

ベースギターの自動採譜システム

5 Z-7

菊地 淑晃 後藤 真孝 村岡 洋一

早稲田大学 理工学部

1. はじめに

本稿では、エレクトリックベースによって単独演奏された音楽を対象にした自動採譜システムについて述べる。従来の自動採譜システム [1][2] や音源分離システム [3][4] では、主にピアノなどの鍵盤楽器や吹奏楽器を扱っており、ベースやギター等の弦楽器はほとんど扱われていなかった。これらの研究では、音高および音色（楽器種）を同定していたが、楽器をどのように弾いたかという奏法の種類は判別していなかった。また、楽譜だけでなく弦楽器に固有のタブ譜を出力するトータルな自動採譜システムは報告されていなかった。

本研究では、ベースのみによって演奏された音響信号を入力とし、5種類の代表的な奏法の判別できる自動採譜システムを実現した。本システムは楽譜・タブ譜・標準 MIDI ファイルの3種類の形式で出力できる。これにより、楽譜の読めないベジストや奏法を自分で判断しながら演奏するのが困難な初心者にとっても、奏法付きのタブ譜があることで自動採譜結果を有効に活用できる。

2. 実現上の課題と解決法

1. 奏法を判別する特徴量の検討

奏法の違いを周波数の変化として捉えるために、各奏法で演奏したサンプルに対して周波数解析を行ない、特徴量として何が有効かを検討した。その結果、発音時のパワー、発音直後に周波数が下降・上昇・安定する時間変化、定常部での基本周波数と高調波のパワー比、定常部の有無、前音との無音間隔に奏法の違いが現れていた。そこで各音についてこれらの特徴量を抽出し、5種類の奏法を判別する。

2. 音の高さと音符の長さの同定

一般に、発音時刻直後は周波数変動が大きいいため、音の高さはそれに続く周波数変動が小さい音の定常部で求める。音高は、まず定常部の各時刻において基本周波数と高調波成分との関係から音高の候補を出し、次にその時間的な連続性を考慮して同定する。一方音符の長さは、発音時刻からこの定常部の終りまでの期間とする。

3. 楽譜・タブ譜を出力するための基準拍と運指の決定

楽譜・タブ譜を出力するためには、音符の長さを16分音符の分解能に量子化する必要がある。そのため、全発音時刻の間隔のヒストグラムから4分音符の長さに相当する基準拍を求める。また、タブ譜とはど

の弦のどのフレットを弾けば良いかを示す楽譜である。運指は一意に決まるものではなく複数存在するが、ここではできるだけ指の移動量が少なくなるものを1つ出力する。

3. 処理

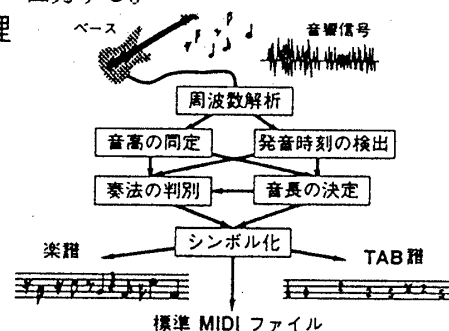


図1: 処理の流れ

本システムの処理の流れを図1に示す。ベースの演奏を入力とし、周波数解析を行なう。音高の同定、発音時刻の検出、音長の決定を行なった後、奏法の判別を行ない、最後にシンボル化を行なう。

3.1 周波数解析

解析法には、基本ウェーブレット関数を Gabor 関数としたウェーブレット変換を用いた。ウェーブレット変換は FIR バンドパスフィルタのフィルタバンクを用いて実現した。バンドパスフィルタの中心周波数は、平均率音階の半音ごとに F0 から E9 に設定した。

3.2 音高の同定

ある時刻において、F0 から E9 の各音に対して基本周波数および高調波成分におけるパワーを調べる。そして、これらのパワーの合計が最も大きくなる基本周波数を、その時刻で鳴っている音高とする。その音高が 50ms 以上持続する場合に音の定常部とみなし、そうでなければノイズとして無視する。また、波形全体のパワーの合計値が、近傍の極大値に対して 0.2 以下の比率のときには無音区間とみなす。

3.3 発音時刻の検出

音が鳴る際にはパワーが急激に増大するので、これを発音時刻として検出する。そのために、パワーの時間変化の1次微分を求め、ある閾値を越える時刻を発音時刻とする。

3.4 音長の決定

音長を決定するには、各発音時刻と音の定常部との対応をとる必要がある。そこで上記のように求めた両者のうち、間隔が 15ms 以下のものを一つの音とみなして対応をとる。対応のとれない発音時刻は、次の奏法の判別において音長のないミュートと判断される。

