

Augoustinos Tsiros: 3rd Year PhD student, Edinburgh Napier University

Working title: Morpheme a multidimensional sketching interface for the control of concatenative synthesis. An investigation in the similarities between of auditory and visual features.

1. Introduction & Purpose

Technological developments in feature extraction, classification, modeling, and data mapping enable us to experience and interact with modal and amodal information in novel ways. These developments have increased our abilities to comprehend and assimilate information and offer many interesting opportunities for interaction. These developments however have also given rise to questions regarding the design of cross-modal mappings. Are there underlying principles based on which crossmodal associations could be made in objective terms? This thesis supports that examining similarities between auditory and visual perception and cognition could lead to a better understanding of the multimodal nature of perception, and expose common underlying principles that underpin all sensory modalities. Based on such principles, multisensory associations could be designed in more objective terms, leading to representations that derive their significance and expressive power from the sensory experience, rather than established through additional learning and convention. There is need for systematic empirical studies to identify areas of convergence between the sensory modalities and apply this knowledge to the design of multimodal interaction. This thesis' contribution consists of three components (i) forming a theoretical framework for audio-visual association, (ii) gathering empirical evidence related to audio-vision, and (iii) design and implementation of technology.

2. Morpheme Interface

In the context of this research, a prototype has been developed named *Morpheme*. *Morpheme* is a multidimensional sketching interface that allows the control of a concatenative synthesis by sketching on a digital canvas. *Morpheme* performs statistical analysis on a digital sketch developed by a practitioner, and it uses the data extracted from the canvas as the target feature for the selection of audio units and the control of the synthesis parameters. The aim is to enable synthesis of sound and the expression of compositional intention by providing perceptually meaningful visual description of sound attributes, in order to enable interaction with concatenative synthesis for creative purposes (e.g. sound design, electroacoustic composition). (for video demonstration of the interface, please follow the link : <http://inplayground.wordpress.com/morpheme/>).

3. Methodology

This work draws on the studies of computer music, human computer interaction, psychology, theoretical studies on perception as well as on contemporary approaches for creative sound and music practice. Two multidimensional mappings between audio and visual features have been constructed, each consisting of five audio-visual feature correlates. The associations between each auditory and visual feature pairs for each mapping is based on the findings of previous empirical work (not conducted by the author) which investigated audiovisual feature correlation. The distinction between the two mapping is that one is volumetric and spatial, while the other is based on the estimation of colour attributes. Two studies have been designed to test the following two hypotheses:

- 1) A number of structural and conceptual similarities between auditory and visual percepts could suggest how to make associations between auditory and visual stimuli in objective terms.
- 2) Achieving an intuitive audiovisual mapping should allow practitioners to use the system with minimal training, create a strong impression of causal relationships between visual and auditory elements, and aid in conceptualizing and articulating sound designs in visual terms.

3.1 Research Questions

- 1) Which association between two multidimensional auditory and visual feature sets would be considered as the most intuitive, and why?
- 2) How does synthesis of sound through the descriptions of sound in visual terms could affect the human-computer dialogue, and the creative output that derives from the interaction?

3.2 Studies

Study 1: Exploring audio-visual correlations: Participants are exposed to a brief audio-visual stimuli that has been created using the Morpheme interface, and different audio-visual feature associations. Participants are asked to rate the perceived similarity between auditory and visual elements.

Study 2 : Comparing Chromatic and Achromatic Mappings: Participants test two mappings using the system and then they have to rate the each mapping.

Study 3 Inferential potential of an audiovisual mapping: Participants are exposed to a brief audio stimuli that has been created using the Morpheme interface, the two mappings and a variety of audio corpus. Participants are asked to sketch the sound they heard. Their sketches are then compared to the sketch that created the sounds.

Study 4: Evaluation of the interface: Participants are given the interface to test at home and a week later a follow up semi-structured interview is conducted.

4. My questions for the panel:

- 1) Do you consider that the two hypothesis are worth testing, and coherent with the research questions?
- 2) Do you think that the studies I propose will allow me to test the hypothesis and answer the research questions?
- 3) Which of the studies that I am proposing you would consider as the least effective and why?
- 4) Do you think it is a shortcoming for my research that I have arrived to the design of the interface and the two mappings without having conducted preliminary studies, but mainly by drawing on the literature?