and John Klima at Maga Festival / Silos COntentor Criativo, Caldas da Rainha, 2018 (photo by Susana Valadas, courtesy of Grémio Caldense)*. Bottom: *studio setup of Sá's audio-visual instrument.*

From an earlier study, we coined the term *fungible mapping* to describe an audio-visual mapping exhibiting medium fit. It combines synchronised and non-synchronised components, exhibiting complexity enough to be confusing. In our study, participants were aware of a causal relationship, and aware of not distinguishing the base cause and effect relationships. As they could not segregate converging and diverging information, their sense of causation extended to the mapping as a whole. The study was greatly motivated by the iterative development of an audio-visual instrument,[51] which combines an acoustic string instrument and 3D software that operates based on the acoustic input (Figure 10.4).[52] [53] [54] It clarified how the instrument could confound the cause-effect relationships in spite of using a 3D engine – a technical platform intended to maximise the audio-visual fit. It also enabled extrapolations into the physical setup. In performance the relation between physical gesture and instrument output is sometimes synchronised and at other times not. Additionally, two stereo audio pairs crossed in space blur the relation between the visible sound emitters on the screen and the corresponding sounds emitted through the loudspeakers.

Figure 10.5 - Performance of Atau Tanaka's group, Sensors Sonics Sights, 3 Legged Dog, New York 2007.

Another example of medium fit in our own creative work can be found in the performances of Tanaka's group, *Sensors Sonics Sights.*[55] [56] The trio uses sensor-based digital musical instruments, capturing performer gesture to modulate 3D imagery and synthesised sound (Figure 10.5). Two members play sound and one plays image, with the connection between the media taking place through the traditional ensemble practice of synchronising by eye contact and gesture. The audience senses a causal connection between the performers' gestures and the sonic and visual outputs. Nevertheless, the nature of the instruments confounds the base cause and effect relationships. There is no technological connection between sound and image, but points of sensory unison convey perceptual binding. As perception doesn't segregate the elements that produce a sense of causation from the elements that do not, the feeling of causation extends to the audio-visual relationship as a whole. The practice of small ensemble chamber music performance, therefore, is extended to audio-visual performance.

Another way of creating medium fit can be found in many live coding events, where textual programming onstage generates music and/or visuals as it is written live in performance. The code is usually projected on a screen so that people can see the process, but the cause and effect relationships are often confounding. Alex McLean sometimes purposefully obscures his code to make it more difficult to read, while still showing some of the activity of the edits.[57] In Thor Magnusson's performances with the *Threnoscope*[58] the digital cause-effect relationships are exposed with a graphic notation system and real-time programming code; yet, even coders won't fully understand the cause-effect relationships because the code is relatively high-level and the system is complex.[59] In other words, medium fit is also compatible with consistent synchrony.

Laptop performances are often criticised for having a low fit between physical gesture and system output. But in watching an audio-visual performance or a film we are driven to perceive – and imagine - connections between the sounds and images, even when their fit is low and perceptual binding is weak. Often one can extrapolate meanings from video images of one thing coupled with sounds from something completely different, even if there is no synchrony.

Performer's position relatively to the image

We propose an additional parameter for characterising the performer's physical position relative to the image. It is useful to distinguish the following three types of arrangement:

- **Integrated** means that the image and the performer's physical body form a single visual scene, as happens when an image is projected upon a performer.
- **Separated** means that the image is separated from the performer, who is nevertheless visible. This type of arrangement can divide attention, or deviate attention from the performer.

- **Hidden** means that the performer is not visible. The audience does not see their agency, but knowing what type of interface is being used can influence how the work is perceived.

Our parametric visualization uses three discrete points to represent these types of arrangement.

The way the physical setup influences attention also depends on speaker placement, lighting, audience distribution and physical architecture. All these details count in the audience's experience. However, in order to facilitate the application of the parametric model, we do not parameterise each aspect independently. Instead, our model includes two high-level parameters – 'semantics' and 'performative arena' - which provide cues about variables that the other parameters do not address.

Semantics

Semantic organisation can be described with respect to causes and concepts, but attention dynamics have intrinsic semantics as well: every experience has meaning, including when we focus on perceptual motion itself.[60] [61] The mental representation of the work as a whole can be considered an endogenous continuity.

The notion of endogenous continuity expands Jeff Pressing's semantic characterisation of sounds,[62] so as to embrace the visual and the audio-visual domains. He distinguished 'expressive', 'informational' and 'environmental' sounds, stressing that these typologies normally overlap. Expressive sounds would include all kinds of music and song. Examples of informational sounds would be speech, alarms, and sonified data. Examples of environmental sounds would include animal calls, wind sounds, and the noises of machinery. We adapt these semantic typologies as follows:

- **Informational Semantics** prompt causal percepts, shifting attention to a meaning.
- **Expressive Semantics** means that the focus of attention is upon a central target.
- **Environmental Semantics** means a focus upon a context or environment.

These three semantic dimensions can be quantified independently. We can quantify the informational dimension by assessing our conclusiveness about a cause or meaning. It might be useful to look at semiotics, where the notions of icon, index and symbol characterise different types of relation between signifier and meaning.

But these notions do not suffice to quantify expressive and environmental semantics, which also exist when the informational load is low. Regardless of concepts and interpretations, the more attention that is focused on a specific target, the less it spreads through the environment, and vice-versa. Our parametric visualisation model also takes advantage here; the expressive and environmental dimensions of a work can be represented in a single axis. As such, two axes suffice to represent the three types of semantics.

When analysing a creative work, we might consider the semantics of interaction, sound, image, audio-visual relationship and physical setup. One can assess the semantics of each element, and estimate their relative weight in the global meaning of any particular work.

