

事例データに基づくエレキギターの表情付けシステム 「Guitar-Case Maker」

水本直希^{†1} 北原鉄朗^{†2} 片寄晴弘^{†1}

近年、DTM といった音楽環境の構築が容易になったこともあり、音楽の表情付けに関する研究が盛んである。しかし、既存研究ではピアノにおける表情付けがほとんどであり、ギターやバイオリンといった、演奏者が発音後にピッチやダイナミクスを制御できる持続系楽器の表情付けについてはあまり扱われてこなかった。本研究では、ポピュラー音楽において使用頻度の高いエレキギターを対象とし、MIDI データにおけるエレキギターの表情付けを行うシステムの開発を目的とする。本稿では開発したプロトタイプシステムの概要と、その動作結果について報告する。

Guitar-Case Maker: Case-based Performance Redering System for Electric Guitar

NAOKI MIZUMOTO,^{†1} TETSURO KITAHARA^{†2}
and HARUHIRO KATAYOSE^{†1}

The recent increases of non-professional people who create their own musical pieces have demanded technologies for rendering expressive musical performances. Most of the existing studies for performance rendering, however, targeted the piano's performances, so they did not deal with sustained-tone instruments, the pitch and dynamics of which can be continuously changed. Focusing on the electric guitar as a sustained-tone instrument, we have been developing a performance rendering system for electric guitar solo phrases. In this paper, we describe the design and implementation of our system and report preliminary results.

^{†1} 関西学院大学 理工学研究科

KGraduate School of Science and Technology, Kwansai Gakuin University

^{†2} 日本大学 文理学部

1. はじめに

エレキギターは、現代のポピュラー音楽において重要な位置づけにある。そのため、DTM^{*1}を用いたアマチュアによる音楽制作においても、エレキギターの音色が使われることが少なくない。しかし、エレキギターは、自然な演奏を DTM により表現するのが困難な楽器として知られており、自然な演奏を自動あるいは半自動的に生成する計算機技術が望まれている。

エレキギターの演奏生成が難しいのは、エレキギターではチョーキング、ハンマリング・オン、プリング・オフといった様々な奏法が使用され、それらをピッチの細かな連続的制御により、適切に表現し分ける必要があるからである。これまで、演奏の自動生成（演奏表情付けという）に関する様々な研究^{1)–8)}が行なわれてきた。しかし、そのほとんど^{1)–6)}がピアノを対象としており、エレキギターで多様される発音後のピッチの連続的制御は扱っていなかった。一方、発音後にピッチやダイナミクスを制御可能な楽器（持続系楽器という）を扱った研究として、サクスを対象とした SaxEx⁷⁾、バイオリンを対象とした清水らの研究⁸⁾がある。しかし、これらの研究で提案された技術が、エレキギターに見られる多種多様な奏法に対して有効かどうかは検証されていない。また、エレキギターを対象とした研究として中西らの研究⁹⁾があるが、ペロシティの制御のみを扱っており、ピッチの連続的制御は扱っていなかった。

本研究では、エレキギターのソロフレーズを表す MIDI データに対して、ピッチバンド情報を付加することで、自然なピッチ遷移を実現するシステムを開発する。具体的には、様々な旋律にピッチバンド情報があらかじめ付与されたデータからなる事例データベースをあらかじめ用意し、入力された MIDI データに旋律の特徴に近いものをこのデータベースから探索し、探索された事例に付与されているピッチバンド情報を利用する。事例データベースとして特定のギタリストによる演奏事例のみを用いることで、そのギタリストの特徴を反映した演奏表情づけが可能になる。

2. 演奏表情付けにおける課題

エレキギターは撥弦楽器の一種であるが、他の撥弦楽器と比べて特徴的な事項として、撥

College of Humanities and Sciences, Nihon University

*1 DTM (Desktop Music): パソコンと電子楽器を MIDI ケーブルなどによって接続し、演奏する音楽や音楽制作行為の総称。

弦して音が鳴っている状態でピッチを変化させる奏法が多用されることがあげられる。特に、いわゆる歪み系のエフェクトを使用したときは撥弦から長時間音が持続するため、このような奏法が頻繁に使用される。

