3R-3

楽譜の自動譜めくりシステムのための演奏位置推定方法

小倉 淳　　松島 俊明

東邦大学理学部 情報科学科

1. はじめに

演奏に合わせて表示する楽譜を自動的に切り替える自動譜めくりシステムの研究を行っている.

自動譜めくりシステムと同様の処理が要求されるシステムには自動伴奏システムがあるが，これらのシステムでは，演奏から抽出されたピッチ情報と演奏の元となる楽譜情報を，単旋律であればDPマッチング[1]，和音を含む場合では HMM を用いる例[2]など，高精度でのマッチングを行なうために様々な方法が試みられている.

このように自動伴奏システムが音符レベルでの演奏位置推定が必要であるのに対し，自動譜めくりシステムでは小節レベルでの演奏位置の推定ができれば良い．また，和音を含む演奏から高速かつ高精度にピッチ抽出を行なうことは現時点では困難なため，ピッチ抽出誤りの影響を受け難い方法が好ましい.

そこで，精度よりも抽出速度を重視したピッチ抽出方法[3]の採用を前提として，多少のピッチ抽出誤りがあっても演奏に追従できる方式の検討を行なっている．即ち，第1段階では音高の出現頻度のヒストグラムの比較により演奏中の小節候補を幾つか選出し，第2段階ではそれらの候補に対して音の旋律的な出現パターンをより詳細に比較する2段階マッチングにより，小節レベルでの演奏位置推定を試みた.

2. 音高出現頻度の比較による第1段階マッチング

全く同じ演奏が繰り返されていないのであれば，1小節あるいは数小節内に含まれる音符の分布はそれぞれ特有のパターンを持つ可能性が高い．例えば図1に示したように数拍ずつずらしながら一定の長さの区間内に含まれる音符の長さの音高毎の合計(演奏時間の合計)を求めると，図2のような区間毎の音高分布を得ることができ，そのパターンは区間毎に異なることが分かる．この音高別演奏時間の割合を区間の音高出現頻度と呼ぶことにする.

演奏者に提示する楽譜情報(以後，楽譜情報と呼ぶ)は既知であるため，予め楽譜情報の音高出現頻度の表を作成しておき，演奏者による演奏情報(以後，演奏情報と呼ぶ)から作成した一定時間内の音高出現頻度と比較を行い，類似度の高いもの上位3個を演奏位置の候補として選出する.

この方法では，旋律や和音構造に関する情報を比較の際に利用することはできないが，各音高の累積時間を用いるためインパルス的に出現するピッチ抽出誤りの影響を受け難く，様々な尺度が考えられる和音同士の類似性を考慮する必要がないというメリットがある.

A performance position presumption method for the automatic score turning-over system,
Jun Ogura and Toshiaki Matsushima,
Dept. of Information Sciences, TOHO University

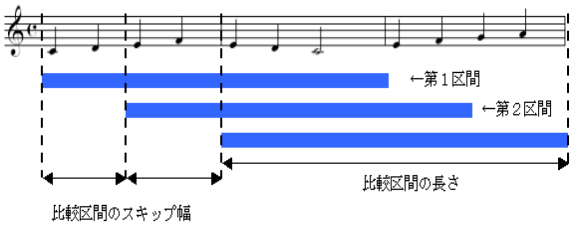

図1：第1段階における比較区間

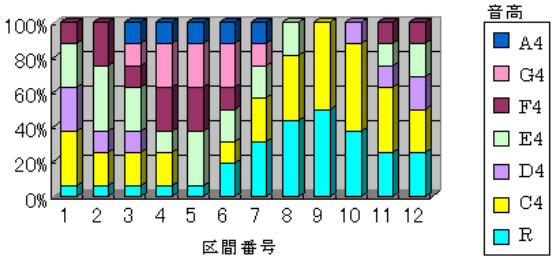

図2：区間別の音高出現頻度の例

3. 音の出現数比較による第2段階マッチング

第1段階では，旋律的な情報を考慮せず，音高の出現頻度のみに着目して候補区間を求めた．そこで，第2段階では旋律的な情報も利用して候補区間の絞込みを行なう.

まず，楽譜情報の候補区間内および一定時間内の演奏情報を数拍単位の小区間に分割し，小区間内での音の出現数を比較する．小区間をオーバーラップさせながら区間全体での出現数の変化を調べることで，旋律的特徴が反映されるように考慮した．即ち，小区間内における楽譜情報の音高の出現数 $\mathbf{n}$ および演奏情報の小区間の長さに対応した一定時間内での出現数 $\mathbf{m}$ をそれぞれ $\mathbf{n}=(n_{C3},n_{C\#3},n_{D3},\cdots,n_{C6}),\mathbf{m}=(m_{C3},m_{C\#3},m_{D3},\cdots,m_{C6})$ (ただし n_P は音高 P の音の出現数)とすると，小区間

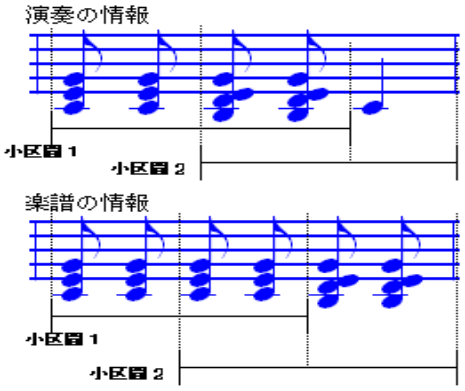

図3：第2段階での小区間

Copyright ©2011 Information Processing Society of Japan.
All Rights Reserved.

同士の距離: $d(\mathbf{n},\mathbf{m}) = \sum_P |n_P - m_P|$ を全ての小区間 i に関して求めた距離: $D = \sum_i d_i(\mathbf{n}_i, \mathbf{m}_i)$ を求める.

