Reverbalize: A Crowdsourced Reverberation Controller

Prem Seetharaman
Northwestern University
EECS Department
prem@u.northwestern.edu

Bryan Pardo
Northwestern University
EECS Department
pardo@northwestern.edu

ABSTRACT

One of the most commonly-used audio production tools is the reverberator. Reverberators apply subtle or large echo effects to sound and are typically used in commercial audio recordings. Current reverberator interfaces are often complex and hard-to-understand. In this work, we describe Reverbalize, a novel and easy-to-use interface for a reverberator. Reverbalize uses crowdsourced data to create a 2-dimensional map of adjectives used to describe reverberation (e.g. "underwater'). Adjacent words describe similar reverberation effects. Word size correlates with agreement for the definition of a word. To use Reverbalize, the user simply clicks on the descriptive adjective that best describes the desired effect. The tool modifies the sound accordingly. A text search box also lets the user type in the desired word.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: User-centered design; H.5.5 [**Sound and Music Computing**]: Signal analysis, synthesis, and processing

Keywords

Human computation; audio engineering; interfaces

1. INTRODUCTION

In the physical world, reverberation is created by the reflections of sound off of the solid surfaces (e.g. walls) of an enclosed space. These reflections result in a decaying series of echoes that modify the sound's loudness, timbre, and perceived spatial characteristics. In the digital realm, reverberation (reverb) can be simulated using networks of delays and gains to create a decaying series of echoes. The tool that does this is called a reverberator.

The reverberator is one of the most commonly-used audio production tools. Reverberation can make the audio sound as if it were recorded in a different acoustic environment (e.g. change a recording made in a sound booth to sound like one recorded in a cathedral), and are used to increase the pleasantness of the sound.

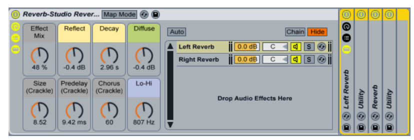

Figure 1: A standard parametric reverberator.

Figure 1 shows the interface of a typical professional quality parametric reverberator. While the "effect mix" and the "decay" dials may make intuitive sense, other dials like "predelay" have no obvious meaning to the average person. This is because the controls are conceptualized in terms of the underlying processes used to create the reverberation effect, as opposed to perceptually relevant terms. Musicians and producers often conceptualize the effect of reverb using words like "deep" or "spacious". Much of audio production involves bridging the gap between artistic goals and the manipulation of the available audio tools.

2. ENGINEERING CONCEPT

Our concept is to make an interface for reverberation based on natural language adjectives a layman would use to describe reverberation. This interface should be easy to navigate and also show the relationship between words (e.g. is "deep" more like "spacious" or "warm"?). The goal is to provide clear affordances that allow natural navigation through the set of options. We do this by building the interface from a crowd-sourced data set of natural language descriptors provided by the SocialReverb project.

Using data from SocialReverb, we create a 2-D map (see Section 5) that visualizes the relationships between descriptive words provided by the crowd. By associating these words with the parameters used to manipulate a reverberator, we create a new kind of reverberation controller, shown in Figure 2. Users explore the space of reverberation options by clicking and dragging around the map, using the words as a guide. When a word is highlighted, the associated reverb is applied to the sound. The slider to the right lets the user filter out words that don't show strong agreement across the population about the effect. If the user would rather quickly jump to a concept, they can type a word in the search box.

MM'14, November 3–7, 2014, Orlando, Florida, USA.
ACM 978-1-4503-3063-3/14/11.
http://dx.doi.org/10.1145/2647868.2654876 .

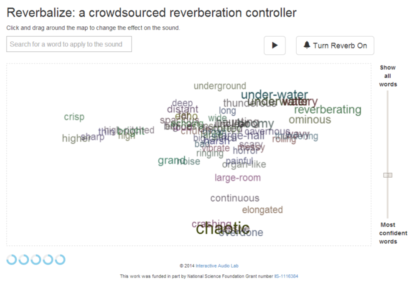

Figure 2: The proposed interface.

3. PROPOSED DEMONSTRATION

The demonstration is implemented using the Web Audio API and runs in a Chrome browser. The demonstration will consist of a laptop computer connected to our web server. All audio processing is implemented locally in the browser, so there are no latency issues. The user will be able to try the concept-map controller shown in Figure 2 and listen to the effect over a pair of provided high-quality headphones. We will also provide a standard interface for comparison.

