Student	Isabelle Kernhof
Examiner	Prof. Dr. Guido Schuster
Subject Area	Sensor, Actuator and Communication Systems

Towards a Vinyl Sound Emulation Utilizing Machine Learning Methods

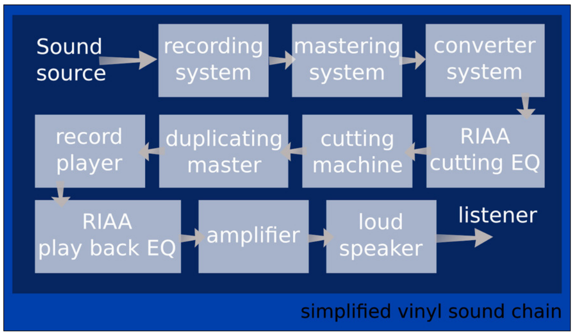

Figure 1: Abstraction of the vinyl signal chain from the sound source to the listener.
Own presentment

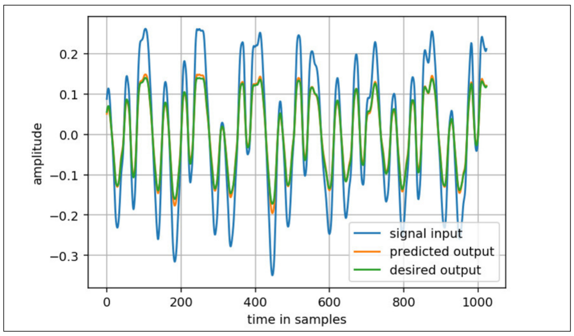

Figure 2: Example of learning non-linearities with a simple recurrent network architecture.
Own presentment

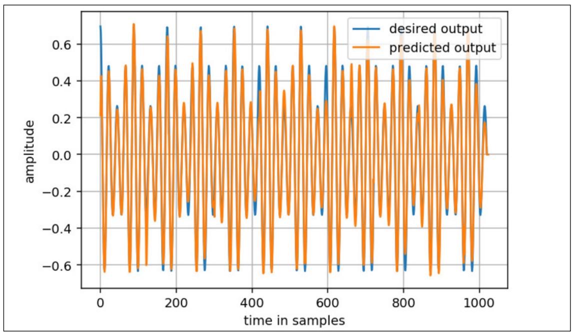

Figure 3: Example of learning time-variant effects with a bidirectional LSTM network architecture.
Own presentment

Introduction: Last years music revenue report of the USA indicated a growth rate of the streaming sector of more than 25%. With a revenue growth rate of 13% for the vinyl sector there is an obvious market demand not only for streaming, but also for authentic analogue sound.
Generating a link between these both demands is the main motivation of this project: Creating authentic analogue vinyl sound for digital audio streaming applications. To obtain this goal, this work utilizes state-of-the-art machine learning methods to emulate the analogue vinyl sound.

Approach: With the main objective being the emulation of the vinyl sound, an End-to-End supervised machine learning approach is pursued. The block diagramm in figure 1 of the vinyl signal chain illustrates the complexity of the underlying system. To be able to cope with this complexity a detailed analysis of the learning data (test vinly recordings) is performed to identify characteristics of the vinyl sound. Furthermore the results of the data analysis facilitate the generation of synthetic vinyl data beyond the recorded test signals – offering the possibility to shift the data generation into the digital domain and evaluate the vinyl emulation performance for single effects of the vinyl sound individually.

Result: Based on the available vinyl recordings, several characteristics could be identified and digital effect processors for individual vinyl effects are developed. A procedure to align the recorded vinyl data represeting the output with the digital input data is presented.
A range of supervised machine learning methods are utilized, tested and compared to learn characteristics of single vinyl sound effects: (1) Deep Neural Networks (DNNs), (2) DNNs extended with Convolutional Neural Networks (CNNs) and (3) DNNs extended with Long Short-Term Memory (LSTM).
For musical data modified with synthetic vinyl effects, the LSTM models performed the best. Two outputs of such systems are presented in figure 2 and 3. Figure 2 shows the network output signal after 20 epochs of training with a solo accordion piece for the vinyl distortion effect. Figure 3 visualizes the network output after 200 epochs of training for a simplified vinyl rumble effect. In general, it could be shown that it is possible to learn non-linear and time-varying audio effects comparable to the vinyl sound utilizing supervised machine learning.
The combination of the individual vinyl characteristics into one interdependent system and training it with a large dataset from real vinyl recordings will be the scope of further research - bringing us one step closer to real-time vinyl emulation applications used for streaming audio.