



# SIGMUS 第14回 研究発表会 質疑記録

1996年2月24日 情報処理学会 会議室 (東京・芝浦)

## (1) GA を用いたピアノ演奏の局所テンポ解析

山内敏彰, 野池賢二, 野瀬隆, 乾伸雄, 小谷善行, 西村恕彦  
(東京農工大)  
記録: 藤島 琢哉 (ヤマハ), 鈴木 孝 (東京高専)

Q: 堀内 (千葉大) 学習のやり方として遺伝的アルゴリズム (以下 GA) を選んだ理由は?

A: 手法の基本は重回帰分析だ。GA は説明変数を選ぶためだけに使用している。GA だと重回帰分析とじかに組み合わせやすく、都合がよい。ニューラルネットだと後で分析ができない。

Q: GA の遺伝子組み替えなどの側面も含めて大事だということか。

A: GA は他の分野でも用いられている。コンピュータのパワーを用いるのに向いている。予想以上に強力である。

Q: 言葉は悪いが、結果オーライということか。

A: そういう面もある。

Q: サンプルデータの拍子や調性はどのように配慮したか? たとえば3拍子の曲も学習は4拍子で行なうなどとなるとまずいと思う。

A: 被験者が知らない曲を選んだ。被験者は大学生で、プロではない。調性や拍子はばらばらとなるよう選曲した。

Q: 平賀 (図書館情報大) 音高については絶対音高でやっているのか。

A: そうである。

C: 調をそろえると結果が違ってくるかも知れない。

Q: 演奏者はどの程度の習熟度か? 素人だと弾くたびに違うということはないか。

A: 10数年の習熟度である。研究室で短時間で演奏した場合と、2週間かけた場合がある。問題だったかも知れない。

Q: 橋本 (早稲田大) 説明変数として積を取っている理由は?

A: 2次まで利用しようとしたからである。

Q: えられた結果について説明できるか。

A: 重回帰分析は結局は近似である。音高は全て正にとったので、積は AND の意味になる。一方の値が高くて、もう一方の値が低いとだめ、というようなことだ。

Q: 音高も他の積で評価すべきではないか。係数は1次と2次でどう違うのか。1次が支配的とかいったことはあったか。

A: 説明変数どうしの相関が高いと重回帰分析では具合が悪いので、その場合は、どちらか一方を選んで使うようにしている。これは手作業でやっている。GA を使うのではない。

C: 2つの積ではなく、1つの変数の巾(2次, 3次, ...) でも良かったのではないか。

C: 平賀 2変数の積を見たいという発想もわかるが、... 一変数の巾だと、結局テラー展開したのと同じ結果になるのだろうか。

C: 橋本 多少は違うが大体同じ結果が出るはずだ。

Q: 南高 (カシオ) 単旋律以外はどうか。

A: 評価要素の選択が重要になると思う。多旋律の場合だとテンポの扱いが難しくなる。メロディを抜き出して使うとか、細かい音符を選ぶということになるだろう。合奏になるとまた難しいだろうと思う。

Q: 平賀 未知曲 (学習には使わず、システムでテンポゆらぎをつけた曲) を人間が弾いた例はないのか。

A: ない。未知曲は弾いてもらっていない。

## (2) 長唄における唄と三味線の付き方

藤井誠, 矢向正人 (九州芸工大)  
記録: 藤島 琢哉 (ヤマハ), 鈴木 孝 (東京高専)

Q: 平賀 分析3で、ファラミの 패턴のようなものは、西洋音楽の逸音などの考え方と類似点があるか?

A: データをさらに検討したい。三味線は装飾的な音が多く、基本音と装飾音に分解ができる。唄から三味線の基本音をとりだすといったことともつながってくると思う。

Q: 32分音符刻みで分析しているとのことだが、楽譜を分析しているということは、演奏を追っているわけではないということか。

A: 演奏は反映されていない。三味線の方はあまり揺れはなく、32分音符刻みでよいと思う。唄の方はずいぶん違う。人によって揺れが大きい。

Q: 実際の曲データからの検証は?