If we think of interaction in terms of cognitive effort, predictable, clearly perceivable interface behaviours provide a large amount of information about how the system should be interacted with, making the interaction effortless. A system that does not depend on real-time control is effortless as well, but the information content of the interaction is low, as the audience does not perceive to what extent the performer influences system output. Furthermore, perceiving effort implies interpreting causes and meanings. Perceived effort tends to attract attention, supporting expressive semantics.

Sounds and images have informational content whenever they evoke something beyond themselves. Symbolic systems such as programming code might provide a large amount of information if one understands the code, and very little if one does not. In other cases, the informational dimension can support the expressive or the environmental dimensions. For example, a piano recording leads us to imagine a piano and a pianist, and the recording of singing birds evokes a natural environment. Semantic categorisation might be less obvious in narrative film; it is useful to draw from Leo Braudy's distinction between 'closed frame' and 'open frame.'[63] A wide shot provides all the information necessary to interpret the image, leading attention to focus within the limits of the frame. Conversely, a close-up does not describe a scene, leading imagination to build what is not seen within the frame. In other words, a wide-shot of a landscape conveys expressive semantics, and a close-up of a person or object conveys environmental semantics.

The dynamics of sound and image have their own semantics, be it expressive or environmental. To parameterise these dynamics, we use the taxonomy of continuities and discontinuities. Steady and progressive continuities create a sense of environment because they fulfil expectations; attention can draw us towards the context/environment. Ambivalent discontinuities also leave attention under individual volition, but they entail more pathos; they attract more attention, reinforcing the expressive dimension of the work. Radical discontinuities make expressive semantics very strong: they prompt automatic attention, monopolizing conscious awareness.

The informational load of the audio-visual relationship is proportional to its ecological fit, determining the strength of perceptual binding. The binding is informed by concepts of causation,[64] which have informational load by definition. A high level of audio-visual fit provides a large amount of information, enabling integrated representations. Medium fit conveys a sense of causation, but the informational load is not very high as one does not understand the base cause-effect relationships. Finally, low fit provides little information about causation, making the binding weak.

The semantics of the physical setup are also important in performances and interactive installations. The central position of a performer, a sound source placed next to them or a spotlight over them will have the effect of directing attention to a central target, conveying expressive semantics. Conversely, the distribution of sound and light sources in space will emphasise the environment.

A performer's position relative to moving image is equally influential. A hidden performer confounds the extent to which the work is created in real-time, which means a decrease in informational semantics. For example, Ikeda rarely takes the stage, and is instead an invisible performer. The informational content of his work relies not on the performative gesture, but in the data information content of the sound and image – in his case often invoking a kind of information overload.

With a separated arrangement, the semantics of the physical setup are more expressive when the focus is on the performer, and more environmental when the focus is upon the spatial relation between his body and the moving image. We might consider the different arrangements of DJ/VJ setups as an example of this. If the focus is on a DJ onstage, the interaction between sound and image find an expressive arrangement. If, on the other hand, the DJ and VJ are not centre stage but is behind the dance floor in the DJ booth, the visuals and music might create a more environmental effect.

Finally, the semantics of an integrated arrangement can be expressive or environmental. They are environmental whenever the visual output functions like a stage scene, and expressive whenever the physical scale of the work equals that of the human body. In Chikashi Miyama's *Modulations,* the performer's body movements create movement in a particle system (Figure 10.6). The performer's position, coincident with the centre point of the image, with projection as much on the body as on the screen behind him create an expressive intensity. As the piece evolves and the particle system expands beyond the bounds of the body, with the sound spatialised around the audience, the same audio-visual material becomes environmental.

Figure 10.6 - Chikashi Miyama performing Modulations in the Kubus, ZKM, 2016. Photo: Chikashi Miyama. Used by permission.

Sometimes, the semantics of the system output are highly expressive due to sonic and visual discontinuities, while the integrated arrangement brings an environmental quality to the work. An example is in Ikeda's *Superposition*[65] where the performers are in front of a large visual projection, surrounded by multiple video monitors. Another example is in a performance by Metamkine at the Lausanne Underground Film & Music Festival (2012).[66] The performers sit in front of a large projection, using a Super 8 camera, colour filters and various devices to create a multitude of light effects and noises (Figure 10.7). In contrast with these works, Guy Sherwin created a series of silent performance-films where an integrated arrangement creates expressive semantics, and the image exhibits continuity. In these works, called *Man with Mirror* (1976-2011),[67] he forged a kind of *exquisite corpse* by exploring the relation between his physical body and the changing angles of its reflection on a live-manipulated mirror.

