SYNTHESIZING FLEXIBLE, COMPOSITE HIERARCHICAL STRUCTURE FROM MUSIC DATASETS

Ilana Shapiro

UC San Diego

ilshapiro@ucsd.edu

ABSTRACT

Music is an innately hierarchical system, comprising semantic levels such as formal structure segmentation, disjoint motif repetition, functional harmonic contour, and melodic contour that are informed by music theory. Historically, researchers in the music information retrieval community have focused on developing analyses for single levels in this hierarchy. Existing research has addressed neither (1) how to combine arbitrarily many levels of structure analyses into a single unified model and (2) how to extract a representative such structure from a corpus of music, rather than a single piece. In this work, we propose a novel data structure called the *semantic temporal graph* that captures both the semantic (i.e. music theoretic) relationships between levels of the hierarchy, as well as the temporal relationships between the elements of adjacent-level analyses. Furthermore, given a corpus of such graphs derived from individual pieces, we introduce a method rooted in stochastic optimization to derive a representative graph encoding the music dataset's overall structure.

1. INTRODUCTION

Music is both composed and comprehended within a framework of intrinsic hierarchical structure. Automatic identification of musical structure, also known as *music structure analysis*, continues to be a major interest to both musicologists and the MIR community. Research thus far has focused on the automatic contiguous segmentation (both flat and hierarchical) of musical form [1–10], the detection of disjoint repeating motifs [11, 12], and more recently on harmonic [13], functional harmonic [14], and melodic [15–17] contour extraction. All of these tasks have been proposed in annual competitions of the Music Information Retrieval eXchange (MIREX) [18–20], which provides a standard format for their output.

To our knowledge, all existing research addresses a single aspect of the compositional hierarchy, such as motif extraction, or melodic contour. There is no notion of how to unify

1.1 Our Contributions

2. RELATED WORK

3. ANALYSIS FORMATS

3.1 MIREX Standard Formats

3.2 Parsing

4. ABSTRACT REPRESENTATION

4.1 Semantic Temporal Graph

5. SYNTHESIS

6. CONCLUSIONS AND FUTURE WORK

7. REFERENCES

[1] O. Nieto, "Discovering structure in music: Automatic approaches and perceptual evaluations," Ph.D. dissertation, New York University, 2015. [Online]. Available: https://www.proquest.com/openview/09f67403121bcbc7d2ee431985bf0568/1

[2] C. Hernandez-Olivan and J. R. Beltran, "Musicaiz: A python library for symbolic music generation, analysis and visualization," *SoftwareX*, vol. 22, p. 101365, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352711023000614

[3] C. Wang, J. Hsu, and S. Dubnov, "Music pattern discovery with variable markov oracle: A unified approach to symbolic and audio representations," in *Proceedings of the 16th International Society for Music Information Retrieval Conference, ISMIR 2015, Málaga, Spain, October 26-30, 2015*, M. Müller and F. Wiering, Eds., 2015, pp. 176–182. [Online]. Available: http://ismir2015.uma.es/articles/78_Paper.pdf

[4] O. Nieto and T. Jehan, "Convex non-negative matrix factorization for automatic music structure identification," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2013, Vancouver, BC, Canada, May 26-31, 2013*. IEEE, 2013, pp. 236–240. [Online]. Available: https://doi.org/10.1109/ICASSP.2013.6637644

[5] O. Nieto and J. P. Bello, "Music segment similarity using 2d-fourier magnitude coefficients," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 664–668.

[6] B. McFee, O. Nieto, M. M. Farbood, and J. P. Bello, "Evaluating hierarchical structure in music annotations," *Frontiers in Psychology*, vol. 8, 2017. [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01337

[7] B. McFee and D. P. W. Ellis, "Learning to segment songs with ordinal linear discriminant analysis," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2014, Florence, Italy, May 4-9, 2014*. IEEE, 2014, pp. 5197–5201. [Online]. Available: https://doi.org/10.1109/ICASSP.2014.6854594

[8] B. McFee and D. Ellis, "Analyzing song structure with spectral clustering," in *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27-31, 2014*, H. Wang, Y. Yang, and J. H. Lee, Eds., 2014, pp. 405–410. [Online]. Available: http://www.terasoft.com.tw/conf/ismir2014/proceedings/T073_319_Paper.pdf