A: 楽譜なしの曲の場合は楽譜起こしてやってみたい。今回、楽譜があるのは大正時代に採譜されたものである。

Q: 三味線の方が揺れないというのは、流派ごとに習うからだろうか。

A: 三味線は、楽譜でなく手の動きで覚える面が多いので差が出にくい。

Q: 三味線と唄の流派は対応するのか。

A: 対応する。

Q: 平田 (NTT) 分析3.4でのパターンマッチのアルゴリズムについて、最長マッチなのかどうか。同パターンが2回出現した場合のカウントのしかたなどを教えて欲しい。

A: 4は2音の動きに注目して、窓幅を決めて、1回でも現れたらカウントする。「ドシ」に注目していて、ドシドシシというのに出会った場合は、音域的対応として見る。

Q: exact の出現回数を数えるということは、音をはさんだものとしてもまた重複して数えるのか。

A: そうだ。

Q: 資料のグラフの意味だが、微分してどうしてこうなるのかわかりにくい。

A: わかりにくかったと思う。絶対値をとっているのだから。

## (3) 長唄三味線の旋律シミュレーション

鳥海英宏, 矢向正人 (九州芸工大)  
記録: 藤島 琢哉 (ヤマハ), 鈴木 孝 (東京高専)

Q: 後藤 (早稲田大) マルコフモデルを使うなら、パターンからパターンの遷移確率が最大のものを常に選ぶのではなく、生成のたびに、確率に従っていろいろな旋律が出てくる方が面白いのではないか。

A: それは検討している。

Q: 旋律は生成できるとして、できたパターンの意味は何になるのか。

- A: これは中間目標で、最終目標は音組織のパターンを集めるということである。
- Q: さらに上位のルールを抽出するなど、遷移確率の利用法は考えているか。
- A: 今回はまずシミュレーションしてみたということである。窓幅を広くとる、つまりマルコフモデルの次数を上げると、冗長な感じがよく出る。狭くする、つまり次数を下げると、実際にはないキビキビした感じになって面白い。パターンの組合せを確かめたかった。現実には近づけるといよりは、作曲ということだ。
- Q: 貝塚(早稲田大) マルコフモデルの生成だと、初期値によって出力されるものや収束が違うと思うが。
- A: 実験ではあらゆる初期値を選んでパターンを生成してみている。
- Q: 藤島(ヤマハ) OHP で示された譜例3が、予稿集と異なっているようだ。
- A: 予稿集では、学習させた曲がそのまま出てきてしまったので、より面白い旋律例に差し替えた。訂正をお願いします。
- Q: 音高と音価には相関があると思うが、今回のシミュレーションでは配慮したか。
- A: 今回は相関は配慮していない。両者を独立に学習し生成したものを組み合わせた。今後の課題である。
- Q: 曲によってスケールが異なると思うが、音高の学習の際、そろえたのか?
- A: そろえていない。曲ごとにスケールは異なるが、楽譜通り、絶対音高で学習させた。長唄ではさまざまなスケールがつぎつぎ交替するのが普通で、これで不自然ではない。
- Q: 平田 マルコフモデルは、イリアック組曲以来、30年の昔から利用されている。今邦楽で用いるメリットはどこにあるのか。マルコフモデルは構造を捨象してしまうもので、その限界を見極めずに使っているとしたら意味がよくわからないのだが。
- A: 西洋音楽のソナタなどがモチーフの反復変形という形であるのに対し、邦楽は、有限パターンの組み合わせであるので意味があると考えている。邦楽のパターンが完全に統計的に扱えるかは、わかっていない。一つの試みである。
- Q: 何を目標としているかにもよるが...新しい組合せの generate ならよいのだが、メカニズムを知りたいのだとしたら、今の手法では限界があるのではないかと。
- A: 規則があるとは思いますが、ゆるいものだと思う。
- Q: ルールがあるのは認めるが、このモデルでは相過ぎる。
- A: それはその通りで、メカニズムについては別に研究したいと考えている。
