UNCOVERING BIOKINETIC INTERFACING

BEN JARRETT

Unitec
Auckland, New Zealand

ben@metier.co

Keywords: Biokinesis, Abstracted Musical Form, Neural Networks, PWM Controlled Electromagnets, Deep Cognition, Deep Pattern Recognition, Music-Form-Emotion, Intuitive Interfacing

The following paper details a research tangent explored within the scope of a larger project exploring a new method of interfacing the author has termed biokinetic:

> *A biokinetic interface is a new form of dynamic and ever changing device that enables instrumental mediation between one or more users and an electronic component.*

This research project employs an audio media player as the vehicle for validating the concept. In order for this interface to act in a nominally universal manner, the interface itself is visually abstracted, with the intention that anyone may approach the device and sense how to use it, and intuit what is happening, without recourse to agreed-upon historical symbols such as the play triangle, and the stop square.

Abstracted form would appear to have no underlying logical inference from which users can navigate the system, however in prior work, the author provisionally demonstrated – that for certain types of music at least – people appear to have a consistent and quantifiable abstracted visual language associated with specific music. It is the author's intent to use this visual language as the basis for interfacing.

1. INTRODUCTION

Current digital interfacing is digital, generally ignoring the particular qualities and sensitivities of each unique situation. This linear progression works precisely because it is indifferent – iTunes plays music we want at the volume we set: it does not matter how sensuously we stroke the play button: it plays. If we angrily stab the stop button: music ceases. The (emotional) state of the user – as evidenced by the quality of gesture and specific music chosen – is neglected in the design of most media players. It is the author's contention that current digital music players disconnect listeners from meaningful interaction from a product that is ostensibly designed to facilitate emotional connection.

Despite design theoreticians positing that intuitive interfaces should be responsive and physically dynamic (Norman, 2007), domestic audio electronics have yet to engage in this emergent field. A simple media player was chosen a suitable vehicle for this design investigation due to both the ease of data capture, and more specifically the quality of the data. There has been much research upon the emotional connection with music; from Kate Hevner's early experiments onwards (Hevner, 1936) and, modern computational methods of analysis such as MIRToolbox[1] have enabled sophisticated insight into audio data.

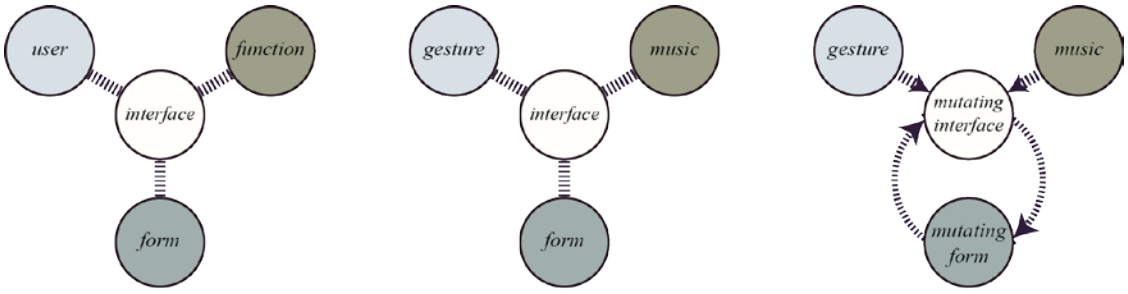


Fig. 1 Biokinetic interface logic

1.1. BIOKINESIS DEFINITION

A biokinetic interface is a new form of dynamic and ever changing device that enables instrumental mediation between one or more users and an electronic component. Eschewing touchscreens, a biokinetic interface uses physical input from the user to control the component and responds giving the user feedback by moving all or part of itself. The basis of the movement is a combination of user input and the component's function.

