Pattern discovery and music similarity with compositional hierarchical model

Matevž Pesek[1], Aleš Leonardis[1,2], Matija Marolt[1]
[1]University of Ljubljana, Faculty of Computer and Information Science
[1]Centre for Computational Neuroscience and Cognitive Robotics, University of Birmingham

{matevz.pesek,ales.leonardis,matija.marolt}@fri.uni-lj.si

Abstract

We present a compositional hierarchical model for symbolic music representations (SymCHM). The SymCHM is a deep architecture model with a transparent multi-layer structure. The model can be used for pattern discovery, as well as music music similarity. It can learn a set of representative repeated patterns of individual works or larger corpora in an unsupervised manner, relying on statistics of pattern occurrences. A learned model contains representations of patterns on different layers, from the less complex on lower layers to the longer and more complex on higher layers.

Its transparent nature enables insight into the found patterns, while the inference process with hallucination and inhibition mechanisms enables the search for pattern variations. The SymCHM is general in the way it represents the learned patterns and can be employed for various tasks, including pattern finding, similarity and novelty estimation and composer identification.

For evaluation, we focused on the Discovery of repeated patterns & sections task proposed by Mirex community. We evaluated the model in the Mirex 2015 evaluation campaign and on the JKU PDD dataset.

SymCHM

The SymCHM [1] is built around the premise that the repetitive nature of patterns can be captured by observing statistics of occurrences of their sub-patterns, thus providing a hierarchical representation of a symbolic music representation. Due to the possible overlap and mutual inclusion of patterns, a hierarchical approach is very suitable for this task, as it can concurrently provide multiple pattern hypotheses on several levels of complexity.

The SymCHM is an extension of the compositional hierarchical model (CHM), we previously introduced for a variety of audio-related tasks [2]. The structureand learning mechanisms of the model remained similar, while its input waschanged so that it encodes a set of note events N, where each note event is defined by its onset time N_o and pitch N_p.

Input layer

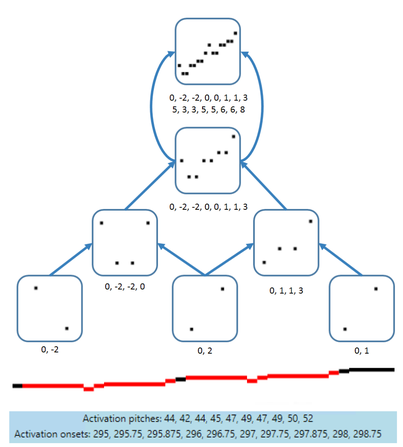
The input layer of the model is a symbolic representation of the music signal, consisting of a set of pitches, each defined by an onset and an offset. It contains a set of atomic parts (pitches), which are activated (is present in the signal) at any MIDI location and any given time. Similar to the original model, where any time-frequency representation can be used for the input layer, any two-dimensional representation - in this case pitch-time representation - can be used as an input to the SymCHM.

Model Structure

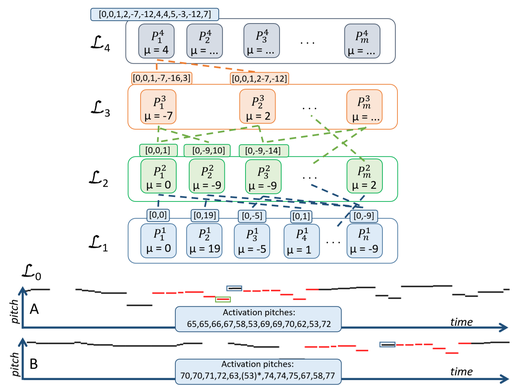
The SymCHM is used in a two-stage manner. During the first, the build stage, the model is developed layer-by-layer. By composing atomic L_0 parts, the model first produces compositions of two L_1 events (pitches). We apply the same approach to each subsequent layer, by composing the parts from the previous layer, where each part represents a sequence of melodic patterns.

To retain the compositions which cover the most information in the input layer, a statistical approach is employed. Based on the compositions' occurrence, the learning process retains the compositions which are **more frequently activated**.

The parts are observed as *patterns* whereas the activations are treated as *pattern occurences*.



Pattern Discovery



The input layer corresponds to a symbolic music representation (a sequence of pitches). Parts on higher layers are compositions of lower-layer parts - depicted as connections between parts, parameter μ is given in semitones.

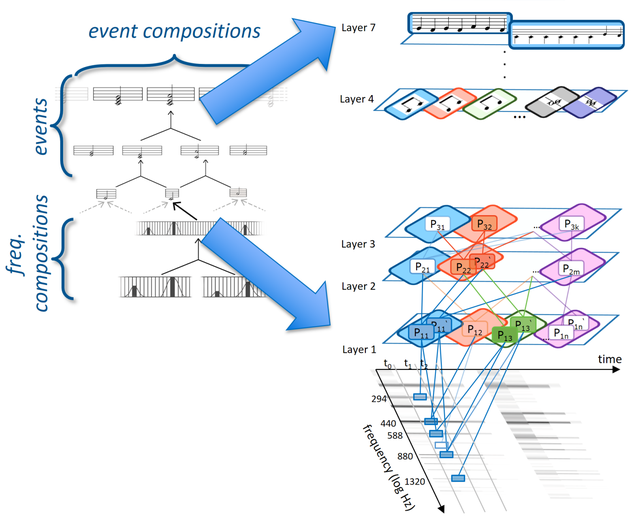
The part structure is displayed above each part in the Figure, represented by a sequence of pitch values relative to the central part (e.g. [0,0,1] for part P^1_2). A part may be contained in several compositions, e.g. P^1_n is part of compositions P^2_2 and P^2_3. The entire structure is transparent, thus we can observe the entire subtree of part P^4_1. A part activates, when (part of) the pattern it represents is found in the input.

As an example, P^4_1 activates twice (input A and B), however there are differences in the found patterns. Pattern A is positioned 5 semitones higher than B, pattern B is missing one event (green rectangle), and the pitch of one event (blue rectangle) differs between the two patterns.

Applications

The CHM and SymCHM can be used for a variety of MIR tasks. The CHM was previously applied to the audio transcription task.The SymCHM was applied to the Repeated Patterns and Sections task at Mirex 2015, and an extension of the model SymCHMMerge was also applied to Mirex 2016. Due to the general applicability, the model can be used for other tasks, such as inter-opus pattern discovery with applications pattern finding, similarity and novelty estimation and composer identification, using symbolic data as an input.

The model can also be applied to other domains, such as chemistry and bioinformatics. The current work in progress includes applications in chemistry by employing the SymCHM for pattern discovery in mass spectrometry data.



1. Matevž Pesek, Aleš Leonardis, and Matija Marolt. SymCHM: A Compositional Hierarchical Model for Pattern Discovery in Symbolic Music Representations, Mirex 2015
2. Matevž Pesek, Aleš Leonardis, and Matija Marolt. A compositional hierarchical model for music information retrieval. In Proceedings of the International Conference on Music Information Retrieval (ISMIR), 131–136, Taipei, 2014.

Naložba v vašo prihodnost
OPERACIJO DELNO FINANCIRA EVROPSKA UNIJA
Evropski socialni sklad
Univerza v Ljubljani
Fakulteta za računalništvo in informatiko