SymCHMMerge - Hypothesis Refinement for Pattern Discovery with a Compositional Hierarchical Model

Matevž Pesek[1] and Aleš Leonardis[1,2] Matija Marolt[1]

[1] University of Ljubljana,
Faculty of Computer and Information Science
{matevz.pesek,matija.marolt}@fri.uni-lj.si,
[2] University of Birmingham
School of Computer Science
Centre for Computational Neuroscience and Cognitive Robotics
a.leonardis@cs.bham.ac.uk

Abstract. This paper presents our compositional hierarchical model for pattern discovery in symbolic music SymCHM and its extension SymCHMMerge. The model uses unsupervised learning to build a hierarchical structure reflecting the patterns present in the input music representation. It can generate several competing hypotheses on the found patterns, so a pattern selection approach must be used to extract the most salient patterns. The extended model refines the identified pattern candidates and merges competing hypotheses, thus improving the model's performance.

Keywords: pattern discovery, compositional modelling, music information retrieval, symbolic music representations

1 Introduction

The discovery of repeated patterns is a well-researched task in different domains, including computer vision, bioinformatics and music information retrieval (MIR) [5,6,9]. In MIR, the discovery of repeated patterns aims to find repetitions, which represent one of the more significant aspects of a music piece [1,6].

Repeated pattern discovery is different to standard pattern matching [2]. While pattern matching finds the place of a searched pattern within a dataset and usually has a clear quantitative relation between the query and a match, discovery of repeated patterns finds locations of multiple similar sequences of data in the dataset. In the pattern discovery task, the patterns may partially overlap or be subsets of another pattern by definition. A variety of approaches have been proposed for pattern discovery in music in the past years, e.g. [3,4]. Among the most recent, Velarde and Meredith [9] expand on the previously introduced approach to melodic segmentation [10] using Haar wavelets for filtering. Lartillot [5] introduced the PatMinr algorithm which uses an incremental one-pass approach to identify pattern occurrences.

In the paper, we present a compositional hierarchical model for pattern discovery in symbolic music (SymCHM) and its extension SymCHMMerge, which uses a hypothesis refinement algorithm to improve the model's performance.

2 Symbolic Compositional Hierarchical Model

The SymCHM [8] 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 [7]. The structure and learning mechanisms of the model remained similar, while its input was changed so that it encodes a set of note events $N \in \mathcal{S}$, where each note event is defined by its onset time N_o and pitch N_p.

The model contains several layers of parts (compositions), where each part represents a melodic sequence of events. A part is composed of two or more parts from lower layers, with parameter μ governing the relative distance (in semitones) of parts in a composition. Parts on the first layer thus represent simple intervals, while parts on higher layers encode longer melodic patterns. A four layer model is depicted in Figure 1. The model is learned in an unsupervised manner from a set of melodic lines. Learning is based on statistics of pattern occurrences, by incrementaly adding parts that cover additional patterns (thus increasing the cumulative coverage) in the training set layer by layer. More details on the model's structure and learning is described in [7, 8].

2.1 Inference

A learned model captures (relatively encoded) repetitive patterns in the training set. Locating patterns in a new piece is performed through *inference* of the model, which calculates *activations* of parts in the input. *Activation* of a part represents a concrete occurrence of the pattern it represents in the input. An activation has three components: *location* and *onset time*, which map the relative pattern onto a specific set of pitches within the input sequence of events (thus making it absolute), and *magnitude*, representing its strength. A part can simultaneously have multiple activations at different locations, which indicates multiple occurrences of the represented pattern in the input representation. Activations may be exact or approximate, where in the latter case additional mechanisms of hallucination and inhibition enable the model to find patterns with deletions, changes or insertions.

