

sonic.snapshot

Ilia Viazov

Musikhochschule Lübeck
Anschützstraße 7a, Lübeck, Germany
iliaviazov@gmail.com

Abstract

This work explores new forms of human–machine improvisation through the interaction with a sonic self-image of the performer, projected into acoustic space via spatial sound techniques. A recorded improvisation on a Serge Modular Synthesizer is recontextualized into a structured sonic archive, transforming gestures into a material that can be revisited in performance. Previously fluid and momentary material becomes a compositional field, allowing the performer to navigate and reveal hidden sonic relationships during live improvisation. This approach shifts focus from linear performance to an unfolding dialogue with remembered and restructured sound. Spatialisation is treated as an active compositional process through which sound becomes dynamic and relational. The projected sonic material unfolds across the acoustic environment, creating a shifting field in which gestures are continuously recontextualized. The sonic material moves, overlaps, and transforms, allowing spatial relations to emerge and dissolve over time and underling its semantic meaning. This approach emphasizes sound as a latent presence that can be sculpted, displaced, and reinterpreted in real time—inviting the performer to engage with space as a responsive and evolving partner. The work contrasts the fluid, unstable nature of analog synthesis with a more reflective performative control, where musical decisions emerge through an interplay between intuition and structured re-engagement with past material.

1 INTRODUCTION

Live-Coding is an improvisational and compositional performative practice that investigates the computer as a performative instrument and interface (Blackwell et al., 2022). Within this paradigm, the computer is not only a tool but a screen-based instrument that mediates between logical structures and embodied musical gestures (Magnusson, 2006).

The project *sonic.snapshot* proposes an approach to improvisation as an interaction with previously recorded material, reorganized into a structured sonic memory. The system captures a performance and transforms it into a systemized dataset, creating a hybrid mode where real-time decision-making coexists with the reactivation of past gestures.

The software environment *tusa*, or *Tidal Unit for Sonic Activities* (Ilia Viazov, 2026), provides the technical framework. Built around the Tidal Cycles live coding environment (McLean, 2026), it integrates corpus-based concatenative synthesis, machine learning, and spatialisation into a unified system, drawing on the FluCoMa toolkit for feature extraction and data-driven audio querying (Tremblay et al., 2021). Its role remains supportive—the artistic focus lies in how the performer engages with the resulting sonic archive.

2 FROM IMPROVISATION TO SONIC ARCHIVE

Recent live coding practices have increasingly incorporated machine learning as an active component of musical performance, extending traditional deterministic interaction toward responsive and adaptive interfaces. Machine learning enables the organization, transformation, and generation of sonic material in ways that are not strictly predefined, opening new possibilities for engaging with sound as data (Fiebrink and Caramiaux, 2018).

In AI-assisted systems, performers can engage with models that evolve during performance. In *Pandora’s Dream*, Celeste Betancur explores real-time training and interaction with machine learning models, allowing the system to adapt dynamically and generate material (Betancur Gutierrez, 2023). Such approaches shift the paradigm from fixed tools toward co-creative systems, where outcomes emerge from a continuous negotiation between human intention and computational processes.

Corpus-based concatenative synthesis provides a structurally grounded approach within this context. Rather than generating new material, it reorganizes existing recordings by segmenting, analyzing, and recombining them according to feature-based descriptors (Schwarz, 2012), maintaining continuity between recorded and performed sound.

This descriptor-based organization contrasts with approaches that structure sonic material through learned latent representations. RAVE, for instance, trains a variational autoencoder to compress and reconstruct audio, producing a continuous latent space shaped by the network’s training rather than by perceptually legible axes (Caillon and Esling, 2021). In *sonic.snapshot*, by contrast, the feature space is built from directly interpretable descriptors—loudness and spectral shape—so proximity between fragments corresponds to audible similarity rather than a learned embedding: every point in the space is a real fragment of the original performance, not a synthesized approximation.

In this sense, the system continues sample-based sequencing rather than departing from it. Conventional sample banks organize material categorically—by file name or tag—requiring the performer to already know what a sample is before retrieving it. By organizing fragments through continuous, feature-based descriptors in-

stead, *sonic.snapshot* keeps the basic logic of triggering discrete recorded units while making the space between them intelligible: the performer navigates by ear and by relation rather than by label.

3 SOFTWARE OVERVIEW

sonic.snapshot is performed using *tusa* (Tidal Unit for Sonic Activities), a terminal-based performance environment built around Tidal Cycles (McLean, 2026). *tusa* assembles a small ecosystem of tools into one continuous performance surface: a Haskell interpreter for evaluating Tidal Cycles patterns (The Glasgow Haskell Team, 2024; McLean, 2024), a lightweight text editor for shaping and storing material (Free Software Foundation, 2024), a searchable documentation pane for quick reference mid-performance/mid-practice (Charmbracelet, Inc., 2026), and a SuperCollider server handling audio synthesis and processing (McCartney, 1996)—all held together within a single terminal multiplexer session (Marriott, 2024). The result is a self-contained instrument that can be launched from a single command.