4. NOVELTY AND RELATED WORK

This is the first controller we are aware of that builds an interface for an audio production tool using a word map built from crowdsourced data as the interaction paradigm.

Commercial reverberation tools attempt to simplify the interface with a list of presets (saved settings given a name like “warm church long”). Preset names are chosen by experts and may not have meaning to laymen. Further, a simple list gives no clue about the relationship between presets. Some commercial tools (e.g. iZotope Ozone) have hundreds of presets, making the preset list nearly as difficult to navigate as the original interface. Our map-based interface using labels learned from the crowd addresses these issues.

Convolution reverberation tools (e.g. Altiverb) provide lists of locations (e.g. ”Carnegie Hall”) where the impulse response of a space was taken. While users can imagine the effect a location may have on a sound, many words used to describe reverberation (e.g. “warm”) are not particular location names. Our map includes these more natural descriptors.

There is some prior work in building audio controllers using verbal descriptors. A 2-dimensional audio equalizer was implemented in [4]. The interface lets users explore different equalizations in a 2-D map. However, it uses only four labels (“bright”, “warm”, “tinny”, “dark”) provided by the authors, rather than hundreds of words crowdsourced from a large group of users. Also, the interface is for equalization, rather than reverberation. [2] creates a similar interface and employs more descriptors of higher complexity (“warm hi-cut”, “less sizzly”). However, it also implements an equalization tool, rather than a reverberation tool, and the labels are chosen by the authors. Amateur musicians considered some of these terms too technical. Our system circumvents this issue by crowdsourcing terms used by laymen.

5. CROWDSOURCING A CONCEPT MAP

The concept map is built using data collected by the SocialReverb (socialreverb.org) project. To date, the project has collected 3388 descriptive words from 658 users describing 256 different reverberation effects (reverbs). Participants are recruited on Amazon Mechanical Turk. When teaching SocialReverb the participant listens once to an audio sample, and then again to the same sample with some reverb applied. They are then asked to type in words of their own choice that describe the effect of the reverb. They are then shown a list of words provided by other users, and asked to select the words they agree also describe this reverb. Each time a word is used to describe a reverb, we add the reverb to the set of reverbs associated with that word.

Each reverb is represented by a vector of five measurements: reverberation time, echo density, clarity, central time, and spectral centroid ([3]). Once we find all reverbs associated with a word (e.g. “warm”), we use the averages of the five measure as the word definition. This gives a five-dimensional definition that can be used to set a reverberator to modify a sound to make it “warm” (or some other word).

In the data, some words (e.g. “warm”) are applied broadly to a wide variety of reverberation effects and others (“boomy”) are more specific and are associated with a narrow range of reverbs. We can measure the consistency of a word definition by measuring the variance of the associated reverbs along each of the measured dimensions. If variance is low, the reverbs associated with the word have similar qualities. If the variance is high, many different sorts of reverb are associated with the word.

The map contains only words which have been used to label at least 15 instances of reverberation. For each such word, we calculate a definition and measure its consistency. This places each word in a 5-D space. We project this 5-D space onto two dimensions, using multidimensional scaling [1]. On the resulting 2-D map, word position is the center of the distribution of reverberation settings associated with the word. Each word is scaled using its variance. Larger words indicate greater consistency of definition.

6. VIDEO LINK

A video describing and presenting Reverbalize can be found at `http://prem.seeth.org/static/reverbalize.mp4`.

7. ACKNOWLEDGEMENTS

This work is supported by National Science Foundation Grant 1116384.

References

[1] Ingwer Borg and Patrick JF Groenen. *Modern multidimensional scaling: Theory and applications.* Springer, 2005.

[2] S Mecklenburg and J Loviscach. “subjEQt: Controlling an equalizer through subjective terms”. In: *CHI’06 Extended Abstracts on Human Factors in Computing Systems* (2006).

[3] Zafar Rafii and Bryan Pardo. “A Digital Reverberator Controlled through Measures of the Reverberation”. In: *Northwestern University, Technical Report* (2009).

[4] AT Sabin and B Pardo. “2DEQ: an intuitive audio equalizer”. In: *Proceedings of the 7th ACM conference on Creativity and Cognition* (2009).