- Q: 橋本 音価パターン図の方には距離が入れているが、パターン安定度の定義の方は結局アドホックにしていたように見える。その尺度定義の妥当性はどうか。安定パターンという呼び方がおかしいのではないかと。ダイナミクスがあって、そのパターンが生じやすいというのならいいのだが。
- A: まだ統計はとっていない。距離と関係してくるかどうかも含めて研究してみたい。
- (4) すべてのプレーヤーが対等なジャズセッションシステム I. システムの全体構想と分散環境での実装  
後藤 真孝, 日高 伊佐夫, 松本 英明, 黒田 洋介, 村岡 洋一 (早大)  
記録: 小坂 直敏 (NTT)
- Q: 堀内(千葉大) 楽譜についてはどこまで与えているか?
- A: テーマのメロディとコード進行を与えている。伴奏はその場で生成している。
- Q: 計算機は自発的に演奏者にジェスチャーを送るのか?
- A: 現在の実装では、シナリオ中の各部の最大繰り返し回数に達すると送るだけである。まだ実現していないが、人間のソロを計算機が低調だと判断したときなどに、計算機の方から次の人のソロ演奏部へ移るようジェスチャーを送るようにすることなどを考えている。
- Q: 人間の奏者が(ベースとドラムスの)計算機のモニターを2カ所見るのはつらいのではないかと。
- A: 一般にシナリオ内の各部分の終り付近でジェスチャーが重要となる。全ての演奏時間ジェスチャーを見続ける必要はない。
- Q: 平田(NTT) サウンドサーバが RMCP パケットを受けとって MIDI 楽器を制御するとき、だいたい何 msec 程度ずれるのか?
- A: サウンドサーバでパケットをバッファリングし、それらのタイムスタンプで再スケジューリングしたのが割り込みで読み出されるので、事前に届いたパケットに関してはほとんどずれない。実際には、演奏情報は音が鳴るほぼ一拍前には相手に届くようにしている。
- Q: 一拍前とは 300msec 程度か?
- A: テンポによるが、数百 msec 程度。
- Q: 平賀(図情大) タイムスタンプは絶対時刻が書かれているか? それともビートに対する相対時刻か?
- A: 絶対時刻が書かれている。
- Q: 平田(NTT) RMCP のパケットの遅延時間が 1msec と短いので実時間でできないのか。
- A: 実時間でできる。
- Q: バッファリングのためのデッドラインの時刻が一拍分ということか?
- A: そうではない。現在時刻の情報はすぐ処理する。デッドラインはなく、単に事前に予約しておくようなものだ。
- Q: キャンセルはできるか?
- A: 現在は必要がないため、その機能はない。
- Q: パケットの総数はプレーヤーの個数に比例するか?
- A: 演奏中の音数などにも依存するため、比例関係かどうかは一概に判断できない。単純には、デモビデオ中の各演奏音について RMCP パケットが一つ発生すると思えばよい。あとはジェスチャー等の各動作が変わるときやビート時刻にもパケットが流れる。プロトキャストにより、1対1のコネクションの場合に見られるような指数関数的なパケットの増加が避けられるのは大きな利点である。
- Q: 小坂(NTT) ジェスチャーを検出するアルゴリズムは?
- A: 論文中にもあるが、背景画像との差分を二値化して x 軸方向のヒストグラムを求め、そのピークから頭の位置を検出する。体を傾けたのは頭の x 座標の変化でわかる。また手で頭を指すのは、頭の x,y 座標に大きな変化がなく、その左右の領域に大きな変化があることでわかる。
- Q: アニメーションするときのコンセプトは?
- A: 最終的には人間の演奏する姿により近づけていくことを考えている。現在は、ベーシストの運指では、ポジションの変化が少なく、開放弦を使用できるように割り当てている。ドラマーでは、各打楽器について右手と左手の優先動作と専用動作を割り当てている。例えば、ライドシンバルは右手が専用に叩くなど。
- Q: ピアニストおよび第3者のシステムの評価はどうなっているか?

