

ピアノ演奏自動評価支援のための音違い検出手法

Wrong key detection method for automatic evaluation supporting on piano playing

山田 和広[†]
Kazuhiro Yamada

中平 勝子[†]
Katsuko T Nakahira

1. はじめに

教育機関等でのピアノ弾き歌い学習では、多くの学習者に1人の指導者による対面授業が行われる。しかし、学習者の熟練度によって個別指導が必要になるため、学習者1人当たりの指導時間が少なく、教育の質が低下するという問題がある。そこで、不足する指導を補完するための研究が行われている。そのうちの 하나가、eラーニング教材を活用した指導法である^[1]。これによって、特に学習者の歌唱技能が向上する事が分かった。しかし、学習者のピアノ演奏については技能の向上がほとんど見られず、eラーニング教材による指導の限界とされている。そのため、指導者から直接伝授しなければならない教育とそうでない教育を分離し、後者を演奏自動評価支援システムによる自習を組み合わせることにより、教育効果の向上が期待される。本稿では、ピアノ演奏自動評価支援システムに組み込むための音違い検出手法を提案する。学習者の音間違いを検出するためには、ピアノ音高の推定、演奏開始位置と終了位置の検出、打鍵タイミングとリズムの正規化が必要となるが、本稿ではピアノ音高の推定を対象とする。

2. ピアノ音高の検出手法

音高推定法は、様々な手法が提案されており、くし形フィルタ^[2]、逆ノッチフィルタ^[3]等がある。本稿では、オクターブの異なる音高でも推定可能、計算量が少ない点から、逆ノッチフィルタ(以下、INF)を用いた音高推定法を応用する。

INFは、特定の周波数成分を抽出するフィルタである。以下に、INFの伝達関数を示す。

$$H_N(Z) = \frac{1}{2} \left\{ 1 - \frac{r_A + \alpha z^{-1} + z^{-2}}{1 + \alpha z^{-1} + r_A z^{-2}} \right\} \quad (1)$$

ここで、 α は抽出する周波数 f から決定されるフィルタ係数であり、以下の式で定義される。

$$\alpha = -(1 + r_A) \cos\left(2\pi \frac{f}{f_s}\right) \quad (2)$$

ただし、 f_s は音声信号のサンプリング周波数、 r_A は通過帯域幅を決定する定数($0 < r_A < 1$)である。音高推定には、音高に対応する周波数を通過させるINFを並列に接続し、INFのパワー値を調査した後、パワー値が閾値以上のINFに対応する音高を推定音高とする処理を行う^[4]。

しかし、本手法には音高推定された音と隣り合っている音高(以下、隣接音高)を除外する制約条件が存在する。こ

の制約条件をつける理由は、楽音の調波成分の前後もパワー値が大きくなるため、この影響を受けやすい直近の音を推定音高として選択されるのを防止するためである。しかし、実際には隣り合った音を同時に弾くことがある楽曲が存在するため、この制約条件を適用した場合、正しく音高推定ができない。

この問題を解決するため、隣接音高の音高推定を次のように行った。まず、従来手法で検出された音高候補を P_n 、そのパワー値を V_n (n はいずれもノート番号)と置く。この時、 V_{n-1} の音を V_n と同じ強さで弾いていた場合、 $V_n \cong V_{n-1}$ が成立すると仮定し、 $V_n \gg V_{n-1}$ の場合にのみ音高推定から除外することで、隣接音高の音高推定を行う。

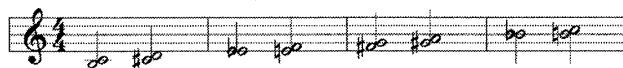
3. 実験

提案手法が有用であるかを確認するための予備実験を行った。入力波形には、電子ピアノで図1の旋律を演奏した録音波形を用いた。波形は44.1kHzモノラルで記録した。また、音高推定間隔は441サンプル(0.01s)おきに行う。なお、この録音は電子ピアノの音声出力から直接録音したものであるため、無音区間から決定する閾値の代わりに適切な閾値を手動で設定している。

結果は図2の通りである。この図では、横軸は時間、縦軸はノート番号を示す。なお、ノート番号は60をピアノ中心のCの音と定義し、半音変化する毎に1増減する。従来手法では隣接して弾いている音のうち、パワー値が大きい音のみが推定音高として選択される。それに対し、提案手法では隣接音高を両方推定音高として選ばれたことが分かる。このことから、隣接音高推定における本手法の有用性を確認することができた。

続いて、実際のピアノ楽曲に対して提案手法を適用し、音高推定を行った。アコースティックピアノで演奏した楽曲を録音し、音高推定を行う。録音時の記録フォーマットは44.1kHz、モノラルである。また、音高推定間隔は441サンプル(0.01s)おきに行う。演奏楽曲はバイエル60, 88, 106番の3曲である。そのうち、バイエル60番の旋律の一部を図3に示す。

3曲における音高推定結果のうち、バイエル60番の結果の一部を図4に示す。また、各楽曲における楽譜音を検出できた割合を表1に示す。推定できた割合は3曲の平均で83.8%であり、抽出すべき音高については高い精度で推定できた。



