OPHIDIAN AND THE UNCANNY VALLEY

Dr David Bessell

ICCMR
School of Performing Arts
Music Department
University of Plymouth
Drake Circus
Plymouth
Devon
PL4 8AA

ABSTRACT

Physical modelling is one of the most flexible and exciting of recent developments in software synthesis. New developments in synthesis methods are often accompanied by the emergence of new paradigms for the organisation of sound in composition. Here connections are drawn between Dr Masahiro Mori's concept of the 'Uncanny Valley' and tension and relaxation structures in timbrally varied musical phrases generated by physical models. This practice based study explores some of the possibilities with a flute physical model in an etude titled 'Ophidian'. In addition some further examples are discussed of the compositional possibilities of hybrid instruments made from mismatched components of other models. Parallels are made between ideas related to the uncanny valley and ways of categorising and organising synthesized sounds, which are more or less closely related to conventional acoustic instruments. The implications of these ideas for physically modelled instruments are tested in the piece Ophidian. The possibilities are indicated for a more generalised application of these ideas to other forms of synthesis.

1. INTRODUCTION

Dr Masahiro Mori is not normally considered a household name in the world of music, but he may prove to be an important one in the world of physical modelling. The reason is contained in the intriguing phrase 'the uncanny valley' coined by Dr Mori in a paper in 1970, [2] which was originally intended to describe some of his speculative theories on human psychological reactions to robotic design. This paper proposes that concepts associated with and arising from this theory may well prove useful to composers when working with unfamiliar materials, which nevertheless resemble closely conventional acoustic resources. This hypothesis is then tested in a practice based application of the ideas explored.

2. THE ORIGINAL THEORY

To reduce Dr Mori's ideas to their bare bones, the theory is, that if emotional response were plotted against similarity to human behaviour and appearance in robots, then the result would not be a steadily increasing curve of positive response. Rather there would be a peak of positive emotional response before the point of completely human look and behaviour is reached, followed by a deep trough of negative response at the point of close similarity, before rebounding to a positive response as completely convincing resemblance to human form is reached. This trough or valley, occurs at the point where the observer sees something which closely resembles the human but is just slightly unreal enough to seem eerie or unsettling.

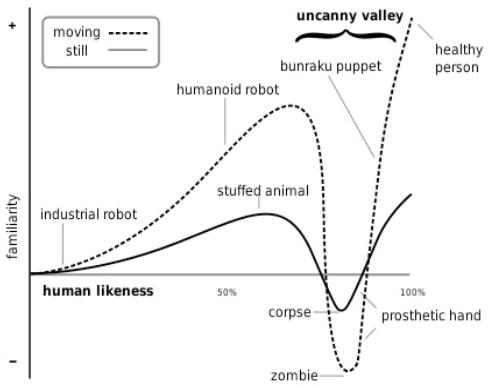

Figure 1. Mori's diagram of the uncanny valley [7] (trans MacDorman and Minato) [5]

Although this theory of the Uncanny Valley has become popular in the world of robotics, science fiction, computer games and computer animation it is still primarily a speculative theory and is not backed by a significant body of validating research. Indeed such research which does exist for example in the work of Karl MacDorman, [5] seems to suggest that the

phenomenon may be a good deal more complex than it at first seems, with multiple conflicting and contributing factors. For the purposes of this paper however ideas arising from it are used in an intuitive manner as an imaginative springboard in the piece 'Ophidian' for physical modelled flute. Informing both the organisation of the timbral material used and aspects of the structure of the piece.

3. MAPPING THE THEORY ON TO SOUND

If we begin mapping some of these concepts from robotics on to the world of electronic sound we might begin somewhat like this. An electronic sound synthesised with a technique such as Frequency Modulation with its typically highly 'unnatural' symmetrical spectrum might be placed far to the left of the graph, with low familiarity and low similarity to typical real world acoustic situations. Then somewhat further along we might place basic subtractive synthesis and so forth, [fig. 2] but the really interesting area comes with the question, what would lie in and around the area of the uncanny valley itself? There is more than one possible answer but in the piece Ophidian physical modelling is proposed as an uncanny synthesis technique par excellence. Of particular use to the composer wishing to explore this area is the attribute of a well realised physical model that allows a variety of subjectively real and near real timbres and musical behaviours to be executed.

