MIREX 2014 ENTRY: MUSIC SEGMENTATION TECHNIQUES AND GREEDY PATH FINDER ALGORITHM TO DISCOVER MUSICAL PATTERNS

Oriol Nieto and Morwaread M. Farbood
Music and Audio Research Lab
New York University
{oriol, mfarbood}@nyu.edu

ABSTRACT

This extended abstract describes the pattern discovery submission to MIREX 2014 of an algorithm that uses music segmentation (or music structure analysis) techniques and a refined greedy method in order to identify the repetitive musical patterns of a given music piece, either represented symbolically or with an actual audio file. We obtain a harmonic representation of the input and compute the Self Similarity Matrix in order to extract the most prominent paths (or repetitions) that appear on it. The audio features are pre-filtered using standard music segmentation techniques, and the paths are extracted using a greedy algorithm that allows us to obtain results, in the audio domain, comparable to other techniques that operate at a symblic level exclusively. This algorithm is meant to run on both audio and symbolic representations, and extract either monophonic or polyphonic patterns. Its implementation is open source and available for public download [1].

1. INTRODUCTION

Pattern discovery is a reasonably novel task in MIREX that aims at automatically discovering the repeated patterns (sections, themes, or motives) that occur in a specific piece. Traditionally, in the literature, this task has been approached from a symbolic perspective (c.f. [1–3, 6]), even though algorithms working at an audio level have also been presented (e.g. [8,9]). This problem can be additionally divided based on the type of patterns to be retrieved: monophonic or polyphonic. Last year we submitted an algorithm that treated this task from a music segmentation technique, allowing us to work both in the symbolic and audio domains, identifying monophonic and symbolic patterns [5].

Building upon this idea, this year we submitted an improved version of that algorithm (**NF1**) that adds a novel greedy path finder method to it, still being able to work with audio and symbolic representations, and retrieving monophonic and polyphonic patterns. This submitted method is described in [7], and we slightly modified it for MIREX in order to compete in the four different flavors of this task.

2. AUDIO FEATURES

As specified in the MIREX website [2], the input to the symbolic task is a CSV file that contains the BPM information and onset times for each note with their respective MIDI pitches. For the audio version, a one-channel audio file sampled at 11025Hz and synthesized from a midi file (also called dead pan audio) is used as input, along with the BPM data. In this case, a spectrogram is computed with a Blackman-Harris window of 44dB, whose length is determined based on the BPM of the piece, and its hop size is half the window length. A chromagram is then computed either from the symbolic representation or from the spectrogram, and the rest of the algorithm remains identical for both the symbolic and audio flavors of the task.

3. IDENTIFYING THE PATTERNS

Given a chromagram, we take the key-transposition invariant self-similarity matrix as described in [4]. We further process the matrix by applying a diagonal filter in order to enhance the paths or repetitions. Once we have this representation, we extract these paths by applying a greedy algorithm presented in [7]. This algorithm traverses half of the diagonals of the matrix and assigns a score that in order to determine the relevance of the path. If this score is higher than a given threshold, the path is stored, and later grouped into occurrences of the same pattern depending on its location within the matrix.

The results of this process on the JKU Patterns Development dataset are shown in Table 1. The evaluation metrics in the Table are the standard metrics for this task, and they are explained in the MIREX page for this task. As a comparison, the results for the previous version of the algorithm submitted to MIREX in 2013 are shown at the bottom of the table. The evaluations of this year's submission are significantly better in all the metrics and versions

[1] https://github.com/urinieto/MotivesExtractor/tree/MIREX2014

[2] http://www.music-ir.org/mirex/wiki/2014:Discovery_of_Repeated_Themes_%26_Sections

