LEA

LEONARDO ELECTRONIC ALMANAC

VOL 19 NO 3 VOLUME EDITORS **LANFRANCO ACETI, STEVE GIBSON & STEFAN MÜLLER ARISONA** EDITOR **ÖZDEN ŞAHİN**

Live visuals have become a pervasive component of our contemporary lives; either as visible interfaces that re-connect citizens and buildings overlaying new contextual meaning or as invisible ubiquitous narratives that are discovered through interactive actions and mediating screens. The contemporary re-design of the environment we live in is in terms of visuals and visualizations, software interfaces and new modes of engagement and consumption. This LEA volume presents a series of seminal papers in the field, offering the reader a new perspective on the future role of Live Visuals.

LIVE VISUALS

LEONARDO ELECTRONIC ALMANAC, VOLUME 19 ISSUE 3

Live Visuals

VOLUME EDITORS
LANFRANCO ACETI, STEVE GIBSON & STEFAN MÜLLER ARISONA

EDITOR
ÖZDEN ŞAHİN

. .Sabancı. .
Üniversitesi

KASA GALERİ

Goldsmiths
UNIVERSITY OF LONDON

Operational & Curatorial Research in
Contemporary Art, Design,
Science & Technology

NYUSteinhardt
Steinhardt School of Culture, Education, and Human Development

---

NYUSteinhardt
Steinhardt School of Culture, Education, and Human Development

MUSIC AND PERFORMING ARTS PROFESSIONS

Ron Sadoff, Director

Music Technology
B.M., M.M., Ph.D.
Including a new 3-Summer M.M.

Immersive Audio, Computer Music, Informatics,
Cognition, Recording and Production

Music Composition
B.M., M.M., Ph.D.

Concert Music, Jazz, Film Scoring,
Electro-Acoustic, Songwriting

- **Study with a premier faculty** who are active in the local and international music field,
  including Juan Pablo Bello, Morwaread Farbood, Phil E. Galdston, Paul Geluso,
  Tae Hong Park, Kenneth Peacock, Agnieszka Roginska, Robert Rowe, S. Alex Ruthmann,
  Ronald Sadoff, David Schroeder, Mark Suozzo, and Julia Wolfe

- Work within a **large and enriching university environment** in the heart of New York City

- Have access to **state-of-the-art facilities** including the James L. Dolan Music
  Recording Studio, one of the most technologically advanced audio teaching facilities
  in the United States

- **Collaborate** with an outstanding variety of department performance groups, along
  with choreographers, visual artists, writers, filmmakers, and scholars in other fields

- Take advantage of **special courses** offered abroad and during the summer

Visit **www.steinhardt.nyu.edu/music** or call **212 998 5424** to learn more.

NEW YORK UNIVERSITY

NEW YORK UNIVERSITY IS AN AFFIRMATIVE ACTION/EQUAL OPPORTUNITY INSTITUTION.

Leonardo Electronic Almanac
Volume 19 Issue 3

When Moving Images Become Alive!

"Look! It's moving. It's alive. It's alive… It's alive, it's moving, it's alive, it's alive, it's alive, it's alive, IT'S ALIVE!"

Frankenstein (1931)

Those who still see – and there are many in this camp – visuals as simple 'decorations' are living in a late 19th century understanding of media, with no realization that an immense cultural shift has happened in the late 20th century when big data, sensors, algorithms and visuals merged in order to create 21st century constantly mediated social-visual culture.

Although the visuals are not actually alive, one cannot fail to grasp the fascination or evolution that visuals and visual data have embarked upon. It is no longer possible to see the relationship of the visual as limited to the space of the traditional screens in the film theater or at home in the living room with the TV. The mobility of contemporary visuals and contemporary screens has pushed boundaries – so much so that 'embeddedness' of visuals onto and into things is a daily practice. The viewers have acquired expectations that it is possible, or that it should be possible, to recall the image of an object and to be able to have that same object appear at home at will. The process of downloading should not be limited to 'immaterial' digital data, but should be transferred to 3D physical objects. [1]

Images are projected onto buildings – not as the traditional trompe l'oeil placed to disguise and trick the eye – but as an architectural element of the building itself; so much so that there are arguments, including mine, that we should substitute walls with projected information data, which should also have and be perceived as having material properties (see in this volume "Architectural Projections" by Lukas Treyer, Stefan Müller Arisona & Gerhard Schmitt).

Images appear over the architecture of the buildings as another structural layer, one made of information data that relays more to the viewer either directly or through screens able to read augmented reality information. But live visuals relay more than images, they are also linked to sound and the analysis of this linkage provides us with the opportunity "to think about the different ways in which linkages between vision and audition can be established, and how audio-visual objects can be composed from the specific attributes of auditory and visual perception" (see "Back to the Cross-modal Object" by Atau Tanaka).

iPads and iPhones – followed by a generation of smarter and smarter devices – have brought a radical change in the way reality is experienced, captured, uploaded and shared. These processes allow reality to be experienced with multiple added layers, allowing viewers to re-capture, re-upload and re-share, creating yet further layers over the previous layers that were already placed upon the 'original.' This layering process, this thickening of meanings, adding of interpretations, references and even errors, may be considered as the physical process that leads to the manifestation of the 'aura' as a metaphysical concept. The materiality of the virtual, layered upon the 'real,' becomes an indication of the compositing of the aura, in Walter Benjamin's terms, as a metaphysical experience of the object/image but nevertheless an experience that digital and live visuals are rendering increasingly visible.

"Everything I said on the subject [the nature of aura] was directed polemically against the theosophists, whose inexperience and ignorance I find highly repugnant. . . . First, genuine aura appears in all things, not just in certain kinds of things, as people imagine." [2]

The importance of digital media is undeniably evident. Within this media context of multiple screens and surfaces the digitized image, in a culture profoundly visual, has extended its dominion through 'disruptive forms' of sharing and 'illegal' consumption. The reproducibility of the image (or the live visuals) – pushed to its very limit – has an anarchistic and revolutionary element when considered from the neocapitalistic perspective imbued in corporative and hierarchical forms of the construction of values. On the contrary, the reproducibility of the image when analyzed from a Marxist point of view possesses a community and social component for egalitarian participation within the richness of contemporary and historical cultural forms.