撥弦後にピッチを変化させる奏法は、「チョーキング」「ハンマリング・オン」「プリング・オフ」「スライド」「ビブラート」「グリッサンド」「アーミング」といった名前で分類される。以下、エレキギターで多用される特殊奏法を概説する。

- チョーキング：弦を指板と平行に押し上げたり押し下げたりして音高を変化させる
- ハンマリング・オン：滑らかに音高を上げる場合に用い、弦を指板に叩き付ける様に押さえ付けて音を出す
- プリング・オフ：滑らかに音高を下げる場合に用い、左手の指で弦を弾いて音を鳴らす
- スライド：押弦した指を滑らせることによって、始点と終点の音を途切れさせずに滑らかに音高を変えていく
- ビブラート：弦を上下、又は左右に微妙に動かして音を震わせる
- グリッサンド：一音一音を区切ることなく、隙間なく滑らせるように流れるように音高を上げ下げする
- アーミング：ビブラート・ユニットを使用し、弦を伸縮させることによって音程を連続的に変化させる

これらの特殊奏法が用いられる条件と各特殊奏法の演奏の仕方は、ある程度決まってくるものの、どちらも一意ならず、ギタリストによる違いが存在し、それがギタリストの個性となっていると考えられる。そこで、本研究では、単にギター演奏として自然な演奏を生成するだけでなく、ユーザが指定したギタリストの個性をできるだけ反映した演奏を生成するようにする。

2.1 問題設定

表情付けが行われていない MIDI データを入力とし、ギタリストの演奏特徴が反映された MIDI データの生成を出力とする。本システムの表情付けの対象は単旋律であるギターソロのみとし、ピッチバンドと音量の2つを扱うこととした。

2.2 表情付けのアプローチと課題

表情付けを行う際に用いられるアプローチとして、大きく jPop-E¹⁰⁾ のようなルールベースと Kagurame¹¹⁾ のような事例ベースの2つに分けられる。ルールベースは、音楽に関する知識や概念をルールという形で記述し、そのルールに基づいて音楽を生成する。この手法は、楽曲の特徴と演奏表情との関連付けなどを行うことによって演奏表情を作り出すことが

でき、専門家が注意深くルールを設計すれば、品質の高い演奏を生成することができる。しかし、エレキギターのピッチ制御は、ピアノ演奏に比べて学術的な分析や理論化がなされておらず、ギター演奏の専門家が持つ暗黙的な知識をルールとして書き下すことは、必ずしも簡単ではない。また、本研究では、ユーザが指定したギタリストの演奏の個性を模写することを目指しており、ギタリスト毎にその個性を反映するルールを書き下さなければならないという点でも現実的でない。一方、事例ベースは、あらかじめ見本となる演奏事例を多数用意し、その事例を利用して演奏を生成する方法である。この方法には、利用する演奏事例集を切り替えることで、生成される演奏の特徴を切り替えることができるというメリットがある。特定のギタリストによる演奏のみを演奏事例として用いることで、そのギタリストの演奏の個性を反映した演奏を生成することが可能である。

事例ベースには、処理単位（例えば、入力楽譜の1フレーズ）毎に条件に合致する事例を見つけ出して、その事例の特徴を直接反映させるインスタンスベースと、何らかのモデル（例えば確率モデル）を想定し、事例からそのモデルパラメータを推定して、そのモデルを元に演奏を生成するモデルベースが考えられる。前者は、条件に合致する事例が見つからない場合に何も演奏表情が付与されない可能性があるという問題がある（データスパースネス問題という）。また、後者には、精度よくモデルパラメータを推定するには、大量のデータが必要である。また、出現頻度の低い特徴的な演奏が、統計的推定によって埋もれてしまうという問題がある。

本研究では、特定のギタリストの演奏事例を大量に用意するのは困難であることから、インスタンスベースの方法を採用する。データスパースネス問題に対しては、事例が見つからないときに条件を段階的に緩くするという方法を取る。