1 MIRToolbox for MATLAB by Lartillot, Toivainen & Eerola *https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox*

1.2. BIOKINETIC OBJECTS

Biokinetic objects are practically constructed from three interlinking forms: *dynamic*, *base* and *object*. *Object form* simply denotes the form of the physical form of the component itself, *base form* is the underlying structure for a specific class or function (such as a kettle boiling, or a particular song being played), and *dynamic form* is temporal and situation specific (such as gesture input, and instant data feedback).

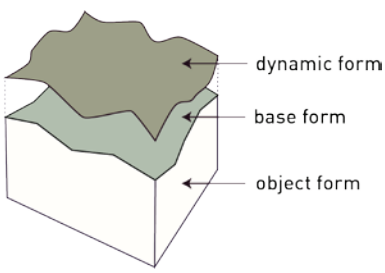
In the following paper a brief overview of visual perception and abstracted form is put forth, followed by deeper examination of the three interlinking forms: *dynamic*, *base* and *object* including different modalities and the use of neural networks. Finally, a tablet-based prototype is discussed, before a conclusion being drawn.

2. VISUAL PERCEPTION AND PATTERNS

The concept of universal human visual languaging is not a new one, although it appears very difficult to achieve in practice – body language being notoriously different across cultural boundaries, for example. The lowered wave of 'come here' in Eastern cultures being read as 'go away' among Westerners constitutes a common misunderstanding. Hans-Georg Gadamer refers to the crux of this position as *wirkungsgeschichtliches Bewußtsein – historically effected consciousness* (Gadamer, 1960). This referentialist argument of the primacy of the environment in our (visual) cognition, which is also echoed within our knowledge about differing cultural influences in music cognition, makes a strong case for unique individual bias in cognition.

However, if we relax the definition to only 'reading and understanding', then in practice, and with practice, we can name dynamic visual patterns that are coherent across all cultures – namely from the natural environment itself. The inherent subconscious understanding of the way clouds ceaselessly swirl, the way fire licks around a branch and the languid splat when fat raindrops hit the ground *are* part of universal human experience. Arguably, a wizened gaucho from the pampas in Argentina, and a young babysitter in Manila have the same physical reality when experiencing these singularities. Of course the reflexive and cognitive dimension of these events within recipient's consciousness will be likely distinct and different, but the event itself is con-

Fig. 2 Biokinetic object typology



sistent and universal. These natural states surround us – even in the city - and are coherent, meaning that cohesive changes in pattern occur together. For example, the storm surge increases as the rain and wind both intensify.

A quick examination (!) of our current understanding of visual perception, and how it relates to cognition is germane at this point to this discussion. In broad strokes, there are two main schools of thought (Gregory 1997): firstly the *direct perception* model as promulgated by J. J. Gibson. He essentially states that our visual perception requires little or no knowledge – our eyes receive ambient light bouncing off object surface and we then process this information within the brain.

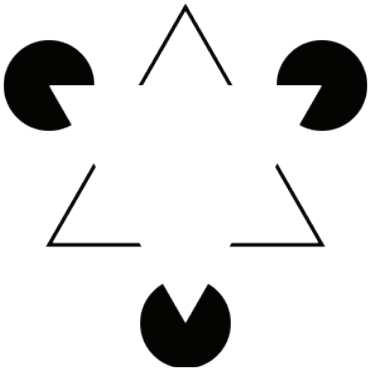
> *Direct perception means that the information in our sensory receptors is enough to perceive anything. We do not need any higher-level cognitive processes to mediate between our sensory experience and our perception (Sternberg 1997).*

Alternatively, the great polymath Hermann von Helmholtz (Helmholtz 1866) posited that visual perception arises out of unconscious inductive inference: we mentally access prior world knowledge and map it into our visual perception in order to process the limited and ambiguous imagery our eyes process. Richard Gregory further states that perceptions are hypotheses permitting us to sense the whole of objects by prediction – this is called amodal perception.

