2L-5

階層構造による音楽リズムの分析と生成

笠原 俊一† 橋本 周司†
† 早稲田大学理工学部応用物理学科

1 はじめに

コンピュータにより支援された音楽の作成やインタラクティブパフォーマンスにおいては、人間が演奏したリズムから人間が意図した音符列を推定し、これを元に、リズムを生成する事が必要となる場合がある。

リズム推定は、一般に発音時刻列推定によって行われるが, 上記の目的の為には、時系列情報のみでなく、リズムの構造的な特徴の推定が必要と考えられる。そこで本稿では、リズムの階層構造を定義し、リズムの構造推定とそれを用いた時系列推定を行うことで推定精度の向上を試みる。さらに構造推定後の構造情報利用の一例として、リズムアレンジへの展望を述べる。

2 従来手法の課題

人間が演奏するリズムの物理的発音時刻は、音符列による正規の発音時刻より、ズレや抑揚によって揺らぐ。そのような揺らぎを吸収するために、分解能を予め指定しておき、閾値処理によって音符列を推定する事が考えられるが、分解能の不足や、過分割、平行移動に弱い、などの問題がある。提案手法はリズムの階層構造を仮定する事で、可変分解能での音符列の推定を行い、固定分解能での問題の解決を試みる。

3 リズム階層構造の定義

リズムの重要な性質の１つとして、群化と階層性が挙げられる。類似パターン、時間的近接、時間的規則性によって群化され、群と群がさらにまとまり、上位の群を形成するという、階層的体制化が心理学の観点からも報告されている [1]。また、リズムを構文木と捉える手法も知られている [2]。

本研究では、リズムをフレーズの階層構造として定義する。ここでフレーズとは、１周期の時間を１としたときの、発音音符 (N;Note) または休符 (R;Rest) の発生時刻に関する系列である。例えば、フレーズ [0(N) 0.25(R) 0.5(N) 0.75(N)] は、リズム (0 0.5 0.75) を表す。フレーズの各要素はさらにフレーズを内包することができる。図1に示されるように、[0(R) 0.25(R)[0(R) 0.5(N)] 0.5(N) 0.75(N)[0(N) 0.333(N) 0.66(N)]] はリズム (0.125 0.375 0.5 0.75 0.833 0.916) を表す。

システムには予め分割パターンD (0 0.5),(0 0.333 0.666),(0 0.25 0.5 0.75) を用意しておき、分割パターンの各要素にNまたはRを割り当てる事によりフレーズを生成する。なお、今回の研究では音の発音時刻のみに着目し、扱うリズムは一定の周期構造を持つリズムとした.

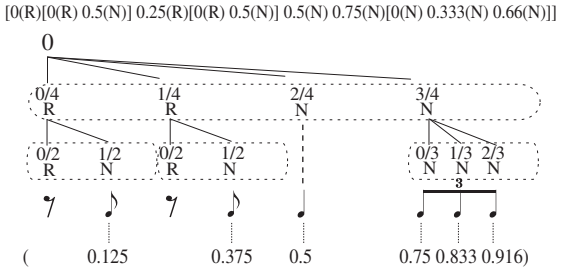


図 1: リズム階層構造の例

4 リズム階層構造探索

階層構造によるリズムの発現の仮定を基に、逆問題として、観測されたリズムから階層構造を推定するアルゴリズムを説明する。

4.1 調和リズム関数

演奏者が、図 1 のようなリズム構造を意図していた場合、観測系列から休符や大局的な構造を推定する必要がある。その為に、調和リズム関数 (Rhythm Harmonic Function) を定義する。

観測リズム系列 $(X|x_1, x_2, ...x_i, ...x_n)(0 \leq x_i < 1)$ の調和リズム関数の構成法を、注目するリズム（サブセットリズム）の長さを 1 として以下に説明する。

発音時刻の揺らぎを吸収する為の分散値 σ をもとに、打点 x_i が持つ k 倍調和リズム関数 $e_k(x_i, t)$ を次の様に定義する。

$$e_k(x_i, t) = \sum_{h(k)} \exp\left[-\frac{\{t - (x_i \pm h)\}^2}{2\sigma^2}\right]$$

ただし、$k = 2$ のとき $h = 0, 1/2$, $k = 3$ のとき $h = 0, 1/3, 2/3$, $k = 4$ のとき $h = 0, 1/4, 2/4, 3/4$, $k = 6$ のとき $h = 0, 1/6, 2/6, 3/6, 4/6, 5/6$, $k = 8$ のとき $h = 0, 1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8,$ とする。

それぞれの k 倍調和リズム関数を各打点について次のように掛け合わせ、観測リズム X の k 倍調和リズム関数 $E_k(X, t)$ を得る。

$$E_k(X, t) = \prod_i e_k(x_i, t)$$

Analysis and Generation of Musical Rhythms with Hierarchical Structure

† Shunichi KASAHARA (kasa@shalab.phys.waseda.ac.jp)
† Shuji HASHIMOTO(shuji@waseda.jp)
Department of Applied Physics, Science and Engineering, Waseda University (†)