The digital live visuals – with their continuous potential of integration within the blurring boundaries of public and private environments – will continue to be the conflicting territory of divergent interests and cultural assumptions that will shape the future of societal engagements. Reproducibility will increasingly become the territory of control generating conflicts between *original* and *copy*, and between the layering of *copy* and *copies*, in the attempt to contain ideal participatory models of democracy. The elitist interpretation of the aura will continue to be juxtaposed with models of Marxist participation and appropriation. [3]

Live visuals projected on public buildings and private areas do not escape this conflict, but present interpretations and forms of engagements that are reflections of social ideals. The conflict is, therefore, not solely in the elitist or participatory forms of consumption but also in the ideologies that surround the cultural behaviors of visual consumption.

Object in themselves, not just buildings, can and may soon carry live visuals. There is the expectation that one no longer has to read a label – but the object can and should project the label and its textured images to the viewer. People increasingly expect the object to engage with their needs by providing the necessary information that would convince them to look into it, play with it, engage with it, talk to it, like it and ultimately buy it.

Ultimately there will be no need to engage in this process but the environment will have objects that, by reading previous experiences of likes and dislikes, present a personalized visual texture of reality.

Live visuals will provide an environment within which purchasing does not mean to solely acquire an object but rather to 'buy' into an idea, a history, an ideology or a socio-political lifestyle. It is a process of increased visualization of large data (Big Data) that defines and re-defines one's experience of the real based on previously expressed likes and dislikes.

In this context of multiple object and environmental experiences it is also possible to forge multiple individualized experiences of the real; as much as there are multiple personalized experiences of the internet and social media through multiple avatar identities (see "Avatar Actors" by Elif Ayter). The 'real' will become a visual timeline of what the algorithm has decided should be offered based on individualized settings of likes and dislikes. This approach raises an infinite set of possibilities but of problems as well.

The life of our representation and of our visuals is our 'real' life – disjointed and increasingly distant from what we continue to perceive as the 'real real,' delusively hanging on to outdated but comfortable modes of perception.

The cinematic visions of live visuals from the 19th century have become true and have re-designed society unexpectedly, altering dramatically the social structures and speeding up the pace of our physical existence that constantly tries to catch up and play up to the visual virtual realities that we spend time constructing.

If we still hold to this dualistic and dichotomist approach of real versus virtual (although the virtual has been real for some time and has become one of the multiple facets of the 'real' experience), then the real is increasingly slowing down while the virtual representation of visuals is accelerating the creation of a world of instantaneous connectivity, desires and aspirations. A visuality of hyper-mediated images that, as pollution, pervades and conditions our vision without giving the option of switching off increasingly 'alive' live visuals. [4]

The lack of 'real' in Jean Baudrillard's understanding is speeding up the disappearance of the 'real' self in favor of multiple personal existential narratives that are embedded in a series of multiple possible worlds. It is not just the map that is disappearing in the precession of simulacra – but the body as well – as the body is conceived in terms of visual representation: as a map. These multiple worlds of representations contribute to create reality as the 'fantasy' we really wish to experience, reshaping in turn the 'real' identity that continuously attempts to live up to its 'virtual and fantastic' expectations. Stephen Gibson presents the reader with a description of one of these worlds with live audio-visual simulations that create a synesthetic experience (see "Simulating Synesthesia in Spatially-Based Real-time Audio-Visual Performance" by Stephen Gibson).

If this fantasy of the images of society is considered an illusion – or the reality of the simulacrum, which is a textual oxymoron at prima facie – it will be determined through the experience of the *live visuals becoming alive.*

Nevertheless, stating that people have illusory perceptions of themselves in relation to a 'real' self and to the 'real' perception of them that others have only reinforces the idea that Live Visuals will allow people to manifest their multiple perceptions, as simulated and/or real will no long matter. These multiple perceptions will create multiple ever-changing personae that will be further layered through the engagements with the multiple visual environments and the people/avatars that populate those environments, both real and virtual.

In the end, these fantasies of identities and of worlds, manifested through illusory identities and worlds within virtual contexts, are part of the reality with which people engage. Although fantastic and illusory, these worlds are a reflection of a partial reality of the identity of the creators and users. It is impossible for these worlds and identities to exist outside of the 'real.' This concept of real is made of negotiated and negotiable frameworks of engagement that are in a constant process of evolution and change.

The end of post-modernity and relativism may lead to the virtuality of truism: the representation of ourselves in as many multiple versions – already we have multiple and concurrent digital lives – within the world/s – ideological or corporate – that we will decide or be forced to 'buy into.'

It is this control of the environment around us and us within that environment that will increasingly define the role that live visuals will play in negotiating real and virtual experiences. The conflict will arise from the blurred lines of the definition of self and other; whether the 'other' will be another individual or a corporation.

The potential problems of this state of the live visuals within a real/virtual conflict will be discovered as time moves on. In the end this is a giant behavioral experiment, where media and their influences are not analyzed for their social impact *ex ante facto*; this is something that happens *ex post facto*.

Nevertheless, in this ex post facto society there are some scholars that try to understand and eviscerate the problems related to the process of visuals becoming alive. This issue collects the analyses of some of these scholars and embeds them in a larger societal debate, hinting at future developments and problems that society and images will have to face as the live visuals become more and more alive.

The contemporary concerns and practices of live visuals are crystallized in this volume, providing an insight into current developments and practices in the field of live visuals.

This issue features a new logo on its cover, that of New York University, Steinhardt School of Culture, Education, and Human Development.

My thanks to Prof. Robert Rowe, Professor of Music and Music Education; Associate Dean of Research and Doctoral Studies at NYU, for his work in establishing this collaboration with LEA.

My gratitude to Steve Gibson and Stefan Müller Arisona, without them this volume would not have been possible. I also have to thank the authors for their patience in complying with the guidelines and editorial demands that made this issue one that I am particularly proud of, both for its visuals and for its content.

My special thanks go to Deniz Cem Önduygu who has shown commitment to the LEA project beyond what could be expected.

Özden Şahin has, as always, continued to provide valuable editorial support to ensure that LEA could achieve another landmark.