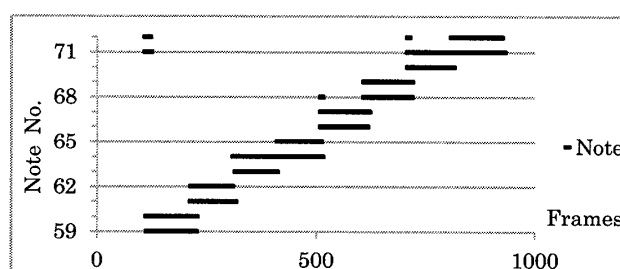
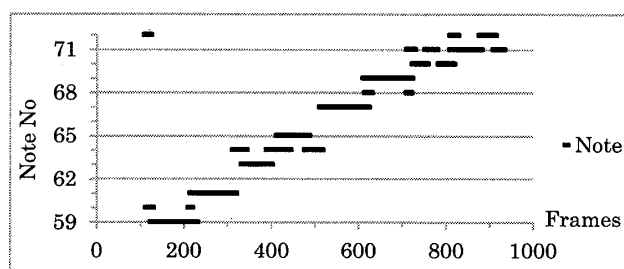
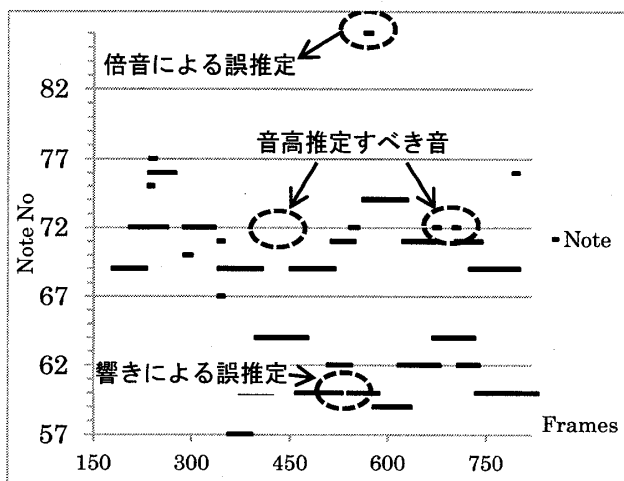


図2. 従来手法(左図)と提案手法(右図)の予備実験の音高推定結果比較



図3. 演奏楽曲(バイエル 60 番先頭 4 小節)

図4. 演奏楽曲の音高推定結果
(バイエル 60 番先頭 4 小節)

しかし、以下の問題点が明らかとなった。

- 1) 高調波成分の周波数が抽出される
提案手法では高調波成分のうち3倍音までを考慮しているためである。低音ではより多くの高調波成分があるため、これが影響していると考えられる。
- 2) 隣接音高が抽出される
ピアノ打鍵時に本来の音の基本周波数よりその前後の周波数のパワー値が大幅に高くなるためである。
- 3) オクターブ違いで同時に打鍵している場合に、その音高が抽出されない
提案手法では音高推定された音の高調波成分に相当する音高を除外するため、オクターブが異なる音は音高推定されない場合がある。
- 4) 音の響きが残っている場合、その音も推定する
鍵盤から手を離しても音の響きが残っており、その周波数のパワー値が高くなるためである。
- 5) 低音の音高推定できない
低音では高調波成分が非常に大きく、基本周波数成分

表1. 楽譜音を検出できた割合

楽曲番号	音符数	推定数	割合
60 番	149	132	0.886
88 番	360	340	0.944
106 番	478	327	0.684
計	987	799	0.838

を抽出できなかったと考えられる。また、基本周波数が非常に低いため、本実験の音高推定間隔が不適切であったことも考えられる。

4. おわりに

本稿では、学習者のピアノ演奏データから音間違い検出のための方法を述べ、第1段階となる音高推定の方法について検討した。また、実際の楽曲を用いて入力音と教師データとの比較を行い、有用性を確認することができた。しかし、ピアノ演奏の録音信号を用いた音高推定は、低音の高調波成分による誤判定、オクターブ違いでの音高推定の失敗など、様々な問題が起りえることが明らかとなった。

今後は、実験で明らかとなった問題点を解決し、音高推定の精度を高めると共に、演奏リズムの推定を行えるようにしたい。

謝辞

本研究の一部は科学研究費基盤 C(代表・深見友紀子、10283053 および 18500742)の助成を受けて行われた。

参考文献

- [1] 深見友紀子, 中平 勝子, 赤羽 美希, “ピアノ弾き歌いによる e ラーニング教材の効果”, 京都女子大学発達教育学部紀要, vol.6, pp.35-46, 2010
- [2] 三輪多恵子, 田所 嘉昭, 斎藤 努, “零出力に注目したくし形フィルタによる音階検出”, 電学論(C), vol.J118-C, no.1, pp.57-64, 1998
- [3] 干場圭太郎, 川村 新, 飯國 洋二, “ノッチフィルタを用いた音高推定”, 信学技報, SIP2007-16, pp.31-36, 2007
- [4] 信太 誠, 内田 剛, 川村 新, 飯國 洋二, “逆ノッチフィルタを用いた音高推定法のしきい値設定に関する検討”, 信学技報”, SIP2009-8, pp.39-43, 2009