3. エレキギターに対する演奏表情付け手法

本章では、前章の議論に基づいて設計した演奏表情づけ手法について述べる。

3.1 処理の概要

本手法では、2章で述べたように、ギターソロの単旋律の楽譜を入力とし、それにピッチの揺らぎ（ピッチバンド）を付与した MIDI データを出力する。個々の音符を処理単位とし、音符ごとに条件を満たす事例を事例データベースから見付け出し、その事例に付与されているピッチバンドを対象楽曲に付与する。具体的には、対象楽曲の旋律（音符列）を N_{t-k} とすると、各 t に対して、対象音符 N_t の近傍 N_{t-k}, \dots, N_{t+k} の特徴が一致する旋律断片を探し出し、その旋律断片に付与されているピッチバンド情報を利用する。特徴が一

致する旋律断片が見つからなかった場合は、考慮する近傍の範囲 k を小さくし、複数見つかった場合は、 k の値を大きくして、事例が 1 つだけ見つかるまで事例探索を繰り返す。

3.2 特徴抽出

対象楽曲の各音符 N_t に対して、

- 近傍 $k + 1$ 個の音符 N_{t-k}, \dots, N_{t+k} の音高, 音量, 音長 (但し, 音高は N_t の音高に対する相対値)
- 用いられている奏法名のタグ情報を抽出する。

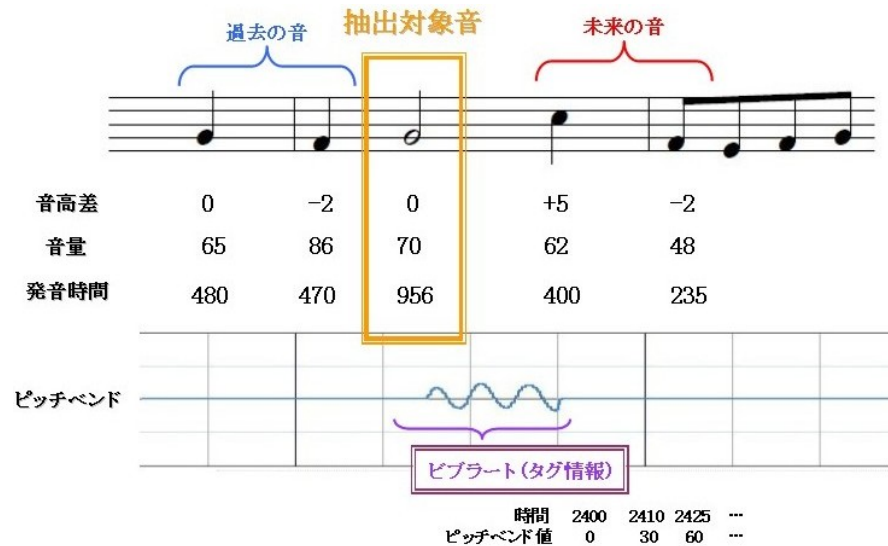


図 1 特徴抽出の例

ピッチベンドに関しては、全データを保持するのは冗長で、また、MIDI データ作成時に製作者が GUI を用いてフリーハンドで入力している場合が多いので、不必要な微妙な揺れが存在する。そこで、MIDI データ中のピッチベンドデータに対して平滑化を行った後、変化量が急激に変わる点 (特徴点と呼ぶ) の時刻とピッチベンド値のみを保持する。

特徴点の抽出は、基本的には平滑化されたデータに対して極大値・極小値を抽出することで行う。しかし、単純に極大値・極小値を求めてしまうと、平滑化によって平坦な区間や一瞬だけ大きく変化している区間を無視してしまう。そこで、この問題を長区間単純移動平均の結果と短区間単純移動平均の結果を合わせることによって解消する。長区間単純移動平均による平滑化によって、ピッチベンドの大まかな旋律の流れが得られるため、平坦な区間の中点あたりを極値として残せる。また、短区間単純移動平均による平滑化によって、一瞬だけ大きく変化している区間を極値として残せる。これら 2 つの異なる区間を持つ単純移動平均を用いることによって、ピッチベンドの特徴点を抽出することができる。