Lanfranco Aceti
Editor in Chief, Leonardo Electronic Almanac
Director, Kasa Gallery

1. 3D printing the new phenomenon will soon collide with a new extreme perception of consumer culture where the object seen can be bought and automatically printed at home or in the office. Matt Ratto and Robert Ree, "Materializing Information: 3D Printing and Social Change," *First Monday* 17, no. 7 (July 2, 2012), http://firstmonday.org/ojs/index.php/fm/article/view/3968/3273 (accessed October 20, 2013).
2. Walter Benjamin, "Protocols of Drug Experiments," in On Hashish, ed. Howard Eiland (Cambridge, MA: Harvard University Press, 2006), 58.
3. " The point here is not to issue a verdict in the debate between Adorno and Benjamin, but rather to understand the debate between them as representing two sides of an ongoing dialectical contradiction." Ryan Moore, "Digital Reproducibility and the Culture Industry: Popular Music and the Adorno-Benjamin Debate," Fast Capitalism 9, no. 1 (2012), http://www.uta.edu/huma/agger/fastcapitalism/9_1/moore9_1.html (accessed October 30, 2013).
4. Paul Virilio, *Open Sky*, trans. Julie Rose (London: Verso, 1997), 97.

Back to the Cross-modal Object

A Look Back at Early Audiovisual Performance through the Lens of Objecthood

by

Atau Tanaka

Goldsmiths, University of London

A B S T R A C T

This paper looks at 2 early digital audiovisual performance works, solo work Overbow and the group Sensors_Sonics_Sights (S.S.S) and describes the compositional and performance strategies behind each one. We draw upon the concept of audiovisual objecthood proposed by Kubovy and Schutz to think about the different ways in which linkages between vision and audition can be established, and how audio-visual objects can be composed from the specific attributes of auditory and visual perception. The model is used as a means to analyze these live audio-visual works performed using sensor-based instruments. The fact that gesture is not the only visual component in these performances, and is the common source articulating sound and visual output, extends the classical 2-way audio-visual object into a three-way relationship between gesture, sound, and image, fulfilling a potential of cross-modal objects.

INTRODUCTION

With the arrival of live visual software running on laptop computers, computer graphics got the equivalent of what musicians had with the laptop – the possibility to bring computer generated media into a live, real time performance setting. It became *de rigueur* that performances of electronic music be accompanied by projected visuals, either in collaboration with visual artists, sometimes as an ad-hoc accompaniment by the house VJ, or in holistic works conceived by a single artist, sometimes where sound and image were generated from a single computer and software. Regardless of musical style or of configuration, the question of linkages between what is seen and what is heard are at the base of how the audience, and critics, consider the work. We are quick to criticize work where the connection between sound and image seems arbitrary, unprepared, or unconsidered. At the opposite extreme, we praise the purity of hyper-minimalist work where sound and image seem to be the same signal, where the visuals become a kind of oscilloscope reading of the sound or some other kind of direct visualization. In between, work with looser association between sound and image are seen as being cinematic, often wondering if one medium has be-

come subservient to the other, either as a soundtrack, or visual accompaniment.

In this implicit value system is the notion that direct association yields work of higher artistic quality. Is association the only objective criterium we can apply to the interrelationship between sound and image in a live audiovisual work? Are there really only 2 elements at play? This paper draws upon the concept of audio-visual "objecthood" proposed by Kubovy and Schutz [5] to think about the different ways in which linkages between vision and audition can be established, and how audio-visual objects can be composed from the specific attributes of auditory and visual perception.

We use Kubovy and Schutz' model as a means to analyze three early live audio-visual works performed using sensor-based instruments that use performer gesture to articulate sound and image in concert. With the inclusion of gesture, we are able to apply Kubovy's example of visual impact and percussion. However, given the fact that the gestural sensor instruments used in the performances described detect far more than impact, instead tracking continuous gesture, we

propose an extension to Kubovy and Schutz that goes beyond linkages of event-based triggering. Finally, with the fact that the gesture is not the only visual component in these performances, and is the common source articulating sound and visual output, we extend Kubovy's 2-way audiovisual object into a three-way relationship between gesture, sound, and image, that fulfills the potential outlined in their paper for "cross-modal objects."

We begin with a summary of Kubovy and Schutz's original article, hilighting aspects of their theory that are relevant to live performance. We then present three audiovisual concert works from the 1990's and 2000's, two of them solo works, and one ensemble. We analyze the composition of these works using Kubovy and Schutz' concepts, and finish by extending their work by going beyond triggering and adding a third, gestural, component to audio-visual objecthood.

AUDIO-VISUAL OBJECTS

Kubovy and Schutz take an experimental psychology approach to investigating audio-visual *objecthood*, that is to say, the attributes that audible and visual information can take on to become discernable entities, and the ways in which audition and vision come together to form cross-modal objects. They cite an earlier paper of Kubovy in which he proposes a working definition of object.

> *A perceptual object is that which is susceptible to figure-ground segregation.* [4]

This situates object formation clearly as a function of sensory perception, where the object is an entity distinct from the environment surrounding it. With this definition, Kubovy skirts phenomenological and philosophical questions on the constitution of objects.

In order understand how audio-visual objects form, Kubovy and Schutz look at linkages between audition and vision, and point out a fundamental duality. According to K&S, vision focuses on surfaces while audition focuses on sources. Light illuminates an object rendering it visible by virtue of light reflecting off the surface of the object. As long as the object is visible, the source of light illuminating it is discounted. A sound object, on the other hand, is a source from which sound emanates. While sound fills the a space and is contextualized by the acoustic properties of that space, by reflecting off surfaces in the space, mechanisms of auditory perception such as the precedence effect allow us to discount surface effects in sound to focus on the sound source itself.

In the tradition of experimental psychology, K&S set out to prove this duality by describing a 'thought experiment.' In the proposed experiment, they seek to identify which attributes in visual and auditory perception are indispensable to distinguish separate objects. They propose that space is indispensable,

because 2 lights shining on surfaces can be perceived as separate visual objects only if they don't overlap in space. On the other hand, they state that color (light frequency) is not indispensable for visual objecthood – two distinct colors that overlap mix to create a new, single color. And two visual objects that are spatially distinct can have the same color and still be perceived as two separate objects.

In audition, on the other hand, frequency (or pitch) is indispensable. Two test tones that play together will be heard as a chord (a superposition of notes) rather than fusing as color does. Meanwhile space is not indispensable – two tones can be played in separate parts of the space, or superimposed one other the other, and still be perceived separately.

While the experiment does not address auditory effects where in fact two tones might combine to be heard as a single sound with richer harmonic partials, it does successfully describe the complementarity of attributes indispensable in vision and audition. One aspect not addressed in the experiment is the temporal dimension. For purposes of the experiment they state that time is assumed to be essential, that is indispensable, to both audition and vision. They go on later in the article to create linkages in time-based audiovisual events – the mallet strike of a marimba – and to separate sound and image components to show that variations in temporal characteristics of one mode (changing visual strike duration) can alter our perception of constancy in the other (an invariant percussion sound).

