

## Aesthetic challenges of sonified video gestures

MICHAEL FILIMOWICZ

*School of Interactive Arts & Technology, Simon Fraser University*

**ABSTRACT:** This commentary responds to Jensenius and Godøy's paper, "Sonifying the shape of human body motion using motiongrams." Here I outline key aesthetic challenges to be addressed in future research in the sonification of video and gesture.

Submitted 2013 May 13; accepted 2013 May 28.

**KEYWORDS:** *gesture, spectromorphology, video sonification, sonomotiongram*

THE strategy of converting time-based matrix information such as video footage into spectrographs is a promising route for the development of new techniques that link visual and auditory mediation. While early efforts focused on what could be described as issues related to scoring sound through novel visualizations (e.g., UPIC or Metasynth, drawing as composing), current applications include information design (e.g., data sonification) and audiovisual media facades in which architectural spaces become ambient and immersive environments. This is not to even mention the vast practices in performance and installation found in art contexts. According to futurologists such as Michio Kaku (2012), one day the walls of our homes, work and public places will be wallpapered with wafer thin sheets of video display, and no doubt audio will be integrated in this transformation of built space into glowing and constantly changing surfaces of moving light and emanated sound. In short, visual music, as a form of ambient and non-narrative audiovisual media, may one day become as ubiquitous as furniture, as originally prophesied in Satie's conception of furniture music.

Attempts at articulating an aesthetics of sound and image have ranged from the Wagnerian *Gesamtkunstwerk* to Chion's *Audiovision*, and in between can be found modernist theories aiming at formal correspondences between sound and image (e.g., Scriabin's system of matching colors for specific notes) or even theories of non-correspondence more common in film theory (e.g., Pudovkin's "asynchronicity" or Bresson's non-duplication of seeing and hearing). Antecedent to these figures of course is Newton and the effort of the *Optics* to create a triadic theory of color for painters analogous to what the musicians had established with the construct of harmony. There is a rich history of ideas to draw on, spanning several hundred years (the earliest date given by Wikipedia for speculations about the color organ alone is 1590),<sup>[1]</sup> for clues as to what may ultimately yield a satisfying aesthetic framework that unifies video input, spectrographic sonic output, and motion-gestural articulation. The online archives of Rhythmic Light, Visual Music and Filmsound.org may prove useful for readers interested in probing deeper into this vast scholarship.<sup>[2]</sup>

A full review of the major themes of this aesthetic history is of course beyond the scope of the current commentary: however in what follows I will draw upon some of the major themes while keeping the focus on spectrographic sonification. What stands out from this history are the ever-changing means by which visual-sonic relationships are generated, the many attempts to formalize a conceptual or aesthetic language amongst practitioners, the shifting terrain of media concerned (from painting to performance to film, video and interaction), and the ever-renewed interest in this creative and technical domain.

### THE SYNAESTHETIC THEME

In reviewing all nine [video clips](#), it is clear that only up-down gestures map in a synaesthetic manner to the resultant sonification. Left-right movements yield a more or less steady-state tonality, while circular or diagonal movements in effect are only articulated by the vertical (up-down) aspect of those movements. This can be negatively understood as an essential limitation of sonomotiongrams (the authors refer to it as "the largest problem"), alternatively this can be positively understood as a basic feature or affordance of this technology as an instrument. As Collopy, Fuhrer, and Jameson note (1999, p. 111), "Any instrument imposes certain features of an aesthetic on its users. One cannot do with a trumpet precisely what one can do with a piano. One cannot achieve in oils the effects that can be produced easily using watercolors.

Thus, an instrument's design both limits expression and favors particular types of expression. Instrument designers must identify which choices are passed to the player. Choices provide the potential for expressiveness, though generally at the cost of added complexity." The mapping of frequency to the y-axis corresponds easily enough to cognitive schema in which the up/down binary overlaps with the sense of "high and low" frequency, since pitch itself is cognitively schematized metaphorically in experience in the vertical spatial dimension (Tolaas, 1991). Thus sonomotiongrams do not offer theremin-like intuitive control of primary sonic parameters (pitch and volume, or frequency and amplitude, mapped to the X-Y axis), but this is also to say that sonomotiongrams are not theremins, when considered as an instrument with its own affordances. Thus one strategy for future development may be to find a second manipulable parameter that can become part of a more sophisticated gestural vocabulary that operates within the specific limitations of sonomotiongrams. In my view, delta (variance in the velocity of a gesture) may offer one such possibility. This is easily captured within the technologies currently employed in the authors' prototype (the Max environment), and may yield a usable and much-needed second primary parameter of gestural control that may enrich the synaesthetic potentialities of motiongrams. Indeed, the steady-state character of left-to-right movements could prove highly useful in producing continuous non-changing sonic material (drones), since no motion at all would cease sound production entirely. The "problematic" nature of sideways movement can, in other words, be just as much a solution as a problem, creating a means to continue a tone without changing it, with stillness producing silences.