例えば, 図3に示した2つの小区間からなる楽譜の場合では, それぞれの音高出現数は, 演奏情報の小区間1は Bb3=2, C4=2, D4=2, Eb4=4, G4=4, 小区間2は Bb3=2, C4=1, D4=2, Eb4=2, G4=2, となっており, 一方楽譜候補の小区間1は C4=4, Eb4=4, G4=4, 小区間2は Bb3=2, C4=2, D4=2, Eb4=4, G4=4,であるので, $d_1 = 6$, $d_2 = 5$となり, この候補区間の演奏情報との距離 D=11 となる. この要領で他の候補区間との距離も求める.

4. 演奏の連続性の利用

第2段階までの処理では区間単位での演奏位置の推定であり, 区間同士の関連性に関しては考慮されていない. そのため, たまたま距離が小さくなった区間が演奏位置の候補として選出される可能性がある. しかし, 演奏は時間の推移にしたがって順次進んでいくものであるから, 演奏の連続性を考慮することで, このような誤りを防ぐことができる.

例えば直前の演奏情報に対する第1段階マッチングで候補として選出された楽譜区間 a の次の楽譜区間 a+1 が, 次の演奏情報に対する第1段階マッチングの候補として選出された場合, この2つの区間は演奏に対する正しい楽譜区間である可能性が他の不連続な区間に比べてより高いと考えることができる.

このような演奏の連続性の影響を反映させるために, 第2段階のマッチングで求めた距離に以下のような補正を行なう.

$$D'_{a+1} = \begin{cases} D_{a+1} - \beta \cdots \text{区間}a\text{が直前の候補だった場合} \\ D_{a+1} \cdots \text{区間}a\text{が直前の候補でなかった場合} \end{cases}$$

ただし β は正の定数

これにより, 連続して候補として抽出された区間の方が優先的に選択されるようになるため, 時間経過を無視したジャンプや, 誤ったマッチングからの復帰が行なえるようになる.

5. 結果とまとめ

旋律部と伴奏部からなる15小節のピアノ曲楽譜を用いて評価実験を行った. 図4(a)に実験に用いた楽譜の一部を示す. 旋律部のみに若干の誤りを含む演奏例1と, 旋律部により多くの誤りを含み, さらに伴奏部にも誤りを含む演奏例2をMIDIで作成し, 正しい演奏位置を検出できるか実験を行った. 第1段階における比較区間の長さを1小節(3拍), 比較区間のステップ幅は3拍(1小節), 第2段階での小区間の長さを2拍, ステップ幅は1拍, 演奏連続性の補正値 β =10 として実験を行ったところ, 表1および表2に示すような結果が得られた.

実験結果からも分かるように, 音の出現頻度を比較するという単純な処理にも関わらず, 比較的良好な結果を得ることができた. 第1段階では低位の候補だったものが第2段階の処理の結果で上位に認識されており, 2段階マッチングの有効性が確認できた.

一方, 第1段階のマッチングで正解区間が候補として選択されない場合も多いことが分かった. その理由としては, 比較する区間内の音数が少ない場合が考えられる. 従って, 旋律音のみで和音等がない場合や演奏者側のミスで音数が大きく減ってしまった場合における第1段階マッチングの区間長やステップ幅の設定方法を検討する必要がある. また, 現在開発中のピッチ抽出システム[3]からの採譜結果を用いた実験も今後の課題である.

参考文献

[1] B. Dannenberg: An On-Line Algorithm for Real-Time Accompaniment, Proc. of ICMC, pp.193-198 (1984)
[2]武田晴登, 西本卓也, 嵯峨山茂樹: HMM による MIDI 演奏の楽譜追跡を用いた自動伴奏, 日本音響会秋季研究発表会講演集, pp.571-572 (2006)
[3]宮坂広純, 松島俊明, 小田弘良: 多重音の基本周波数評価尺度の最小値選択による実時間音高・音源数推定の試み, 第7回情報科学技術フォーラム講演論文集(2), pp.241-243 (2008)

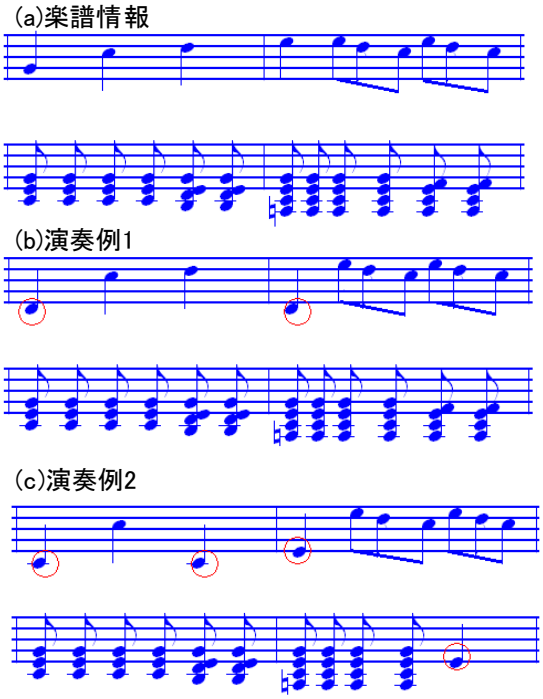

図4：実験用楽譜と演奏例

表1：演奏例1の実験結果(累積正答率)

	順位1	順位2	順位3
第1段階	87%	87%	93%
第2段階	93%	93%	93%

表2：演奏例2の実験結果(累積正答率)

	順位1	順位2	順位3
第1段階	67%	80%	87%
第2段階	87%	87%	87%

Copyright ©2011 Information Processing Society of Japan.
All Rights Reserved.