Embodied Sketching for Neural Synthesis

Giacomo Lepri
InfoMus Lab
University of Genoa
Genoa, Italy
giacomo.lepri@edu.unige.it

Nicola Privato
Intelligent Instruments Lab
University of Iceland
Reykjavik, Iceland
nprivato@hi.is

Thor Magnusson
Intelligent Instruments Lab
University of Iceland
Reykjavik, Iceland
thormagnusson@hi.is

ABSTRACT

Stacco is a novel Digital Musical Instrument (DMI) with embedded magnets and sensors. It attracts and repels magnetic spheres and detects the changes in its magnetic fields upon interaction. It is designed to perform with neural audio synthesis models such as RAVE, in which sound features are represented and distributed within entangled multidimensional sonic spaces. Stacco allows drawing and embedding scores into the interface itself, and by bridging gesture and notation, it overcomes some of the inherent limitations of traditional notational methods as applied to neural synthesis. Our demo will provide the opportunity to try the instrument first hand. Moreover, we will invite participants to compose musical sketches for neural synthesis models, by drawing, embedding and layering tracing paper sheets on top of Stacco's surface.

CCS CONCEPTS

• **Human-centered computing → Sound-based input / output**.

KEYWORDS

Digital Musical Instrument , Neural Audio Synthesis, Embodied Sketching

ACM Reference Format:
Giacomo Lepri, Nicola Privato, and Thor Magnusson. 2024. Embodied Sketching for Neural Synthesis. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3678299.3678358

1 INTRODUCTION

In recent years, the artistic practices incorporating neural audio sythesis architectures [2] have multiplied [4, 6], and with them a series of novel affordances have entered the practice of musicians and composers. Among such are the entangled nature of neural synthesis models' latent space and the arbitrariness of the feature distribution obtained by training these architectures on a given corpus of sounds [3].

Due to this idiosyncrasy, each model may be understood, in performative and compositional terms, only after an a posteriori exploration of the latent space, which is unique to every particular model and mapping.

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678358

To facilitate such exploration we developed Stacco (Fig. 1), an instrument-score [5] specifically designed for the navigation of latent dimensions in neural synthesis models such as RAVE [2].

Stacco attracts and repels magnetic spheres and ferromagnetic objects allowing the performer to control fine musical details and wander freely into unexpected sonic realms.

What makes this instrument an ideal tool for the navigation of neural synthesis models, is the possibility of composing and performing with emmbeddable sketched oval sheets directly on its surface, thus merging latent exploration, gesture and notation on the body instrument.

2 COMPOSING & PERFORMING WITH NEURAL SYNTHESIS

Neural synthesis is a new, exciting synthesis method whose distinctive features open new expressive possibilities as well as novel compositional methodologies. Our work focuses on RAVE [2], a fine-grained latent variable model encoding raw audio into a stream of latent vectors before decoding it back to audio.

RAVE combines a variational autoencoder with an adversarial reconstruction term, distributing the sound features into an n-dimensional latent representation, with realistic details sampled in the decoding phase.

Such architecture offers (i) timbre transfer capabilities from a given audio prompt and (ii) real-time control through the direct manipulation of a multidimensional sound-space. In this paper, we focus on the latter function, thus mapping the model's latent space with control signals used as the inputs of its decoding function.

Among others, RAVE's latent dimensions exhibit the following relevant features:

- Control parameters are deeply entangled: it is impossible to univocally map one input with one particular sonic feature.
- The model's latent dimensions are always active. Consequently, interfaces employing RAVE are ideally designed to offer continuous values rather than discrete booleans.
- The distribution of the features in the latent space is arbitrary. During training, the system autonomously extracts and maps the sound features into the latent space. These relations can be only explored empirically, a posteriori, by interacting with the trained model.

In this paper, we propose a demo on the notion of *embodied sketching*, a practice we developed in order to deal with such affordances in the act of composing and performing with neural synthesis. Our method is based on drawing and layering cardboard ovals and transparent sheets on Stacco (Fig. 1).

Figure 1: Stacco. Under the hood, Stacco features four three-axis magnetometers and as many magnetic attractors, it performs embedded synthesis or can forward its readings to a laptop for more demanding tasks.

Figure 2: Previous Workshop Setup

3 STACCO

Stacco is an instrument-score [7], embedding magnetic attractors that detect variations in the surrounding magnetic fields. It attracts and repels magnetic spheres displaced by the performer, allowing both fine musical control and the emergence of unpredictable interactions out of its interlaced magnetic forces.

Overall, Stacco provides 8 continuous streams of data which relate to the x and y coordinates and strength of the magnets detected. We use Stacco to perform with a series of overlapping RAVE models. In most models, the 8 channels are mapped one-to-one to the latent dimensions of one model, with some doubling the mapping to the remaining latents.

We propose Stacco as an ideal controller to play with RAVE, as it allows to finely control high-dimensional data through simple gestures, provides continuous readings, and simplifies the mapping to RAVE's latent space.

Importantly, Stacco comes with two graphical scores that function as guides for the exploration of its magnetic attractors. By placing additional tracing paper sheets on the surface of the interface, the performer notates positions and gestures within meaningful areas in the latent space as these are discovered.