Version	P_{est}	R_{est}	F_{est}	$P_{\text{O}(.75)}$	$R_{\text{O}(.75)}$	$F_{\text{O}(.75)}$	P_3	R_3	F_3	$P_{\text{O}(.5)}$	$R_{\text{O}(.5)}$	$F_{\text{O}(.5)}$	Time (s)
S + M	57.00	69.06	**61.58**	73.98	44.31	**53.98**	40.86	53.16	**45.44**	62.05	50.06	**53.64**	**235.55**
S + P	58.97	58.83	**53.57**	58.19	44.67	48.94	45.83	47.51	**40.84**	53.67	51.91	52.55	**188.20**
A + M	53.51	62.50	**57.00**	54.02	36.81	**43.03**	30.85	37.56	**33.17**	44.00	37.27	**39.59**	685.11
A + P	56.76	51.26	**50.36**	37.61	27.60	**31.79**	36.00	35.26	**32.48**	44.96	34.42	**38.35**	269.38
2013 NF1 [5]	22.78	32.25	26.57	0.00	0.00	0.00	18.62	32.84	23.48	28.12	22.22	24.81	590.12
2013 NF2 [5]	54.77	53.41	48.06	65.50	51.23	**57.09**	43.51	47.51	38.78	60.30	50.45	**53.38**	475.79
2013 NF3 [5]	21.01	25.94	22.55	13.33	3.33	5.33	11.02	16.92	13.05	33.30	12.66	18.20	**236.36**
2013 NF4 [5]	40.83	46.43	41.43	32.08	21.24	24.87	30.43	31.92	28.23	26.60	20.94	23.18	**196.29**

Table 1. Results of the four different versions of our method submitted to MIREX on the JKU Patterns Development Dataset, averaged across pieces. In the Version column, S stands for Symbolic, A for Audio, M for Monophonic, and P for Polyphonic. The bottom of the table contains the results for a previous version of the algorithm submitted in 2013.

except for the occurrence measures for the symbolic, polyphonic version. Even though the running time in the audio version has increased with respect last year, the symbolic version has become faster to compute. For a more detailed information of this algorithm we refer the reader to [7].

4. REFERENCES

[1] Tom Collins. *Improved methods for pattern discovery in music , with applications in automated stylistic composition*. PhD thesis, The Open University, 2011.

[2] Berit Janssen, W Bas De Haas, Anja Volk, and Peter Van Kranenburg. Discovering Repeated Patterns in Music: State of Knowledge, Challenges, Perspectives. In *Proc. of the 10th International Symposium on Computer Music Multidisciplinary Research*, Marseille, France, 2013.

[3] Olivier Lartillot. Multi-Dimensional motivic pattern extraction founded on adaptive redundancy filtering. *Journal of New Music Research*, 34(4):375–393, December 2005.

[4] Meinard Müller and Michael Clausen. Transposition-Invariant Self-Similarity Matrices. In *Proc. of the 8th International Conference on Music Information Retrieval*, pages 47–50, Vienna, Austria, 2007.

[5] Oriol Nieto and Morwaread Farbood. MIREX 2013: Discovering Musical Patterns Using Audio Structural Segmentation Techniques. In *Music Information Retrieval Evaluation eXchange*, Curitiba, Brazil, 2013.

[6] Oriol Nieto and Morwaread M. Farbood. Perceptual Evaluation of Automatically Extracted Musical Motives. In *Proc. of the 12th International Conference on Music Perception and Cognition*, pages 723–727, Thessaloniki, Greece, 2012.

[7] Oriol Nieto and Morwaread M. Farbood. Identifying Polyphonic Patterns From Audio Recordings Using Music Segmentation Techniques. In *Proc. of the 15th International Society for Music Information Retrieval Conference*, Taipei, Taiwan, 2014.

[8] Oriol Nieto, Eric J. Humphrey, and Juan Pablo Bello. Compressing Audio Recordings into Music Summaries. In *Proc. of the 13th International Society for Music Information Retrieval Conference*, Porto, Portugal, 2012.

[9] Ron Weiss and Juan Pablo Bello. Unsupervised Discovery of Temporal Structure in Music. *IEEE Journal of Selected Topics in Signal Processing*, 5(6):1240–1251, 2011.