> *The notion is that stored-from-the-past potential intelligence of knowledge, is selected and applied to solve current perceptual problems by active processing of kinetic intelligence. The more available knowledge, the less processing is required; however, kinetic intelligence is needed for building useful knowledge, by learning through discovery and testing. (Gregory, 1997)*

Primate brains such as ours have significant lag between reality and the photoreceptors being active in the brain, which Helmholtz discovered in 1850, and has since been verified scientifically by others (Nijhawan, 2002), which may be indicative of this theory being valid (our minds also predictively account for this lag in moving ob-

Fig. 3 Kanizsa Triangle: amodal perception



jects). Additionally this theory also explains our visceral response to optical illusions, which the direct perception model cannot.

If we accept the validity of the second school of thought, it becomes possible to intentionally work within the vein of subconscious experience. Acknowledging that visual knowledge is unpacked more easily, and perhaps more quickly, with stored-from-the-past understanding, the author questioned the limits of visual kinetic intelligence, and posited that if we push visual complexity to a point where cognitive load becomes too great we must then intuit an interface, and attempt to process the information at a subconscious level.

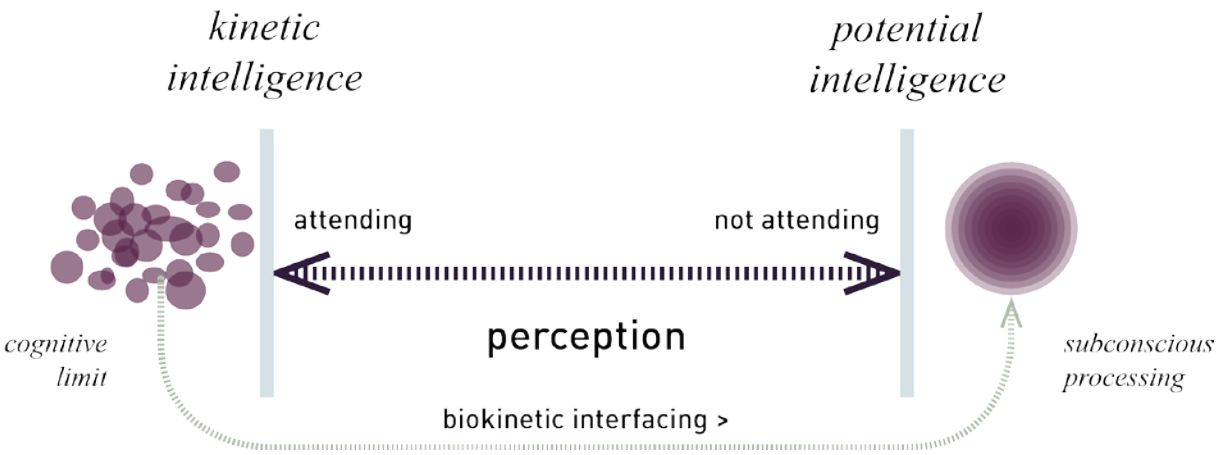


Fig. 4 Kinetic vs. Potential intelligence

Conscious (kinetic) processing via either model of representative (symbolic) form inherits distinct cultural bias therefore the concept of an abstracted interface was adopted as a possible route to a *nominally universal* design, which might be applied to music interfacing.

3. BIOKINETIC INTERFACING – DYNAMIC FORM (AUDIO)

Work at the beginning of this project assumed using multiple natural inputs to create computational algorithms that could emulate the dynamic nature of music from sources such as weather patterns, fluid dynamics, thermodynamics from wind and fire, and landscape forms from geography. However, when roughly tested with a series of simple After Effects created 'hairy interfaces' animations[2] it was discovered that people found simultaneous and competing inputs hard to 'read'.

