Matching Artificial Reverb Settings to Unknown Room Recordings:
A Recommendation System for Reverb Plugins

Nils Peters 1,2, Jaeyeong Choi 1, Howard Lei 1
1 ICSI - International Computer Science Institute, Berkeley, CA, USA
2 CNMAT - Center for New Music and Audio Technologies, UC Berkeley, CA, USA
[nils|jaeyeong|hlei]@icsi.berkeley.edu

Abstract
For creating artificial room impressions, numerous reverb settings exist and are often controllable by many parameters. To efficiently create a desired room impression, the sound engineer must be familiar with all the available reverb setting possibilities.

Although plugins are usually equipped with many factory presets for exploring available reverb options, it is a time-consuming learning process to find the ideal reverb settings to create the desired room impression, especially if various reverberation plugins are available.

For creating a desired room impression based on a reference audio sample, we present a method to automatically determine the best matching reverb preset across different reverb plugins. Our method uses a supervised machine-learning approach and can dramatically reduce the time spent on the reverb selection process.

Our Solution
Often a user has an ideal reverb in mind that he is trying to create by means of an artificial reverb, e.g., a reverb heard in another recording.

Based on a provided reverberant audio file, we propose a system that automatically selects the best matching reverb setting from all of their available reverb options.

Prior Art
To simplify the workflow with artificial reverbs, a few control strategies have been proposed in the past:

Perceptual Parameter Layer
Rather than technical parameter, the user controls perceptual parameter (e.g., room presence, intimacy) that are linked to underlying technical parameter. InCam’s SpatialListen pioneered this approach.

A Personalized Reverb UI
On an individual level, [1] described a method to train a personal interface on top of an artificial reverb. A linear regression model is used to map the reverberated audio samples in less than 3 minutes.

A Unified Reverb Interface
The authors of [2] propose a unified API based on Open Sound Control (OSC) to control different artificial reverbs from one single user interface.

References

Problem: Reverbs Are Complicated

Many parameters
Parameter differ across reverbs
Complicated GUIs
No time for exploring

The Matching Algorithm
Our recommendation system is derived from a GMM-based system historically used in speaker recognition [3] and uses Mel-Frequency Cepstral Coefficient (MFCC) audio features. From each audio file, MFCC C0-C19 along with deltas and double-deltas are extracted, 90 dimensions in total. The window length is 25 ms and frame intervals are 10 ms.

The open-source ALIZE toolkit [4] is used for GMM training, factor analysis implementation, and likelihood ratio computation for the reference audio. Factor analysis seeks to obtain a low-dimensional subspace representing the undesired variations of RIR-processed audio [5], with the RIRs coming from the same reverb preset. The undesired variations would be subsequently removed from each preset-related GMM and reference audio. Factor analysis is computationally intensive and can be done offline. The resulting model is stored and will be recalled when a reverb recommendation is requested.

A Prototype
We created a prototype of the recommendation system for 97 monaural reverb presets. All of those reverb presets are available as IRs in the Web [6]. The system was trained with 1688 reverberant audio files created with those reverb units from 40 anechoic speech files (20 seconds).

Testing Results
Audio tracks of 18 videos from the Flickr video database were extracted. Half of these videos were tagged with Living room, the other half were tagged with Church acoustics.

The videos have a maximum length of 30 sec. and all were captured with consumer video cameras or mobile phones. In some of the footage, either the audio is clipping or an active gain compression notably affects the dynamic.

<table>
<thead>
<tr>
<th>Flicker Video</th>
<th>5 best matching reverbs</th>
<th>5 least matching reverbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Plate Space</td>
<td>1. Plate Space</td>
<td>1. Plate Space</td>
</tr>
<tr>
<td>3. Large Hall</td>
<td>3. Large Hall</td>
<td>3. Large Hall</td>
</tr>
<tr>
<td>8. Studio 1</td>
<td>8. Studio 1</td>
<td>8. Studio 1</td>
</tr>
</tbody>
</table>

Example results of the recommendation system