Towards composer-specific harmony theories: an unsupervised harmonic analysis model with neural hierarchical HSMM

Yui Uehara*, Satoshi Tojo†

Harmony theory plays a central role in music theory, and harmonic analysis is one of the most important techniques in music education and practice. However, the current standard of harmony theory is biased toward the classical era, and the rich features of other eras are not widely inherited or utilized. Furthermore, harmony theory has rules that do not necessarily correspond to reality, while also containing many ambiguous expressions and lacking quantitative and mathematical justification.

This study concerns computational music analysis, aims to realize harmonic analysis that reflects the individual characteristics of composers. To this end, we propose an unsupervised harmonic analysis model, which acquires the individual idioms of composers from data and is not simply a tool that performs harmonic analysis in place of humans. In this sense, this study differs from existing machine-based automatic harmonic analysis, which collects a large amount of human analysis[1] and imitates it.

To achieve practical analysis even without supervision, we propose a hierarchical semi-Markov model, which is based on our prior work[3] and improved to take into account the hierarchical nature of keys and chords. Since the duration of a key can be extremely long, implementing this hierarchy naively would result in high computational costs. However, we applied the idea of previous research[4] to control the transition of higher layers using an activation flag, thereby realizing the modulation without significantly increasing computational costs.

We apply the proposed model to J. S. Bach's "Little Organ Book" and compare it with human analysis. The results demonstrate that the proposed model achieves performance nearly comparable to an existing supervised learning model (AugmentedNet[2]). Thus, our model can perform practical harmonic analysis without supervision, which suggests the possibility of expanding the analysis to other composers and works to obtain composer-specific harmonic idioms.

Acknowledgments

This work was in part supported by grants from JSPS KAKENHI Grant Number 25H01169 and JST ACT-X Grant Number JPMJAX24C6. We used ABCI 3.0 provided by AIST and AIST Solutions with support from "ABCI 3.0 Development Acceleration Use".

*Yui Uehara, Kanagawa University, yuiuehara@kanagawa-u.ac.jp
†Satoshi Tojo, Asia University, tojo_satoshi@asia-u.ac.jp

References

[1] M. Gotham, G. Micchi, N. N. López, and M. Sailor. When in Rome: A Meta-corpus of Functional Harmony. *Transactions of the International Society for Music Information Retrieval*, 6(1):150–166, Nov. 2023.

[2] N. Nápoles López, M. Gotham, and I. Fujinaga. AugmentedNet: A Roman Numeral Analysis Network with Synthetic Training Examples and Additional Tonal Tasks. In *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, pages 404–411, 2021.

[3] Y. Uehara. Automatic mode recognition and harmonic analysis with unlabeled data. *International Journal of Music Science, Technology and Art*, 6(2):17–37, 2024.

[4] K. Wakabayashi and T. Miura. Forward-backward activation algorithm for hierarchical hidden markov models. In *Advances in Neural Information Processing Systems*, volume 25, 2012.