A METHOD FOR DEVELOPING AN IMPROVED MAPPING MODEL FOR DATA SONIFICATION

David Worrall

School of Music,
Australian National University,
Building 100, The Australian National University, Canberra ACT 0200, Australia.
David.Worrall@anu.edu.au

ABSTRACT

The unreliable detection of information in sonifications of multivariate data that employ parameter mapping is generally thought to be the result of the co-dependency of psychoacoustic dimensions. The method described here is aimed at discovering whether the perceptual accuracy of such information can be improved by rendering the sonification of the data with a mapping model influenced by the gestural metrics of performing musicians playing notated versions of the data. Conceptually, the Gesture-Encoded Sound Model (GESM) is a means of transducing multivariate datasets to sound synthesis and control parameters in such as way as to make the information in those datasets available to general listeners in a more perceptually coherent and stable way than is currently the case. The approach renders to sound a datastream not only using observable quantities (inverse transforms of known psychoacoustic principles), but latent variables of a Dynamic Bayesian Network trained with gestures of the physical body movements of performing musicians and hypotheses concerning other observable quantities of their coincident acoustic spectra. If successful, such a model should significantly broaden the applicability of data sonification as a perceptualisation technique.

1. INTRODUCTION

Parameter mapping sonification (PMS) is the most widely used technique for representing multi-dimensional data as sound [1]. PMSs are sometimes referred to as sonic scatter plots [2], [3], nth–order parameter mappings [4], or multivariate data–mappings, in which multiple variables are mapped to a single sound [5]. In this case data dimensions are mapped symbolically to sound parameters: either to physical (e.g. frequency, amplitude), psychophysical (e.g. pitch, loudness) or perceptually coherent complexes (e.g. timbre, rhythm).

Scaletti describes one way of implementing this as a “mapping of each component of a multidimensional data point to a coefficient of a polynomial and then using that polynomial as the transfer function for a sinusoidal input” [4]. Within an overall analogic mapping, symbolic representations such as auditory beacons [5] can be used to highlight features such as new maxima and minima, or absolute reference points, such as ticks to indicate the regular passing of time.

Frysinger provided an overview of the early history of the technique [6], and Flowers highlighted some of its pitfalls, including the observation that “while the claim that submitting the entire contents of ‘dense and complex’ datasets to sonification will lead to the ‘emergence’ of critical relationships continues to be made, I have yet to see it ‘work’” [3]. The main limitation of PMS is thought to be the known co-dependence of psychophysical parameters: linear changes in one domain produce non-linear auditory effects in another. These perceptual parameter interactions can also produce auditory artefacts that obscure data relations and confuse the listener regarding the parametric origin of the effect. A similar scenario occurs in visualisation, such as when parallel lines can appear more or less curved on different backgrounds.

\ There is general agreement among sonification researchers that 'the mapping problem' [3] is the most significant impediment to an otherwise flexible and potentially powerful means of representing information. Despite the enunciation of general heuristics, the problem has essentially remained unsolved. Kramer suggested that, although a truly balanced multivariate auditory display may not be possible in practice, given powerful enough tools, it may be possible to heuristically test mappings to within acceptable limits for any given application [5]. I outlined the historical and paradigmatic nature of this problem [7] and argued [8] that it is related to the problem faced by AI researchers at MIT in the 1960s and '70s trying to build a computational model of behaviour based on representation and predicate calculus. The failure of that approach for all but the simplest cognitive scenarios, so devastatingly critiqued by Dreyfus [9], has resulted in a search for alternate means.

I believe the strong historical alliance between sound synthesis for computer music and a cognitivist approach to artificial intelligence research [10] has both protected the problem from critical exposure to its causal analysis and impeded progress towards empirically verifiable solutions.

1.1. A new paradigm of perception

Recent findings in neuroscience, summarised in [11], indicate that a different paradigm of perception is emerging that involves the role of body gesture in neuronal mirroring. Churchland and others [12][13] have shown the limitations of concentrating on the neural correlates of conscious perception at the neglect of the role they have in servicing behaviour. Specifically, that there is a strong neuronal dynamic feedback loop between an observed phenomena and an organism’s own movements, exploratory and otherwise; that this loop extracts vastly more information about the causal properties of the external world in a given time interval, leading to greater predictive prowess, i.e. skills regarding the causal structure of the world, than could a purely passive system. The critical importance of timing in causal knowledge, as well as efference copy (being aware that a movement is one's own and not the world's), the nonconscious 'analysis' and memory of the movement of other movers, such as in pursue–evade relationships, for example, are encapsulated in the Guillery and Sherman hypothesis that messages to the thalamus and cortex also carry information about ongoing instructions to the organism's motor structures [14]. Consequently, as a developing organism interacts with the world, sensory signals also 'carry' gestural predictions: as an animal learns the consequences of a particular movement, it learns about what in the world will probably happen next, and hence what it might do after that.

