Abstract

Music is a universal culture of humankind that is believed to have existed since the prehistoric ages, and it has become an indispensable part of life for many people even today. The fact that many people spend a great deal of time, effort, and money on music suggests its considerable importance. Music can be discussed from various aspects, but three elements are generally considered important for its composition, especially in Western music: rhythm, melody, and harmony. This study focuses on harmony, particularly in tonal music.

There have been many attempts to analyze the harmonic structure in tonal music. One of the standard methods is to use a distance model to represent the structure of harmony. Broadly speaking, there are theory-based models and statistical/machine learning-based models, each of which has its own advantages and disadvantages. In general, theory-based models are more interpretable and easier to use for human analysis, understanding, and education of music. On the other hand, however, they also have drawbacks: they are not well supported by actual data, they are not easy to fine-tune due to differences in types or genres of music, and they are also not easy to modify or expand the structure without specialized knowledge. We therefore aim to integrate theory-based and data-based methods to address these weaknesses while preserving interpretability.

First, we propose a framework in which we construct, train, and evaluate a variety of distance functions that share the same domain and range as TPS. We define three basic harmonic features (mode, tonic, and degree), and utilize these to define various distance functions by combining them, and then we evaluate all simple combinations exhaustively. We then train and evaluate each function through the task of key estimation. This entire process is to provide a perspective from which to examine how well various features and their combinations can represent harmony using our distance functions, which are themselves distance models, as probes. We show that very simple functions, i.e., functions with a very small number of effective parameters (¥#EP), can achieve around 80¥% accuracy as long as they include all important harmonic features. This accuracy score is greater than that of TPS. On the other hand, we also confirm that some aspects of TPS are indeed well designed. In addition, a learned model is compared to other theoretical and experiment-based models to verify agreement and minor differences.

This framework, however, uses the task of key estimation for a sequence of chord symbols, where the distances between a chord symbol and its candidate interpretations are not taken into account. If we consider TPS as a model that defines the distance between chord interpretations, the distance between a chord symbol and its candidate interpretation is inherently out of scope. However, since basic space, one of the components of TPS, calculates distances by taking into account the importance of each pitch class (PC), this structure can be applied to the calculation of distance between chord symbols and chord interpretations. So we extend the framework one more step to consider the distance between chord symbols and candidate chord interpretations by comparing them on a PC-by-PC basis. The functions are then trained and evaluated in the task of key:degree estimation from a sequence of chroma vectors. This allows the importance (or distance) of each PC to be obtained from the data. Experiments showed that the structure of the basic space and the assigned values were quite appropriate, but the performance can be further improved by adding the distinction of major/ minor scales.

Keywords: Harmony Analysis; Distance Model; Probabilistic Model; Tonal Music; Machine Learning