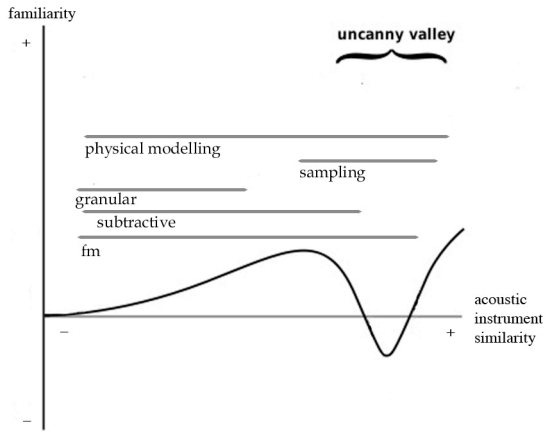

Figure 2. Diagram of the uncanny valley with common synthesis methods superimposed.

3.1. Some caveats and comparisons

Before proceeding with possible realisations of these ideas in the piece itself it should be noted that it would be rash to claim that these concepts from robotics map in a one to one fashion on to the area of musical synthesis. Rather there are what are best described as useful analogies. So for example the feeling of unsettling unease associated by Dr Mori with human like robotics would more accurately translate in musical terms to a kind of fascination associated with a sense of tension. This perhaps might be likened to the effect of some types of horror film or video game which make use of near human like characters, for example the classic zombie films of George Romero which it has been suggested are examples of the uncanny valley in action. [1] These films no doubt create uneasy fascination, but in musical terms a sound falling in to the uncanny valley might translate in to a generalised sense of tension and fascination, rather than unease specifically.

3.2. Implications arising from theoretical mapping

Of course the interaction between mechanisms for tension and resolution have been the mainstay of many musical structures for hundreds of years, if we now have access to a technique for organising and categorising a new form of musical tension and release then this surely must be of significant interest to the composer with technology. This tension and release mechanism can be utilised by the composer in ways which in a more conventional piece would be achieved by the use of tonal hierarchical relationships. In effect psychological dissonance and resolution arising from repeated traversing of the valley of the uncanny serves a purpose similar to musical dissonance in more conventional musical materials.

4. ANTHROPOMORPHISM AND THE MODELLED INSTRUMENT

The anthropomorphic aspect of Dr Mori's theory is also interesting in relation to its possible musical use with regard to physical modelling. This requires a short detour through historical uses of anthropomorphic forms in the visual arts which might fall somewhere on the first positive peak of the graph, to the left of the uncanny valley. Bryant has suggested [1] that in this area we might find various 'cute' or 'loveable' human animal hybrids, children's stuffed toys, the many anthropomorphic Disney characters or half human/animal mythical conglomerates associated with various forms of power, such as the Egyptian falcon headed god Horus or centaurs for example. It is interesting that these positive or powerful anthropomorphic characters often combine aspects of different animal and human parts, cartoon mice who walk upright with human like hands and so on.

Is there an equivalent positive position for physical modelling instruments made out of various more or less recognisable parts of existing instruments? In this category would fall instruments such as the clariflupet, a hybrid of clarinet, flute and trumpet components, or more radically the pizzicato gong. As part of this study these hybrid models have also been realised in Max msp

by the composer. The expectation would be, assuming the uncanny valley theory has relevance, that the timbres produced by such hybrid instruments would tend to fall somewhere on the first peak of the graph, in other words these timbres would have an instantly appealing quality to them, have some recognisable relation to acoustic instrument behaviour, but yet be unrecognisable as a specific instrument. This is potentially important because as Max Mathews has pointed out [6]

> '…the sounds that you produce with a digital process are either, uninteresting, or disagreeable, or downright painful and dangerous. So it's very hard to find beautiful timbres.'

He continues,

> 'The future will lie in better understanding of what sounds or sound sequences turn on the pleasure centre in the human brain, and the answer to finding this out will lie not in the technology of the machines, the instruments, the computers, but rather will lie in understanding how our brains interpret music.'

4.1. Implications for other synthesis methods

In short there are many interesting possibilities leading from an examination of this theory of psychological dissonance and resolution which suggest avenues of exploration for any composer interested in synthesis techniques which can create some semblance of realistic recognisable acoustic behaviour. The list of these might include such things as analysis/additive resynthesis or various forms of sample manipulation in addition to physical modelling. Even techniques such as FM synthesis [2] can be used in this way, for example Phone by Chowning [3] which repeatedly navigates back and forth between more or less recognisable vocal like sounds.

5. IN PRACTICE

The piece presented here, Ophidian, is comparable in this respect to Phoné in that it is mainly concerned with navigating to and fro across the uncanny valley. The model used is a waveguide model based on principles established by Julius Smith [8] and elaborated by Perry Cook [4]. Some details of the implementation of the excitation and angle of the wind jet in the flute model are extensions of the ideas presented by Smith and Cook.