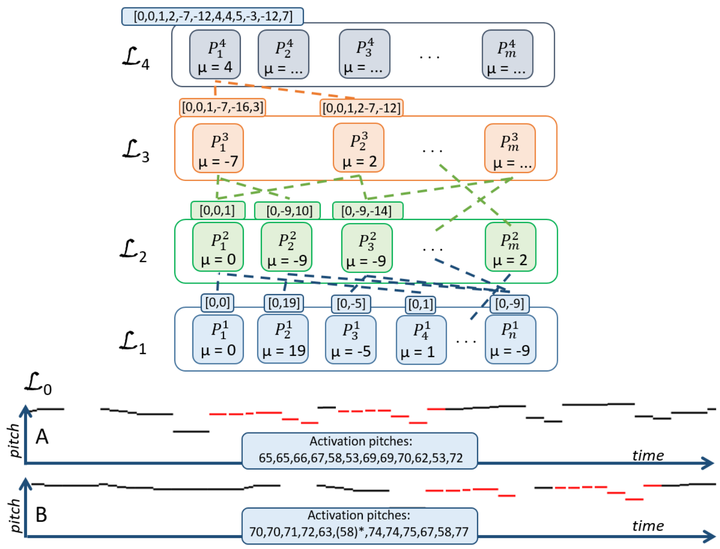

Fig. 1. The symbolic compositional hierarchical model displaying two P_1^4 activations at different locations (denoted A and B). The input layer corresponds to a symbolic music representation (a sequence of pitches). Parts on higher layers are compositions of lower-layer parts with a set relative distance encoded by parameter μ. The sequence of pitches that each part represents is shown above each part, encoded relative to the first pitch (e.g. [0,0,1] for part P_1^2 represents a sequence three pitches, the second is repeated and the third raised by a semitone). A part may be contained in several compositions, e.g. P_1^2 is part of compositions P_1^3 and P_2^3. The entire structure is transparent, thus we can observe the entire subtree of part P_1^4.

2.2 SymCHM and SymCHMMerge

Melodic patterns in a given piece are discovered by observing activations of the learned model's parts, where each activation of a part is interpreted as a *repetition* of the pattern encoded by the part.

To produce a set of patterns in a given input piece, the SymCHM model simply outputs all patterns of sufficient complexity (defined as a combination of occurrence frequency and size), as encoded by activations of parts starting from layer M up to the highest layer. As parts on higher layers are compositions of parts on lower layers, all parts which are subparts of a composition on a higher layer are excluded to avoid redundancy. The selection of parts can be formulated as:

$$\bigcup_{l=M}^{N} \{P_i^l : P_i^l \in \mathcal{L}_l \wedge P_i^l \notin P_j^{l+1}\} \tag{1}$$

Since in our model parts are learned in an unsupervised manner, several learned parts may represent similar or overlapping patterns in the input (e.g.

shifted by a few notes in time). For the pattern finding task, outputting such redundant patterns is undesirable.

In SymCHMMerge we introduce a new pattern selection algorithm that improves selection by merging redundant patterns. Let $\pi(A_L(P_i^n))$ represent a pattern occurrence in the input encoded in the model by P_i^n and discovered during inference by its activation A_L. Π_i^n represents the set of all such pattern occurrences discovered by activations of the part:

$$\Pi_i^n = \bigcup_{A_L} \{\pi(A_L(P_i^n))\}. \tag{2}$$

Two sequences of pattern occurrences a_i and a_j produced by parts P_i^n and P_j^m are taken to be redundant, if they overlap significantly. We express this by calculating the Jaccard similarity coefficient and compare it to a threshold τ_R:

$$a_i = \pi(A_L(P_i^n)), a_j = \pi(A_L(P_j^m)), \quad J(a_i, a_j) = \frac{|a_i \cap a_j|}{|a_i \cup a_j|} > \tau_R. \tag{3}$$

We aim to merge redundant pattern occurrences of two parts, if they frequently produce overlapping patterns. We thus calculate the ratio of such patterns produced by two parts as:

$$\frac{1}{|\Pi_i^n| + |\Pi_j^m|} \sum_{a_i \in \Pi_i^n} \sum_{a_j \in \Pi_j^m} |J(a_i, a_j) > \tau_R|. \tag{4}$$

If the ratio exceeds a threshold τ_M, all redundant pattern occurrences of the two parts are merged.

For evaluation, the thresholds τ_R and τ_M were both set to 0.5, meaning that pattern occurrences produced by two parts had to share at least 50% of events in the input layer and appear together in at least 50% of cases, to be merged.