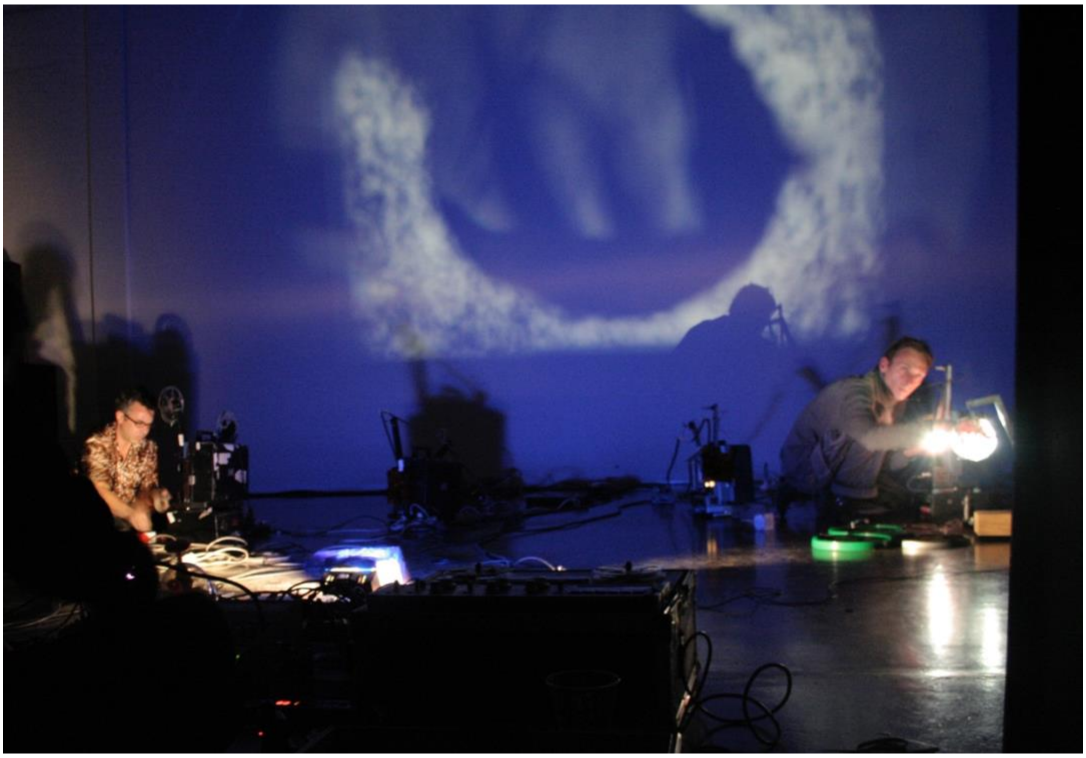

Figure 10.7 - La Cellule d'Intervention Metamkine performing during Kill Your Timid Notion tour, 2008. Photo: Bryony McIntyre. Used by permission.

Performative arena

To complete our parametric model we need a final parameter: one that enables us to summarise how the different semantic dimensions of a creative work intertwine so as to shape a performative arena. The performative arena corresponds to how the work creates the potential space of presence. It can contract and expand, inextricably related to attentional processes. Those processes might depend on the characteristics of the sound, the image and the audio-visual relationship, on the performer's interaction with the system, the speakers' placement, the spatial relation between performer and visual projection, if any, the lighting, the physical architecture and the audience location. Individual predisposition might be influential as well, but it does not depend on the work, and we do not intend to parameterise experience itself. We rather consider verifiable variables, and provide methods to interpret their relationships.

This final high-level parameter complements the other parameters, facilitating the disambiguation of certain aspects. It provides cues about elements that have no direct representation, such as the placement of speakers and the lighting.

We distinguish three types of performative arenas, which are not mutually exclusive:

- **Local arena** means that the work conveys a focus upon the performer. Expressive semantics are dominant.
- **Distributed arena** means that the work conveys a focus upon the environment. Environmental semantics are dominant.
- **Extended arena** means that the work conveys a subjective sense of presence beyond the physical performance space. It requires perceptual cues, which imply informational semantics.

Our parametric visualization model uses three discrete points to represent these three types of arenas.

The local arena relates to Ciciliani's notion of *centripetal* performance tendencies, where the focus is upon the performer.[68] An unequivocal example is when a sound source is placed next to a musician. We can also revisit Sherwin's *Man with Mirror,*[69] in which he integrates his physical image and its reflection on a mirror by using a light projector in a dark room, without any light reflections on the wall. The expressive scale of the work is reinforced with informational load, as the interaction is clearly perceivable. Furthermore, a large visual projection can convey the local arena as well. It happens when the performer is separated from the image and the image shows their interaction with the system. An example is in an expanded cinema performance by Arnont Nogyao (2017),[70] where the visual projection shows his interaction with the modified surface of a vinyl record.

Meanwhile, the distributed arena relates to Ciciliani's *centrifugal* performance tendencies, where the focus is upon space.[71] Niblock's performances serve as examples that create immersive environments. His drone music is played through multiple loudspeakers distributed in space, and our eyes are directed to large visual projections rather than to the performer, who sits in darkness.

While Birnbaum[72] model and Ciciliani's[73] model consider the relationships between the physical and the psychological space of a creative work, they do not address how a work might expand one's presence beyond physical space. Shifting spatial presence beyond physical space requires perceptual cues derived from the informational load of the sound, the image, or the audio-visual relationship. An example of the extended arena is seen in Vasulka's performance dedicated to Nam June Paik,[74] where the imagery recalls real-life situations: trees shaking in the wind, Vasulka playing the violin, Michel Waisvisz playing *The Hands*. The informal semantic dimension of that imagery is strong – by definition, video can emulate how we naturally see the world. That enables a subjective sense of presence beyond the physical performance space.

The local and the distributed arena are often combined. In Lucier's *Music for Solo Performer,*[75] attention is driven to the performer, who sits still in a central position, using his brain interface to activate percussion instruments. At the same time, attention is driven to the environment, because the instruments and the loudspeakers are distributed in the room, amongst the audience. Another, imaginary example would be a performer carrying a TV monitor displaying footage from the surrounding environment. Here, the distributed arena does not require a distributed physical setup. Furthermore, a creative work can combine all three types of arena. For example, Mick Grierson created a 3D composition and improvisation system that behaves like first-person computer games[76]. The user can create, adapt and combine elements with varying physical attributes to produce musical structures. The system does not exhibit pre-determined constraints on the environment, object properties and interactions. It responds to the user behaviour through adaptive algorithms. We can speak of a local arena because the performer is visible and the interaction design highlights their possible skills. We can also speak of a distributed arena because the system uses a multichannel audio system. And finally we can speak of an extended arena because the system looks and behaves like a first-person video game – it extends the sense of presence to the digital world.