3.3 事例探索

入力楽曲中の音符列 N_{t-k}, \dots, N_{t+k} と特徴が近い音符列を探索するため、音符列間の類似度関数を定義する。類似度関数 R_t は、音高差に対する類似度 L_1 、音長に対する類似度 L_2 、弦情報に対する類似度の L_3 の重み付け和

$$R_t = \frac{k^2}{k^2 + 1} \sum_{n=1}^3 w_n L_n$$

で定義する。

選択音の音高差 P_t に対する類似度 L_1 は、抽出対象音の音高差 P'_t と $k + 1$ 個の近傍音の音高差を用いて、

$$L_1 = \prod_{n=-k}^k \operatorname{sech}(P_{t+n} - P'_{t+n})$$

同様に、選択音の音長 G_t に対する類似度 L_2 は、抽出対象音の音長 G'_t と $k + 1$ 個の近傍音の音長を用いて、

$$L_2 = \prod_{n=-k}^k \operatorname{sech}\left(\frac{G_{t+n} - G'_{t+n}}{0.1G'_{t+n}}\right)$$

同様に、選択音の弦情報 F_t に対する類似度 L_3 は、抽出対象音の弦情報 F'_t と $k + 1$ 個の近傍音の弦情報を用いて、

$$L_3 = \prod_{n=-k}^k \operatorname{sech}(F_{t+n} - F'_{t+n})$$

3.4 類似事例の探索

前節で定義した類似度関数を用いて、類似事例を探索する手法について述べる。今、 t 番目の音符 N_t に対して表情づけを行うとすると、次のアルゴリズムによって、類似事例を探索する。

- (1) 近傍音の音符数 k を初期値にする
- (2) 事例データベースの各音符に対して、 N_t との類似度を 3.3 節で定義した類似度関数を使って求め、類似度が高いものから 3 つ選び出す（但し、類似度が事前に設定した閾値を下回るものは除外する）
- (3) (2) で選んだものが全て閾値処理によって除外された場合、 k の値を 1 減じて以上の処理を再び行う
- (4) (2) で選んだものの中で類似度が等しいものがある場合、 k の値を 1 足して以上の処理を再び行う

3.5 特徴の転写

前節の方法により、各音符に対して転写元の事例が決まったら、転写元の MIDI データに付与されているピッチベンド情報を対象音符に転写する。

類似度に基づく特徴転写を行う際、転写元の MIDI データと転写先の MIDI データの音価が異なる場合がある。すると、ピッチベンドといった時々刻々と変化するパラメータは音価に比例して長さを調節しなければならない。本システムは、転写元の MIDI データの音価とピッチベンドの時間変化量との比率が変わらないように転写した。

また、音量変化を付与する際、転写先の MIDI データの対象音のベロシティ V'_{input} は、転写元の MIDI データの対象音の一つ前の音のベロシティ V_{input} 、転写元の MIDI データの対象音のベロシティ V'_{data} 、転写元の MIDI データの対象音の一つ前の音のベロシティを V_{data} を用いて、

