Echobjects: A Context-Aware Sound Installation

Philippe Kocher
philippe.kocher@zhdk.ch
Zurich University of the Arts,
Switzerland

Daniel Bisig
daniel.bisig@zhdk.ch
Zurich University of the Arts,
Switzerland

Echobjects is a sound installation for public spaces consisting of five kinetic objects with integrated microphones and loudspeakers. It alternately records and plays back sounds and thus interacts with its acoustic environment. By storing some parts of the recorded sound, the installation forms a long-term memory so that the sound environment shapes the audio output throughout the whole duration of the exhibition. As a complement, the installation adds additional generated sounds whose qualities are based on a feature analysis of the stored sounds.

Keywords Sound Installation, Audio Feature Analysis, Interaction, Generative Art.
DOI 10.34626/2024_xcoax_027

Description

Echobjects is a sound installation designed for public spaces. It consists of five kinetic objects, each equipped with a loudspeaker and a microphone. The installation oscillates between two alternating phases: a listening phase and a playback phase. During the listening phase, it records the surrounding soundscape, and during the playback phase, it replays the recorded sounds. The installation is autonomous yet open to human interaction. Visitors are encouraged to engage with it by contributing sounds directly through one of the microphones.

By selecting and storing some of the recorded sounds, the installation forms a long-term memory that grows throughout the exhibition. During playback, audio fragments from the entire memory are randomly recombined. This means that the installation interacts with the sur-

rounding soundscape on different time levels: It reproduces not only the most recent inputs but also earlier ones, possibly from a long time ago. In addition, it contributes generative sounds whose synthesis is informed by the analysis data of the recorded sounds.

The installation *Echobjects* illustrates the dynamic adaptability of technical systems, showcasing how they store, retrieve, and manipulate information, generating temporal experiences that engage with past inputs and actions. It continues the authors' earlier works dealing with interactive or context-aware objects (Bisig and Kocher 2015; Bisig, Kocher and Neukom 2018; Kocher, Bisig and Inauen 2019; Unemi, Kocher and Bisig 2023).

Kinetic Objects

The installation consists of five identical kinetic objects, each featuring a 'head' mounted on a stand. Integrated into this head are a loudspeaker and a microphone, which are positioned in opposite directions to each other (fig. 1). Driven by a servo motor, the head rotates along its horizontal axis, allowing either the loudspeaker or microphone to face upward (fig. 2). This rotation occurs simultaneously in all five objects. The orientation of the head visibly indicates whether the installation is in listening or playback mode.

The arrangement of the five objects can vary, depending on available space and the exhibition venue's layout. They can be grouped close together (fig. 3) or arranged in a line along a wall (fig. 4), adapting to the surroundings and the specific situation at the exhibition venue.

Fig. 1. One kinetic object of the sound installation *Echobjects*.

Fig. 2. The kinetic object in listening (left) and playback (right) position.

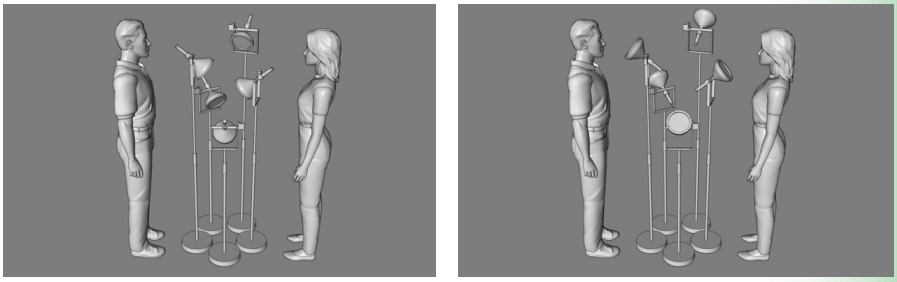

Fig. 3. The kinetic objects in group arrangement, listening (left) and playing back (right).

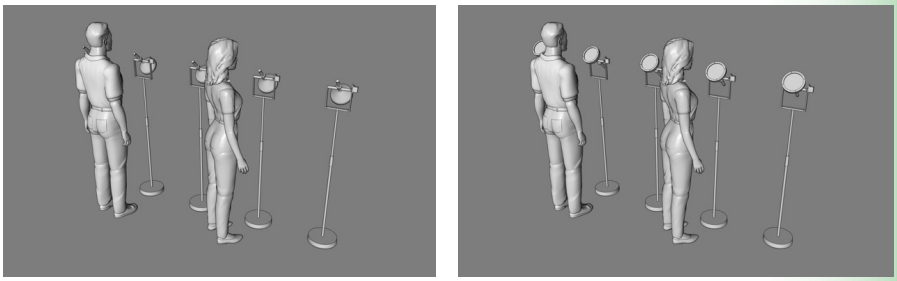

Fig. 4. The kinetic objects in line arrangement, listening (left) and playing back (right).