K&S use the complementarity of aspects essential to vision (surfaces, space) and audition (sources, frequency) to arrive at a Theory of Indispensable Attributes with which they establish how linkages lead to the formation of *gestalts* – wholes that are perceived as more than the sum of their parts – across perceptual modalities. According to K&S, it is through

Gestalt laws of grouping that an otherwise disparate collection of elements might perceptually be grouped together to form the representation of a single object. This could be a group of dots or lines converging through optical illusion to imply a shape or form, or of a group of sounds coalescing to produce recognizable chords or melodies. Perceptual illusions allow us to interpret incipient objects in more than one way – a 3D wireframe of a cube can be interpreted to be viewed from above or below. The rapid alternation of note ranges in Bach's sonatas for solo stringed instruments (violin or 'cello) famously create the illusion of two distinct voices emanating from a single melodic line. K&S refer to these different possible ways to see the same thing as "multiple putative objects." It is through forms of privileged cross-modal binding that our mechanisms of auditory and visual perception deduce an ecological context and decide on a single interpretation of an object.

J. J. Gibson in the 1950s pioneered an ecological approach to understanding visual perception. K&S cite a famous quote of Gibson from 1966 where he describes the sensory elements at play in the perception of fire – sight, sound, heat, odor, and asks how these sensations come together to form the gestalt notion of flame. [2] Gibson concludes that the dynamic at play is not just one of association, but one of discrimination from all the other sights, sounds, and smells that do not specify fire. Gibson's invitation to go beyond direct association to consider ecological discrimination is a key driver in K&S's privileged cross-modal bindings. It is also a fundamental element at play in our discussion of live audiovisual performance – how can we bring to bear the importance of cross-modal binding to avoid ad-hoc superposition of sound and image? At the same time, how do we go beyond the direct association dynamic seen in data visualization of sound and the 'oscilloscope effect'?

K&S propose thinking of binding across perceptual modes by thinking about cross-modal causality. They describe a rather simple case of visible impact relating to a percussive sound, and describe studies of videos as well as abstracted animations of percussion instrument performance with different percussive and non-percussive sounds. They follow Gibson in not staying with simple association, but look to more robust forms of inter-modal binding. K&S talk about *perceptual constancies* – the permanent properties robust to changes in the aspect of a visual or audible object (such as changing light conditions) that allow an object to continue to be perceived as such. In this discussion, they show that human visual perception is robust to spectral flux, and that auditory perception is robust to acoustic flux. As the creative works described below are abstract sound/image works not situated in the ecological setting of real-world constraints, we are interested in going beyond the impact/percussion simplicity of K&S's cross-modal causality and to actually incorporate flux in both sound and image as ways of bringing our abstract audiovisual performance objects to life.

EARLY WORKS

We describe two early solo performance works that combine gestural interaction with a live audiovisual system. The two compositions, *Overbow*, and *Rail*, were produced in 1993 and 1995 respectively, and performed through 2002. They were performed on the BioMuse, an instrument that translates neuron impulses from muscle tension resulting in performer arm gesture by means of the electromyogram (EMG) signal. [3] The BioMuse digitizes four channels of EMG (left and right forearm and triceps) and transcodes the amplitude envelope of muscle tension to a stream of MIDI Continuous Controller data. The incoming EMG data was processed in a computer running an early

physically impossible brass horn like sounds (produced on the TG77), where objecthood is established in a relatively traditional way – by conveying a notion of instrument, by playing notes and creating musical phrases and units. The second element in the sound are continuous walls of sound (sound itself as a surface) on the Wavestation where the 'source' remains the same but where its context (feedback, filtering) shifts. The visual content similarly consists of two types of elements. 3D designer Enno Hyttrek produced a group of 3D models, which as objects, were imported as canvases. Visual artist Kerstin Weiberg produced a series of textures that worked directly as canvases. Creating dynamic MIDI controlled image mapping convolving the 3D objects with Weiberg's textures created dynamic surfaces.

The performance opens with single brass notes and 3D objects appearing onscreen, articulated by simple upward clenches on the forearms. Continuing further up articulates higher register notes and modulates the size and position of the 3D blobs. What initially seems like traditional melodic phrases are in fact pinned down to constant aspects of the sound synthesis that do not transpose linearly with pitch. The modulating oscillator in the frequency modulation synthesis chain is held constant while the carrier oscillator (which determines the perceived pitch) transposes in response to gestural input. This creates a shifting timbral roughness not possible in the physical world as if a brass instrument were somehow able to play different pitches without change of acoustical length. The 3D blobs correspondingly appear in sizes and positions onscreen that are 'melodically' displaced by the EMG gesture, but a constancy with respect to the background canvas pins down the otherwise independent object, deforming it in unexpected and unnatural ways.

The brass-like voice becomes sustained, the first indication that such a clearly articulated musical unit

version of the Max graphical programming environment. [7] At this stage in its development, Max was uniquely a control data processing language and did not include the extensions for video processing such as NATO or Jitter, or for audio signal processing such as MSP which would appear later. This meant that Overbow and Rail used outboard, dedicated hardware/software systems for audio and visual content generation.

Sound in Overbow was generated by two rack-mounted MIDI synthesizers, a Yamaha TG77 to modify sounds by means of high rate frequency modulation, and a Korg Wavestation to interpolate vectorially across a linear space of waveforms and provide analog feedback. On both synthesizers, events were initiated

by MIDI Note messages, articulating individual sound samples, loops, or synthesized melodic notes. These become 'sources' in Kubovy's taxonomy. Different auditory characteristics of these sound sources can be modified at the moment of event articulation, and more importantly while the sound is playing, to go beyond the simple trigger paradigm and allow eventual creative confusion between sources and surfaces. Image was generated on a separate computer dedicated to visuals, controlled over MIDI by the control machine that was running Max. The image synthesis machine ran custom software with code name MIDI Kaleido, written by Eric Wenger, developer of the landscape generation software Bryce, and the image sonification software Metasynth. MIDI Kaleido was a prototype of what would later become Videodelic and functions on

the principle of a 'canvas' – a base image (480×480 pixels) that can then be modified in real time in scale, repetition, brightness, contrast, color saturation, color mapping. This corresponds directly to Kubovy's notion of image functioning off of surfaces. In addition to modification of a single canvas, MIDI Kaleido allows image mapping of a second image onto the first, creating a perturbation of the surface.

