

# 音楽情報科学研究会のページ

SIGMUS Home Page <http://www.etl.go.jp/~sigmus/>

## 第39回 音楽情報科学研究会開催のご案内

(情報処理学会 音楽情報処理研究会 (SIGMUS), 日本音響学会 音楽音響研究会 (MA 研), 電子情報通信学会 音声研究会 共催研究会)

日程: 2001年2月22日(木), 23日(金)

会場: 九州芸術工科大学 (福岡県福岡市, <http://www.kyushu-id.ac.jp/>)

照会先: 平野砂峰旅 (京都精華大学)

e-mail: [hirano@kyoto-seika.ac.jp](mailto:hirano@kyoto-seika.ac.jp) (できるだけこちらへ)

Tel: 075-702-5248, FAX: 075-702-5248

## SIGMUS 第37回 研究発表会 報告・質疑記録

2000年10月12,13日 ATR 知能映像通信研究所

### (1) 人や場を演出する音楽創奏システムの提案

多田幸生 (ATR), 西本一志 (北陸先端大),  
前川督雄, 間瀬健二, 中津良平 (ATR)  
記録: 坂本圭司 (千葉大)

Q: 瀏野 (北見工大) 携帯電話では心臓ペースメーカなどへの悪影響や着信音で、電車内などの使用が問題となっているが、このシステムで同様の問題は起きないか?

A: 携帯電話などの電磁波がペースメーカへ影響を与えるかどうかはまだはっきりとわかっていない。楽器自体はヘッドホンで聞くので、音で周囲に迷惑がかかることはない。楽器を身につけている人同士がやりとりするものである。

Q: 誰が発信したものかわからなければ、自分のファッションの主張にはならないのではないか?

A: 片寄 (和歌山大) どちらの方向から送られてきたものなのかぐらいはわかったほうが良い。

Q: 片寄 EXPOなどで実際の効果を試すための計画は?

A: 今のところない。将来的にはやる予定。

Q: 後藤 (電総研) フレーズとフレーズを音楽的につなぐところの仕組みは?

A: 現在検討中。

Q: 岸田 (立命館大) 身に付けるセンサー一つがフレーズに対応しているのか?

A: 違う。センサはフレーズを作ったり、演奏するためのものである。

### (2) リアルタイム音楽情景記述システム: 全体構想と音高推定手法の拡張

後藤真孝 (電総研)  
記録: 坂本圭司 (千葉大)

Q: 高橋 (IAMAS) 推定精度は?

A: 正解との周波数差が50セント以下は正解とした。数セント単位の評価はしていない。

Q: ピブラートのような小さい周波数変化にも追従するのか?

A: 基本的に追従する。

Q: リアルタイム性は?

A: 低域の時間窓が長いために、実際には入力から300ミリ秒程度遅れて音高が出力される。

Q: バンドパスフィルタをかけることは有効なのか?

A: メロディーとベースを別々に推定するために有効である。

Q: 合唱曲のメロディーを調べることも可能か?

A: 合唱の場合には、複数の声が同程度に優勢になると、本手法では困難なことが多い。合唱曲は現在の評価対象曲には含まれていない。

Q: 瀏野 (北見工大) メロディーがどう繋がっているのかを認識する必要がある。優勢なものをとるだけでは安易では?

A: 本発表では詳細は省いたが、時間軸方向の連続性も考慮している。

Q: 音モデルの自動推定はできているのか? バンドパスフィルタ (BPF) の特性を決めることも含めて。

A: 音モデルの倍音比率の推定を今回実現した。BPFの特性は推定せず、今回は仮定を設けている。

Q: 平田 (NTT) 人間は音楽的知識がなくてもメロディーを聞き分けることが無意識にできているが、そういった処理はどの辺りまでできるのか?

A: それを探るのもこの研究の一つの目標。最終的には高度な音楽的知識も導入する必要があるだろう。

Q: 片寄 (和歌山大) メロディの推移確率などを次の音の推定に利用できるのでは?

A: 事前分布を利用していくことを考えている。

Q: 堀内 (千葉大) 「音楽を理解する」とはどういうことか?

A: ここでは、音楽鑑賞のような高次元レベルではなく、音を分析してどういうまとまりがあるかを理解するレベル。

Q: 人間が実際に音楽を聴くときに、本当にここで認識しているような音楽的要素を聴いて理解しているのだろうか?