1.2. Audio-visual mirror neurones

Recent studies have demonstrated that a mirror neuron system devoted to hand, mouth and foot actions is present in humans. Buccino, Solodkin and Small [15] review this literature and that of the experimental evidence on the role of the mirror neuron system in action understanding, imitation learning of novel complex actions, and internal rehearsal (motor imagery) of actions.

Important for this sonification research, is the finding that actions may also be recognised from their typical sound *alone*. Besides visual properties, it was found that about 15% of mirror neurons, called *audio-visual mirror neurons*, also respond to the specific sound of actions performed by other individuals even if only heard [16]. It has been argued that these neurons code the action content, which may be triggered either visually or acoustically. Phillips-Silver and Trainor demonstrated an early cross–modal interaction between body movement and auditory encoding of musical rhythm in infants [17]. They also found that it is primarily the way adults move their bodies to music, not visual observation, that critically influences their perception of a rhythmic structure [18]. In an experimental study of gestures, subjects of various ages were able, with a high degree of accuracy, on only hearing different individual human's walking and running on various kinds of surfaces, to determine their sex [19]. A consequential inference is that differences in ambulatory action, presumably resulting from relatively small differences in skeletal anatomy, is tacitly 'available' to listeners. Also consequent to these findings is the need for better models of multimodal sensory input, particularly with respect to the integrative functions of vestibulation and proprioception, which some empirical evidence suggests are available to listeners though aural means alone [20], [19].

1.3. Understanding musical structures through gesture

As their ability to understand musical structures such as melodies shows, humans have the capacity to create, transmit, receive, transform and, most importantly for this research, recall certain types of immanent ('mental') objects using sound. A new movement-encompassing action-based approach to the relationship between sound and sensibility began in the 1980s [21], supported by a phenomenology of embodied perception first enunciated by Merleau-Ponty [22] and formally categorised by Todes [23]. Methodologies include the use of abductive as well as inductive inference and are contributing to new perspectives on how to approach relationships between different sensibilities [21], [24]. Truslit studied the body movements of musical performers and suggested they were articulations of inner movements in the music itself [25]. Central to his approach are the notions of dynamics (intensity) and agogics (duration). If the music has the dynamo-agogic development corresponding to a natural movement, it will evoke the impression of this movement. He makes a distinction between rhythmic movement and the inner movement of the music: In contrast to rhythmic movement, which is related to individual parts of the body, the inner movement forms the melody via the vestibular labyrinth of the inner ear and is related to the human body as a whole. Both Nettheim [26] and Clynes [27] also make a connection between music and gravitational movement beneath the musical surface. These ideas are central in to empirical musicology–the systematic analysis of music based on the measurement of musical performance [28].

Studies in empirical musicology, including the mensural study of instrumental performer's gestures, and the neurophysical analysis of 'embodied' instrumental performance in general, is becoming recognised as at least as important for understanding musical ideas as notated structural abstractions (scores) [28][29]. At the same time, there is growing interest in human/machine interfaces, such as those for motion detection, that enable musicians to produce computer-generated sounds under nuanced gestural control [30][31][32]. A technically proficient musician does not normally just play 'sound objects' any more than a reader reads 'word objects', but acts a 'streamer' of events by shaping and integrating past musical actions with those being anticipated–by score-reading ahead for example, in much the same manner as a text reader does by scanning ahead for syntactic 'signposts' in order to dynamically modulate such things as prosody and pronunciation etc, to semantic effect.

2. DEVELOPING A GESTURE-ENCODED-SOUND MODEL (GESM)

The aim of the research described here is to find a means of sonifying multivariate datasets in such as way as to make the information in those datasets available to general listeners in a more perceptually coherent and stable way than with current parameter-mapping techniques.

The research is designed to empirically test whether or not the perception of information through such sonifications can be improved by using a mapping model for the transduction of the datasets to sound synthesis and control parameters that incorporates the gestures of performing musicians.