Onto this base, *tusa* grafts a layer of corpus-based synthesis and spatialisation: recorded audio is analyzed, described, and made searchable in real time through the Fluid Corpus Manipulation Toolkit (Tremblay et al., 2021), turning static recordings into material the performer can query and traverse during play, while an integrated ambisonic encoding and beamforming chain allows the retrieved material to be positioned and moved through space as it sounds.

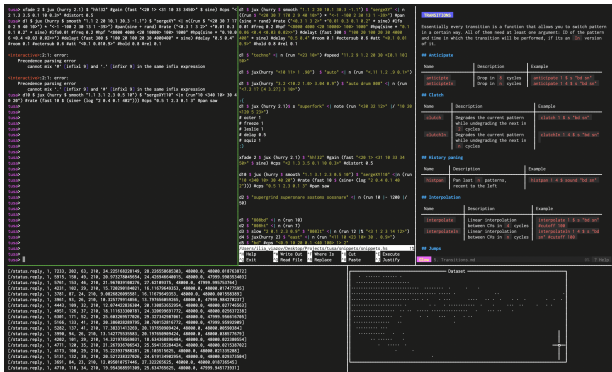


Figure 1: Tidal Unit for Sonic Activities

4 PERFORMING WITH SONIC MEMORY

At the core of *sonic.snapshot* is a recorded improvisation on a Serge Modular Synthesizer. The recording is segmented into individual grains at each detected onset, yielding a dataset of 5,279 fragments. Each fragment is described along two dimensions—loudness and spectral shape—extracted and organized using the FluCoMa toolkit (Tremblay et al., 2021), and indexed within a k-d tree, a space-partitioning structure that recursively divides the feature space along alternating dimensions. During performance, fragment retrieval is performed via k-nearest-neighbour (KNN) search over this structure rather than by exhaustive comparison against the entire dataset: a naive KNN search requires a linear scan of all fragments per query, whereas a use of k-d tree reduces the expected search cost. This is decisive for maintaining low-latency responsiveness at performance tempo, particularly as the archive grows. This structural choice is what allows the

dataset to function not as a static playback buffer but as a space that can be interrogated in real time.

This transformation changes the status of the sound. A linear temporal event becomes a field of relations defined by similarity and distance in feature space. Individual fragments are no longer bound to their original position in time but exist as elements in a navigable network.

The dataset functions as an externalized memory of the improvisation, preserving expressive qualities while making them accessible in new configurations. Improvisation is extended rather than replaced: the performer interacts with their own past material, mediated through computational organization. Each fragment retains its spectral and temporal identity while its position is determined by feature analysis, allowing traversal through both learned structure and original temporal ordering.

This setup changes the logic of improvisation. The sound is drawn from a fixed recording that defines the limits of what can be accessed while still allowing for continuous re-configuration. The performer operates within this space by shaping trajectories—moving between regions, repeating material, or revealing less prominent areas of the source.

Two modes of interaction condition how gesture is enacted within this system. The performer can navigate the dataset continuously, scrolling through the feature space by hand—an embodied, quasi-physical mode of traversal that recalls the continuous control of analog interaction with the synthesizer. Alternatively, fragments can be queried directly from within the live coding language by specifying target descriptor values, a symbolic and discrete mode of access closer to the abstraction of text-based interaction. The tension between these two modes—continuous, embodied navigation versus discrete, language-mediated query—mirrors the broader tension the piece stages between the physical gestuality of the original Serge performance and its abstraction into data. This tension is not resolved but treated as constitutive of the piece: gesture here is neither fully preserved nor fully abstracted, but held in an unstable position between the two.

Because organization is based on acoustic similarity, transitions often expose relationships not evident in the original improvisation. Abrupt shifts between distant regions introduce discontinuities that become structurally meaningful during performance. These effects are not precomposed but emerge from the interaction between data structure and real-time navigation. Improvisation here involves shaping density, timing, and parameter variation while responding to the material the system generates.

5 SPATIALISATION AS COMPOSITIONAL PROCESS

Spatialisation plays a central role in articulating the sonic archive. In *sonic.snapshot*, it is realized through performance with the IKO loudspeaker, where material is distributed and transformed using beamforming techniques (Zotter et al., 2017). The system projects sound as focused beams that can be steered and modulated in real time.

The spatialisation chain combines Kronlachner’s *ambix_encoder_i8_o3* (Kronlachner, 2024a), an eight-channel encoder for third-order Ambisonics, with *mcfx_convolver24* (Kronlachner, 2024b), a multichannel convolution matrix hosted in SuperCollider via *VSTPlugin* (Institut für Elektronische Musik und Akustik, 2024). The

eight-channel output of Tidal’s SuperDirt is treated as a virtual circular loudspeaker arrangement and encoded into a third-order Ambisonic sound field. This representation is subsequently decoded through a convolution-based beamforming process, producing the 20-channel signal required to drive the IKO. Spatial parameters can be controlled from the same live coding environment that drives the navigation of the sonic archive, coupling spatial behaviour directly to the processes organizing the recorded material.