2 Hairy Interfaces *http://www.metier.co/hairy-interfaces/*

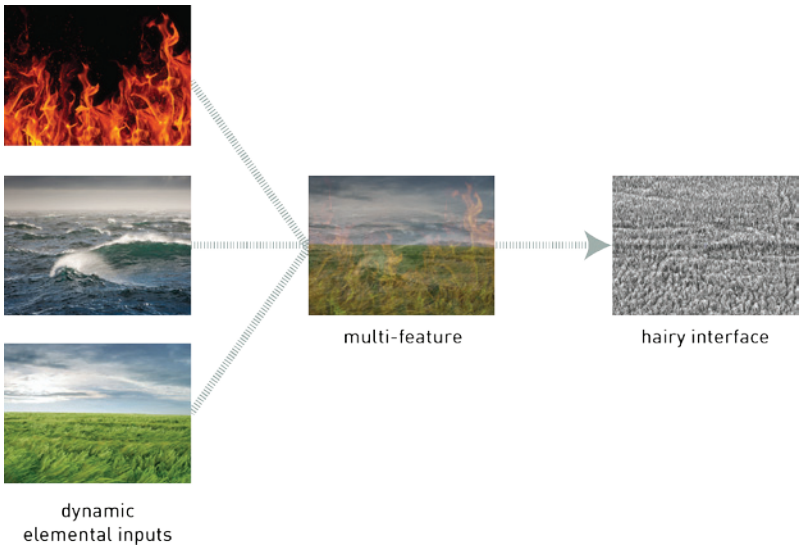


Fig. 5 Multimodal 'Hairy' Interfacing

After consideration, I realized that the best theory to explain why this misperception occurs is because these haphazard combinations constitute a chaotic *visual multi-modal* input. When we remove visual representation of these natural phenomena, and instead use abstracted dynamic form (that is to say using the change in state vs. time as the input), it appears our brains can make a 1:1 intermodal match, but multiple inputs lead to confusion. To put it simply – we shouldn't mix our metaphors.

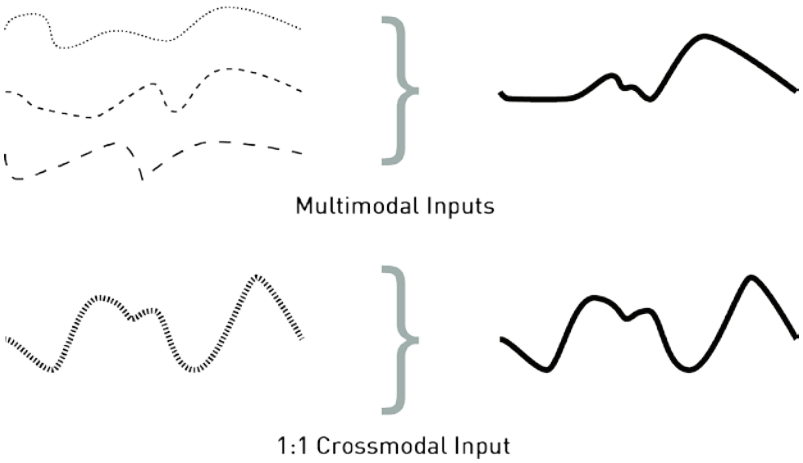


Fig. 6 Multimodal vs. Crossmodal Interfacing

Thusly, we may state for the purposes of this project that for effective biokinetic interfacing within an abstracted vernacular, the following conditions must be observed:

- The *visual modality* translation must be singular
- Users must have deep knowledge of all *visual modalities* presented
- Inputs and outputs must act coherently and consistently

If these conditions are met, then it might be said that these meaningful dynamic patterns project intentional agency. But what of the underlying forms themselves? How might they argue for the sense of thrilling Jazz, or downbeat Electronica?

4. BIOKINETIC INTERFACING – BASE FORM (AUDIO)

What exactly is the shape of sound?

Obviously we can all have different ideas of what colors, memories and emotions music can and does invoke internally. All of these responses are deeply personal, and cannot constitute a unifying constant that we may translate as a coherent, consistent visual language as a basis for interfacing. After looking at several related phenomena such as synesthesia, attempt was made to discover if we do indeed have a deep interrelated visual language associated with specific music. Again, the use of abstracted form was sought; in order to avoid the cultural bias of representative signifiers.