While the three types of arena reflect the three semantic typologies, the inverse is not necessarily true. That is to say, the local arena can be assumed to be expressive, and the distributed area to be environmental. However, an environmental semantic does not presuppose a distributed arena. For example, in Steina Vasulka's performance there is no local arena because the performer is not visible, yet the visual discontinuities and sonic deviations create expressive semantics. Similarly, there can be environmental semantics without the arena being distributed or extended, such as a sonic mass of drones emitted through a single loudspeaker. Also, the semantics of an audio-visual performance can be informational without the arena being extended, as happens in live coding performances. Indeed, the parameters in our model complement each other without redundancy. The high-level ones can be used to assess how any low-level variable informs the meaning of the work, and how the product informs the feeling of presence.

The parameters relevant to expression, sensory dominance and spatial presence
As part of a discussion about expression in audio-visual performance, it may be useful to look at how the amount of interaction effort influences the semantics and dynamics of the work. Our model visualises those expressive dimensions on three axes, *Interaction*, *Dynamics* and *Semantics*. This could play an important, objective role in the development of an idiosyncratic, personal instrument. It was the case in the individual development of our own instruments, mentioned earlier, which we are also going to use in a future collaboration.[77] [78] We will present them in the final section of this chapter so as to illustrate how our model can be used not just in analysis of existing work, but in the planning, composition and creative practice.

Sensory dominance is another dimension for creative manipulation, particularly if one wants the audience to experience complex sonic or visual constructions along with their audio-visual composite. We have sought to bring clarification to the problem by drawing upon knowledge in neuroscience and psychology.[79] [80] Accordingly, the parametric model reveals sensory dominance through the axes 'Audio-visual Fit' and 'Visual Dynamics.'

Highly congruent audio-visual relationships produce integrated perceptual representations, where vision subordinates audition.[81] In fact, visual dominance occurs at the moment we conceptualise causation.[82] [83] Perceptual prioritisation can be inverted, or just become ambivalent when concepts are inconclusive. [84] When the fit is too low to produce any sense of causation, one can bind separate representations of sound and image, but that becomes demanding in terms of cognitive load, hence, attention is likely to focus on making sense of the audio-visual relationship, rather than on the music. Furthermore, sudden visual discontinuities cause vision to subordinate audition, as they attract automatic attention[85]. Importantly, the threshold between ambivalent and radical discontinuities should reflect a particular timescale, since over time, the same behaviours can appear more continuous or discontinuous. The parametric model can be used to represent instantaneous moments as much as summarising large-scale compositional structures.

An important element in any performance is a consideration of the audience's subjective sense of presence. We adopt the notion of spatial presence as a cognitive sense related to semantics and attention. The parameter 'Performative Arena' represents how the mediated space of an audio-visual performance may draw focus on the performer, on the environment, or on imaginary spaces beyond the physical performance space. As a high-level parameter, it also encapsulates information from other elemental parameters. By consolidating them, the performative arena may allow the model to convey forms of experience that a single parameter by itself may not specify. In this way, spatial presence can be inferred from the model as a whole.

Using the model in creative practice

We have shown how the parametric visualisation model can be used to analyse existing audio-visual instruments. It also provides a theoretical perspective from which to create new audio-visual performances, and develop new audio-visual systems [86]. We have applied the model in the development of the audio-visual instrument mentioned earlier in this chapter, which processes 3D sound and image based on an acoustic zither input.[87 88] We also used the model to imagine how a set of musical collaborations would affect the audio-visual performance work as a whole [89 90]. The interaction effort, the visual dynamics, the level of audio-visual fit and the physical setup would remain similar, but the sonic dynamics and the global semantics would be quite variable. The model can put in objective perspective how an audio-visual instrument might create a particular type of subjective experience, and simultaneously be expansive in enabling the discovery of new creative possibilities.

The model can also be used as a compositional tool. As an example, it has served as a communication medium in a remote, long-distance discussion to plan future creative collaborations. In this work, we intend to articulate a new version of Sá's audio-visual instrument and Tanaka's EMG bio-signal instrument. In Figure 10.8 we see a schematization of a performance structured in three sections.

The first section (Figure 10.8 top) has no visual projection; there are only two small lights on stage. The performative arena is both local and distributed, because the audience sits directed to the performers, and the loudspeakers are distributed in the room. The sonic construction explores ambivalent discontinuities, inviting attention to focus on the subtle intertwining of sonic emissions - the zither played with a bow, and the electronic sound textures created with the EMG. Both instruments require medium interaction effort, and there is a fungible relationship between the visible physical gestures and the sonic output: the audience senses causation without understanding the base cause and effect relationships. The recognisable sound of an acoustic string instrument conveys informational semantics, but the sound produced by the EMG does not, and the two are often undistinguishable. Whilst the local arena and the musical volatility convey expressive semantics, the distributed arena and the immersive sonic continuity convey environmental semantics.

In the second section (Figure 10.8 middle) the sonic construction entails radical discontinuities, which prompt automatic attention. These alternate with silences of variable duration; the combination does not really form a pattern, because the music maintains its organic, volatile qualities. While the EMG produces electronic sounds that are silenced in sudden manner, the zither is played with pick and slider, activating pre-recorded urban sounds. The interaction with the audio-visual instrument becomes slightly more effortful when the visual projection comes in: the digital mappings rule out radical visual discontinuities so as to avoid visual dominance, but the performer's attention must spread beyond the sonic construction, even if only occasionally to press a button. The projected image is abstract, and its shape is much smaller than the screen. It dislocates around and in-between the performers, creating a wealth of progressive continuities. The physical bodies are sometimes integrated, and other times separated from that visual shape, forming a wealth of simple, and simultaneously complex visual effects. The audio-visual relationship remains fungible: there are synchronised and non-synchronised audio-visual events, and the global complexity circumvents the human tendency to prioritise causal percepts. Whilst the sonic discontinuities strengthen the expressive dimension of work, the recognisable street sounds strengthen the informational dimension, in a way that extends the performative arena beyond the physical performance space.