Overbow

The compositional content for *Overbow* consists of parallels between sound and image, each medium comprised of two simple groups – the first clearly lending to objecthood, and the latter blurring the distinctions between sound and image where objecthood may be less distinct. The sound is comprised of

> *The metallic timbre becomes a continuous complex oscillator of sorts, and the metal bars become not singular, but a period in a cyclic continuum.*

might lose or contort its own objecthood. The 3D shape smears to fill the dimensions of the canvas and becomes fixed. Short metallic plucked sounds begin a slow, unpredictable yet incessant upward movement in a randomized meandering scale, accelerating and receding as the scale goes higher. This is a transient object, or sprinkling of tiny object prickles, that fades off in the distance just as it establishes itself. Underneath this, the frequency modulation of the sustained brass tones undulates, causing the static pedal to waver under the staccato plucks. The upward scale is not directly controlled by individual arm-hand gestures, but is a semi-automatic process where the speed, random distribution of pitches, density, and the upward arch are shaped by continuous modulations in muscle tension. It is this same muscle tension that modulates more directly the modulation parameters of the brass tones, linking the two distinct musical voices to a single solo performer gesture.

The top of this scale is muted with sustained effort and tension in the arms, effectively choking off the brass voices as well. The visuals use a luminance-add to blot the canvas empty. Arm tension hold this moment, which is finally released by a two-arm gesture that articulates a downward glissando ending in a low impact sound. This series is repeated – preparation, anticipatory glissando of varying length as the performer relaxes more or less quickly the held tension, punctuated by a fist and impactful blow. The 3D blobs from the opening track this musical gesture by appearing out of nowhere, approaching and expanding its jittery manner, filling the screen at the moment of impact, then fading post-explosion by continuing to expand and disappearing by consuming the canvas.

A last dramatic swoop – long descending glissando, teasing approach of the blob, and booming impact results in an explosion that does not subside, but that becomes the sustained underbed carpet of audio

upon which the rest of the piece is made. From within the broadband noise of a time frozen explosion, resonant frequencies become apparent, focused through analog feedback tuning on the Wavestation. All visual elements have been subsumed into the canvas that has by now accumulated, itself, a carpet thickness. Layers obscured by features of surfaces above it swirl underneath, creating surface deformations as resonant feedback frequencies sweep and shift. The last audiovisual mass is no longer an object to behold from a distance, but a mass that envelopes and contains the performer and spectator. The piece ends as this immersion diminished in the distance as if a translucent veil is slowly pulled off to the horizon, leaving an altered perception of the regular sounds of the surrounding environment in the concert space.
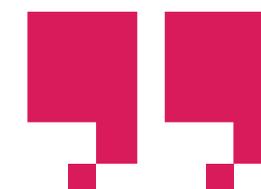
Rail

Rail is a second work for the same system, composed in 1995. The audio component continues to use an outboard rack-mounted MIDI synthesizer, switching to a Kurzweil K2000R for more intergrated and sophisticated control of oscillators and sound samples together in reconfigurable signal processing chains. The visual component maintains use of Wenger's MIDI Kaleido system.

There are just two core elements at play – pure generated tones, square waves and pulse-width modulation (PWM) of square waves to sweep a range of duty cycles, and metal samples, sound recordings of transients from metallic impact, looped, filtered, and time-stretched. The palette of images is entirely in black and white, with black and white bands of varying width and proportion corresponding to the square waves and PWM, and close up photography of metallic surfaces, steel, iron, and lattices.

The piece begins from silence with barely visible tension in the left forearm, causing no gross arm move-

ment but resulting from just slight clenching of the hands results in swells in low frequency square waves just at the threshold of hearing in both frequency and amplitude. The visuals parallel the sound in a cathode ray tube oscilloscope-like manner, with white horizontal bands dividing the screen fading in and out of view. The unstable nature of the EMG signal causes tentative pulsation in the audio and flickering in the image.

After a number of repetitions of swells with the left arm, the right forearm enters in a more visible but restrained fashion, mapping tension in the extensor muscles to tone frequency. With a outward gesture of the right arm, the tone swells take on a sweeping up and down glissando bringing to life the flickering horizontal bars as they change width, phase, and sweep down the screen like an old television out of vertical synch. A second EMG sensor on the right inner forearm, on the anterior flexor comes in to modulate the duty cycle of the tone. Wrist rotation during the outward gesture, then controls the pulse width of the square wave as the frequency sweep is articulated, all while the left arm continues to pulse the amplitude. This creates a single gestural-audiovisual object that fills the space with low frequency, unlocalizable sound and flickering black and white bands, inextricably linked to effort and movement.

In another section of the piece, metal strikes are introduced. Simple thresholds on the left arm EMG are set to trigger events when a clench of the fist and percussive strike are articulated. This sounds single samples of sharp metal strikes and at the same transforms the inside of the white bars to take on a metal trellis texture. The right arm muscle enters into coordination, with differences in tension at the moment of the left arm strikes changing the selection of metal sample. A certain number of samples are reversed, allowing the sample selection to play a series of forward and backward metal sounds, performed by synchronized gestures on the arms. The envelope of the metal sample modulates zoom and rotation in the graphical bars making what previously were horizontal flutters begin to pop out towards the audience, at spinning while zooming. Increasing activity and tempo in the striking gestures superposes the strikes, creating a density of percussive sound with interlocked bars at differing dynamic zooms and spin angle, creating trellises of trellised bars.

The expansion and deformation of metal turns it from a percussive object (single shot sample, single bar trellis) into an oscillator like permanent sound, recalling the original pure tones of the beginning. The metallic timbre becomes a continuous complex oscillator of sorts, and the metal bars become not singular, but a

period in a cyclic continuum. Eventually the 'duty cycle' of the sustained metal waveform fills to 100%, leaving no black between bars, filling the canvas with metal texture, filling the acoustic space with infinitely time stretched percussive instant. The metal object becomes surface, and similar to the close of Overbow, a wall of noise, this time digital, metallic, infinite, where shifting layers expose internal melodies.