The motivic material presented at the start of the piece is developed to a degree, in a somewhat conventional manner. Aspects of the first two phrases are substituted for each other in a brief homage to techniques used in Syrinx by Debussy. This developmental process quickly breaks down however and the pitch material attains a kind of relative stasis with limited motivic development, which throws the emphasis now on timbral development. This timbral articulation takes the form of repeated traversing of the uncanny valley, recognisably plausible acoustic flute fragments quickly transform in to near flute like distortions and back again. A series of phrases characterised by tension and relaxation of the timbral aspect of the flute model defined by its relation to the uncanny valley are created, which form the primary focus of the piece.

This timbral articulation of phrases is intended to make use of the property of tension, which it is suggested is the defining aspect of the traversing of the uncanny valley in a musical context. There is a perceptual game of hide and seek at work in which the audience is challenged to categorise the flute like sounds as convincingly real or fascinatingly off kilter in a rapid alternation of timbral transformation elicited by transformation of the model in real time. The listener may be attracted by the novelty of the timbral distortions of the flute model but this fascination also comes with a sense of resolution of tension when the sound becomes convincingly flute like. At times this happens even within the timbral evolution of a single note, which may for example begin realistically only to transform in some unlikely fashion.

Timbral transformations are deliberately limited to a continuum between realistic representations of the flutes sound and behaviour (conventional playing, breath noise, realistic multiphonics and whistle tones) and various near real flute related sounds which are hard to define but might include, extensions of the range, unrealistic whistle tones, and unlikely overblown and multiphonic like techniques. Sounds which are clearly not easily related to the flute although possible with the model architecture are generally not used, except at the end of the piece where the range drops so low and is combined with harmonic overblowing that the sound becomes dissociated completely from the recognisable flute family and is more akin to something related to an electronic didgeridoo. This final departure is intentionally intended to remove us from the area of the uncanny valley and produce the equivalent of a cadential gesture. Rather interestingly this cadential gesture not only attempts to make use of the relaxation of tension associated with the avoidance of the uncanny valley by moving to the left of Mori's graph but it also has some of the aspects of an interrupted cadence in that it finishes with an ambiguously abstract timbre, something like a further question which has not been answered.

6. PERFORMANCE IMPLEMENTATION

In the performance the midi score is read by the computer and the performer controls aspects of the

model which generate the timbral transformations. Model parameters are controlled with a video games controller, which is also used to spatialise the performance in four channels. The performer also governs aspects of tempo, thus placing him/her somewhere between the interpretative role of the conventional performer and conductor. The physical model is realised in Max/msp, the score in Sibelius and Cubase directs traffic via rewire. All sounds are generated in real time by the interaction of the computer score, performer and flute model.

7. CONCLUSION

The theory of the uncanny valley and the associated ideas and concepts which it throws up are fascinating terrain. They seem to offer a tantalising glimpse in to the psychological aspects of human concepts of beauty, ugliness and artistic rightness. Despite the lack of a significant body of verifying research the persistence of this theory and its migration to other contexts along with the questions raised indicate that it may well act as a useful set of signposts to creators in many artistic fields. Its real value in the field of music may well be not in its scientific verifiability but in the ideas it engenders and the responses it seems to elicit at a pragmatic or even subconscious level from composers and audiences.

8. REFERENCES

[1] Bryant, D. "The Uncanny Valley: Why are Monster-Movie Zombies so Horrifying and Talking Animals so Fascinating." Glimpses, Written Nonfiction. 2005. http://www.arclight.net/~pdb/nonfict…ny-valley.html

[2] Chowning, J. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation". Journal of the Audio Engineering Society 21 (7), 1973.

[3] Chowning, J. "Phoné" Wergo, 2012-50, 1993.

[4] Cook, P. "Real Sound Synthesis for Interactive Applications", A.K. Peters Press, Massachusetts, 2002.

[5] MacDorman, K. F. "Androids as an experimental apparatus: is there an uncanny valley and can we exploit it?", *CogSci-2005 workshop: Toward Social Mechanisms of Android Science, 108-118*, 2005.

[6] Mathews, M. "Keynote" *Music Technology Summit,* http://www.youtube.com/watch?v=Mi9ZikB_BUA

[7] Mori, M. "Bukimi no tani" (the uncanny valley), *Energy, 7, 33-35,* 1970.

[8] Smith, J. O. "A Basic Introduction to Digital Waveguide Synthesis" https://ccrma.stanford.edu/~jos/swgt/swgt.html