Audio Processing

During the listening phase, the microphones of the five kinetic objects capture ambient sound, which is then stored in the computer's memory. At the beginning of the playback phase, the most recently recorded sounds are immediately replayed. This echoing of the recorded sounds illustrates the principle of this sound installation, which consists of listening to the environment, memorising the sounds, and playing them back. Gradually, more and more older sounds from memory are re-inserted during the playback phase. The duration of the listening and playback phases varies depending on the acoustic activity of the environment; the more activity detected, the shorter these phases become. The listening phase lasts 10–60 seconds, and the playback phase takes two to five times longer than the listening phase.

The structure of the audio processing is shown in fig. 5. It consists of recording, storing, analysing and preparing sounds for playback. All sound analysis procedures are performed using the FluCoMa toolkit (Tremblay et al. 2022). The audio processing takes place in the following steps. After each listening phase, the computer slices the freshly captured audio based on spectral analysis, starting a new slice whenever there is a significant change in the analysis data (Foote 2000). Those slices whose analysis data prominently diverges from average analysis data are selected and copied to the long-term audio slice memory.

Simultaneously, the audio slice memory is constantly analysed. Older slices that are too similar to newer ones are removed. An automatic clustering algorithm categorises stored slices by spectral similarity. Concurrently, a generative synthesiser produces sounds informed by the analysis data of the recorded sounds. The synthesis process attempts to complement the recorded sounds by generating audio with differing characteristics, aiming to 'fill the gaps' in the analysis statistics of the recorded sounds.

During the playback phase, the audio concatenation engine selects slices from different audio memories to form a continuous audio stream. It uses analysis data, particularly the categorisation into clusters, to select the slices for playback and juxtapose different sound qualities. To ensure variety, the engine randomly chooses different playback modes,

such as mixing, looping, or freezing the slices. The resulting audio output is distributed to the five objects' speakers.

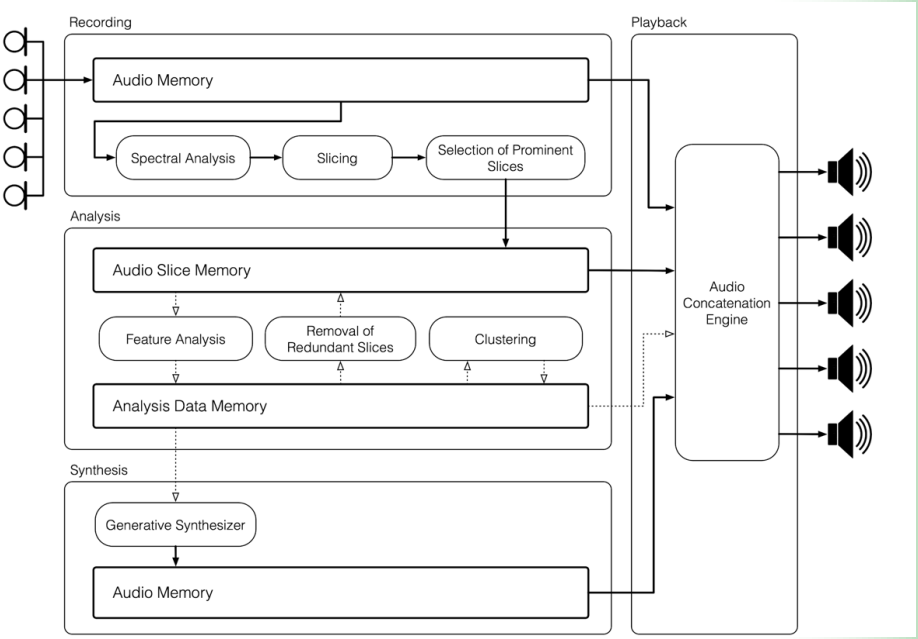

Fig. 5. Flowchart of the audio process (solid lines: audio signals, dashed lines: audio feature data, rectangular shapes: memory buffer, rounded shapes: processing operations).

References

Bisig, Daniel, and Philippe Kocher. 2015. "DRIFT – Virtual Sand in Augmented Space". In *Proceedings of the Generative Art Conference*, Venice, 51–64.

Bisig, Daniel, Philippe Kocher, and Martin Neukom. 2018. Watchers. Directionality as Metaphor for Interaction with a Generative Installation. In *Proceedings of the Generative Art Conference*, Verona, 94–105.

Foote, Jonathan. 2000. Automatic Audio Segmentation Using a Measure of Audio Novelty. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, New York, 452–455. https://doi.org/10.1109/ICME.2000.869637

Kocher, Philippe, Daniel Bisig, and David Inauen. 2019. Playing the Piano: Understanding Algorithmic Music Through Interaction. In *Proceedings of the Generative Art Conference*, Rome, 317–331.

Tremblay, Pierre Alexandre, Owen Green, Gerard Roma, James Bradbury, Ted Moore, Jacob Hart and Alex Harker. 2022. *The Fluid Corpus Manipulation Toolbox*. Zenodo. https://doi.org/10.5281/zenodo.6834643

Unemi, Tatsuo, Philippe Kocher, and Daniel Bisig. 2023. Greedy Agents and Interfering Humans. In *Proceedings of the Conference on Computation, Communication, Aesthetics & X*, Weimar, 363–367. https://doi.org/10.34626/xcoax.2023.11th.363