S_S_S

Sensors_Sonics_Sights (S_S_S) is a later project (2003–2008) that differs from *Overbow* and *Rail* in a number of ways. While it retains use of the BioMuse and EMG in live sound/image performance, it is not a solo work, but a trio ensemble. The trio consisted of the author on the BioMuse, Laurent Dailleau on Theremin, and Cécile Babiole on ultrasound sensors. The group developed a series of six audiovisual compositions that were performed in different combinations, in different contexts, from galleries to festivals, clubs to concerts halls, during the five year life of the group.

Figure 2. Sensors_Sonics_Sights performing in the round at Festival Musique Action, Vandoeuvre-les-Nancy, France, 2005. © Atau Tanaka, 2005. Used with permission.

There were significant developments in both technology and the currency of audiovisual performance in the 10 years since Overbow and Rail. The interactive systems continued to be programmed in Max, but now benefitted from the extensible software libraries MSP for audio signal processing, and Jitter for matrix mathematics, Open GL 3D graphics, and video display. With MaxMSP/Jitter the live audiovisual performance setup for many artists in this period became entirely

software based. All three members in S_S_S use Max-MSP/Jitter at the heart of their interactive systems, but contrary to the software's possibility to "do everything," each group member's patch is distinct, and is programmed to carry out only the musical tasks that member plays as part of the ensemble. Each patch is different, no one member generates both sound and image, there is no automatic visualization of sound, and there is no network communications connecting the three subsystems.

Instead, all communication takes place in non-technological channels, through eye contact, and gestural coordination amongst the performers. This use of traditional techniques for musical ensemble performance are called into play to execute digital, audiovisual compositions. In the group, all three instruments are gestural instruments. Alongside the futuristic nature of the BioMuse is the Theremin, invented in the 1920's and one of the original, most iconic gestural musical instrument. Dailleau's use of of the Theremin is both directly as an analog sound source, as well as through MSP-based pitch-tracking to use the Theremin as control interface for digitally generated sounds. Babiole's ultrasound instrument takes distance measuring sensors of the same type found on STEIM's The Hands,[3] but mounted in a stationary fashion so as to sense hand movement in front of the laptop. Babiole's Jitter patch translates hand movements to graphics animation. In this musical trio, then, two musicians (Tanaka, Dailleau) play sound and one musician (Babiole) plays image. Correspondance between sound and image is compositional in the choice of materials for each piece, and based on ensemble communication at the time of execution.

Le Loup is one composition often played as the opening piece in an S_S_S set. It opens with Dailleau alone, bringing in, from silence, a rich, thick bass pedal. It is not static, but is a long, slow ostinato melody, not dissimilar to a 'passacaglia' or 'chaconne.' On second iteration of the minute-long figure, Babiole enters with slow movements of her hands, articulating the appearance of a group of thin parallel white lines in the projection. The lines are broken and do not traverse the whole of the screen. The regularity of the group of lines, each with their break off point start to imply a shape through absence. [1] The implied emergence of
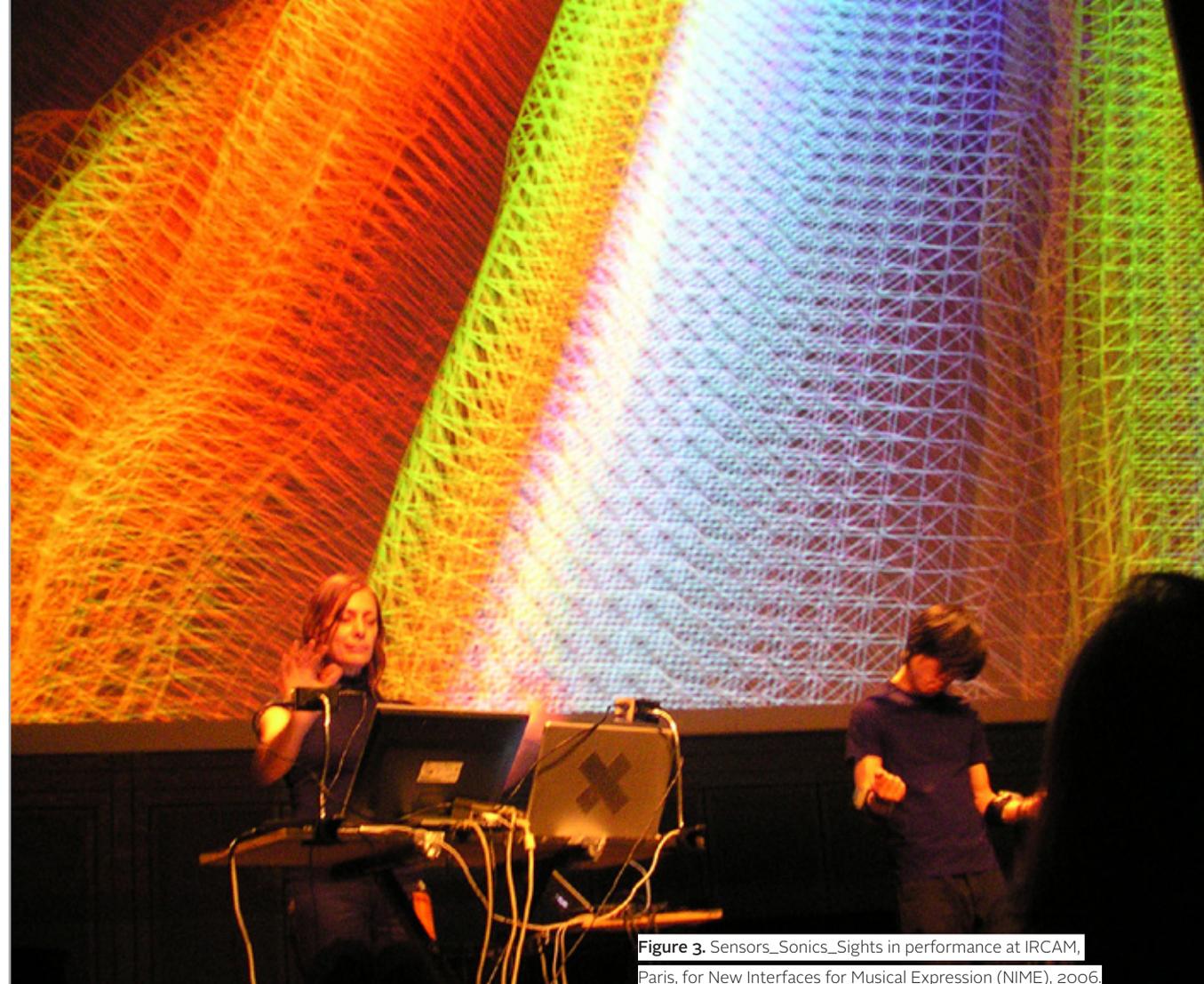
Figure 3. Sensors_Sonics_Sights in performance at IRCAM, Paris, for New Interfaces for Musical Expression (NIME), 2006. © Atau Tanaka, 2006. Used with permission.