$$V'_{input} = V_{input} \frac{V'_{data}}{V_{data}}$$

とする。このようにすることで、転写先の MIDI データに表情付けが行われる際に、転写元の MIDI データの音量変化が転写先の MIDI データに反映される。

4. 実装・実験

3 章で述べた特徴抽出法と特徴転写法に基づき、プロトタイプシステムを開発し、2 つの実験を行った。

4.1 システム実装

本システムでは、様々な表情付けを実現するため、GUI 上で以下のパラメータを操作できるようにした（図 2）。

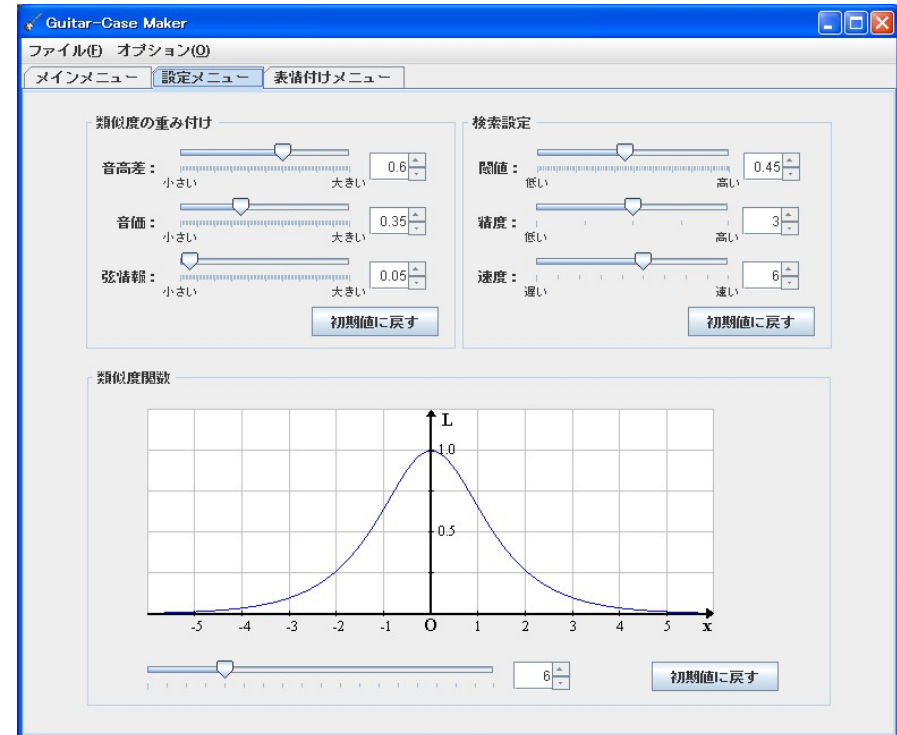


図 2 Guitar-Case Maker の GUI 画面

- 類似度の重み付け
 - 類似度の重み係数 w_1 類似度を計算する際の音高情報に対する重み
 - 類似度の重み係数 w_2 類似度を計算する際の音価情報に対する重み
 - 類似度の重み係数 w_3 類似度を計算する際の弦情報に対する重み
- 検索設定

- 検索アルゴリズムの閾値 検索における、転写候補の類似度の下限
- 検索アルゴリズムの精度 検索における、近傍音の音符数 k の初期値
- 検索アルゴリズムの速度 検索における、検索範囲の制限

- 類似度関数 $f(x)$ (但し、設定可能範囲は $2 \leq x \leq 20, 0 \leq f(x) \leq 1$)

操作できるパラメータは大きく類似度の重み付け、検索設定、類似度関数 $f(x)$ の設定の3つに分けられ、それぞれ独立にシステムの設定を変更することができる。

4.2 実験

プロトタイプである本システムの有用性を検証するために、いくつかの実験を行った。データベースを構成する事例データとしてギタリストの演奏を再現した MIDI データを用意したが、他に MIDI ギターを用いて直接ギタリストの演奏データを取る方法が考えられる。しかし、この方法では全ての楽曲に対して演奏情報を記録しなければならず、データの収集は困難を極める。本研究では、MIDI ギターで得られる程度のデータの質を確保するため、市販されている MIDI データや、ギタリスト本人からも認められるほど質の高い MIDI データを事例データとして採用し、B.B. キング、ブライアン・メイ、ラリー・カールトンの3名のギタリストのデータベースを構築した。

4.2.1 タグ情報の有用性検証実験

既存の MIDI データからデータベースを構築する際、データの一部に対して、既存の研究では用いられていない奏法に関するタグ情報を付加した。しかし、タグ情報を利用することが表情付けに有用であるかは不明である。本節では、タグを利用しない場合の表情付けの結果と比較することによってタグ情報が有用であるかを検証する。

ブライアン・メイの楽曲6曲において、タグ情報が付加されていないものとタグ情報が付加されたものの2種類を用意した。前者の6曲をグループ A、後者の6曲をグループ B とし、新たに用意した未知のフレーズを含む楽曲 X と楽曲 Y の2曲に対して、グループ間で同一演奏箇所での表情付けの結果を比較する。システムに関するパラメータは、抽出する近傍音の個数 k 、転写する際に再検索を実行する回数 n 、類似度の閾値 t の3つがあるが、本実験では、 $(k, n, t) = (3, 3, 0.05)$ とした。

実験によって生成されたデータの一部を図3に示す。図3は (a) がグループ A の楽曲 X に対する結果 (b) がグループ B の楽曲 X に対する結果 (c) がグループ A の楽曲 Y に対する結果 (d) がグループ B の楽曲 Y に対する結果である。グループ A の結果とグループ B の結果をそれぞれ比較すると、いずれの楽曲においてもタグ情報を用いているグルー

