

Title	和声距離モデル構築のための効果的な特徴の組み合わせ方の研究
Author(s)	山本, 紘征
Citation	
Issue Date	2025-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/20081
Rights	
Description	Supervisor: 上原 隆平, 先端科学技術研究科, 博士

氏 名	山本 紘征
学 位 の 種 類	博士（情報科学）
学 位 記 番 号	博情第 562 号
学 位 授 与 年 月 日	令和 7 年 9 月 24 日
論 文 題 目	Exploring Effective Feature Combinations to Define Harmonic Distance Models
論 文 審 査 委 員	上原 隆平 北陸先端科学技術大学院大学 教授
	緒方 和博 同 教授
	平石 邦彦 同 教授
	北原 鉄朗 日本大学 教授
	東条 敏 亜細亜大学 教授

論文の内容の要旨

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

論文審査の結果の要旨

音楽は人類に普遍の文化であり、先史時代から存在していたと考えられており、現代においても多くの人の生活に欠かせないものとなっている。音楽は様々な側面から論じることができるが、特に西洋音楽においては、リズム・メロディー・ハーモニーの3つの要素が音楽を構成する上で一般的に重要と考えられている。音楽情報処理においては、和音のコードシンボル列に対して妥当な調・度数の解釈を与えることは重要な研究課題である。このため、和音間の距離を数値的に与え、その和音列の距離の和が最小であるものが最も妥当な解釈であるとする理論 **Tonal Pitch Space (TPS)** が知られている。しかしながら、この「距離」は **TPS** の提案者 **Fred Lerdahl** が恣意的に定義したものであり、この数値が人間の心理的な距離を反映しているかどうかは明らかではない。本研究では、逆に、和音列の最適解釈を正解データとして与えて機械学習を行って、そこから最適な解釈を導く距離関数を構成することを目指した。

TPS は、(i)五度圏、(ii)調内の和音距離、(iii)調内の各音の重要度などを基礎素材として和音間距離を構成するが、これと対比するために本研究においても(i)～(iii)を基礎素材とし、これらを足し算・かけ算する方法を全列挙することで **1540** 通りの関数を定義し、調推定タスクを通して各関数を学習・評価した。このプロセス全体の目的は、距離関数（それ自体が距離モデルである）をプローブとして用いて、様々な特徴とその組み合わせが和声をどの程度適切に表現できるかを検証するための視点を提供することである。

実験の結果、オリジナルの **TPS** を上回る **80～90%** の **accuracy** を実現することができた。また、オリジナルの理論の和算の定義も比較的良い性能を示すこと、二和音間の進行は逆方向と対称に考えても影響が小さいこと、短調へは同主調より平行調に近いことが実証された。本研究は、**TPS** の理論的な側面と、実データに基づく最短経路を学習する統計モデルを組み合わせた画期的な試みであると考えられ、学術的に貢献するところが大きい。よって博士（情報科学）の学位論文として十分価値あるものと認めた。