In the third section (Figure 10.8 bottom) the image increases in size, with progressive continuity; it ends up covering the whole screen. The large-scale projection over the performers becomes a reactive stage scene, and this new type of integrated arrangement reinforces the environmental qualities of the work. The semantics of the sonic construction are strongly environmental as well. The zither is dribbled, activating sounds of nature, and these merge with electronic continuities produced by the EMG. The soundscape is dense, and rich in ambivalent discontinuities. Attention is invited to focus on the wealth of sonic details, while the multiple emissions interlace like a braid, emerging and submerging from each other. The audio-visual relationship remains fungible, but overall, there is a decrease in informational semantics because the sensorial complexity makes the cause and effect relationships now definitely indistinguishable – the audience should not even try to understand, just feel.

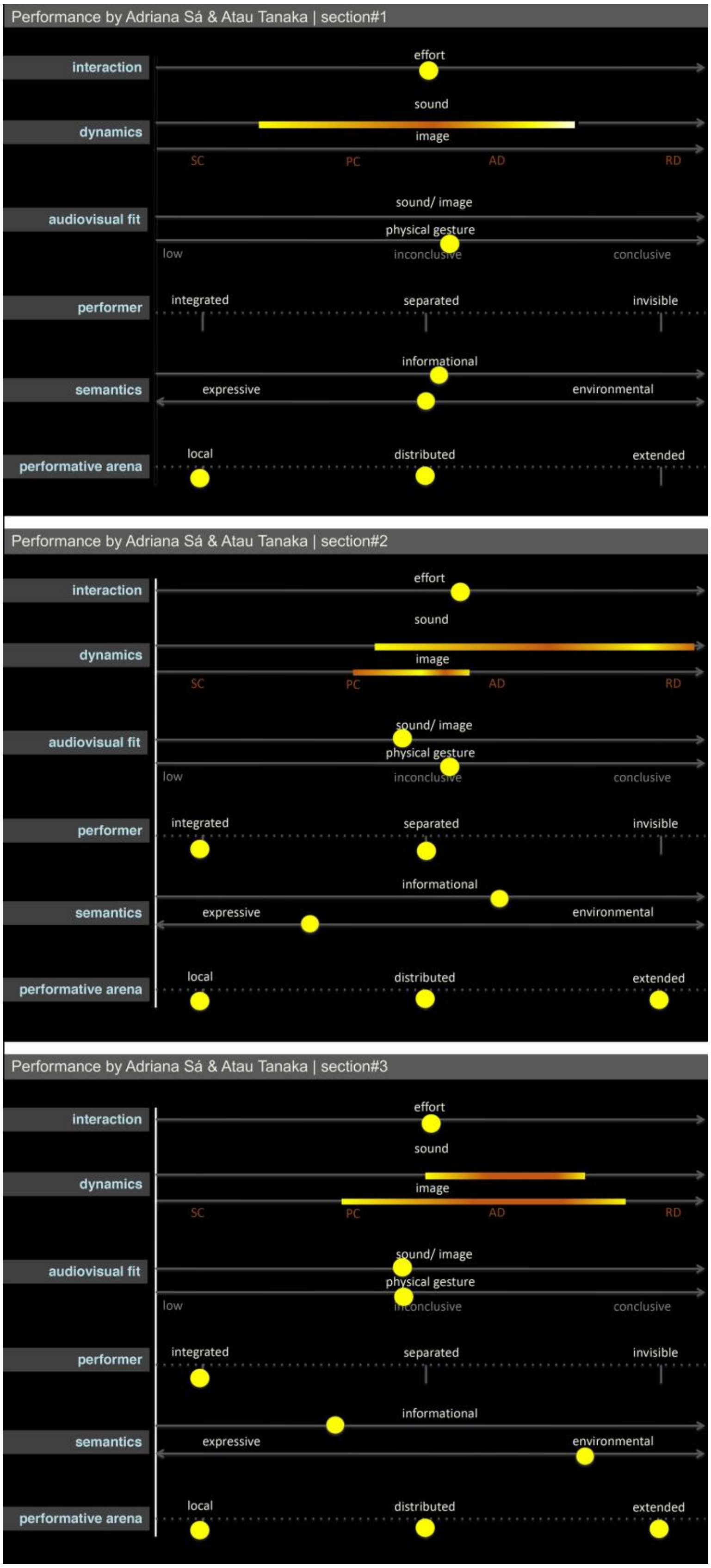

Figure 10.8 - Top: first section of the performance. Middle: second section of the performance. Bottom: third section of the performance.

Conclusion

We have presented a parametric visualization model as a means to represent in abstract form audio-visual artworks. Each parameter has been described separately, and in their interdependent relationship with other parameters. Examples from the history of audio-visual art have been used to illustrate the parameters. Furthermore, we showed that the model provides an operational means to analyse the relationship between sonic expression, sensory dominance and spatial presence. That relationship is crucial in any audio-visual performance language, regardless of its particular sphere of creative concerns.

Clearly, the model is very useful as an analytical tool, applicable to any technical platform, aesthetical approach and physical setup. By using this tool in analysis, we are able to discuss together a disparate series of works using a common framework.