4.1. BACK PROPAGATED NEURAL NETWORKS AND ABSTRACTED MUSIC IMAGING

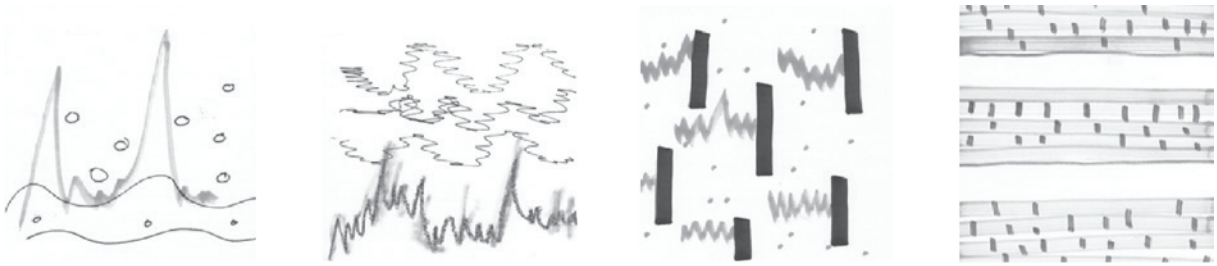


Fig. 7 User One: Abstracted images drawn to music

A sample set of 40 people were provided with a variety of drawing instruments and asked to draw four - abstracted – black and white images, while listening to four different styles of music. The 160 individual images were then digitized, and loaded them into a Back Propagated Neural Network (BPNN); which designed to find common patterns within each musical image group. New images are then inputted and ideally the BPNN is able to assign the new image to the correct group.

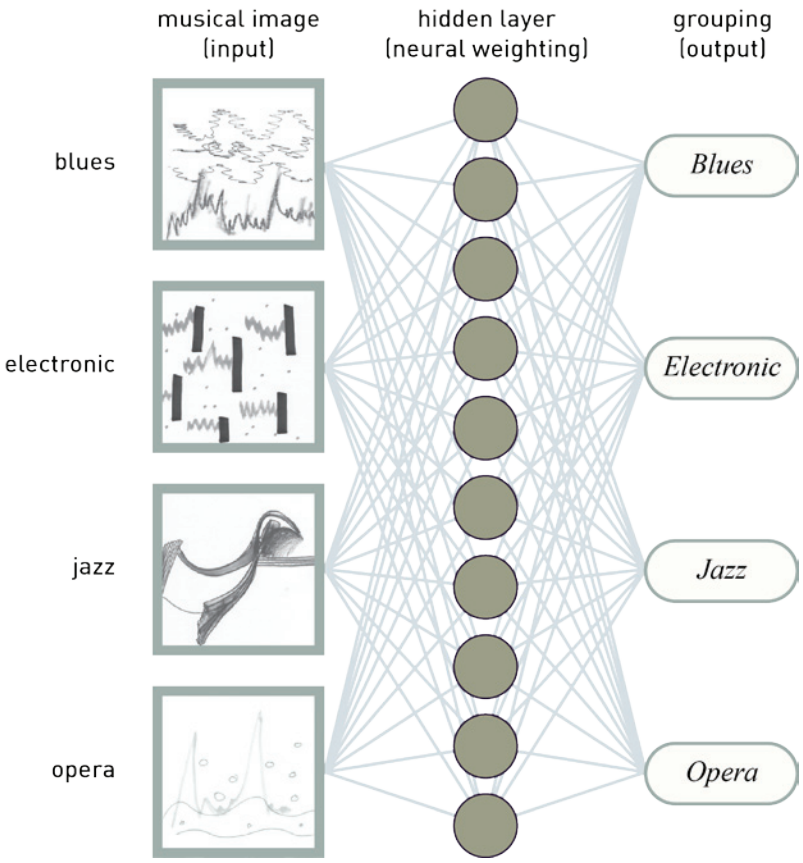


Fig. 8 Back Propagated Neural Networks for abstracted music imaging

Given that it is a small sample with very complex datasets, there does appear a trend towards a common and innate abstracted visual language associated with specific music types, if not all. In the validation set, the following results were recorded, where 25% is random chance:

- Opera 49.2%
- Blues 56.1%
- Electronic 100.0%
- Jazz 47.6%

Other BPNNs were run where drawings that were visually inconsistent or overtly representational were removed first, and the network was able to sort images with a much higher degree of accuracy. More research is presently been done in this area with the aim of validating the position with much larger sample sets collected from a greater diversity of users, which may be automatically analyzed and collated.

5. BIOKINETIC INTERFACING – OBJECT FORM

Fig. 9 Final Object – Biokinesis using magnetite; controlled via a grid of PWM electromagnetic coils

The final intended object form of this investigation is a physical interface that utilizes Pulse Width Modulation to dynamically control a grid of electromagnets. The resultant flux superposition governs the behavior of ferrous powder (magnetite) that is the manifestation of the visual interface. Whilst any given superposition of the magnetic flux is repeatable, the chaotic corporeal qualities of the powder reflect music's emotional complexity. For the purposes of this paper we shall put aside this final output to concentrate on a tablet based simulacra that manifests many biokinetic behaviors.

6. BIOKINETIC TABLET PROTOTYPE

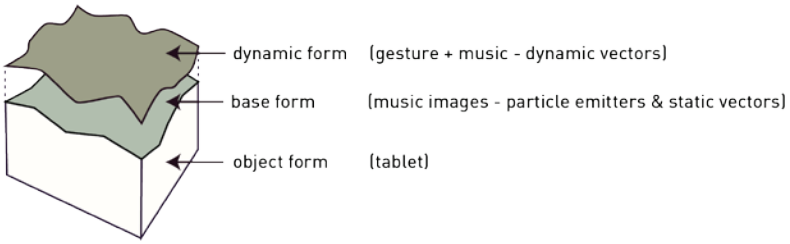


Fig. 10 Biokinetic tablet overview

This tablet-based interface currently utilizes Memo Akten's wonderful library: MSA Fluid[3] and coded within C++ / openFrameworks. This particle library responds to vector based inputs.

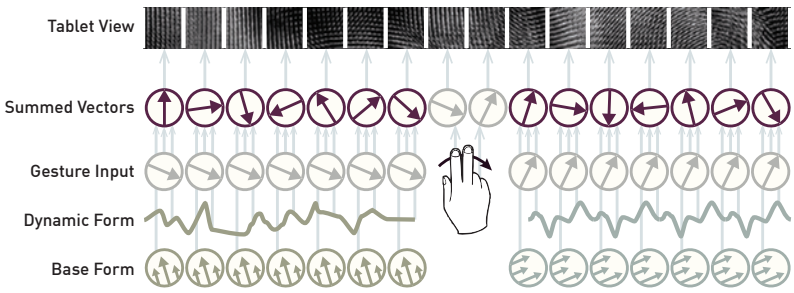


Fig. 11 Biokinetic tablet inputscoils

3 MSA Fluids Library by Memo Akten *http://www.memo.tv/ofxmsafluid/*

When music is playing dynamic vector forces are taken from two sources: *gesture input*, and *base form (image vectors)*. The *base form* variable is currently derived from the images drawn in the Neural Network experiments. The images are transformed into an of array particle emitters, each with an associated vector, and placed within the application for use as constant *base form*. The *dynamic form (music input)* as the name suggests is derived from the music currently playing. At present this input is simply acquired from discreet audio bands and parsed to the particle emitters; in time it is projected to use a multiplicity of computationally derived musical data such as dissonance, brightness, attack slope, etc. to affect both emission and vector angles. The *dynamic form* dynamically adjusts parameters (number of particles, emission speed, etc.) within the MSAFluid library emulating the desired 1:1 crossmodal input (music : fluid dynamics).

