

多重音を含む音楽音響信号のピッチ推定を用いた演奏表情抽出のための楽譜との対応付け

酒造 祐介[†] 武田 晴登[†] 片寄 晴弘^{††}

[†] 関西学院大学大学院理工学研究科 ^{††} 関西学院大学理工学部

1 はじめに

名演奏家による情緒溢れる演奏の実際のテンポや音量の制御を計測したデータ(以下,演奏表情)は,自動演奏や演奏解析,そして音楽検索における特徴量として利用できる.そのデータベース化の効用は大きい.本稿では,演奏曲の楽譜が既知であるという条件下で,CD等に収められている演奏家の演奏録音から演奏情報を精密かつ効率的に取得する手法について検討する.

従来研究では,演奏されている各音の発音時刻と音高情報がMIDI(Musical Instrument Digital Interface)によって記録されている場合は,演奏表情の抽出データベースが作成されている[1].しかし,演奏を録音した音響信号とその演奏曲の楽譜との対応付けによる演奏表情抽出については,従来から研究[2, 3]が進められているが,これらの手法では和音として演奏される音の発音時刻のずれなど,各音の細かい変動を求むることができなかった.一方,自動採譜を主な目的としている多重音の基本周波数推定手法[4, 5]には音響信号からの演奏音の発音時刻や音高の推定をする潜在的な能力がある.そこで,本稿では多重音解析技術を用いて音響信号からの演奏された各音の発音時刻の変動を含んだ演奏表情を効率的に求める手法を検討する.

2 手作業を含む音響信号と楽譜との対応付け

2.1 音響信号と楽譜との対応付け

2.1.1 手作業を含む効率的な対応付け作業

和音の各音の発音時刻のずれや音の大きさを精密に抽出するには,音響信号に含まれる各音の発音時刻,音高,強さ,消音時刻を推定しなければならない.これらの情報を求めるには多重音解析が必要であるが,一般に楽器音の多重音解析では,演奏楽器音に関する先験的知識が与えられない限り,原理的に誤推定を避けることはできない.

我々は多重音解析を全て自動化するのではなく,多

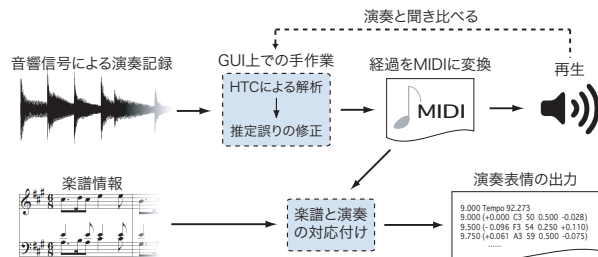


図1: 演奏表情抽出の手順

重音の解析結果に手作業で修正を加えて信頼性の高い情報を求める.実際,楽器音に対するあらゆる先験的知識の用意は困難なので,楽譜情報と手作業を加えてデータを作成する方が現実には有効であると考えられる.以下では,人間が行う修正作業を効率的に行えるようにするために,目と耳で推定結果を確認しながら作業を行えるインターフェイスを備えたソフトウェアについて検討する.

2.1.2 音響信号と楽譜との対応付けを行うソフトウェア

図1に示す処理により,音響信号と楽譜との詳細な対応付けが可能になり,演奏表情抽出が実現できる.

まず,演奏表情を抽出したい部分を切り出した音響信号とその部分の楽譜情報を用意する.次に,音響信号に多重音解析技術を用いて音響信号内の各楽音の発音時刻,音高,強さ,消音時刻を推定する.その推定結果をGUI上での手作業で修正を施す.このとき,随時修正結果をMIDIに変換して元音と聞き比べながら作業を行うことで,効率的に修正が行える.十分な修正を施した後,楽譜情報を対応付けて,修正を施されたデータと楽譜情報の差分を演奏表情として抽出する.

2.2 音響信号からの演奏表情の推定の手順

2.2.1 音楽音響信号のからの楽音の情報の抽出

音響信号に既存の多重音解析技術を用いて音響信号内の各楽音の情報を推定する.各楽音の情報とは,本研究で演奏表情の特徴量としている発音時刻,強さ,消音時刻を指す.

亀岡らによる多重音解析技術「調波時間構造化クラスタリング(HTC)」[4]は,時間周波数平面上に分布するエネルギーを音響イベントごとにモデルフィッティングを行っており,その結果得られる推定結果のモデルの音響イベントが演奏の各音と一致していれば,各

Alignment of Polyphonic Musical Audio Signal and Score Using Multi-Pitch Estimation for Extracting Performance Deviation

[†] Yusuke MIKI

[†] Haruto TAKEDA

^{††} Haruhiro KATAYOSE

Graduated School of Science and Technology, Kwansai Gakuin University ([†])

School of Science and Technology, Kwansai Gakuin University (^{††})

音の発音時刻，消音時刻，パワーが得られる．この手法は，楽器に関する特殊な情報を事前に与える必要がないので，本研究での音響信号から各楽音の音響特徴量を同時に推定する目的に使用することができる．

