

Distant Memories: A Co-Creative Approximation to Urban Soundscape Memory

Anna Gatzoura*

Social and Responsible Computing Group
Universitat Pompeu Fabra
Barcelona, Spain
anna.gatzoura@gmail.com

Abstract

Distant Memories is a human-AI interactive multichannel sound performance that experiments with the semantic relations of soundscape memories, as evoked through listening, and their evolution as a function of the evolution of time. It is based on an unsupervised machine learning algorithm that analyses a corpus of field recordings and transforms them to a latent space, to capture their characteristics and similarities, which then uses in real time to identify the sound segments that will be used in the performance, further modified and mixed by the artist. Shaped by both the creative decisions of the human and the actions of the algorithm, exploring the physical and latent spaces, this performance presents every time unique spatial soundscape compositions.

1 INTRODUCTION

Distant Memories is a multichannel sound performance that explores the idea of urban soundscape memory, through the use of Artificial Intelligence. From one side, it is inspired by the cognitive processes of human memory and the creation of sound objects, as a result of attention and repetitions over time, while from the other, it questions the semantic connotations evoked through listening, shaped by our expectations and past experiences.

Rather than trying to reconstruct, or generate based on prompted characteristics, specific soundscapes, this performance proposes a creative human-AI collaboration (Thelle and Pasquier, 2021), where the initial and final sonic spaces are defined by the aesthetic decisions and the performative actions of the human, who interacts with a computational model that operates on a latent space. This model, first analyses a corpus of field recordings to then, in real time, select the compositional elements to be used in each moment, influenced by time, similarity and randomness, asking the artist to perform on them, resulting every time in unique compositions, even based on the same sound corpus.

This work aims to contribute towards the exploration of small data and collaborative performances for non-symbolic music, especially soundscape composition (Truax, 2008), a yet less explored domain in the area of computational creativity.

2 CONCEPTUAL DESIGN

Gilles Deleuze describes the real as consisting of two registers: the “actual” and the “virtual”, with the later being an expression of the potential. However, for Deleuze the virtual goes beyond mere probability, it forms an expression of the non-actualised tendencies of real existents, existing together with the real (Cox, 2006). Similarly, sonic memory is considered as an internal, virtual space, since sound objects exist only as part of individual perceptions, balancing between the past and our expectations of the future, affected by, and in parallel, with the real.

John Cage, uses silence and indeterminacy to approximate the virtual, as an impersonal situation that permits him to approach the essence of the sonic flow and explore its creative potential, without imposing any personal perceptions or aiming for a predefined final piece (Kostelanetz, 2003). In Distant Memories the virtual space of sonic memory is formulated as a latent space, existing as a parallel description of the sonic attributes captured through listening and recording. The use of an unsupervised algorithm introduces a controlled indeterminacy in the performance, since the artist selects the sonic material, but it is the model that decides which parts will be used and when (Fiebrink and Caramiaux, 2016). Similarly to Cage, this approach removes personal intentions and sets the emphasis on the creation process, in this case embodied by the algorithm, rather than on a final piece.

The memory of a soundscape, and the initially created cognitive sound object associated to it, gets amplified or modified over time as a result of the occurrence of other, relevant or interfering, sound events (Snyder and Gregg, 2011). To capture this evolution, following the distinction of time into *pulsed time* (the time between divisions in music, defined as the tempo or rhythm) and *non-pulsed time* or *duration* (the time when the emphasis is on the sonic matter itself, without a compositional objective) by Max Neuhaus (Cox, 2006), this work introduces the term *distanced time* or *temporal distance*. Going beyond the expected duration of a sound, temporal distance defines a virtual dimension capturing the time needed to register a sound event and steadily observe the evolution of the related sound object (Solomos, 2018). This parameter is a key element of the performance, since it is used to control when and how the elements of the composition evolve.

3 TECHNICAL DESIGN

At the core of Distant Memories is a machine learning algorithm (implemented in Max using the FluCoMa library (Tremblay et al., 2021)) that for each performance first preprocesses a selected corpus of field recordings. It analyses the contained segments, extracts the Mel Frequency

*<https://anna-gatz.webnode.page>

Cepstral Coefficients (MFCCs) and then performs a dimensionality reduction, applying the UMAP algorithm (McInnes et al., 2018), transforming the sound descriptions to a two-dimensional latent space.

