Evaluation of AI Reverberation on Guitar

Manni Chen
School of Creative Media, City University of Hong Kong
manni.chen@my.cityu.edu.hk

PerMagnus Lindborg
School of Creative Media, City University of Hong Kong
pm.lindborg@cityu.edu.hk

Abstract

Artificial intelligence (AI) technology has been applied to create a range of different sound effects. However, perceptual evaluations of such applications have not yet been fully explored in research studies. This paper focuses on reverberation and we report results from objective measurements and perceptual listening tests in two stages (N=10, N=32). Participants rated samples of solo guitar phrases, taken from a musical context, and processed by either an AI software or by a professional sound engineer. The analysis aims to predict the salient modifications made to reverberation by AI and humans. This study contributes to our understanding of AI in studio music production. Future work will expand to other audio effects, such as compression and equalization.

Introduction

As AI technology is applied in sound effects plug-ins, machines start to undertake part of sound work. However, existing publications that refer to evaluation on these emerging intelligent tools are missing. Therefore, our research project aims to compare sound effects made by AI and human sound engineers. In this paper, we focus on reverberation. Previous studies have explored perception of reverberation effects e.g. amount of reverberation and early decay time (De Man et al. 2017). However, there appears to be no previous study into the perceptual effects of AI-based reverberation.

In this connection, in order to deepen the understanding of AI audio effects, evaluations should be conducted based on musical features and perceptual listening tests.

Testing Task

In the present study, we concentrate on reverberation and guitars. We selected six samples from our own recordings and a free online mixing resource (Senior 2011) regarding electric and acoustic guitars, in each case containing a musical phrase with a duration of 15 to 20 seconds. From each sample we generated three versions: one with reverberation added by the *smart:reverb* processor (Sonible 2020); one with reverberation added by a trained sound engineer; and one without any reverb, i.e. the original. This constitutes the set of 18 stimuli used in the study. We directly obtain the AI-version, automatically generated by the Sonible plugin. The human engineers were required to manipulate the samples using only reverberation plugins that they were familiar with through commonly available software (such as Pro Tools). They were not allowed to edit the samples or use any other audio effects. Followingly, the 18 stimuli (6 samples x 3 versions) were explored through objective measurements and perceptual listening evaluations.

Objective Evaluation

Objective measurements were carried out using the MIR Toolbox (Lartillot et al. 2008), extracting several features including RMS, brightness, centroid, mirolloff, decaytime, and mel frequency cepstral coefficients (MFCC). For an explanation of these features, see the reference above. Objective characteristics extracted from the audio signals were compared, and differences in parameters between musical versions were observed. Through a selection process, the features entered a regression model as predictors.

Subjective Evaluation

First, we conducted an initial online experiment (N=10) using QuestionPro presented in Chinese language. Participants evaluated the perceived reverberation of the 18 stimuli, presented in randomised order in a repeated-measures design. Each was evaluated on 15 semantic rating scales (cf. Pulkki and Karjalainen 2014, 353), labelled 混响量 (*hùnxiǎngliàng*, "reverb amount"), 温暖感 (*wēnnuǎngǎn*, "warmth"), 明亮感 (*míngliànggǎn*, "brightness"), 粗糙感 (*cū cāogǎn*, "roughness"), 空气感 (*kōngqìgǎn*, "airiness"), 甜美感 (*tiánměigǎn*, "sweetness"), 纵深感 (*zòngshēngǎn*, "depth"), 尖锐感 (*jiānruìgǎn*, "harshness"), 喜爱度 (xǐàidù, "likeability"), 清晰度 (*qīngxīdù*, "clarity"), 混响尾巴长度 (*hùnxiǎngwěibāchángdù*, "length of reverb tail"), 湿度 (shīdù, "wetness"), 距离感 (*jùlígǎn*, "distance"), 厚度感 (*hòudùgǎn*, "thickness"), and 扩散感 (*kuòsàngǎn*, "diffusion").

We reduced the 15 scales into two orthogonal dimensions using PCA (Principal Component Analysis), spanned by four constructs, labelled 清晰度 (*qīngxīdù*, "clarity"), 湿度 (shīdù, "wetness"), 扩散感 (*kuòsàngǎn*, "diffusion") and 温暖感 (*wēnnuǎngǎn*, "warmth"). We then conducted an online experiment (N=32) as described above, this time using only these four semantic scales. The subjective measurements (i.e. perceptual experiment tasks) will be added as predictors to an enlarged regression model. Statistical analysis is ongoing at the time of abstract submission and more results will be presented at the conference.

References

De Man, Brecht, Kirk McNally, and Joshua D. Reiss. 2017. "Perceptual evaluation and analysis of reverberation in multitrack music production." *Journal of the Audio Engineering Society,* 509

Lartillot, Olivier, Toiviainen, Petri, & Eerola, Tuomas. 2008. "A Matlab Toolbox for Music Information Retrieval." In *Data Analysis, Machine Learning and Applications*, 261–268. Springer.

Pulkki, Ville, and Karjalainen, Matti. 2014. *Communication acoustics: an introduction to speech, audio and psychoacoustics*. New York: Wiley.

Senior, Mike. 2011. *Mixing secrets for the small studio*. Burlington, MA: Focal Press, https://www.cambridge-mt.com/ms/mtk/

Sonible. 2020. smart:reverb, https://www.sonible.com/

Biographies

Manni Chen is a currently a PhD student in the School of Creative Media, City University of Hong Kong. She was a sound engineer working in the audio industry before, now her research interests focus on AI sound production.

PerMagnus Lindborg is an Associate Professor in the School of Creative Media, City University of Hong Kong. His research interests in sound perception and spatial design are anchored in computer music. He founded the *Soundislands Festival* (2013–17) and serves on the Board of International Computer Music Association, as Music Coordinator and Regional Director for Asia-Oceania. He is director of the Data Art for Climate Action Conferene (DACA) which will take place in January 2022.