### THE CORRESPONDENCE THEME

Conceptually, the most contradictory aspect of sonomotiongrams is the allocation of energy (event density or intensity) in the resulting visualization. Natural acoustic events possess the largest amount of energy in the lowest bands, with energy typically diminishing in the upper harmonic ranges. Sonomotiongrams possess no such distribution of intensity ranges, and therein may lie an avenue of exploration for enhancing audience reception or understanding of gestural and sonic connections. Indeed, the authors' Figure 4(b) shows a complete inversion of natural acoustic phenomena, in that the visual bands weakest in intensity are distributed toward the bottom of the matrix, topped by the strongest bands (thickest brightest pixels). This suggests that perhaps an intermediary stage might be worth exploring, situated between Stage A (translation of a gesture into a motiongram) and Stage B (treatment of motiongram as a spectrograph), in which a reprocessing of the motiongram into an energy distribution more indicative of natural acoustic events (diminishing higher partials, strong lower registers) might make for a smoother transition to eventual sonification, and perhaps might yield an aesthetically stronger result (noting the general uniformity of the sounds produced by the current oscillator bank).

### THE NOISE THEME

Are we listening to music, sound or noise? A question first explored in Luigi Russolo's *Futurist Noise Manifesto* of 1913,[3] the authors are intriguingly ambivalent on this point, both in their stated aims and textual formulations. On the one hand, they clarify that sonification is not music making, and note that sonification can just as well serve informationally interpretive purposes as well as creative compositional practice. The video clips skew in the direction of the latter of course, and there is little elaboration on the possibilities of more "forensic" (e.g., information design) possibilities in their current research. And yet, despite the demonstrated leaning toward the music side of application, the authors radically separate motion sonification and traditional music making in their demonstration clips that do not combine the fiddler and the sonomotiongram's electroacoustic output. The acoustic recording and sonified output are rendered as two separate clips when, given the real-time nature of the sonomotiongram process, the fiddler could, rather easily, play in tandem (with or against, improvising to) the output of the sonification process. Further examples of this ambivalence or contradiction (between noise and music, and potentially noise-friendly music) are the clips sonifying the fiddler's feet clogging patterns and the electronic music "jam" footage in another clip. The mixing of acoustic and electroacoustic processes is a long-established performance tradition, and it would be helpful to this reader to see these two domains brought together more clearly in a real-time context in order to better gauge the qualities of the sonomotiongram technique in an actual performance context which combines real-time acoustic and gestural material.

## THE SYNCHRONY THEME

The real-time nature of sonomotiongrams connects its concerns with the over-arching concept of synchronization or even, to use Chion's term, "synchresis" (hearing-seeing). Here I believe that the authors' discussion of "shape" would perhaps be better articulated as an engagement with gesture. As they note, perception of shapes is "holistic" as is, of course, perception of gesture. This holistic feature is precisely what allows us to speak of this gesture, then that gesture, and so on. The use of the term "shape" seems oddly static, if only because the term suggests a figure within a spatial frame rather than an evolving temporal feature, whereas "gesture" combines spatial formation with a temporal envelope. A gesture is a contained and identifiable movement (a gesture-shape), and it is within this notion of movement that the authors may find an engagement with Denis Smalley's spectromorphology (1997) germane. In his now-classic article, Smalley articulates a schema for understanding sonic change which, while intended to describe spectral transformations in general, can perhaps work almost as well as a general language of gestural movement that can be applied to real-time performances of sonomotiongrams. The chart below is easily translatable into the kind of gestures demonstrated in the online demonstration clips of sonomotiongrams, and need not be limited to the structuring or description of the electroacoustic timbral output alone:

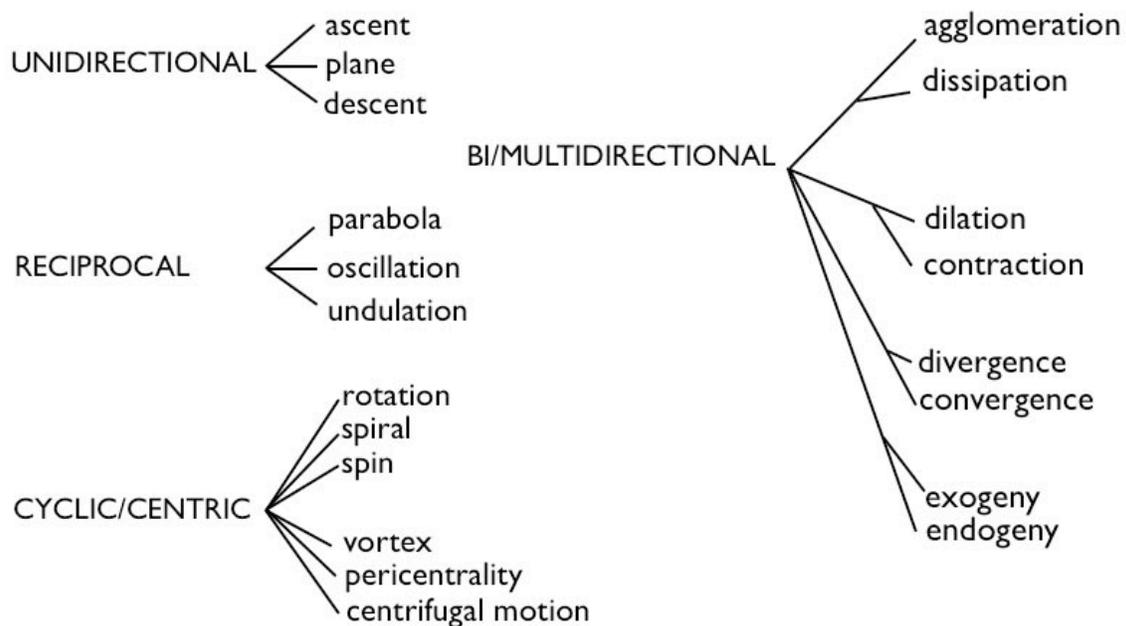


Fig. 1. Diagram from Denis Smalley's (1997) article on spectromorphology[4]

Conceivably, if not ambitiously, one can imagine a research project in which a database of specific gesture shapes are mapped to spectromorphologies in the output stage of the sonomotiongram process, more or less guided by Smalley's taxonomy which has been widely applied in the electroacoustic community.

## CONCLUSION

The sonomotiongram, while very much related to the pre-existing tradition of visual music, of course intersects with the newer computer-centric forms of interactive performance media. The history of visual music is rich in technological and aesthetic dead-ends (when was the last time any of us attended a performance of the color organ?), as well as major discoveries that have become well-assimilated into an array of common practices. The history of digital forms is highly volatile, with platforms coming and going and many works and applications disappearing in the tumultuous process of creative destruction that is the continuous development of the computational in general. The spectrograph and live video feed, however, appear (for now) to be mainstays of electronic infrastructure. Gesture, we should remember, is a

form of communication that can be widely apprehended by diverse audiences, whereas coding in a programming language is more like insider trading or the speech of experts to other experts. While the tools are digital, common understanding is analog. Since the authors make several references to “higher-level affective and aesthetic features” or “high-level, aesthetic and affective music-shape relationships,” we should be mindful of the role that analogic understanding plays in making the connection to these “higher” cognitive, affective and aesthetic dimensions that can be experienced by the viewer-listener. This indeed is the major challenge of this kind of work, since very few visual-sonic parameters map so clearly as ‘up/down screen = high/low pitch.’

### NOTES

[1] [http://en.wikipedia.org/wiki/Color\\_organ](http://en.wikipedia.org/wiki/Color_organ)

[2] See for example:

- <http://www.centerforvisualmusic.org>
- <http://rhythmiclight.com/archives/bibliography.html>
- [filmsound.org](http://filmsound.org)

[3] [http://www.artype.de/Sammlung/pdf/russolo\\_noise.pdf](http://www.artype.de/Sammlung/pdf/russolo_noise.pdf)

[4] Diagram published online at: <http://www.ceiarteuntref.edu.ar/blackburn>

### REFERENCES

Collopy, F., Fuhrer, R., & Jameson, D. (1999). Visual music in a visual programming language. In: *IEEE Symposium on Visual Languages*. Tokyo, Japan, pp.111-118.

Kaku, M. (2012). *The Physics of the Future*. Anchor.

Smalley, D. (1997). Spectromorphology: explaining sound-shapes. *Organised Sound*, Vol. 2, No. 2, pp. 107-126.

Tolaas, J. (1991). Notes on the origin of some spatialization metaphors. *Metaphor and Symbolic Activity*, Vol. 6, No. 3, pp. 203-218.