プ B の結果の方が適切な表情付けを行っていた。例えば、グループ A の場合、特徴転写を行う際に抽出対象音の近傍の音符情報から計算される類似ランクによって事例データが検索されるので、表情付けとして違和感のある箇所が存在する。しかし、グループ B の場合、事例データはデータベースにあるタグ情報に関する事例データに対してのみ検索されるため、グループ A より正確に表情付けが行われている。ただ、生成されたデータに幾つかの問題点がある。本実験で用意した楽曲数が少なく、タグ情報に関する事例データが少ないために、転写を行う検索結果が無い場合が発生し、タグ情報が付与されていない箇所が大半であったのでピッチベンドの変化の違いが少なく、曲の全体的な印象は余り変わらないものとなってしまった。また、楽曲中ではタグ情報が付加されていない音符が大半を占めるにも関わらず、ほとんどの箇所ですべての事例データを転写してしまっているため、生成されたデータのピッチベンドの値が常に変化し、音高が不安定になり、まともな演奏とは言えないものになってしまった。

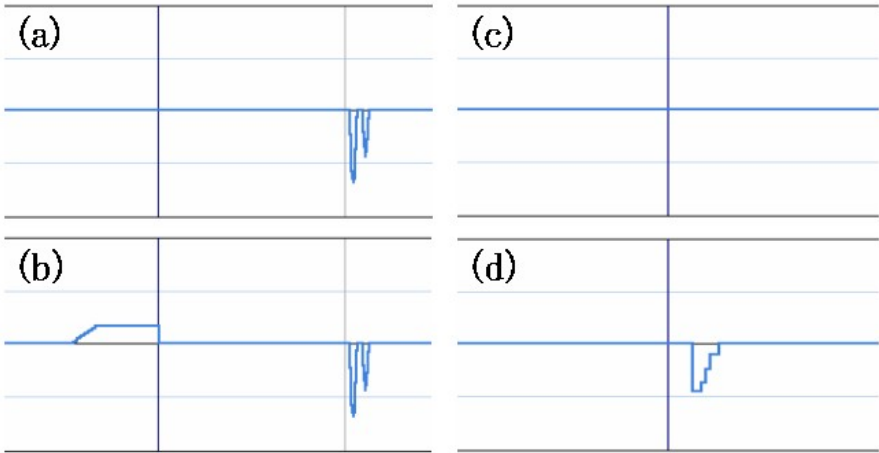


図3 タグ情報の有用性検証実験結果の一部

4.2.2 データ量の違いによる性能比較実験

本研究では、60曲程度のギタリストの演奏を模写した MIDI データを表情付けの生成に使用した。しかし、楽曲数やジャンルなどの異なる条件下の MIDI データを用いた場合に

においても同様の振る舞いをするのが望ましい。本節では、提案したシステムが表情付けに用いるデータ量の違いによって結果がどれほど異なるかを検証する。

4.2.1 項の実験で使用したタグ情報が付加されていないブライアン・メイの楽曲 6 曲と、それら以外のブライアン・メイの楽曲 60 曲を用意した。前者の 6 曲をグループ C、前者と後者の楽曲を合わせた 66 曲をグループ D とし、新たに用意した未知のフレーズを含む楽曲 X と楽曲 Y の 2 曲に対して、表情付けの結果を比較する。システムに関するパラメータは、抽出する近傍音の個数 k 、転写する際に再検索を実行する回数 n 、類似度の閾値 t の 3 つあるが、本実験では、 $(k, n, t) = (3, 3, 0.05)$ とした。

実験によって生成されたデータの一部を図 4 に示す。図 4 は (a) がグループ C の楽曲 X に対する結果 (b) がグループ D の楽曲 X に対する結果 (c) がグループ C の楽曲 Y に対する結果 (d) がグループ D の楽曲 Y に対する結果である。グループ C の結果とグループ D の結果をそれぞれ比較すると、いずれの楽曲においてもデータベースにある事例データの量が多いグループ D の結果の方が適切な表情付けを行っていた。例えば、楽曲 X の音符数は 54 個、楽曲 Y の音符数は 34 個あるが、いずれもグループ D の方が、入力データに対して表情付けが行われている箇所が少ない。これは、検索する事例データの中に表情が付いていないものが増えたためである。通常の演奏では、全ての音符に対してピッチベンドによる表情付けが行われているとは考えにくく、表情付けがない箇所が適度見られることは、より自然な表情付けが行えていると証拠だと考えられる。