2.1. Integration of extant psychoacoustic knowledge

The first task for developing the GESM is to integrate extant knowledge and techniques into a common software environment, namely *SoniPy,* an existing data sonification framework of this researcher's design [33]. While fine motor gestural data of the type required appears scant, there is well-known psychoacoustic research data such as inverse filters, as well as more recent and comprehensive work [34] that should prove useful in building a relational database of psychoacoustic predicates and variables.

2.2. Stimulus generation and response-data collection

In order to generate gesture data, professional musicians will play datasets with known information content that have been transcribed to music notation. The quality and consistency of this notational 'stimulus' to the performers is a significant variable in the gesture-data generation process. So, a high-quality automatic notation-generation component will be incorporated into *SoniPy* and a sonification-dataset to notation control-file parser written.

Algorithmically generated test datasets within controlled gamuts and correlations will be stochastically generated and converted to musical notation for the performing musicians. Sound and gesture stimulus-response data from the 'reactions' of these musicians to the test data will be collected with both microphone and motion-capture (MOCAP) technology.

2.3. Design and test a DBN

Guided by the extant psychoacoustic findings (2.1 above), an analysis of common statistical features, such as auto- and cross-correlations between parametric variables will made of the collected response data. This analysis will be used to suggest initial separations and settings of potential latent variables of a DBN. This DBN, implemented in *Sonipy,* will be used to iterate over the initial test data to within acceptable convergent and divergent tolerances.

2.4. Implement DBN as GESM and test efficacy

Two sets of sonifications of the test data will be prepared– identical in all respects except that only one will be encoded with reference to the GESM. Empirical double-blind testing with subjects having a range of musical literacy will then be used to establish the degree of verisimilitude of the speculation on the basis of which this research is undertaken.

3. SPECULATIVE CONCLUSION

While the empirical studies of performer gestures and gesturally-controlled computer-music performance are of relevance to this investigation, the former are largely deficient for our purposes in being more analytical than generative, and the latter, in being more concerned with the musical affects than in an empirical evaluation of the roles that the gestures themselves play in producing musical coherence. However, it is the complex semi-tacit mindful body-integrating functionality of musical performance that the research described here seeks to capture, model and then apply to the synthesised 'sounding' of large multivariate datasets, in such as way as to make the information in those datasets available to general listeners in a more perceptually coherent and stable way than is currently the case.

Although the methodology of the project is quite advanced, the empirical work will be undertaken. When funding is secured. If eventually proved successful, such a model should broaden the applicability of parametrically-mapped data sonification as a perceptualisation technique.

4. REFERENCES

[1] D.R. Worrall, “An introduction to data sonification,” in R. T. Dean (ed.), *The Oxford Handbook of Computer Music and Digital Sound Culture*, Oxford: Oxford University Press, 2009.

[2] J.H. Flowers, D.C. Buhman and K.D. Turnage, “Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples,” in *Human Factors*, Volume 39, 1997, pp. 341-351.

[3] J.H. Flowers, “Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions,” in *Proceedings of the First Symposium on Auditory Graphs*, Limerick, Ireland, July 10, 2005.

[4] C. Scaletti, “Sound synthesis algorithms for auditory data representation,” in G. Kramer (ed.), *Auditory display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings, Volume XVIII. Reading, MA: Addison Wesley Publishing Company, 1994, pp. 223-251.

[5] G. Kramer, “Some organizing principles for representing data with sound,” in G. Kramer (ed.), *Auditory display: Sonification, Audification, and Auditory Interfaces,* Santa Fe Institute Studies in the Sciences of Complexity, Proceedings, Volume XVIII, Reading, MA: Addison Wesley Publishing Company, 1994, pp. 185-221.

[6] S.P. Frysinger, “A brief history of auditory data representation to the 1980s,” in *Proceedings of the First Symposium on Auditory Graphs*, Limerick, Ireland, July 10, 2005.

[7] D.R. Worrall. Sonification and Information: Concepts, instruments and techniques. Unpublished PhD thesis, University of Canberra, 2009. Available for download at http://erl.canberra.edu.au/public/adt-AUC20090818.142345/

[8] D.R. Worrall. Parameter mapping soniculation and the perceiving body, in *Proceedings of The 16th International Conference on Auditory Display,* Washington, D.C, USA, 2010.

[9] H. Dreyfus, *What computers still can't do,* Cambridge, MA: MIT Press, 1992.

[10] P.M. Todd and G. Loy.(eds.). Music and connectionism. Cambridge, MA: MIT Press, 1991.

[11] P.S. Churchland, “A neurophilosophical slant on consciousness research,” in V.A. Casagrande, R. Guillery, and S. Sherman, eds. *Cortical function: a view from the thalamus*. *Progress in Brain Research* , Volume 149, Amsterdam: Elsevier, 2005.