A: 研究を進めるための作業仮説として導入したが、適切であるかは常に再検討しながら研究する必要がある。どこまで人間の理解に近づけるかは、今後の研究次第。

(3) バービーブーン：音符レベルでユーザ意図を把握して編曲を行う事例ベースシステム

平田圭二 (NTT), 青柳龍也 (津田塾大)  
記録：坂本圭司 (千葉大)

- Q: 後藤 (電総研) 結果の曲を聞くだけでは、システム動作が期待通りか判断するのが難しい気がするが、どう開発しているのか?  
A: プログラム中に print 文を埋め込み、トレースして確認しながら開発する。  
Q: 生成された結果を聴いたときの平田さんご自身の印象は?  
A: 何とかしなければと考えている。  
Q: メロディーとぶつかることについて考慮したか?  
A: 最大下界 (glb) を求める部分で工夫をする必要がある。音楽性を考えた glb と最小上界 (lub) が必要であると考えた。  
Q: 片寄 (和歌山大) メロディーパートと伴奏パートを分ければ、もっとわかりやすいのでは?  
A: ここではメロディーと伴奏が渾然一体となっている曲を対象としているため、難しい。

(4) 携帯端末に適した情報表示・操作方式の検討

増井信彦, 宮本勝, 小沢英昭 (NTT)  
記録：石毛大悟 (千葉大)

- Q: 荒木 (豊田中研) 魚眼ビュー (fish-eye view) との比較を行っているのか?  
A: その方法は自分の注目したい情報だけを効率よく探すのであれば有効であるが、今回の方法は概要情報や他の情報との関連性を見ながら情報を見るという前提で行った。実際には比較する必要があるだろうが、一番最初のステップで比較評価する対象には、今は入れてない。  
Q: 淵野 (北見工大) 個別の情報だけ知りたい人と、その情報が絶対層の中でどの階層にあるかまで知りたい人がいるが、それを考えたときの比較は難しいのではないのか?  
A: おっしゃる通り。人によって見方が違うので、どのような表示情報にこの方針を使うと良いかを言えればよかったが、まだ言い切れてない。  
Q: いろいろなタイプの評価を並列して行うのか?  
A: できれば両方に対応したいが、それを考慮しつつ、情報によってどちらに適しているかという方向に持って行きたい。  
Q: 理想としてはどの評価方法でも良いような表示方法を提案して行くのか?  
A: 特別な端末やデバイスを使ったりすれば解決できると思うが、その前に、画面表示や、あらかじめあるボタンの操作などで問題を解決できるところまで解決しておくという考え方もある。  
Q: 平田 (NTT) 表示レベル 3 の時に、表示レベル 2 の内容を削れるか?  
A: 今はできない。そういう需要に対応できるような作り方が必要だと考えている。

(5) 作曲者のメンタルスペースの外在化による作曲支援環境の研究

網谷重紀, 堀浩一 (東大)

記録：石毛大悟 (千葉大)

- C: 片寄 (和歌山大) 手を広げすぎている。詳細化してからのほうが良いのでは?  
C: 水口 (シャープ) アウトラインプロセッサのようなシステムと比較検討すると良いのでは?  
Q: 西本 (北陸先端大) アイコンを動かした時の重み変更の自動化はできないのか?

- A: 現在製作中。自動化の予定。  
Q: 複数解が出てくると思うが、どのように単一の解にするのか?  
A: 旋律、リズム、調の各々について各フレーズ間の相対距離をシステムが計算し、再配置後のフレーズ間距離が 3 つの音楽的特徴でどれが一番近いかを見て重み付けをする。  
Q: 最初の空間はどれに重みを置いて配置しているのか?  
A: 全部イーブン。  
Q: フレーズしか配置されていないと空間の基底がわからない。そうすると大きな穴が存在するかも知れないが、その穴がどういう意味を持っているかを提示する方法は?  
A: なんらかの逆計算により、位置からフレーズを提示することができるのではないかとという考え方もある。  
C: 双対尺度法により音楽的特徴を見ていると空間の基底が見えてくる。それにより空間の位置の意味が見えてくるという方法がある。  
Q: 平野 (京都精華大) 支援システムとの相性は作曲のやり方に依存すると思うが、作曲方法をどの程度の人数から調べたか?  
A: 身の回りの作曲をする友人 10 名。  
C: 片寄 音色の情報も周波数表現したら大体見えてくる。アイコンベースで貼って行き、拡張してゆくという作曲ツールが発表されているので、それらを踏まえた面白と思う。  
C: 大野 (NTT) アイコンのところにいろんな情報を表示することで、外在化した知識をより再利用できるのでは?  
Q: 上田 (阪大) そのツールを使うことで、良い曲ができるのか? 早くできるのか?  
A: 作曲がしやすくなることを目的としている。  
Q: よくプロの作曲家は「埋もれてしまうフレーズはいらない」と言うが?  
A: その通りだと思うが、埋もれてしまったことを悲しむ人もいる。提示するタイミングなど、時間に依存するところは今後考えたい。