During the performance the temporal distance is set as the time between the moments on which the sound elements in the composition change. It takes random values between 20 and 40 seconds, empirically selected to permit listeners to enter an attention state, while also perceiving the sequence of changes. On those moments three latent descriptions (acting as virtual sonic points) are generated. The first point moves randomly in the latent space, while for the other two the x coordinate is the normalised difference between two consecutive temporal distance values and the y coordinates are functions of $1/e^x$ and $-\log x$, respectively, both having 30% probability to pick a random value. Those trajectories were chosen as a way to balance between the occurrence of random elements, permitting the creation of new sound objects, and the repetition of acoustic patterns amplifying prior experiences.

For each of those points the three “most relevant memories”, being the sound segments with the latent nearest neighbours, are retrieved using a KDTree structure. From the resulting nine sound sources, three are used without modifications, while to the rest are automatically applied effects related to key elements of the human sonic memory: repetition, echo and fading, formulated as functions of the actual temporal distance, properly scaled. (For instance, delay times range in 1-5% of its value, while pitch modifications depend on its ratio to the total performance time until that moment.)

Finally, the intensity, spatial distribution and mix in the three-dimensional space are controlled in real time by the artist (through the use of ambisonics), depending on the characteristics of the selected segments, the performance space and the desired emphasis. Starting from a circular, symmetric arrangement, with 40° separation, the nine sound sources get moved in space, with additional effects (such as filters and reverb) potentially applied to them, as a way to transition among recognisable soundscapes, blurry textures and intense moments.

Conceived for an eight speakers’ circular configuration, common in electroacoustic and acousmatic music, this performance generates unique and varied soundscapes in every moment and point in the space, asking listeners to choose their position as a way to experiment different listening experiences. An overview of the process is presented in Figure 1, where the left side refers to the actions of the artist and the right the algorithm. In addition, the upper part refers to the preprocessing phase, while the lower presents the actions taking place in real time.

A personal sound corpus, containing four hours of field recordings from Barcelona, is used. Given the semantic dimensions of soundscape memory that this work aims to explore, soundscapes of places of historic importance, places marked by recent social transformations and places where the urban architecture determines the sonic profiles, have been selected.

AUTHOR DECLARATIONS

This work was partially supported by the MSCA Ramon Llull-AIRA postdoctoral programme (grant agreement number: 101126667).

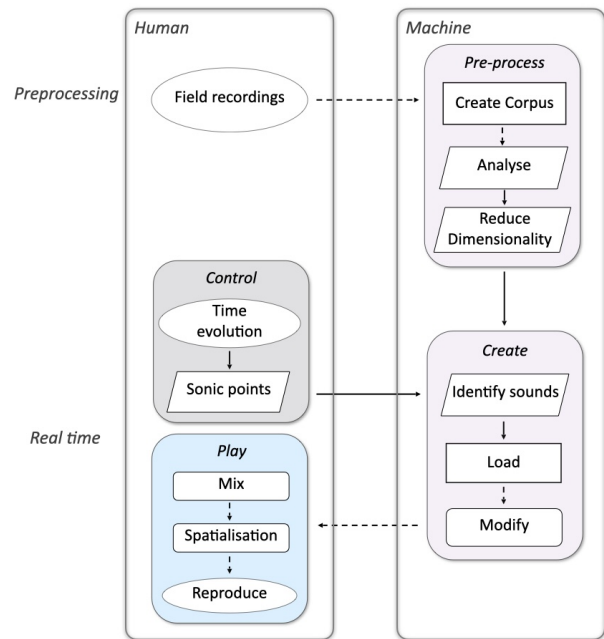


Figure 1: Process overview. Ovals represent inputs and outputs, rectangles processes, parallelograms computational processes and rounded rectangles the processes that modify sounds. The arrows highlight information flows while the dashed arrows refer to audio signals.

REFERENCES

- Cox, C. (2006). From music to sound: Being as time in the sonic arts. *Sonambiente Berlin*, pages 214–23.
- Fiebrink, R. and Caramiaux, B. (2016). The machine learning algorithm as creative musical tool. *arXiv preprint arXiv:1611.00379*.
- Kostelanetz, R. (2003). *Conversing with Cage*. Routledge.
- McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
- Snyder, J. S. and Gregg, M. K. (2011). Memory for sound, with an ear toward hearing in complex auditory scenes. *Attention, Perception, & Psychophysics*, 73(7):1993–2007.
- Solomos, M. (2018). From sound to sound space, sound environment, soundscape, sound milieu or ambiance. . . . *Paragraph*, 41(1):95–109.
- Thelle, N. J. W. and Pasquier, P. (2021). Spire Muse: A Virtual Musical Partner for Creative Brainstorming. *International Conference on New Interfaces for Musical Expression*. <https://nime.pubpub.org/pub/wcj8sjee>.
- 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.
- Truax, B. (2008). Soundscape composition as global music: electroacoustic music as soundscape. *Organised Sound*, 13(2):103–109.