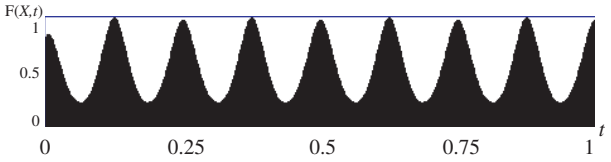


図 2: リズム (0.125 0.375 0.5 0.75 0.833 0.916) の調和リズム関数 (Normalized)

それぞれ k 倍調和リズム関数の重み付き和を X の調和リズム関数とする。

$$F(X,t) = \sum_{k=2,3,4,6,8} a_k E_k(X,t)$$

この関数により、観測系列がどのようなフレーズを基に発現しているか判別困難である場合でも、”暗 ”に意図された階層上位のフレーズを推定する事が可能となる。a_k は k が整数倍の関係にあるときに、ピーク発生の重複を考慮し、決定すべきであるが、今回はすべて 1 とした。

4.2 フレーズの推定

観測リズムの調和リズム関数に、システムに用意された各分割パターンの分割位置の値 $D = (d_1, d_2, ...dj, ...d_m)$ を代入し、その平均をパターン選択評価値 $H(X,D)$ とする。

$$H(X,D) = \frac{1}{m}\sum_j F(X,d_j)$$

この値が最も大きい分割パターンによって、観測リズムを分割する。

4.3 探索アルゴリズム

上記の様に選択された分割パターンに沿って、観測リズムをサブセットリズムに分け、分けられたサブセットリズムが、先頭にのみ打点を持つか、又は打点を持たない場合、そのサブセットが階層構造の終端フレーズとなる。また、分割可能な場合、同様のアルゴリズムによって終端を得るまで繰り返し、最終的にフレーズによる階層構造を得る。このアルゴリズムにより、階層的に分割を行うので、大局的リズムの優先推定、可変分解能での閾値処理が可能になると期待できる。なお、音符有無の閾値処理、休符位置の推定、評価値の更新、階層構造終端判別については省略する。

5 時系列推定実験

5.1 実験条件

今回は観測リズムから階層構造を推定し、それから発現する時系列を推定時系列とする時系列推定実験を行った。実験データは BPM=120、1 小節のジャズ、サンバ、シャッフル、バウンス等を含むリズム 5 種類を被験者 5 人が任意の方法で演奏したリズムを観測したものを用いた。意図したリズムに完全一致したもの

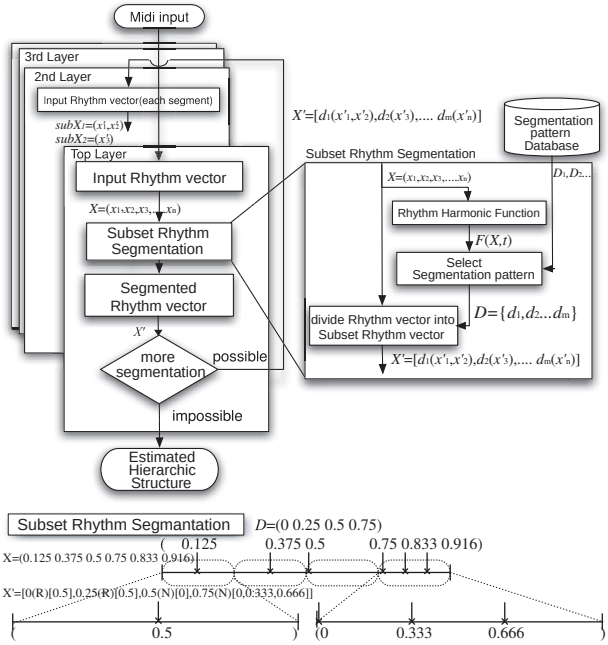


図 3: 階層構造探索アルゴリズム

のみを正解をして、認識率を求めた。比較対象は、固定分解能 (48) での閾値処理とした。また、諸関数に用いる分散値 σ は 0.02 とした。

5.2 実験結果

全体の認識率は提案システム 77.2%、固定分解能 27.6%であり、時系列認識で有効性が見られた。特に提案手法により、観測系列の平行ズレや、大局的なリズムの誤認識に対して改善を確認した。また、構造推定により、拍子推定の可能性も確認する事が出来た。しかし、発現するリズムは同一でも階層構造が一意でない例が見られた。

6 リズムアレンジへの展望

階層構造を用いる事により、大局的リズムアレンジ、イーブンフィールからバウンスフィールへの変換、単純リズムの複雑化など、様々なアレンジが可能になる事が確認されたので、構造推定と合わせたインタラクティブパフォーマンスへの利用を考えていきたい。

7 おわりに

本稿では、階層構造を仮定したリズム認識を提案し、その有効性を実験により確認した。今後は階層におけるマルコフ性を用いたより音楽的な階層性の推定と、リズム階層性の利用法を検討したい。
http://www.phys.waseda.ac.jp/shalab/music-j.html

参考文献

[1] 末富大剛, 中島 祥好:リズム知覚研究の動向, 音楽知覚認知研究,1998, Vol. 4,No. 1, 26-42

[2] Ronguet-Higgins, H.C.:*Mental Processes*, The MIT Press(1987).