(6) メディアインスタレーション “Movement” のインタラクションについて

平野砂峰 (京都精華大)

記録：石毛大悟 (千葉大)

- Q: 堀内 (千葉大) 変形されて出てきた映像と自分とのインタラクションが実現できるようなループがあると良いのでは?  
A: それについては考えている。わかってくると完全に制御できなくても予想ができるようになる。思い通りに動かせようと逆に面白みが欠けると考えている。  
C: 楽器が自分に返してくるというインタラクションがあった方がよりおもしろくなると思う。  
Q: 高橋 (IAMAS) アイコンを見えなくするとか、作品として考えて取り替えられるようにした方が作品としての抵抗がなくなると思うが、その可能性はあるか?  
A: 幾何学的な形ではなく楽器の形にするという方法もある。アイコンの上に画像がかかるとアイコンを変化させることは、間にもう一台機械を入れるとできるだろうが、今のシステムだとつらい。あまりわかりやすくするよりも、逆に抽象的な形にしておきたいと考えている。  
C: 映像をキャンバスとして考えれば、人の振る舞いによって絵画の作成をしようとするのを踏まえたインタラクションができると良いかと思う。  
A: そのような方向もある。だが、自分の姿や動きが影響されるのがおもしろくてインパクトがあるというのがねらいの一つであるので、シンプルに形にした。これを一つの例として、いろいろな発展型があると思う。

(7) L<sup>A</sup>T<sub>E</sub>X ソースの解析による編集ツールの提供

吉田慎, 瀧野昌 (北見工大)

記録: 平野砂峰旅 (京都精華大)

Q: 大野 (NTT) ファイルを編集していくと構文も変わっていくので、それに応じて色分けのしかたも変わることが期待されるが、現在はどのようにしているのか?

A: 編集しながらの色分けの変化にはまだ対応していない。

Q: 外部プログラムにせずに Emacs を拡張しても良いのではないか。

A: 瀧野 Emacs の方言をむやみに増やしたくない。現在のソフトを改良していき、必要性が確認できた後に、Emacs の拡張を考えたい。

Q: 間瀬 (ATR) L<sup>A</sup>T<sub>E</sub>X でひとつのファイルサイズが 600kBytes というのは現実的ではないと思うが。

A: 瀧野 自分の場合は、600kBytes のファイルを原稿としてやりとりすることがある。

C: 色分けの仕方とミスの防ぎ方の関連とかについてだと HCI としてコメントしやすいのだが。

A: 瀧野 まずは、ツールを作って提供してから詰めて行きたい。

Q: (質問者不明) 構文解析をしなくても、マクロだということがわかれば良いのではないか。そうすれば複雑な処理が軽減でき処理が速くなると思う。

A: 入れ子のマクロの中の文章のスペルチェックを行う場合などを考えるときちんと展開する必要がある。

(8) クリックابلムービーとそのオーサリングツール

藤本英明, 吉岡信夫 (大工大), 大村皓一 (宝塚造形芸大)

記録: 平野砂峰旅 (京都精華大)

Q: 瀧野 (北見工大) 動画の場合は、ミスクリックしやすいと思うがそのときの対策はどうなっているのか。

A: クリックابل領域以外の部分をクリックするとムービーが一時停止するようになっている。

Q: 間瀬 (ATR) クリックابلムービーは、他にも事例があるので参考にして欲しい。だが、XML にした点は評価できる。また、先の質問にもあったミスクリック時にムービーが停止することが、妥当かどうかについては検討して欲しい。

A: ミスクリック時にプレイヤーが停止するかどうかは、メニューなどで選択できるようにしたいと思う。

Q: 増井 (富士通研) SMIL 規格に関して何か具体的な調査を行ったか?