The model has another useful function: it can work as a compositional tool. In the final section of the chapter, we showed how it has been applied in planning a new duo performance by the authors. The representation serves as a kind of score for sections of the new work. It serves to create a common terminology to tie together performance practice on two very different instruments – an electrified zither, and a muscle EMG instrument, and two distinct forms of output – digital audio processing of acoustical signals connected to 3D computer graphics, and pure sound synthesis.

The model has been useful to set out an overall performance structure. It can be further used to score the piece with more detail, leaving an open space for the choice of sonic and visual materials, audio-visual mappings and individual timings; each parameter can summarise several aspects of the work, and one can also use the model to analyse each aspect independently. This is particularly useful because we desire to rely on a grounding structure, and simultaneously believe in real-time motivations for expression.

Beyond our creative work, the model is potentially useful to any audio-visual practitioner. It can be used to analyse existing instruments, create new audio-visual performances, and develop new audio-visual systems.

Endnotes

[1] D. Birnbaum, R. Fiebrink, J. Malloch, and M.M. Wanderley, "Towards a Dimension Space for Musical Devices". In *Proceedings of the International Conference on New Interfaces for Musical Expression,* Vancouver, BC, Canada (2005).
[2] Thor Magnusson, "An Epistemic Dimension Space for Musical Devices". In *Proceedings of the Conference on New Interfaces for Musical Expression,* Sydney, Australia (2010).
[3] Marko Ciciliani and Zenon Mojzysz , "Evaluating a Method for the Analysis of Performance Practices in Electronic Music". In *Proceedings of the Second International Conference for Live Interfaces* (2015).
[4] Salome Voegelin, *Listening to Noise and Silence: Towards A Philosophy Of Sound Art* (New York, London: Continuum, 2010).
[5] Francisco López, official Web Site, "Against the Stage" (February 2004), http://www.franciscolopez.net/stage.html (accessed February 7, 2020).
[6] J.V. Draper, D.B. Kaber, & J.M. Usher, "Telepresence". In *Human Factors 40*, no. 3 (1998): 354–375.
[7] Thomas W. Schubert, "A New Conception of Spatial Presence: Once again, with Feeling". In *Communication Theory 19*, no. 2 (2009):161–187
[8] Paul Dourish, *Where the Action Is: The Foundations of Embodied Interaction* (Cambridge, MA: MIT Press, 2004).
[9] A. Sacau, J. Laarni, T. Hartmann, "Influence of Individual Factors on Presence". In *Computers in Human Behavior 24, 5*, Elsevier Science Publishers B. V. (2008), 2255-2273.
[10] Marko Ciciliani and Zenon Mojzysz, "Evaluating a Method for the Analysis of Performance Practices in Electronic Music". In *Proceedings of the Second International Conference for Live Interfaces* (2015), ISBN 978-989-746-060-9.
[11] Simon Emmerson, *Living Electronic Music* (Hampshire, U.K.: Ashgate Publishing Limited, 2007).
[12] Simon Emmerson, *Living Electronic Music* (Hampshire, U.K.: Ashgate Publishing Limited, 2007).
[13] Michel Chion, *Audio-Vision: Sound on Screen*. Trans. C. Gorbman (New York: Columbia University Press, 1994).
[14] Merilyn Brakhage, "On Stan Brakhage and Visual Music," *Vantage Point* (January 31, 2008), http://vantagepointmagazine.wordpress.com/2008/01/31/on-stan-brakhage-and-visual-music
[15] Salomé Voegelin, *Listening to noise and silence: Towards a philosophy of sound art* (New York, London: Continuum, 2010).
[16] Pierre Schaeffer, *Traité des Objets Musicaux: Essai Interdisciplines* (Paris : Éditions du Seuil, 1966).
[17] Jeff Pressing, "Some Perspectives on Performed Sound and Music in Virtual Environments". In *Presence 6, no. 4* (1997:8).
[18] Meghan Stevens, *Music and Image in Concert*. Music and Media, 2009, 3.
[19] S. Sinnett, C. Spence, S. Soto-Faraco, "Visual dominance and attention: The Colavita Effect Revisited". In *Perception & Psychophysics 69, no. 5* (2007), 673-686.
[20] Tod Machover, "Beyond Guitar Hero - Towards a New Musical Ecology." In *RSA Journal*, London (2009).
[21] Joel Ryan, "Some Remarks on Musical Instrument Design at STEIM", Contemporary music review, Vol. 6 issue 1 (1991), 3-17.
[22] D. Birnbaum, R. Fiebrink, J. Malloch, and M.M. Wanderley, "Towards a Dimension Space for Musical Devices". In *Proceedings of the International Conference on New Interfaces for Musical Expression,* Vancouver, BC, Canada (2005).
[23] Thor Magnusson, "An Epistemic Dimension Space for Musical Devices". In *Proceedings of the Conference on New Interfaces for Musical Expression,* Sydney, Australia (2010).
[24] Adriana Sá,, "Designing Musical Expression", *Proceedings of xCoAx – Conference on Computation, Communication, Aesthetics & X,* Lisbon, Portugal (2017). ISBN 978-989-746-128-6.
[25] Phill Niblock's *Movements of People Working* - https://www.youtube.com/watch?v=mCKtqsy9gcY and https://www.youtube.com/watch?v=LJtIZskdHOc
[26] Alvin Lucier *Music for Solo Performer* - http://daily.redbullmusicacademy.com/2017/05/alvin-lucier-music-for-solo-performer
[27] 1986 interview with Ev Grimes for Yale's Oral History of American Music

[28] Jeff Pressing, "Cognitive Complexity and the Structure of Musical Patterns". In *Generative Processes in music - The Psychology of Performance, Improvisation, and Composition*, ed. by J. Sloboda, New York, Oxford University Press (1987), 129-178.