[9] C. Finkensiep, M. Haeberle, F. Eisenbrand, M. Neuwirth, and M. Rohrmeier, "Repetition-structure inference with formal prototypes," in *Proceedings of the 24th International Society for Music Information Retrieval Conference, ISMIR 2023, Milan, Italy, November 5-9, 2023*, A. Sarti, F. Antonacci, M. Sandler, P. Bestagini, S. Dixon, B. Liang, G. Richard, and J. Pauwels, Eds., 2023, pp. 383–390. [Online]. Available: https://doi.org/10.5281/zenodo.10265305

[10] K. A. Sidorov, A. Jones, and A. D. Marshall, "Music analysis as a smallest grammar problem," in *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27-31, 2014*, H. Wang, Y. Yang, and J. H. Lee, Eds., 2014, pp. 301–306. [Online]. Available: http://www.terasoft.com.tw/conf/ismir2014/proceedings/T055_226_Paper.pdf

[11] Y. Hsiao, T. Hung, T. Chen, and L. Su, "Bps-motif: A dataset for repeated pattern discovery of polyphonic symbolic music," in *Proceedings of the 24th International Society for Music Information Retrieval Conference, ISMIR 2023, Milan, Italy, November 5-9, 2023*, A. Sarti, F. Antonacci, M. Sandler, P. Bestagini, S. Dixon, B. Liang, G. Richard, and J. Pauwels, Eds., 2023, pp. 281–288. [Online]. Available: https://doi.org/10.5281/zenodo.10265277

[12] C.-i. Wang and G. J. Mysore, "Structural segmentation with the variable markov oracle and boundary adjustment," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 291–295.

[13] T. Chen and L. Su, "Harmony transformer: Incorporating chord segmentation into harmony recognition," in *Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019, Delft, The Netherlands, November 4-8, 2019*, A. Flexer, G. Peeters, J. Urbano, and A. Volk, Eds., 2019, pp. 259–267. [Online]. Available: http://archives.ismir.net/ismir2019/paper/000030.pdf

[14] ——, "Functional harmony recognition of symbolic music data with multi-task recurrent neural networks," in *Proceedings of the 19th International Society for Music Information Retrieval Conference, ISMIR 2018, Paris, France, September 23-27, 2018*, E. Gómez, X. Hu, E. Humphrey, and E. Benetos, Eds., 2018, pp. 90–97. [Online]. Available: http://ismir2018.ircam.fr/doc/pdfs/178_Paper.pdf

[15] J. Salamon, E. Gomez, D. P. W. Ellis, and G. Richard, "Melody extraction from polyphonic music signals: Approaches, applications, and challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 118–134, 2014.

[16] K. Kosta, W. T. Lu, G. Medeot, and P. Chanquion, "A deep learning method for melody extraction from a polyphonic symbolic music representation," in *Proceedings of the 23rd International Society for Music Information Retrieval Conference, ISMIR 2022, Bengaluru, India, December 4-8, 2022*, P. Rao, H. A. Murthy, A. Srinivasamurthy, R. M. Bittner, R. C. Repetto, M. Goto, X. Serra, and M. Miron, Eds., 2022, pp. 757–763. [Online]. Available: https://archives.ismir.net/ismir2022/paper/000091.pdf

[17] Y.-H. Chou, I.-C. Chen, C.-J. Chang, J. Ching, and Y.-H. Yang, "Midibert-piano: Large-scale pre-training for symbolic music understanding," 2021.

[18] M. McCallum. (2017) Structural Segmentation. Music Information Retrieval Evaluation eXchange (MIREX). [Online]. Available: https://www.music-ir.org/mirex/wiki/2017:Structural_Segmentation

[19] T. Collins. (2017) Discovery of Repeated Themes & Sections. Music Information Retrieval Evaluation eXchange (MIREX). [Online]. Available: https://www.music-ir.org/mirex/wiki/2017:Discovery_of_Repeated_Themes_\%26_Sections

[20] Sutashu. (2010) Harmonic Analysis. Music Information Retrieval Evaluation eXchange (MIREX). [Online]. Available: https://www.music-ir.org/mirex/wiki/2010:Harmonic_Analysis