a shape or object in the canvas through surface deformation is similar to the dynamic image mapping seen in *Overbow* and *Rail*. Here it is richer, by taking on 3D geometry, but simpler and more stark, as it described only as a series of very low resolution horizontal scan lines. Finally in the third repetition of the bass ostinato, Tanaka enters with slowly gnarled inward grasping gestures articulating a kind of muted mumbling voice-like sound. As the grasping becomes more intense, the gnarled sound becomes less muted, and it becomes apparent that the sound is not of human voices, but of wolves growling. Through this evolution, Dailleau continues the ostinato. The implied volumetric form deforms the horizontal scan lines more and more in Babiole's. The associations are not literal, but the two musical voices can be seen grossly as the canvas of horizontal scan lines (ostinato) and the volumes pushing through them (wolves).

Another piece, *Crackles*, is built solely on tiny snaps and pops, taking the nostalgic aesthetic of crackles on a vinyl record, and isolating it, making it the main material in the composition, performing the pops from BioMuse, Theremin, and ultrasound gesture. The audio sources used by Tanaka and Dailleau are audio artifacts – perhaps there were samples of vinyl surface noise, alongside data artifacts, and short noises. For the visuals, Babiole uses a matrix of rectangular white pixels. Early in the piece when articulated throwing or tossing gestures from the BioMuse sprinkle the audible field with small crackles, Babiole's pixels appear transiently like digital snow, quickly blown away by the wind. As Dailleau enters with scuttling aerosol swells, the pixels form into an organized matrix, establishing the canvas surface. The piece moves into less ad hoc, more mechanical repetitions, by which point groups of pixels, subdivisions of the canvas in vertical bands,

checkerboard groupings of pixels begin to break apart in the z-axis, popping out of the screen towards the viewer. At the climax of the piece when low frequency noise impulses almost set a rhythm, these panels of pixels are shifting and dancing rigidly in the depth plane, pixelating the surface in large low-res pixels which themselves are composed of the smaller elementary pixels. The near coordination of events ultimately dissolves as the piece reverts back to the lighter sporadic character of the beginning, closing on wisps of pops and wheezes, pixels blown off the canvas in one last breath.

Noise is a piece often used either as the climactic moment in an S_S_S set, else as a strong finale. It uses the principles from the last sections of *Overbow* and *Rail*, of infinitely stretched materials creating a wall of sound and image. This follows noise artist Masami Akita (aka Merzbow)'s original inspiration to delve into the essence of rock music by entering into the drum solo, remove the song around it, and make a whole song, album, or career out of the energy in the solo. **6** Tanaka this time lays out the sonic carpet with a series of oppressive low frequency blips punctuated by repetition of distant implosions. Babiole's Open GL line graphics quickly take shape as wireframe representation of a shifting surface. She cycles through subdued green, blue, orange tones while rotating 3D perspective, creating a three dimensional fly over effect, as Dailleau enters with tortures squeals. The composition has a classical single upward Boléroian arch, building in intensity. As layer upon layer of noise is introduced by Tanaka and Dailleau, Babiole's flat wireframe landscape becomes punctured by mountainous deformations. The green/orange hues and hilly terrain rumbling and pulsing turn the visual surface into a mesh wirefame – hollow but full – rendering of volcanic destruction.

DISCUSSION

While theories of objecthood were not used directly in the conception of *Overbow* or *Rail* or in the ensemble work of Sensors_Sonics_Sights (Kubovy and Schutz's paper hadn't been published yet), their approach to thinking about the relationship between sound and image are useful ways to analyze these performance pieces. More than being an a posteriori theorization of practice, the applicability of Kubovy's framework point out motivations in the creative work which is consistent with Gibson's assertion that sound/image relationships are not only about association. In other aspects, such as the gestural component, we will extend Kubovy's notion of audiovisual objecthood to become a 3-way notion of cross modal object.

Overbow begins with a classical approach to objecthood. The almost metaphoric brass-like sounds and melodic turns recall traditional musical objects – single notes as elementary objects, coalescing into phrases as higher level objects. Hyttrek's 3D blobs are amorphous yet distinct visual objects that jiggle and move as the melodic musical objects scuttle about. Sound as source and visual as surface become apparent in ways consistent with Kubovy's theory as the brass notes recede in distance and approach, and shift around in the stereo field. Meanwhile the 3D rendering of Hytrrek's models highlight variations in surface reflection.

The departure from classic objecthood first become apparent in the timbral synthesis of the brass-like sounds. The modulating oscillators that determine the upper harmonic make up of the instrument sound are fixed and do not transpose with the fundamental. This means that as the main tone plays a melody, certain synthesis parameters are held constant. It is as if some aspect of the sound source is rendered immobile, pulled out of the source and stuck to the background surface in a constant location. While the metaphor is not so attractive, it is as if one walked on some chewing gum, and lifting most of it off the ground

with the sole of the shoe (the melody) while some parts remain stuck to the pavement (the surface).

Kubovy's source/surface distinction becomes more ambiguous as the sounds in *Overbow* lose their note and phrase-based objecthood to become infinitely stretched washes. Meanwhile the visual canvas takes on life with dynamic image mapping that deform surface depth. The stretching of individual musical note into sustained tone then noise make it not so much a surface, but make its acoustic force in the concert space take on volumetric thickness. Likewise, the bulging out into space of the visual canvas make the image-surface object itself volumetric but short of becoming a 'source.' Working in acoustic volumes and implied visual volumes creates a kind of meeting point between sonic sources and light reflecting surfaces where sound objects and visual objects undergo a rapprochement not through association, but by losing constraints specific to their original mediums.