The structured dataset is not only sequenced temporally but positioned and moved across the loudspeaker surface. Individual fragments can be localized, spread, or rapidly displaced, creating dynamic spatial configurations. Spatial control is integrated directly into the live coding process, where panning and modulation functions influence the direction and behavior of the projected sound field.

Ambisonics enables the treatment of space as a continuous domain. Rather than assigning sounds to fixed speaker positions, the system defines trajectories and regions of activity that can be reshaped during performance. Spatial motion can reach high temporal resolution, allowing modulation to operate at rates comparable to spectral processes.

This has direct consequences for perception. Fragments are differentiated not only by sonic features but by spatial behavior. Rapid movements, rotations, and dispersions introduce additional layers of variation, making spatial position an active parameter in recomposing the dataset. In this way, spatialisation functions as a form of synthesis, enriching the perceived structure of the sound. The sonic archive is articulated simultaneously in time, spectrum, and space.

The relationship between the performer’s traversal of the archive and what becomes audible to a listener is only partially perceptible independent of prior knowledge of the source material. What is perceptible, however, is a correlation between feature-space distance and spatial behavior: transitions between acoustically similar fragments tend to produce continuous, localized spatial motion, while jumps between distant regions of the archive tend to surface as abrupt spatial displacements. This correlation offers listeners an audible trace of the underlying navigation, even without access to the original recording, though the significance of any specific fragment as *remembered* material remains available primarily to the performer. Strengthening this audible correspondence—for instance through a more consistent mapping between feature-space distance and spatial parameters—is an ongoing direction for the piece.

6 CONCLUSION

sonic.snapshot presents a concept of improvisation as a dialogue with structured sonic memory. By transforming a recorded performance into a navigable dataset, it enables musical interaction that bridges immediacy and reflection.

The performer engages with past gestures not as fixed artifacts but as dynamic material reshaped in real time. Spatialisation further expands this interaction, embedding the sonic archive within the physical environment and emphasizing relational listening.

Within this framework, improvisation becomes an exploration of temporal layering, where memory and action continuously inform one another. The result challenges the distinction between composition and improvisation,

proposing instead a fluid continuum mediated by technology.

REFERENCES

- Betancur Gutierrez, C. (2023). AI in Live Coding Environments: Pandora’s Dream. In *Proceedings of the International Conference on AI and Musical Creativity (AIMC)*.
- Blackwell, A. F., Cocker, E., Cox, G., McLean, A., and Magnusson, T. (2022). *Live Coding: A User’s manual*. MIT Press.
- Caillon, A. and Esling, P. (2021). Rave: A variational autoencoder for fast and high-quality neural audio synthesis. *arXiv preprint arXiv:2111.05011*.
- Charmbracelet, Inc. (2026). Glow Markdown Reader.
- Fiebrink, R. and Caramiaux, B. (2018). The Machine Learning Algorithm as Creative Musical Tool. In *The Oxford Handbook of Algorithmic Music*. Oxford University Press.
- Free Software Foundation (2024). GNU Nano Text Editor.
- Ilia Viazov (2026). Tidal Unit for Sonic Activities. Repository.
- Institut für Elektronische Musik und Akustik (2024). VST plugin for SuperCollider.
- Kronlachner, M. (2024a). Ambisonic Plug-in Suite.
- Kronlachner, M. (2024b). Multichannel Cross-Platform Audio Plug-in Suite.
- Magnusson, T. (2006). Screen-Based Musical Interfaces as Semiotic Machines. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, pages 162–167.
- Marriott, N. (2024). tmux: A terminal multiplexer. Version 3.4.
- McCartney, J. (1996). SuperCollider, a New Real Time Synthesis Language. In *International Conference on Mathematics and Computing*.
- McLean, A. (2024). Tidal: Haskell package.
- McLean, A. e. a. (2026). TidalCycles official website. Accessed: 29.04.2026.
- Schwarz, D. (2012). The Sound Space as Musical Instrument: Playing Corpus-Based Concatenative Synthesis. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- The Glasgow Haskell Team (2024). GHCi: The interactive GHC. Part of the Glasgow Haskell Compiler, version 9.10.1.
- Tremblay, P. A., Roma, G., and Green, O. (2021). Enabling programmatic data mining as musicking: The fluid corpus manipulation toolkit. *Computer Music Journal*, 45(2):9–23.
- Zotter, F., Zaunschirm, M., Frank, M., and Kronlachner, M. (2017). A Beamformer to Play with Wall Reflections: The Icosahedral Loudspeaker. *Computer Music Journal*, 41(3):50–68.