3.5 奏法の判別

対象とした奏法は、指弾き、スラッピングアップ、ス

ラッピングダウン、ミュート、ハンマリングの5種類である。これらの奏法は表1のような特徴を持つ。

表1: 奏法の特徴

	指弾き	Slapping Up	Slapping Down	Mute	Hammering
基本周波数	下降	安定 上昇	安定 上昇	無	安定
発音時のパワー	弱	強	強	強	弱
基本周波数と高調波の比	小	小	大	無	小
前音との無音間隔	-	-	-	-	短
定常部	有	有	有	無	有

注) - :don't care

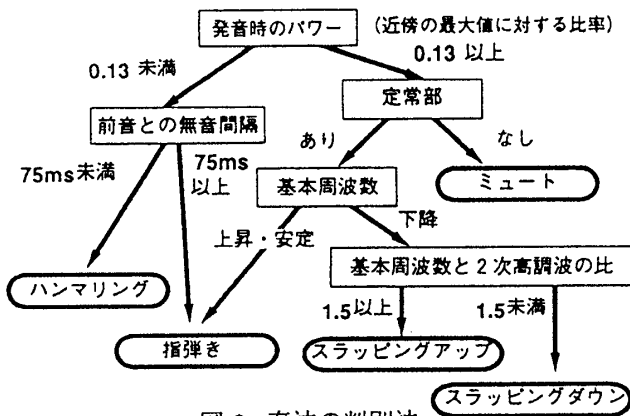


図2: 奏法の判別法

表1の結果に基づき図2の手順で特徴量を抽出することにより奏法を判別する。各特徴量の抽出は以下のように行なう。

- 発音時刻でのパワーと近傍の極大値に対する比率
過去 1000ms の近傍における極大値で、発音時刻でのパワーの合計を割る。
- 基本周波数の安定・上昇・下降
発音時刻から定常部の開始時刻までの期間に注目して判定する。まず、この期間が 30ms より短い場合、発音後すぐに周波数の変動が少ないため安定しているとする。次に、定常部の音高の上下短 3 度の各範囲内にあるすべての音のパワーの合計値を求め、下側のパワーが上側より大きければ上昇しているとし、逆であれば下降しているとする。
- 前音との無音間隔
前音の定常部の終了時刻と発音時刻との差を求める。
- 基本周波数と 2 次高調波の比
それぞれのパワーを、定常部の開始から 100ms の期間合計したものの比である。

4. シンボル化

音高と音長、奏法から楽譜・タブ譜・標準 MIDI ファイルへシンボル化を行なう。その際、各音長を音符に変換するために、4 分音符の長さに相当する基準拍が必要になるが、これを正確に獲得するのは難しい。そこで、全発音時刻の間隔のヒストグラムをとり、400~1333 (ms) 間 (テンポが 45~150 に相当) の最大ピークとして求める簡略化した実装をおこなった。

一方タブ譜を出力する際に必要な運指は、まず最初の音をできるだけ低いフレットとし、それ以降は、前のフレットからできるだけ近いフレットに移動するように決定した。奏法の種類はタブ譜の各音符の下に併記する。

5. 実験結果

実際にベースを演奏し、Line 入力、16bit、44.1kHz でサンプリングした 6 サンプルに対し実験を行なった。本システムの入出力例を図3に示す。タブ譜の下側の文字は奏法を表し、F は指弾き、M はミュートである。2 分音符に相当する長さを基準拍として誤検出したため、得られた楽譜では倍のテンポになっている。



図3: 演奏楽譜と出力結果 (上段:演奏楽譜, 中段:出力楽譜, 下段:タブ譜)

(1) 式 [4] により算出した認識率を表2に示す。

ただし、 Nn_t は入力に含まれる総音符数 (105 音)、 Nn_r は出力に含まれる正しく認識された音符の数、 Nn_w は出力に含まれる音符のうち、誤って認識されたか余分に出力された音符の数である。

$$Rn = \left(\frac{Nn_r - Nn_w}{Nn_t} \cdot \frac{1}{2} + \frac{1}{2} \right) \times 100 \quad (1)$$

表2: 認識率

	認識率 (%)
音高	89.05
音長	88.10
奏法	83.33

6. まとめ

以上、ベースギターを対象にした自動採譜システムについて述べた。本システムは有効に機能し、音高や音長だけでなく 5 種類の奏法も認識することができた。

今後は、本システムの認識率を高め、楽器の種類の制限をなくしていく予定である。複数の楽器が鳴っている演奏を自動採譜できるシステムを目指していきたい。参考文献

- [1] 片寄・井口、知的採譜システム、人工知能学会誌、Vol.5、No.1、pp.59-66(1990-01)
- [2] 長束・才脇・井口、異種楽器を対象にした採譜システム、電子情報通信学会春季大会 D-499、1992
- [3] 戸谷・武石・袖山、異種楽器音の分離に関する研究、電子情報通信学会春季大会 A-180、1989
- [4] 中臺・柏野・田中、音楽音響信号を対象とする音源分離システム、音楽情報科学研究報告 No.1、1993