- A: 両者とも同様に評判がよい。また、以前の我々のシステムよりも格段に良い、との感想を得ている。
- Q: 志村(大阪芸大) 演奏の最初はどのように始めるのか?
- A: ドラマーがスティックでカウントをとって始める。
- Q: 人間同士のセッションでは変化のある部分は事前に予測しながら進めているが、本システムはそれができないのでは?
- A: 次の発表でも述べるが、一応このシステムはプレーヤーの主張度の変化を予測しながら演奏している。

(5) すべてのプレーヤーが対等なジャズセッションシステム II. ベーシストとドラマーの実現

日高 伊佐夫, 後藤 真孝, 村岡 洋一(早大)  
記録: 小坂 直敏(NTT)

- Q: 堀内(千葉大) 相手が盛り上がると一緒に盛り上がる必要はないのか?
- A: ありうる。しかし、この判断がどの辺で見分けられるのかという問題が生ずる。
- Q: (盛り上がるべきかどうかについて) 動的であることはよいと思う。しかし制御が難しくなる。
- A: 今のところシナリオの中でその役割がグローバルには決まっている。
- Q: (システムが) 人間の演奏の意図を真に理解できているか、という疑問が生ずる。音が多いから盛り上がっている、とも限らない。
- A: 確かにパラメータは多く、システムの調整は必要である。
- Q: 演奏者はどの程度このシステムと練習したのか?
- A: 10 回程度。
- C: 人間は学習効果があり、次第に慣れていく。この辺の検討もしてほしい。初見ではどうかなど。
- Q: 彌富(筑波大) ドラムスが主張していてもベースがこれを理解しないかのような演奏があったように思える。これはドラムスの主張を判断するパラメータが 2 つと少ないためか。
- A: そうだ。聞くためのパラメータが少ない点などは改良の余地がある。フォーバースだからドラムスが出てくるなどの対応をしているため、主張している。
- Q: 平賀(図情大) 主導率は相対的なもので、主導率は同じで主張度が高い場合と低い場合があるのか?
- A: そうだ。例えばバラードなどでは、全体で主導率は 33 パーセントで一緒に静かに奏していくこととなる。
- Q: 絶対的主張度というのはないのか?
- A: 目標の主導率を達成して満足したときは主張度の初期値を踏襲する。このとき、他者の主張度から自分の目標主導率を決めることになる。
- Q: 平野(i & i) わざと失敗するなどの摂動を与えたときにどう振舞うか? また、人間がいなくなるとシステムはどう振舞うか?
- A: ピアニストが何も弾かなくとも他の二人は影響を与え合っただけで演奏できる。しかし、主導率という概念が普遍的かどうかはわからない。システム改良のため、今後人間の演奏の振舞いをシステムに与えて分析をするなどの実験を試みたい。
- C: 前の発表で述べられていたように、ピアニストもコンピュータにして分析していくと面白い。
- C: 小坂(NTT) しかし、インタラクティブな系の片側を統制して実験することは難しい。最初の制御された振舞いに対し、相手は実験者が想像できぬ振舞いをするのが普通で、そこまで見越した実験はできない。

- C: 平賀(図情大) しかし、自然言語の対話研究などでこうした研究の気運は盛り上がってきた。来年のIJCAI ではロボット同士のサッカーも企画されている。やはり今後はこうした方向も考えなければならないのだと思う。

- Q: 南高(カシオ) 演奏の生成部について、演奏データはどのように選んでいるか?
- A: 複数のモード分けされたパターンがあり、モードを決定した後はその中で頭から順に選んでいく。乱数などは用いていない。
- Q: コード進行のデータがあるのか?
- A: そうだ。その制約を満たしたものを選ぶ。
- Q: 志村(大阪芸大) 演奏全体のクライマックスはシナリオに書かれているのか?
- A: シナリオには書かれていない。偶発的に起こる。
- C: ベースがやや音程が低いのでは。
- A: 音色の問題かもしれない。

(6) 計算機によるピアノ伴奏

小川 大典, 戴 岡, 五十嵐 滋(筑波大)  
記録: 松島俊明(東邦大学)

- Q: 堀内(千葉大) 0.5 秒の遅れの制約はきつい。もっと応答が早いものがあるのでは?
- A: 五十嵐(筑波大) 昨年 10 月の時点では 0.3 秒を切るのは難しいと聞いている。強弱によってレスポンスが変わってくるため。
- A: 彌富(筑波大) 0.2 秒の遅れに設定することもできるが、レスポンスが保証されない。一律 0.5 秒の遅れに設定にすればレスポンスが保証されている。
- Q: リハーサルはシステムと一緒にやるのか?
- A: 伴奏パートを鳴らさずに行っている。
- C: 参照した過去の文献があったら載せておいて欲しい。
- A: 五十嵐 他の文献は参考にせず、完全に独自の知識で行っている。
- Q: 平田(NTT) テスト曲を選んだ基準は何?
- A: 各システムに適した曲の特徴を考慮した。またシステムを作成した人の希望を取り入れた。
- Q: 昔のシステムでこれらの曲を演奏してみたか?
- A: システムが完成したばかりなので、まだやっていない。

(7) 大域結合カオスのコンピュータミュージックへの応用

平野砂峰旅(i & i Multimedia Planning Division)  
記録: 松島俊明(東邦大学)

- Q: 小坂(NTT) 虫の声みたいな音のデモと他の音のデモでは音源が異なるのか?
- A: 音源は同じ FM。うなりの効果のため音色が違ってくる。
- Q: カオスを使ったのは、ランダムな状態から周期的な状態までをコントロールしたかったからか?
- A: 1 つはそう。単一のパラメータでコントロールできる所が良い。カオスの変遷とグルーピングの効果で、自動的に音色が変わってゆく所が面白い。
- Q: ランダムとカオスの違いがはっきりしない。ランダムではだめなのか? なぜカオスなのか?
- A: アルゴリズム的に美しい。ある特定の部分だけ取り出せば他のアルゴリズムでも実現可能とは思いますが、簡単な式で実現できる方法なので面白いと思う。