[12] B. Libet, “Unconscious cerebral initiative and the role of conscious will in voluntary action”, in *Behavioral and Brain Sciences*, 8, 1985, pp. 529–566.

[13] A. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain,* Orlando Fl. Harcourt, 2003.

[14] R.R. Llinas, *I of the vortex: From neurons to self,* MIT Press, MA: Cambridge, 2001.

[15] E.C. Kohler, M.A. Keysers, U. L. Fogassi, V. Gallese and G. Rizzolatti. “Hearing sounds, understanding actions: Action representation in mirror neurons,” in *Science*, 297 (5582), 2002, pp. 846–848.

[16] G. Buccino, A. Solodkin and S.L. Small. “Functions of the Mirror Neuron System: Implications for Neurorehabilitation,” in *Cognitive and Behavioral Neurology*, Volume 19, Number 1, 2006.

[17] J. Phillips-Silver, and L. J. Trainor, “Hearing what the body feels: Auditory encoding of rhythmic movement,” in *Cognition 105*, 2007, pp. 533–546. Amsterdam: Elsevier.

[18] M. Wilson and G. Knoblich, “The case for motor involvement in perceiving conspecifics,” in *Psychological Bulletin*, 131(3), 2005, pp. 460–473.

[19] R. Bresin and S. Dahl, “Experiments on gestures: walking, running, and hitting,” in D. Rocchesso and F. Fontana (eds.), *The Sounding Object*, 2003. Accessed on 25 October 2008 at http://www.soundobject.org.

[20] F. Varela, E. Thompson and E. Rosch, *The Embodied Mind,* Cambridge, MA: The MIT Press, 1991.

[21] N. Cumming, “The sonic self: musical subjectivity and signification,” in *Advances in semiotics,* Bloomington, Ind: Indiana University Press, 2000.

[22] M. Merleau-Ponty, *The phenomenology of perception,* C. Smith (trans.) of *Phénoménologie de la perception*", Paris: Gallimard, Oxford: Routledge & Kegan Paul, 1945/1962.

[23] S. Todes. *Body and World*, Cambridge, MA: The MIT Press, 2001.

[24] H.R. Maturana and F. J. Varela, *The tree of knowledge: the biological roots of human understanding,* Boston: New Science Library, 1987.

[25] B.H. Repp, B.H. 1993. "Music as motion: a synopsis of Alexander Truslit's (1938) *Gestaltung und Bewegung in der Music*," in *Psychology of Music*, Volume 12, Number 1, 1993, pp. 48–72.

[26] N. Nettheim, "How musical rhythm reveals human attitudes: Gustav Becking's theory," in *International Review of the Aesthetics and Sociology of Music,* Volume 27 Number 2, 1996, pp. 101– 122.

[27] M. Clynes, S*entics: the touch of emotions*. New York: Anchor Press, 1977.

[28] E.F. Clarke, "Empirical Methods in the Study of Performance," in E. Clarke and N. Cook, (eds.), *Empirical musicology: Aims, methods, prospects,* Oxford: Oxford University Press, 2004, pp. 77-102.

[29] R. Pelinski, "Embodiment and Musical Experience," in *Transcultural Music Review* Nr 9, 2005. Available at http://www.sibetrans.com/trans/trans9/pelinski-en.htm Accessed 7 June 2009.

[30] T. Winkler, "Making motion musical: Gestural mapping strategies for interactive computer music," in *1995 International Computer Music Conference*. San Francisco: International Computer Music Association, 1995.

[31] G. Paine, "Gesture and musical interaction: interactive engagement through dynamic morphology," in *Proceedings of the 2004 conference on New interfaces for musical expression,* Hamamatsu, Shizuoka, Japan, 2004, pp. 80–86.

[32] G. Paine, "Towards Unified Design Guidelines for New Interfaces for Musical Expression," in *Organised Sound*, Volume 14 Number, 2009, Cambridge UK: Cambridge University Press, pp. 142-155.

[33] D.R. Worrall, "Overcoming software inertia in data sonification research using the SoniPy framework," in *Proceedings of the Inaugural International Conference on Music Communication Science*, Sydney, Australia, December 5-7, 2007.

[34] D. Cabrera, S. Ferguson, and E. Schubert. "Psysound3: Software for acoustical and psychoacoustical analysis of sound recordings", in *Proceedings of the Thirteenth International Conference on Auditory Display*, Montreal, Canada, 2007.