5. ま と め

本稿では、エレキギターに対する演奏表情付けを目指し、それを実現するための特徴抽出法、特徴転写法、及びプロトタイプシステムを開発した。システムのパラメータを調節することで様々な表情付けが可能となったが、その制御には問題点がいくつかあり、また表情付けがされた MIDI データの出力結果にも改善の余地を残す結果となった。今後は、より自然でかつギタリストの個性が反映できるよう、改良を重ねていく予定である。

参 考 文 献

- 1) 星芝貴行, 堀口進: 複数の演奏データを用いた標準的演奏の導出, 日本音響学会音楽音響研究会資料, MA96-13, pp.31-38 (1996)
- 2) 水谷哲也, 田中崇之, 五十嵐滋: 構造機能に基づく演奏法則と演奏創造, 人工知能学会大会 (2003)
- 3) 彌富あかね, 五十嵐滋: 音楽構造分析を用いたピアノ演奏の表情付け, 第 50 回情報処大会 (1995)

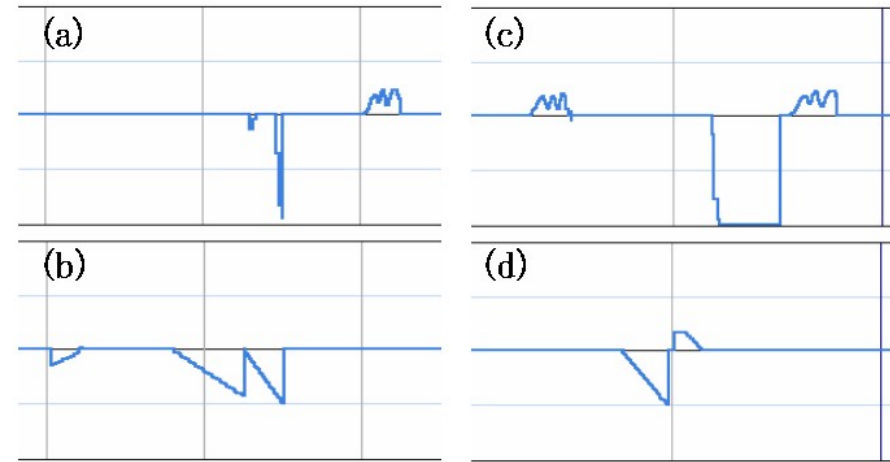


図 4 データ量の違いによる性能比較実験結果の一部

- 4) 小田彦彦, 白川健一, 村上豊, 梶川嘉延, 野村康雄: 演奏者情報加味したピアノの自動演奏システムの構築～ニューラルネットワークによる局所部分に於ける演奏特徴の抽出～, 情報研報, pp.7-12 (1995)
- 5) 木野弘樹, 梶川嘉延, 野村康雄: 特定ピアノ演奏者の特徴抽出による自動演奏システムの構築, 第 59 回情報処大会 (1999)
- 6) 森口徳崇, 山梶雄一郎, 森田慎也, 三浦雅展: ピアノ演奏を対象とした感情付与システムの基礎的検討, 日本音響学会音楽音響研究会資料, pp.7-12 (2008)
- 7) Josep Lluís Arcos, Ramon Lopez de Mantaras, Xavier Serra: SaxEx: a case-based reasoning system for generating expressive musical performances, International Computer Music Conference (1997)
- 8) Satoshi Shimizu, Saori Yoshinaga, Koichi Furukawa: Feature Extraction of Violinists by Rough Sets, The 19th Annual Conference of the Japanese Society for Artificial Intelligence (2005)
- 9) 中西正洋, 門田暁人, 松本健一, 井上克郎: ギターの演奏情報の抽出と分析, 第 59 回情報処大会 (1999)
- 10) Mitsuyo Hashida, Noriko Nagata, Haruhiro Katayose: jPop-E: An Assistant System for Performance Rendering of Ensemble Music, NIME, pp.313-316 (2007)
- 11) 鈴木泰山, 宮本朋範, 西田深志, 徳永健伸, 田中穂積: Kagurame Phase I 事例ベースの演奏表情生成システム, 第 57 回情報処大会, pp.13-14 (1998)