This practice, that is, embodied sketching, greatly facilitates the process of exploration, composition and performance with RAVE, offering an engaging way to understand the latent distribution of each model and to find meaningful areas at will.

4 DEMO DESCRIPTION

Building from Andrersen's work [1] we propose to explore the potential of embodied sketching for neural synthesis in a hands-on demo, where participants are invited to try out Stacco, design scores oval cardboard sheets, and play with them.

At the beginning of the activity, we will introduce the main features of Stacco and let participants freely explore the instrument and the related neural synthesis models.

Participants will be then provided with a set of materials (e.g. drawing tools, cardboard, clips and pins, twin on wire, etc.) to stimulate the creative making of new scores (Fig. 3). Two instances of Stacco bottom layer, with different engravings will be available, and the demo will involve the use of transparent as well as thick cardboard ovals, so that participants will be able to layer them on top of the instrument-score (Fig. 3).

We suggest for the demo an overall duration of 2 hours, so to have as many conference attendees as possible exploring and enjoying the instrument.

The technical setup for the demo will be rather simple: one large table with two electrical power sockets. We will bring the instruments, the workshop materials and two pairs of headphones (one for each instrument) for the participants.

We believe that the proposed demo will be valuable in several ways. First, it will provide the participants with an embodied, intuitive understanding of a new, cutting-edge technology such as neural synthesis.

Second, it will delve into this year's theme, exploring an art-informed and holistic approach to the design of sound-based interfaces involving artificial intelligence.

Third, getting feedback from the audience will allow us to gather new knowledge on performing with neural synthesis, on how to improve our magnetic scores, and on the compositional and design strategies that might develop around this novel technology.

REFERENCES

[1] Kristina Andersen. 2013. Making magic machines. In *10th European Academy of Design Conference*.

[2] Antoine Caillon and Philippe Esling. 2021. RAVE: A variational autoencoder for fast and high-quality neural audio synthesis. *arXiv preprint arXiv:2111.05011* (2021).

[3] Nicola Privato, Giacomo Lepri, Thor Magnusson, and Einar Torfi Einarsson. 2024. Sketching Magnetic interactions for Neural Synthesis. In *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'2024*. ICST - Zurich University of the Arts / Zürcher Hochschule der

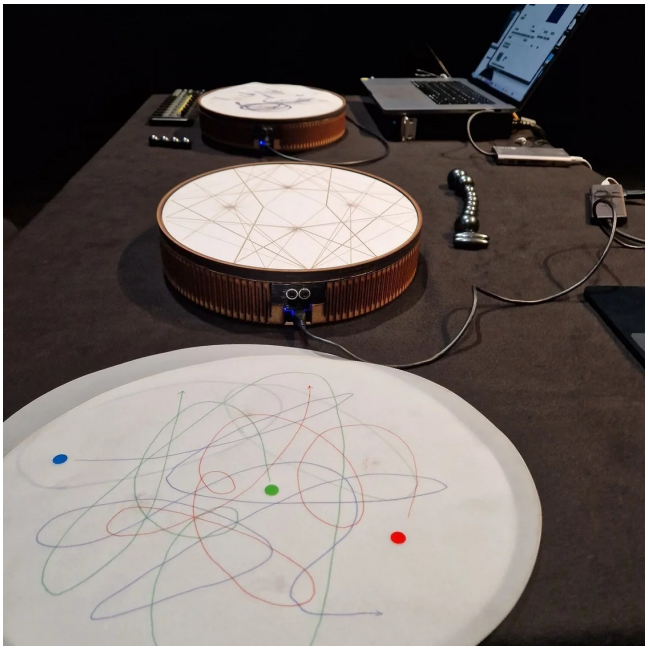

Figure 3: Example of notation layer embedded in the instrument.

Künste (ZHdK) Zurich, Switzerland, Boston, Massachusetts, USA.

[4] Nicola Privato, Thor Magnusson, and Einar Torfi Einarsson. 2023. Magnetic Interactions as a Somatosensory Interface. In *Proceedings of the 2023 International Conference on New Interfaces for Musical Expression (NIME'23)*. Mexico City.

[5] Nicola Privato, Thor Magnusson, and Einar Torfi Einarsson. 2023. The Magnetic Score: Somatosensory Inscriptions and Relational Design in The Instrument-Score. In *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'2023*, Anthony Paul De Ritis, Victor Zappi, Jeremy Van Buskirk, and John Mallia (Eds.). Northeastern University, Boston, Massachusetts, USA, 36 – 44.

[6] Victor Shepardson and Thor Magnusson. 2023. The Living Looper: Rethinking the Musical Loop as a Machine Action-Perception Loop. In *Proceedings of the 2023 International Conference on New Interfaces for Musical Expression (NIME'23)*. Mexico City.

[7] Enrique Tomás and Martin Kaltenbrunner. 2014. Tangible Scores: Shaping the Inherent Instrument Score. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Goldsmiths, University of London, London, United Kingdom, 609–614. https://doi.org/10.5281/zenodo.1178953