[29] Performance by Joel Ryanand Evan Parker - https://www.youtube.com/watch?v=IQ4DqRgtbHc

[30] Performance by Steina Vasulka - https://www.youtube.com/watch?v=9mg1weOHWSs

[31] Steina Vasulka´s Violin Power series - http://www.vasulka.org/Steina/Steina_ViolinPower/ViolinPower.html

[32] Video about Martin Howse - https://vimeo.com/54006161

[33] Adriana Sá, "How An Audio-Visual Instrument Can Foster the Sonic Experience", *Live Visuals for Performance, Gaming, Installation, and Electronic Environments*, *Leonardo Electronic Almanac vol. 19, No. 3,* MIT Press (2013).

[34] Eric I. Knudsen, "Fundamental Components of Attention". In *Annual Review of Neuroscience 30, no.1* (2007), 57-78.

[35] CD released by Experimental Intermedia Label

[36] La Monte Young and Marian Zazeela's *Dream House* - https://www.youtube.com/watch?v=3ahgq-zVQLc

[37] Gary Hill's film *Black and White Text* - https://www.youtube.com/watch?v=bg1O3NcPwBg

[38] Thomas Wilfred's *Lumia* - https://www.youtube.com/watch?v=ojVX8FWYc4g

[39] Phil Niblock's *Movements of People Working* - https://expcinema.org/site/en/dvd/phil-niblock-movement-people-working

[40] Steina Vasulka's *Heraldic View* - http://www.fondation-langlois.org/html/e/page.php?NumPage=481

[41] Ryoji Ikeda's *Superposition* - https://vimeo.com/49873167

[42] Michel Chion, *Audio-Vision: Sound on Screen*. Trans. C. Gorbman (New York: Columbia University Press, 1994).

[43] M. Schutz, and M. Kubovy, "Causality and Cross-Modal Integration". In *Journal of Experimental Psychology: Human Perception and Performance 35, 6* (2009), 1791–1810.

[44] M. Kubovy and M. Schutz, "Audio-Visual Objects". In *Review of Philosophy and Psychology 1, 1,* Springer Science and Business Media B.V. (2010), 41–61.

[45] S.W. Brown and M. Boltz, "Attentional Processes in Time Perception: Effects of Mental Workload and Event Structure". In *Journal of Experimental Psychology: Human Perception and Performance*, 28 (2002): 600-615.

[46] Adriana Sá, Baptiste Caramieux and Atau Tanaka, "The Fungible Audio-Visual Mapping and its Experience". In *Journal Of Science And Technology Of The Arts 6 No. 1* (2014): 85-96. doi:http://dx.doi.org/10.7559/citarj.v6i1.131

[47] Norman McLaren's *Dots* - http://vimeo.com/15919138

[48] Steina Vasulka's *Noise Fields* - http://www.fondation-langlois.org/html/e/page.php?NumPage=483

[49] William Moritz, *The Dream of Color Music, And Machines That Made it Possible*. Animation World Magazine, 1 April 1997. http://www.awn.com/mag/issue2.1/articles/moritz2.1.html (Accessed August 30, 2016)

[50] Tarik Barri's *Versum* - http://tarikbarri.nl/projects/versum

[51] Audio-visual instrument by Adriana Sá - http://adrianasa.planetaclix.pt/research/practiceOverview.htm

[52] Adriana Sá, "How An Audio-Visual Instrument Can Foster the Sonic Experience", *Live Visuals for Performance, Gaming, Installation, and Electronic Environments*, *Leonardo Electronic Almanac vol. 19, No. 3,* MIT Press (2013).

[53] Adriana Sá, "Repurposing Video Game Software for Musical Expression: a perceptual approach". In *Proceedings of NIME - New Interfaces for Musical Expression* (London, 2014): 331-334.

[54] Adriana Sá, "Designing Musical Expression", *Proceedings of xCoAx – Conference on Computation, Communication, Aesthetics & X,* Lisbon, Portugal (2017). ISBN 978-989-746-128-6.

[55] Performance by *Sensors Sonics Sights*- https://babiole.net/en/sss-sensors-sonics-sights/