3 Evaluation

We evaluated the model for discovery of repeated themes and sections in symbolic monophonic music pieces. Since we are searching for patterns within a given piece (and not across an entire corpus), a 7-layer model was built independently for each piece and inferred on the same piece. For corpus analysis, the model can be trained on multiple pieces. Results on the JKU PDD dataset are presented in Table 1.

Compared to the basic pattern selection approach, the improved algorithm achieved significantly better results (Friedman's test: $\chi^2 = 7.2, p < .01$). It yields better results for the F_{1est} establishment measure which indicates an improvement of the algorithm's ability to discover at least one occurrence of a pattern, tolerating for time shift and transposition [1], due to the increased recall of the established patterns. On the other hand, occurrence measures $F_{1occ(c=0.75)}$ and $F_{1occ(c=0.5)}$, that evaluate the algorithm's ability to find all occurrences of the established patterns, have dropped for almost 5%. We attribute this drop to a higher number of established patterns, for which the occurrence measure is calculated. Finally, the absolute precision, recall and F scores significantly increased due to pattern merging.

Table 1. Comparison of SymCHM and SymCHMMerge for discovery of repeated patterns on the JKU PDD dataset. The establishment measure F_{1est} evaluates the algorithm's ability to find at least one occurrence of each pattern shifted in time and pitch. F_{occ} evaluates the extent of the algorithm's ability to find all occurrences, with inexactness tolerance threshold (c factor). P, R and F_1 estimate the match of found patterns with the ground truth in an exact manner.

	P_{est}	R_{est}	F_{1est}	$P_{occ(c=.75)}$	$R_{occ(c=.75)}$	$F_{1occ(c=.75)}$
SymCHM	67.92	45.36	51.01	93.90	82.72	86.85
SymCHMMerge	67.96	50.67	56.97	88.61	75.66	80.02
	$P_{occ(c=.5)}$	$R_{occ(c=.5)}$	$F_{1occ(c=.5)}$	P	R	F_1
SymCHM	78.53	72.99	75.41	25.00	13.89	17.18
SymCHMMerge	83.23	68.86	73.88	35.83	20.56	25.63

References

1. Tom Collins. 2015:Discovery of Repeated Themes & Sections - MIREX Wiki, 2015.
2. Tom Collins, Jeremy Thurlow, Robin Laney, Alistair Willis, and Paul H. Garthwaite. A Comparative Evaluation of Algorithms for Discovering Translational Patterns in Baroque Keyboard Works. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 3–8, Utrecht, 2010.
3. Darrell Conklin and Christina Anagnostopoulou. Representation and Discovery of Multiple Viewpoint Patterns. In *Proceedings of the 2001 International Computer Music Conference*, pages 479–485, Cuba, 2001.
4. Darrell Conklin and Mathieu Bergeron. Feature Set Patterns in Music. *Computer Music Journal*, 32(1):60–70, mar 2008.
5. Olivier Lartillot. Submission to MIREX Discovery of Repeated Themes and Sections. In *10th Annual Music Information Retrieval eXchange (MIREX'14)*, Taipei.
6. David Meredith, Kjell Lemstrom, and Geraint A. Wiggins. Algorithms for discovering repeated patterns in multidimensional representations of polyphonic music. *Journal of New Music Research*, 31(4):321–345, dec 2002.
7. 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)*, pages 131–136, Taipei, 2014.
8. Matevž Pesek, Urša Medvešek, Aleš Leonardis, and Matija Marolt. SymCHM: a compositional hierarchical model for pattern discovery in symbolic music representations. In *11th Annual Music Information Retrieval eXchange (MIREX'15)*, pages 1–3, Malaga, 2015.
9. Gissel Velarde and David Meredith. Submission to MIREX Discovery of Repeated Themes and Sections. In *10th Annual Music Information Retrieval eXchange (MIREX'14)*, pages 1–3, Taipei, 2014.
10. Gissel Velarde, Tillman Weyde, and David Meredith. An approach to melodic segmentation and classification based on filtering with the Haar-wavelet. *Journal of New Music Research*, 42(4):325–345, dec 2013.