Fachgebiet Audiokommunikation

Prof. Dr. Stefan Weinzierl

Magister-/Master-/Diplomarbeit

Comparison of supervised, unsupervised and semi-supervised learning of feature spaces for music similarity computation

Music similarity computation is an advanced and complex topic, which is linked to music science, cognition and computer science (Volk et al. 2016). Especially in recent years, with the rising need for more efficient recommendation systems, the research for more comprehensive music similarity systems in the field of Music Information Retrieval (MIR) has gained momentum. In the context of deep learning, several approaches have dealt with the question of whether feature engineering or feature learning have more potential in pinpointing similarities and differences between tracks of various genres, and also rhythmic and spectral properties (Nam et al. 2018).
An especially interesting study about similarity of world music tracks uses feature learning to achieve good results (Panteli et al. 2016). However, the study is only performed on small datasets with relatively noisy labels about the musical qualities (as is usually the case). Therefore, there is a need for more research on different kinds of methods for similarity computation, namely supervised, semi-supervised and unsupervised methods, since the two first ones might help to enrich or support the labels which are already available.
This thesis aims to develop such methods in a deep learning context. Working with variational autoencoders (VAE), convolutional neural networks (CNNs) and mixtures of both, the goal is to investigate which combinations can provide improved similarity ratings in comparison to baseline (unsupervised) methods such as t-SNE and k-Means. The tests shall be performed on established datasets in the MIR community as well as on real-world data, with the goal of pinpointing differences and eventually developing a scalable, real-time system. The results shall be evaluated with the help of listening tests where the perceived similarity between specific tracks will be correlated with the results of the machine learning algorithms in order to assess their validity.

Literatur

Anja Volk, Elaine Chew, Elizabeth Hellmuth Margulis & Christina Anagnostopoulou (2016): Music Similarity: Concepts, Cognition and Computation, Journal of New Music Research, 45:3, 207-209.
Nam, J., Choi, K., Lee, J., Chou, S. Y., & Yang, Y. H. (2018): Deep learning for audio-based music classification and tagging: Teaching computers to distinguish rock from Bach. *IEEE Signal Processing Magazine*, *36*(1), 41-51.
Panteli, M., Benetos, E., & Dixon, S. (2016, August): Learning a feature space for similarity in world music. ISMIR.

Voraussetzungen

Gute Kenntnisse der Signalverarbeitung und des Maschinellen Lernens (inkl. Neurone Netzwerke); Python.

Betreuung

Roman Gebhardt, rmngebhardt@gmail.com
Athanasios Lykartsis, EN107, athanasios.lykartsis@tu-berlin.de