統計学的および音楽理論的モデルに基づく音楽自動補完の研究

Academic Year 2019
Graduate School of Advanced Mathematical Sciences
Resume of Doctoral Dissertation

Automatic Music Completion through Statistical and Musicological Modeling

Frontier Media Science Program

Christoph Matthias Wilk

1 Research Goal

This dissertation proposes automatic music completion as a new class of music information problems, which are characterized as tasks of automatically generating complete music pieces from any incomplete fragments of music. These fragments can belong to multiple levels of musical abstraction, such as notes, harmonies, and musical keys. Therefore, one can interpret automatic music completion as a generalization of several conventional music information problems, including automatic melody generation and harmonization (generating harmonies from melodies).

The goal is to turn any musical idea of a user into music pieces, allowing users to quickly explore new ideas, as well as enabling musically inexperienced users to create their own music. This principle is applicable to a wide variety of music. As a representative compositional discipline, this dissertation presents mathematical models and algorithms for the automatic completion of four-part chorales, allowing users to freely constrain the melodies of four voices as well as the underlying harmony progression.

2 Summary of the Chapters

Chapter 1 discusses the main contributions of this dissertation, and provides an overview over related research. The contributions include (1) the formulation of automatic music completion as a music information problem, (2) multiple mathematical models and

algorithms to solve this problem for four-part chorales, (3) proposing parameterized modeling of harmony as a new approach to automatic music generation, and (4) presenting several evaluation methods as well as their application to the models and algorithms of this dissertation.

In Chapter 2, automatic music completion is motivated as a fundamental principle for music composition assistance with a focus on user input. In contrast to previous research, this dissertation follows the new direction of allowing user input to be as free as possible. This freedom implies the challenge to handle input of any size (a few notes, melodic fragments, almost complete melodies, no input, or multiple note candidates to choose from), as well as the goal to provide as many modes of input as possible (notes, multiple melodies, harmonies, rhythm, musical keys, abstract tuning parameters etc.).

The chapter also argues that four-part chorales were chosen as a representative compositional discipline for automatic music completion, because rules for chorale composition pose a challenging problem and are applicable to a wider range of musical styles. Furthermore, the harmonic structure of chorales is complex, and the basis of most styles of Western tonal music. Lastly, the large amount of music theory that exists for chorale composition makes it easier to evaluate generated music (writing four-part chorales is a common task for students of music composition, and can be evaluated by teachers quite objectively).

The chapter finally discusses several design principles, such as explicit modeling of musical concepts, as opposed to implicit modeling through neural networks, in order to grant users more influence on the music generation process.

Chapter 3 presents a musicological analysis of the problem of automatic music completion. It discusses the fundamental structure of harmony progressions that has to be captured by mathematical models of harmony, as well as the rules that constrain the composition of inter-dependent melodies in four-part chorales.

Chapter 4 formulates automatic music completion as a constrained optimization problem based on optimization hypotheses derived from music theory. Since the goal of free user input entails the possibility of input that violates rules of music theory, these rules

are interpreted as loose constraints, and treated probabilistically (probability of violating rules). This leads to the optimization hypothesis that optimal solutions satisfy these loose compositional constraints as best as possible, while being additionally constrained by user input.

The musical models of this dissertation were designed in a modular fashion. A complete model of four-part chorales consists of a harmony model (hidden structure) and a voicing model (observed notes).

Chapter 5 presents two harmony models. The first model is based on learnable n-gram probabilities, a concept of natural language processing, adapted to harmony based on the similarity between harmonic structure and language grammar. The second model proposes a completely new approach to generating music, based on parameters derived from music theory. These parameters quantify music-theoretic properties of harmonies, and allow users to individually specify what kind of harmony progression they want to be generated.

Chapter 6 presents two voicing models. The first model handles the high combinatorial complexity of four voices by assuming probabilistic independencies based on human understanding of music composition. It combines statistical learning from data with some heuristics to account for more complex music-theoretic principles. The second model reduces the need for heuristics and is independent from the number of voices. It is based on several trainable factors that each capture a part of the musical context a composer would consider.

Chapter 7 introduces algorithms to solve the problem of automatic music completion by treating it as a graph search problem. It also presents several methods to considerably improve the performance of such algorithms in the context of music, both with respect to computation speed and the quality of generated music.

Chapter 8 discusses the experimental evaluation of models and algorithms proposed in this dissertation. It presents objective metrics derived from music theory, which can be used to evaluate generated voicings. Results with respect to these metrics show that the voicing models come quite close to the famous composer Bach, and compare favorably

against a state-of-the-art deep learning model for imitating Bach. The parameters of the parameterized harmony model are also validated quantitatively.

In addition, the chapter discusses two subjective evaluation experiments in which participants used automatic music completion systems with their individual input of musical ideas. Both systems (based on the respectively first and second models of harmony and voicing) received positive feedback from participants, confirming their usability and the subjective quality of generated music.

Lastly, the possibly most significant evaluation was conducted by a professional composer, who analyzed generated music and rated its quality highly, while also providing insight on a few remaining weaknesses of the music models, which can lead to further improvement.

Chapter 9 concludes the dissertation with a summary and an outlook regarding future research.

重要な音楽専門用語の和訳

Harmony – 和声（和音の音楽理論的な概念）

Key – 調（音楽の和声構造、中心音と音階に関わる概念。例：短調、長調）

Harmonization – 和声付け（旋律にハモる和音を付けること）

Four-part chorale – ４声コラル（音楽理論の四声体を中心とした作曲法による曲）

Voice – 声部（同時にそれぞれの旋律を演奏するパート）

Harmony progression – 和声進行（和音の時系列）

Voicing – ヴォイシング（和声に従って声部に音符を割り当てること）