Flexible visualization of sound collections

Gerard Roma
CeReNeM
University of Huddersfield
groma@hud.ac.uk

ABSTRACT

This submission proposes a talk on content-based visualization of sound collections. Previous works have covered visualization of sounds using content-based descriptors, and visualization of collections using dimensionality reduction algorithms. Combining both provides a flexible framework for devising novel web audio interfaces.

1. INTRODUCTION

The availability of the digital sampler had a huge impact in the history of electronic music, and audio sampling is a basic practice of many musical genres. A sampler is in itself a computer with a specialized interface for music production. Hence, as computers became ubiquitous, many musicians switched to general purpose computers for sampling. Some advantages of general purpose computers are easier file management via the desktop metaphor, and a large screen. In this sense it is a bit strange that so many software samplers (e.g; the EXS24 sampler that ships with Logic) emulate vintage hardware. Alternative approaches for playing samples on general-purpose computers have evolved from the traditions of granular and concatenative sound synthesis [5]. Within this context, content-based descriptors developed in domains such as automatic speech recognition or music information retrieval can be leveraged for indexing, browsing and visualizing sounds.

2. MAPPING SOUND COLLECTIONS

This talk will summarize recent research by the author and collaborators on visualization of sound collections, with the aim of enabling novel interfaces for creative applications on general-purpose computers, phones and tablets. There are two main aspects to visualizing a sound collection. One is the location of each sound in a 2D (sometimes 3D) display. This has been demonstrated by several contributions to the web audio conference (WAC) [4, 1]. Typically, a dimensionality reduction algorithm is used to project audio feature statistics extracted from an audio segment to the display. In a recent work, we compared several algorithms and demonstrated their application to web-based multitouch interfaces [2]. An example is shown in Figure 1.

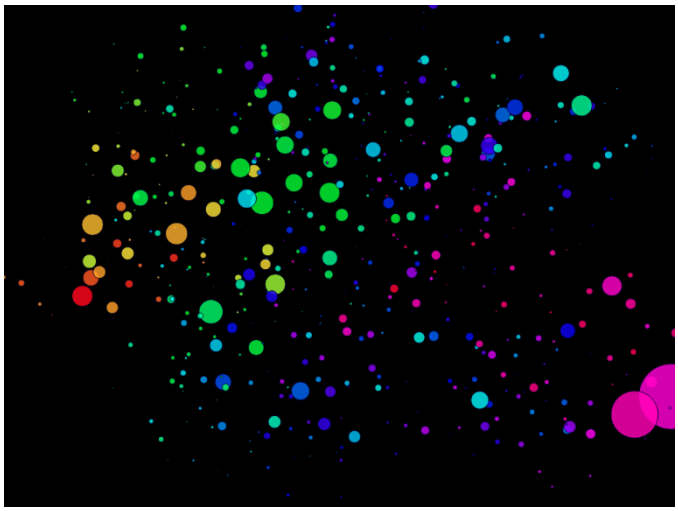

Figure 1: Sound collection scatterplot.

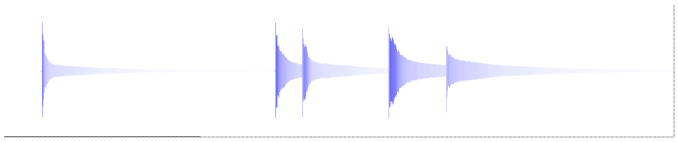

Figure 2: Waveform visualization.

Another issue is how to represent each sound. For this task, a library for audio visualization of descriptors was presented at WAC 2018 [3]. Multiple combinations of plot type and descriptors are possible. Figure 2 shows an example using a wave plot where luminance is controlled by RMS amplitude.

As a follow-up, this talk will describe further work on a flexible framework for combining dimensionality reduction and sound visualization. An example using spectral shape descriptors is shown in Figure 3.

Web Audio Conference WAC-2019, December 4–6, 2019, Trondheim, Norway.

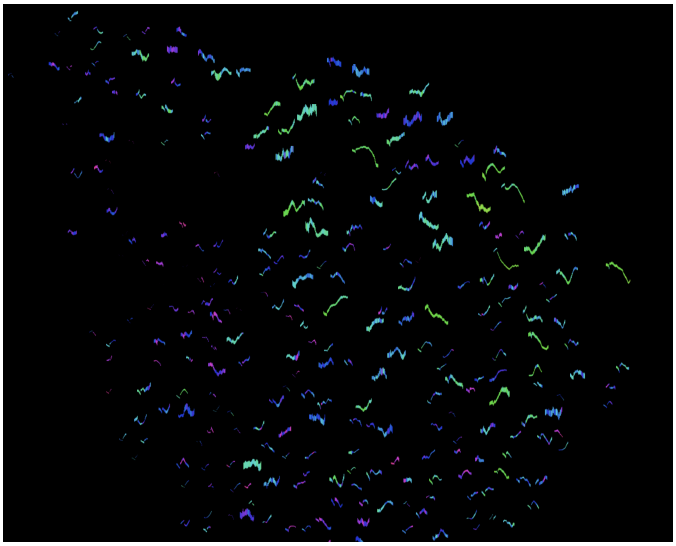

Figure 3: An example sound map.

3. ACKNOWLEDGMENTS

This research was part of the Fluid Corpus Manipulation project (FluCoMa), which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 725899).

4. REFERENCES

[1] F. Font Corbera. Freesound explorer: make music while discovering freesound! 2017.

[2] G. Roma, O. Green, and P. Tremblay. Adaptive mapping of sound collections for data-driven musical interfaces. In *NIME – The International Conference on New Interfaces for Musical Expression*, Proceedings of the conference on New Interface for Musical Expression (NIME), 3 2019.

[3] G. Roma, O. Green, A. Xambó, and P. Tremblay. A javascript library for flexible visualization of audio descriptors. In *Proceedings of the 4th Web Audio Conference, WAC-2018*, 9 2018.

[4] G. Roma and X. Serra. Music performance by discovering community loops. In *Proceedings of the Web Audio Conference (WAC)*, 2015.

[5] D. Schwarz. Concatenative Sound Synthesis: The Early Years. *Journal of New Music Research*, 35(1):3–22, 2006. cote interne IRCAM: Schwarz06b.