[56] Atau Tanaka, “Back To The Cross-Modal Object: A Look Back At Early Audiovisual Performance Through The Lens Of Objecthood”, *Live Visuals for Performance, Gaming, Installation, and Electronic Environments*, *Leonardo Electronic Almanac vol. 19, No. 3,* MIT Press (2013).
[57] Adriana Sá, Joel Ryan, Edwin van der Heide, Atau Tanaka, Andrew McPherson, Thor Magnusson, Alex McLean, Miguel Carvalhais & Mick Gierson, “Live Interfaces: Seeds of Debate”. In *Proceedings of INTER-FACE: ICLI 2014*, (2015):14-28.
[58] Thor Magnusson, “Improvising with the Threnoscope: Integrating Code, Hardware, GUI, Network, and Graphic Scores”, Proceedings of the Conference of New Interfaces for New Expression, London (2014).
[59] Adriana Sá, Joel Ryan, Edwin van der Heide, Atau Tanaka, Andrew McPherson, Thor Magnusson, Alex McLean, Miguel Carvalhais & Mick Gierson, “Live Interfaces: Seeds of Debate”. In *Proceedings of INTER-FACE: ICLI 2014*, (2015):14-28.
[60] John Dewey, *Art as Experience* (New York: Penguin Putnam. Harvard (18th ed.),1980)
[61] Mark Johnson, *The Meaning of the Body*: *Aesthetics of human understanding* (Chicago: University of Chicago Press, 2007).
[62] Jeff Pressing, “Some Perspectives on Performed Sound and Music in Virtual Environments”. In *Presence 6, no. 4* (1997).
[63] Leo Braudy, *The World in a Frame* (Garden City: Anchor Books, 1977).
[64] M. Kubovy and M. Schutz, “Audio-Visual Objects”. In *Review of Philosophy and Psychology 1, 1,* Springer Science and Business Media B.V. (2010), 41–61.
[65] Ikeda’s *Superposition* - https://vimeo.com/49873167
[66] Performance by Metamkine - https://www.youtube.com/watch?v=xWwVvICGeR4
[67] Guy Sherwin’s *Man with Mirror* - https://vimeo.com/31609396
[68] Marko Ciciliani and Zenon Mojzysz, “Evaluating a Method for the Analysis of Performance Practices in Electronic Music”. In *Proceedings of the Second International Conference for Live Interfaces* (2015), ISBN 978-989-746-060-9.
[69] Guy Sherwin’s *Man with Mirror* - https://vimeo.com/31609396
[70] Performance by Arnont Nogyao - https://vimeo.com/242720751
[71] Marko Ciciliani and Zenon Mojzysz, “Evaluating a Method for the Analysis of Performance Practices in Electronic Music”. In *Proceedings of the Second International Conference for Live Interfaces* (2015), ISBN 978-989-746-060-9.
[72] D. Birnbaum, R. Fiebrink, J. Malloch, and M.M. Wanderley, “Towards a Dimension Space for Musical Devices”. In *Proceedings of the International Conference on New Interfaces for Musical Expression,* Vancouver, BC, Canada (2005).
[73] Marko Ciciliani and Zenon Mojzysz, “Evaluating a Method for the Analysis of Performance Practices in Electronic Music”. In *Proceedings of the Second International Conference for Live Interfaces* (2015), ISBN 978-989-746-060-9
[74] Steina Vasulka’s performance - https://www.youtube.com/watch?v=9mg1weOHWSs
[75] Alvin Lucier’s *Music for Solo Performer* http://daily.redbullmusicacademy.com/2017/05/alvin-lucier-music-for-solo-performer
[76] Mick Grierson, “Noisescape: A 3d Audiovisual Multi-User Composition Environment”. In *The World as Virtual Environment*, ed. J. Birringer, T. Dumke, K. Nicolai (Dresden: Trans-Media-Akademie Hellerau 2007), 160-168.
[77] Adriana Sá, “Repurposing Video Game Software for Musical Expression: a perceptual approach”. In *Proceedings of New Interfaces for Musical Expression 2014* (London, 2014): 331-334. http://nime.org/proceedings/2014/nime2014_343.pdf
[78] Atau Tanaka, “Back To The Cross-Modal Object: A Look Back At Early Audiovisual Performance Through The Lens Of Objecthood”, *Live Visuals for Performance, Gaming, Installation, and Electronic Environments*, *Leonardo Electronic Almanac vol. 19, No. 3,* MIT Press (2013).
[79] Adriana Sá, "How an Audio-Visual Instrument Can Foster the Sonic Experience". In *Live Visuals*, *Leonardo Electronic Almanac* vol. 19 No. 3, MIT Press (January 2013): 284-305.
[80] Adriana Sá, *A Perceptual Approach to Audio-Visual Instrument Design, Composition and Performance*. *Research Online*, Goldsmiths, University of London: http://research.gold.ac.uk/19431/

[81] Merilyn Boltz, "The Cognitive Processing of Film and Musical Soundtracks". In *Memory & Cognition 32* (2004): 1194-1205.
[82] Charles Spence, "Explaining the Colavita Effect", *Progress in Brain Research 176,* ed. N. Srinivasan, Elsevier (2009): 245-58. doi: 10.1016/S0079-6123(09)17615-X.
[83] C. Robinson, W. Barnhart, S. Rivera, "Auditory Stimuli Slow Down Responses and First Fixations: Support for Auditory Dominance in Adults". In *37th Annual Meeting of the Cognitive Science Society*, Pasadena, California (2015). DOI: 10.13140/RG.2.1.1564.5286.
[84] Adriana Sá, Baptiste Caramieux and Atau Tanaka, "The Fungible Audio-Visual Mapping and its Experience". In *Journal Of Science And Technology Of The Arts 6 No. 1* (2014): 85-96. doi:http://dx.doi.org/10.7559/citarj.v6i1.131
[85] Adriana Sá, "How An Audio-Visual Instrument Can Foster the Sonic Experience", *Live Visuals for Performance, Gaming, Installation, and Electronic Environments*, *Leonardo Electronic Almanac vol. 19, No. 3,* MIT Press (2013)
[86] Adriana Sá, *A Perceptual Approach to Audio-Visual Instrument Design, Composition and Performance*. *Research Online*, Goldsmiths, University of London: http://research.gold.ac.uk/19431/
[87] Audio-visual instrument developed by Adriana Sá - http://adrianasa.planetaclix.pt/research/practiceOverview.htm
[88] Adriana Sá, *A Perceptual Approach to Audio-Visual Instrument Design, Composition and Performance*. *Research Online*, Goldsmiths, University of London: http://research.gold.ac.uk/19431/
[89] Audio recordings from the instrument developed by Adriana Sá - http://adrianasa.planetaclix.pt/research/ArpeggioDetuning.htm#music
[90] Adriana Sá, *A Perceptual Approach to Audio-Visual Instrument Design, Composition and Performance*. *Research Online*, Goldsmiths, University of London: http://research.gold.ac.uk/19431/