Finally the user's gestures are also used as dynamic inputs. User gestures are used in two ways: firstly there is simple visual feedback that shows the user that the gesture is being read by the machine. Secondly, it is intended that the nature of the gesture is also captured and used to mediate the response of the interface. For example, if the average gesture is fairly gentle in nature, and then the user gestures very aggressively, then the magnitude of the vectors will increase proportionally, mimicking the input. In this way, in concert with the music itself and the volume set we may give the interface a sense of responsive animation.

Fig. 12 Biokinetic tablet screen captures

Furthermore, when the user gestures to change songs, abstracted vector versions of the album cover art appear combined with the abstracted musical images for each track. This highly abstracted representation is the excep-

tion that proves the rule. The author found it difficult to create a logical interfacing segue between audio tracks and albums. Early prototyping revealed the users require a consistent and static image that they can repeatedly intuit. Merely abstracting the existing cover art worked is a simple solution that appears to have deeper subconscious connection for users. One possible explanation why the use of highly abstracted cover art appears to work in this context is that it is the artwork is produced artists and designers that are generally trying to convey the essence of the album. This currently feels like a somewhat temporary solution (and is a manual intervention!), and more experimentation and design is been taken within this area in order to get better results as well as automate the process.

7. CONCLUSION

> *Asynchronic events and disproportions caused by the continuously changing matrix-worlds of personalized, globalized, user-generated content are relegating form back into the orders of the unknown, back into exformations, where the encoding patterns of form are only just emerging – in interactive scenes. (Fasser, 2008)*

While conventional interface design seems converge on simple and efficient interaction, utilizing an increasingly standardized set of conventions, the author believes that alternate paths exist. Biokinesis is potentially one such path, seeking to utilize deep pattern cognition and natural mappings, allowing users to intuit an interface subconsciously. Intuitive interfacing is certainly not a new concept, and this form of interfacing is problematic for many, if not most, HCI applications. With the rise of the internet-of-things and the ubiquitous nature of screen-based interactions, one might feel that a new, nuanced physical screen-less interaction for controlling appropriate sense-data is both aesthetically desirable, and cognitively necessary in domestic environments to avoid technological overload.

In a more immediate sense, work is been done to release the tablet prototype as a iOS application that leverages iTunes to act as a media player. Amongst other efforts, further work is also being pursued to understand the relationship between music and form using web-based applications called ‘Interface Design Machines’ that

are able to collate far bigger data sets from a much wider target group, and automatically parsed this data into the neural network.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Cristiaan de Groot, Dr. Nigel Yee, Noel Zeng, Peter Koraca, Nova Paul, and particularly Taura Grieg for assistance with work associated with this paper.

REFERENCES

Fasser, Manfred. *Pre-Specifics – Some Comparatistic Investigations on Research in Design and Art*. ed. V. Buhlmann & M. Wiedmar, Germany: JRP Editions, 2008.

Gadamer, Hans Georg. *Truth and Method* New York, USA: Bloomsbury Academic, 2004.

Gregory, Richard. *Knowledge in Perception and Illusion, Philosophy Transcription: Royal Society London*. Retrieved from http://www.richardgregory.org/papers/knowl_illusion/knowledge-in-perception.html 1997

Helmholtz, Herman von. *Concerning the perceptions in general*. In *Treatise on physiological optics, vol. III, 3rd edn*. New York: Dover, 1866.

Hevner, Kate. Experimental studies of the elements of expression in music. *American Journal of Psychology*, vol. 48, 1936.

Nijhawan, Romi. *Trends in Cognitive Science: Neural delays, visual motion and the flash lag effect*. United Kingdom: Elsevier Science Ltd, 2002.

Norman, Donald. *The Design of Future Things*. New York: Basic Books, 2007.

Sternberg, Robert J. *Cognitive Psychology*. 2nd Ed. United Kingdom: Harcourt Brace College Publishers, 1996.