

## 短時間フーリエ変換を用いた高精度 MIDI 変換の検討 High-Precision MIDI Conversion Using Short-Time Fourier Transform

○古橋優希<sup>1</sup>, 岸本誠也<sup>2</sup>, 大貫進一郎<sup>2</sup>

\*Yuki Furuhashi<sup>1</sup>, Seiya Kishimoto<sup>2</sup>, Shinichiro Oonuki<sup>2</sup>

Abstract: MIDI, widely used in the field of music information processing, is an international standard for converting musical instrument performance information such as pitch, dynamics, and onset timing into numerical values, and for recording and communicating the data. It enables a wide range of applications, such as automatic transcription from performance information and extracting the performer's emotions. However, when it comes to technology for converting audio signals from musical instruments into MIDI, there remain challenges in accurately extracting and converting pitch, dynamics, and onset timing.

音楽情報処理の分野において広く用いられる MIDI(Musical Instrument Digital Interface)とは、演奏情報である音程、強弱、発音タイミング等を数値化して記録、通信するための国際共通規格である。MIDI を用いることで、演奏情報から自動採譜<sup>[1]</sup>や演奏者の感情抽出<sup>[2]</sup>等、多様な応用が可能である。しかし、楽器演奏のオーディオ信号を MIDI に変換する技術においては、音程、強弱、発音タイミングのいずれも正確に抽出・変換するという点には課題が残されている。

本報告では、ピアノ演奏のオーディオ信号を対象に、短時間フーリエ変換<sup>[3]</sup>を用いて時間周波数特性を解析する。短時間フーリエ変換は、信号を指定した長さの区間であるフレームに分割し、各フレームごとにフーリエ変換をする。時間的に変化する周波数成分を解析する手法であり、指定するフレームによって時間分解能と周波数分解能がトレードオフの関係にあるという特徴がある。図1では短いフレーム時の解析結果、図2では長いフレーム時の解析結果を示し、ピアノを弾いたタイミングを赤線で示した。短いフレームでは時間分解能は高く、発音タイミングを精密に推定できるが、周波数分解能が不足するため音程の識別が不安定になりやすい。一方で長いフレームを用いた場合は周波数分解能が高く、音程を精密に推定できるが、時間分解能が低下するため発音タイミングを精密に推定することが難しい。そこで、短いフレームと長いフレームの短時間フーリエ変換を組み合わせることで MIDI 変換を用いる手法を提案する。この手法により、時間分解能と周波数分解能それぞれの利点を活かし、音程、強弱、発音タイミングに対する高精度な MIDI 変換を実現させることを目的とする。

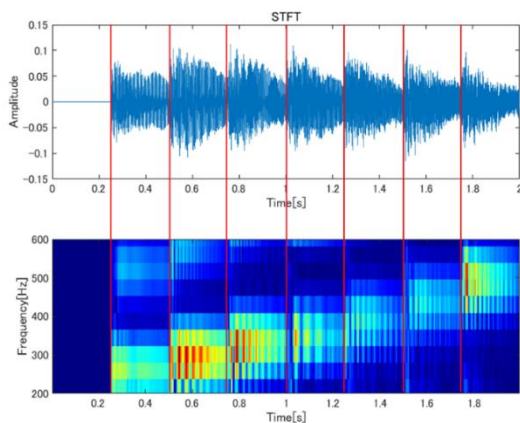


Figure 1. 解析結果(短いフレーム時)

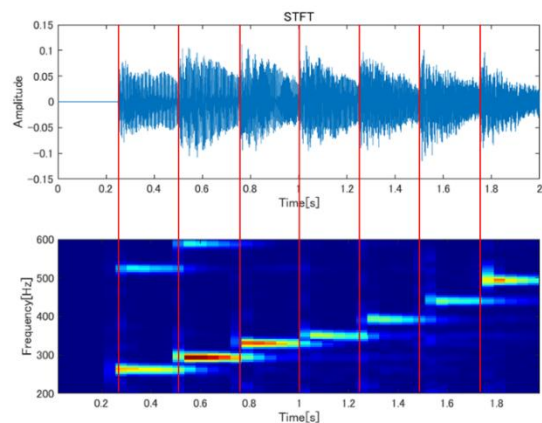


Figure 2. 解析結果(長いフレーム時)

参考文献

- [1] 嵯峨山茂樹, 亀岡弘和: 自動採譜技術の展望, 日本音響学会誌, Vol.64, No.12, pp.715-720, 2008 年.
- [2] 玉木明和, 矢鳴虎夫, 加藤清史: MIDI 信号を用いた演奏音楽からの感情抽出の試み, 九州工業大学研究報告, No.68, pp.57-63, 1996 年.
- [3] 小野順貴: 短時間フーリエ変換の基礎と応用, 日本音響学会誌, Vol.72, No.12, pp.764-769, 2016 年.

1: 日本大学・学部・電気 2: 日本大学・教員・電気