

信頼される AI の基盤技術
2020 年度採択研究者

2021 年度 年次報告書

吉井和佳

京都大学 大学院情報学研究科
准教授

人と AI の同化に基づく能力拡張型音楽理解・創作基盤

§ 1. 研究成果の概要

第二年次は、当初の計画通り、人間との親和性・利便性を念頭に置いたうえで、「(1)信号から記号への変換技術」に継続的に取り組むとともに、「(2) 音楽知識の記号表現技術」に着手した。具体的には、ポピュラー音楽の音響信号を対象として、これまで培ってきた各種音楽要素の解析技術を洗練化させ^{1),3)}、複数の音楽要素間の依存関係を考慮したマルチタスク学習にも取り組んだ⁴⁾。これにより、音楽の三大要素であるメロディ・ハーモニー・リズムを高速かつ高精度に一挙に解析する究極的なシステムの開発に向けて着実に前進した。また、人間と共通の音楽理解様式・知識の実現のため、第一年次では、認識と生成が表裏一体であるという観点のもとで VAE を基礎とした研究を進めてきたのに対し、第二年次はさらに、音楽要素の文法構造が音楽の本質であるという観点のもとで、それらを言語モデルとして形式知へ変換することにも取り組んだ。さらに、Web 上にアップロードされたピアノ演奏の音響信号の自動解析の洗練化についても取り組んだ^{2),5)}。近年、一般ユーザの創作活動では、自らが作曲した楽曲だけではなく、自らが演奏したものを公開することが盛んになっている。特に、主要コンテンツであるピアノ演奏への対処が不可欠である。これらの研究に一貫して、人間の可読性を重視して、ピアノロールではなく、離散化された楽譜形式を高精度に出力することに成功した。

また、「(3)協調的な記号操作技術」に関しても、当初計画より前倒して検討を開始した。具体的には、演奏難易度を考慮したポピュラー音楽のピアノ演奏への自動編曲に向けて予備的な検討を行った。もとの楽曲がポピュラー音楽であっても、ピアノでソロ演奏したいというニーズは根強いが、ピアノ編曲が入手できない場合が多く、運良く入手可能でも、奏者の演奏技術・好みに見合うとは限らない。前述した背景に加えて、自らの好きなポピュラー音楽をピアノで弾きたいということが、ピアノを始める動機となっていることも多く、音楽初等教育を支える基盤技術となることが期待される。

【代表的な原著論文情報】

- 1) Ryoto Ishizuka, Ryo Nishikimi, Kazuyoshi Yoshii, “Global Structure-Aware Drum Transcription Based on Self-Attention Mechanisms”, *Signals*, Vol. 2, No. 3, pp. 508-526, 2021.
- 2) Kentaro Shibata, Eita Nakamura, Kazuyoshi Yoshii, “Non-Local Musical Statistics as Guides for Audio-to-Score Piano Transcription”, *Information Sciences*, Vol. 566, pp. 262-280, 2021.
- 3) Ryo Nishikimi, Eita Nakamura, Masataka Goto, Kazuyoshi Yoshii, “Audio-to-Score Singing Transcription Based on a CRNN-HSMM Hybrid Model”, *APSIPA Transactions on Signal and Information Processing*, Vol. 10, No. e7, pp. 1-13, 2021.
- 4) Takehisa Oyama, Ryoto Ishizuka, Kazuyoshi Yoshii, “Phase-Aware Joint Beat and Downbeat Estimation Based on Periodicity of Metrical Structure”, *International Society for Music Information Retrieval Conference (ISMIR)*, pp. 493-499, 2021.
- 5) Yuki Hiramatsu, Eita Nakamura, Kazuyoshi Yoshii, “Joint Estimation of Note Values and Voices for Audio-to-Score Piano Transcription”, *International Society for Music Information Retrieval Conference (ISMIR)*, pp.278-284, 2021.