

Sketching Sounds: Using sound-shape associations to build a sketch-based sound synthesiser

Sebastian Löbbers^{1*} and George Fazekas¹

^{1*}Centre for Digital Music, Queen Mary University of London, U.K., s.lobbers@qmul.ac.uk

Abstract—Digital synthesisers are an integral part of modern music, but their complex controls can make it difficult to realise sound ideas in a straightforward way. This extended abstract gives an overview of a research project that aims to harness research into cross-modal associations between musical timbre and shapes to develop an intuitive control interface that can produce sound from a visual sketch input.

I. BACKGROUND

Cross-modal associations between shapes and sounds have been researched extensively in a theoretical context [1], but only little research has been conducted on how they could be used for sound synthesis or retrieval. Recent advancements in deep learning for sketch recognition, in particular Google’s *QuickDraw!* project [2], can inspire new mapping architectures for sketch-driven sound applications as demonstrated by Engeln et al. [3].

II. RESEARCH OVERVIEW

This research is centred around human participant studies and can broadly be divided into two parts that contribute to the development of a sketch-based sound synthesiser. On one side, perceptual studies are conducted to find out the different ways in which humans represent timbre through simple visual sketches and, on the other side, interface design and system usability studies are needed to investigate how this synthesis system could be incorporated into music production.

A first study [4], where twenty-eight participants were asked to sketch their associations with ten different sounds, showed that a mixture of *abstract* (lines, shapes etc.) and *realistic* (objects like musical instruments or scenes like ocean waves) representations can be expected if no restrictions are imposed. Quantitative analysis produced significant correlations between visual and audio features mainly in *abstract* sketches that align with existing sound-shape association research. A second evaluation study showed that participants matched these sketches to their related sound significantly higher than the random baseline, suggesting that at least some sound characteristics can be communicated through simple visual representations.

These findings informed a series of three interface design studies with the purpose of finding a setup that guides users towards simple, abstract sketches while maintaining a high level of perceived expressivity. The resulting interface was used to collect sketch representations of a synthesiser dataset by Hayes and Saitis [5] from eighty-eight participants. These sketches were fed into a deep learning (DL) classifier that was

pre-trained on abstract sketches¹ from the *QuickDraw!* dataset to distinguish *noisy* from *calm* sketches. The model was then in-cooperated into a first functional prototype seen in Figure 1.



Figure 1. Screenshot of the SketchSynth prototype implemented to run in a browser. See <https://youtu.be/ca1LYn8Yy-g> for a demonstration video.

III. FUTURE WORK

In the next step, the ability of the prototype to produce appropriate sounds from a sketch input will be evaluated through a user study while continuing to refine and extend the DL architecture and further explore correlations between visual and audio features. A key point of this research is to find out to what extent general sound-shape association can inform the cross-modal mapping and how this system could adapt to individual representational styles to become more robust and nuanced in a music production context.

ACKNOWLEDGMENT

EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology (EP/L01632X/1).

REFERENCES

- [1] Adeli M, Rouat J, Molotchnikoff S. *Audiovisual correspondence between musical timbre and visual shapes*. *Frontiers in human neuroscience*, 2014, p.352.
- [2] Ha D, Eck D. *A Neural Representation of Sketch Drawings*. arXiv preprint arXiv:1704.03477, 2017.
- [3] Engeln L, Le NL, McGinity M, Groh R. *Similarity Analysis of Visual Sketch-based Search for Sounds*. *Audio Mostly*, pp. 101-108.
- [4] Löbbers S, Barthel M, Fazekas G. *Sketching sounds: an exploratory study on sound-shape associations*. arXiv preprint arXiv:2107.07360. 2021.
- [5] Hayes B, Saitis C. There’s more to timbre than musical instruments: semantic dimensions of FM sounds.

¹ Circle, Square, Triangle, Squiggle, Zigzag and Line were chosen