ただし，HTC も他の多重音解析技術と同様に誤推定は不可避であり，速弾きの箇所，小音量の楽音や低音についての解析精度が低い場合がある．このような場合に起きやすい誤推定として楽音の未検出，発音時刻や消音時刻のずれ，実際には演奏されていない音の誤認識がある．

2.2.2 手作業による多重音解析結果の補正

多重音の解析結果と演奏曲の楽譜とを人間が比較し修正を加えて，元の演奏により近い音響信号の MIDI 変換を行うことを考える．解析の誤りを修正する作業を効率的に行うために，時間周波数表現による音響信号の可視化表現，及び，推定結果を MIDI ファイルにして再生することで，比較を容易にする．

音響信号の可視化については，Specmurt 法 [5] による表示で手作業の負担を軽減させることができる．Specmurt 法はスペクトログラムの中で基本周波数成分以外を抑圧低減して，基本周波数が対数スケールで強調表示できる．スペクトログラムでの表示では倍音成分などが表示されて，基本周波数などの情報が目視では認識しづらいが，Specmurt 法を用いることで発音時刻や消音時刻を修正するときの操作性を向上させることができる．

2.2.3 変換した MIDI と楽譜との対応付け

以上の処理で修正された結果を MIDI ファイルに変換し，楽譜情報との対応付けを行う．和音における発音順序の交替を許した DP(Dynamic Programming) マッチングを用いて，音響信号から得られた MIDI ファイルと楽譜を各音について対応付ける．対応付けた後に詳細な演奏表情を求める．

3 システムの実装と動作検証

3.1 システムの実装

音響信号を入力として演奏表情抽出を行うために，図 2 に示す GUI を備えたソフトウェアを実装した．主な機能は以下のとおりである．

- 図 2 左上のスライダによる修正したい箇所の表示
- specmurt の結果と HTC の解析結果を時間周波数平面上で重ね合わせおよび選択表示し，目視による音の脱落や挿入の修正
- 修正結果を保存した MIDI ファイルと元の音響信号との聴き比べによる音の誤りの検出
- Specmurt のみの表示や Specmurt の色合いの調性等による時間周波数平面上での表示の調整

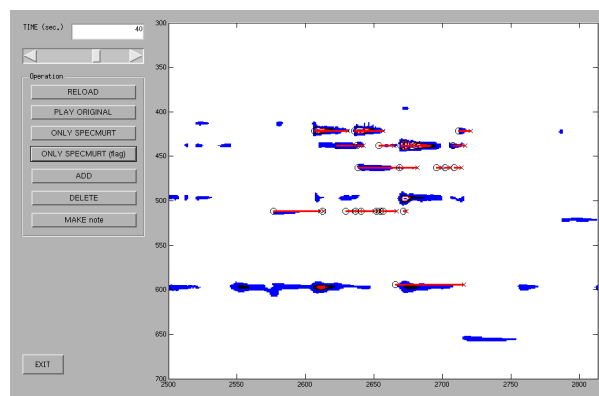


図 2: 多重音推定結果の誤り修正作業を行う GUI

- 得られた結果の note 形式 [1] での保存

3.2 音楽 CD からの演奏表情データの抽出例

本システムを用いて著名なピアノ演奏家，Alfred Denis Cortot による Chopin 作曲の Etude op.10-3 の CD に収録されているピアノ演奏から演奏表情を抽出した．これは 1933 年のもので，演奏者は他界しており新しい MIDI データ作成による演奏表情抽出はできない．専門的な音楽教育を受けていない作業者による約 7 時間の手作業で 90 秒の演奏表情を抽出することができた．その結果のデータは，テンポの変動などが元の演奏を反映しているように感じられた．演奏表情抽出が可能であることは検証できたが，手作業の労力をさらに軽減する必要がある．

4 おわりに

本稿では音響信号と楽譜との対応付けを効率的に手作業を加えながら行う手法を検討した．今後は，音の強弱に関する情報を MIDI のペロシティ値として推定し，RWC 研究用音楽データベースを対象に演奏表情のデータベースの作成を行う予定である．

謝辞

本研究では，システムの実装において東京大学の亀岡弘和氏，斎藤翔一郎氏，嵯峨山茂樹教授に提供して頂いた HTC および specmurt のプログラムを使用した．

参考文献

- [1] Toyoda, K. *et al.*: Utility System for Constructing Database of Performance Deviations, *Proc. of IS-MIR* (2004).
- [2] Widmeret G. *et al.*: In Research of the Horowitz Factor, *AI Magazine*, pp. 111–130 (2003).
- [3] Soulez F. *et al.*: Improving Polyphonic and Poly-Instrumental Music to Score Alignment, *Proc. of IS-MIR* (2003).
- [4] 亀岡弘和ほか: 調波時間構造化クラスタリング (HTC) による音楽音響特徴量の同時推定, *情処研報*, MUS-61-12, pp. 71–78 (2005).
- [5] 高橋佳吾ほか: 対数周波数逆畳み込みによる多重音の基本周波数解析, *情処研報*, MUS-53-13, pp. 61–66 (2003).