Kubovy and Schutz evoke causality as a relationship stronger than simple association. If a visual event, such as an impact, were seen to cause an audio phenomena, in their example, a percussive sound (or vice versa if some event in the sound were perceived to cause some change in image), the objecthood of the audiovisual pair is seen to be stronger. In the performance works described here, there is a third element in addition to the sound and the image – there is the performer's gesture that is captured by the interactive system. In simplest terms, the gesture, as 'controller' is the common cause of perturbations in both sound and image. However, interaction, it is said in HCI research, should not be seen as a one way street, and rich expressive musical interaction, according to research in the field of NIME, should be systems that give back to the performer as much as they take from the performer. **9** The feedback section towards the end of *Overbow* is a classic situation from electric

guitar and amplifier feedback where performer gesticulation is not just volitional, but reactive. In the case of the EMG based performance, shaping gestures 'tunes' the feedback and keeps the audiovisual system from going out of bounds. Far from being just a controller, the performer just responds to the behavior of the media he is producing within the very space he inhabits. The feedback, then, is not just sound output feeding back into sound input, but is a systemic process where unstable audiovisual projection instigated the performer to adjust posture and gesture, which in turn tunes audiovisual output. The audience is caught inside this large feedback loop, and from its multilateral causality, begins to construct a high level gesture-audio-visual cross modal object.

This recalls Gibson's ecological view of causality where causality is not just a question of what causes what, but helps us to understand one thing in the perceptual absence of the other, or of one element without the other. This helps us infer situations in environments where we have a priori knowledge. Experimental media performances are situations where the audience has little a priori knowledge, or where a priori assumptions maybe usurped. Instead of a known environment, each piece or composition creates a new environment, or ecology, and within the duration of its performance, must establish the environment, and familiarize the spectator sufficiently with it for them to make Gibsonian distinctions of presence, absence, linkage and causality.

Rail takes the environmental approach to causality as a means to extend traditional musical notions of objecthood. The base musical object is no longer the pitched note, phrase, or melodic unit, but the ways in which pure tones, in vision and audition, which independently would not constitute musical objects, take on object through linkage and gestural articulation. The visualization-like effect of white band tracking au-

dio pulse width is a form of direct association. Its articulation by performer gesture adds a common single cause to the two media, and with it a cause in their association. If we consider this tripartite relationship as a cross-modal environment, in Gibson's terms, the spectator could perceive incomplete subsets of the complete environment – for example only sound and image, or only gesture and sound, the situation still makes sense to the spectator. The association is not the key relationship, but a by-product of the causality, and the causality is robust to absence. Together, gesture, visual bands, and pure tone frequencies take on a kind of objecthood, a cross modal object that lacks traditional bounds of acoustic impact or visual event described by Kubovy, but that take on a palpable, visceral, near material quality.

S_S_S explodes the notion of gestural-audiovisual object into a multi-user ensemble context. Rather than exploiting the Theory of Indispensable Attributes to distinguish amongst multiple objects that may be generated by multiple players, the three members of the trio constitute together through coordinated ensemble performance, large scale multi-faceted cross modal objects. These are complex objects that may be made up of multiple voices, layers, or masses of particles. At the same time as there is a multifarious complexity, the actual primitives used are highly elemental. Short snippets of sound, the howling of one animal, visual representations in wireframe combine in a spatial dynamic to create rich tangible entities. Their structural evolution, guided through performance gesture result in a series of free-standing pieces where the composition becomes high level cross-modal object.

CONCLUSION

We have applied a concept of audio-visual object proposed by Kubovy and Schutz as a framework for analyzing several early live sound/image performance works. Kuboby's framework itself is built upon principles of environmental psychology established by J. J. Gibson, one where linkages between different perceptual modes is based not just on association, but on causality, as well as the possibility of different incomplete views where absence still allows a sense of the perceptual whole. The extension of association by way of causality and absence is a useful model to enrich relationships between sound and image in audiovisual artworks to go beyond simple effects of syaesthetic visualization and sonification.

The experimental media practice in *Overbow*, *Rail*, and the group S_S_S allow us to push the boundaries of objecthood from the classical impact/percussive event model described by Kubovy. We use traditional musical elements such as notes and melodies as a point of departure to stretch the limits of what might constitute aspects of a cross-modal object. While Kubovy takes time as a given, it is in the temporal nature of gesture, and corresponding evolution in what is seen and what is heard that elemental characteristics of sound such as frequency and image such as deformation coalesce to comprise a form of objecthood that draws upon audition and vision. In this way, we do not seek to identify indispensable attributes that distinguish multiple objects, but discover minimal characteristics that may not establish objecthood in an isolated mode, but contribute to object formation when modes of movement, sound, and image come together.

The gestural, performative nature of the works analyzed provides an opportunity to push Kubovy's framework beyond a binary relationship between sound and image. With their two-way relationship, the visual component of the audiovisual object also becomes the virtual gesture (mallet hit) causing the heard event. In the three-way cross-modality of the pieces presented here, gesture can be cause of sound, image, or both, freeing the sound/image relationship from direct causality. The tripartite audiovisual object uses Kubovy and Schutz's framework and extends it to describe performer-driven live audiovisual works. ∎

REFERENCES AND NOTES

1. D. R. Bradley and H. M. Petry, "Organizational Determinants of Subjective Contour: The Subjective Necker Cube," in *American Journal of Psychology* 90 (1977): 253–262.
2. J. J. Gibson, *The Senses Considered as Perceptual Systems* (Boston, MA: Houghton Mifflin,1966).
3. Volker Krefeld and Michel Waisvisz, "The Hand in the Web: An Interview with Michel Waisvisz," in *Computer Music Journal* 14, no. 2 (1990): 28-33.
4. M. Kubovy, and D. Van Valkenburg, "Auditory and Visual Objects," in *Cognition* 80, no. 1–2 (2001): 97– 126.
5. M. Schutz Kubovy, "Audio-Visual Objects," in *Review of Philosophy and Psychology* 1, no. 1 (2010): 41-61.
6. Edwin Pouncey, "Consumed by Noise," *The Wire* 198 (August 2000).
7. M. Puckette and D. Zicarelli, "MAX-An Interactive Graphic Programming Environment," Opcode Systems, Menlo Park, CA, 1990.
8. A. Tanaka, "The Use of Electromyogram Signals (EMG) in Musical Performance," in *CEC eContact!* 14, no. 2 (2012).
9. A. Tanaka, "BioMuse to Bondage: Corporeal Interaction in Performance and Exhibition," in *Intimacy: Across Visceral and Digital Performance* (In press), Maria Chatzichristodoulou and Rachel Zerihan, eds. (New York: Palgrave, 2012).

↑

la Biennale di Venezia

Arte
Architettura
Cinema
Danza
Musica
Teatro
Archivio Storico

MoCC
Museum of Contemporary Cats

The MoCC Pavillion