

Intelligent Audio Mixing Using Deep Learning

Marco Martínez¹ and Joshua D. Reiss²

¹Centre for Digital Music, Queen Mary University of London, UK, m.a.martinezramirez@qmul.ac.uk

²Centre for Digital Music, Queen Mary University of London, UK

Abstract— We propose a research trajectory in the field of deep learning applied to music production systems such as mixing, mastering, sound design and sound synthesis.

I. BACKGROUND

Multi-track audio mixing is an essential part of music production and with recent advances in machine learning techniques such as deep learning, it is of great importance to carry out research on the applications of these methods in the field of automatic mixing.

Taking into account that audio mixing is a highly cross-adaptive transformation, which, apart from artistic considerations, essentially tries to solve the problem of unmasking by manipulating different audio characteristics such as the dynamics, spatial information, timbre or pitch.

The task of automatic mixing has been researched in present years, where it was approached as an application of adaptive audio effects [1]. This was explored by the development of an intelligent system capable of performing automatic mixing [2], and an extensive understanding of the expert knowledge involved in order to develop a system of this type [3]. Machine learning has been applied to solve punctual tasks such as EQ [4] and dynamic range compression [5], but not to a whole system.

II. RESEARCH

In this research, we explore how can we train a deep neural network to perform audio mixing as a content-based transformation without using standard mixing devices (*e.g.* dynamic range compressors, equalizers, limiters, etc.), and investigating whether an intelligent system is capable of learning the intrinsic characteristics of this transformation and applies them.

In addition, we examine how such a system can perform a goal-oriented mixing task, and if we can guide it with features extracted from a mixdown. Similarly, we research how can a deep learning system use expert knowledge from the field of mixing engineering in order to improve results in an audio mixing task, and also whether we can integrate user interaction as a final fine-tuning of the mixing process.

¹M. Martínez is a PhD researcher with the Centre for Digital Music in the School of Electronic Engineering and Computer Science at Queen Mary University of London. (author e-mail: m.a.martinezramirez@qmul.ac.uk)

²Dr. Joshua D. Reiss (member IEEE, AES) is a Senior Lecturer with the Centre for Digital Music in the School of Electronic Engineering and Computer Science at Queen Mary University of London. (e-mail: joshua.reiss@qmul.ac.uk)

The project will research different deep learning approaches with application to audio signals through the use of state-of-the-art training and architecture models. A very important aspect of the research will fall into this stage, since we will explore how we can understand what the system has actually learned, and if we can use this information to enhance the previous results of the system.

Through this research will also explore different ways to shape and synthesize sounds, because through training and understanding of deep learning networks, several types of test signals could be applied and new and interesting results could be discovered.

One of the main difficulties when applying deep learning to music production is the difficulty in collecting the correct data set. It is for this reason that *The Open Multitrack Testbed* will be of significant importance, with around hundreds of multi-tracks, stems and mixes, we will use this data in the process of training, developing and validation of the system. Correspondingly, we will perform an objective and subjective evaluation of the performance of the system through a series of listening experiments.

The research will focus on the post-production stage and this is where the main impact will occur, as there is a clear need for an intelligent system capable of carrying out the tedious technical tasks and also to be assist the musician or audio engineer in the creative process of audio mixing. In the same way, the research results could be of use in the field of music production, live music performance, generative music, algorithmic composition models, and intelligent music interfaces.

REFERENCES

- [1] V. Verfaillie, U. Zölzer, "Adaptive Digital Audio Effects (A-DAFx): A New Class of Sound Transformations.," *IEEE Transactions on Audio, Speech, and Language.*, no. 14 (5), pp. 1817 -1831, 2006.
- [2] E. Gonzalez, *Advanced Automatic Mixing Tools for Music.*, PhD. Queen Mary University of London, 2010.
- [3] P. Pestana, *Automatic Mixing Systems using Adaptive Audio Effects*, PhD. Universidade Católica Portuguesa, 2013.
- [4] R. Reed, "A Perceptual Assistant to do Sound Equalization.," in *Intelligent User Interfaces 5th Conference*, 2000.
- [5] Z. Man, B. De Man, P. Pestana, D. Black, J. Reiss, "Intelligent Multitrack Dynamic Range Compression," *Journal of the Audio Engineering Society*, vol. 63, 2015.