A: まだ、調査していない。

C: 9月20日に、W3C から SMIL 2.0 規格 (開発コード SMIL-Boston) の最終ドラフトが出て、1か月だけ反論やコメントを受け付けている。(発表当日の10月13日の時点で) 規格制定に貢献することも可能である。来年 SMIL 2.0 が W3C 勧告 (Recommendation) となるが、その後はクリックابلムービーに関する研究・開発成果を、SMIL 2.0 の特定の処理モジュールに対応するソフトウェアとして、アピールできるかも知れない。

Q: 堀内 (千葉大) ソフトウェアが動作しているプラットフォームは。

A: Windows 95, 98, NT で動作している。

Q: 30 frames/sec で動作しているのか。

A: Video for Windows で再生できる速度に依存している。

Q: 時間のズレは、大丈夫なのか?

A: 正確に測定していないが、自分が使っていて問題がないと感じている。

Q: 荒木 (豊田中研) 時刻とその時点でのクリックابل領域を指定する操作が煩雑で使いにくい。現状では矩形を描いた後「取得」ボタンを押し、さらに時刻用の別の「取得」ボタンを押す操作になっている。しかも複数の矩形からなる領域を指定する際には、その都度この操作を繰り返すので、手間がかかっている。これを例えば、複数の矩形を描いてから1回だけ時刻と領域共用の「取得」ボタンを押すようにすれば、使いやすくなるのではないか?

A: 良い提案なので、取り入れて行きたいと思う。

(9) 計算機との合奏データによる人間の演奏モデルの推定

坂本圭司, 堀内靖雄, 市川薫 (千葉大)

記録: 増井誠生 (富士通研)

Q: 後藤 (電総研) 4つのモデルが妥当と帰結しているが、どれが適切は今回の結果からは帰結できないのでは。帰結できるのは、モデルが単純化過ぎるためにフィッティングが不十分であった、ということのように思える。

A: 1拍単位のみでは不十分であるかも知れない。半拍単位や2拍単位などの情報も用いる必要がある。

Q: 平田 (NTT) 楽曲の構造に関係なく、単に1拍前を見るという手法は納得できない。音楽の構造から、依存関係がわかり、精緻なモデリングができるはず。

A: リズムの階層性を考慮に入れる必要がある。

Q: 後藤 テンポ変化する場所は全くのランダムであるのか?

A: 小節の頭 (1拍目) に限定している。

Q: 間瀬 (ATR) 人間がずれを感知できる限界は?

A: 30m 秒と言われているが、正確なことはわからない。

Q: 岸田 (立命館大) 「妥当」という言葉について。人間のモデルに沿っているという意味か。

A: その通りだ。モデルの係数に個人差はあまりなかった。

C: 岸田 データを見て、個人差があるかと思ったため質問した。

(10) 相互作用を考慮した人間の協調演奏モデルの推定

石毛大悟, 堀内靖雄, 市川薫 (千葉大)

記録: 増井誠生 (富士通研)

Q: 上田 (阪大) リズムやテンポが存在するかはわからないのでは。3つの音でみれば、結果がよりはっきりわかるのでは。4音が適当かも知れない。2音では問題あるのでは?

A: その通り。今後、検討を進めたい。

A: 堀内 過去のものを使いたい。直前のずれに対する変化量を採用している。だから人間が敏感に受け取れるものを採用しているということ。

Q: 平田 (NTT) モデルは単純なので、閾値を解析的に求められないのか?

A: 動的にずれを変えているので難しい。

Q: 実験に使った素材は離散的なのでは? 人間の耳では、もっと連続的な推移で判断しているのでは?

A: 今後試していきたい。

C: 田辺 (ガクタスソフト) (1) 過去のコンテキストの中で決まる量。音符の間隔では決まらない。時間のずれは、テンポではなく、フレーズの変化と理解すべき。一定のテンポの中で、拍がゆれながら、演奏が進んでいくもの。実際の音楽では、フレーズの区切れで、時間が合わせている。

(2) シーケンサでの問題: シーケンサは、自分の演奏を聴いてはいない。他人の演奏も聴いていない。独奏者の演奏をフィードバック、独奏者の演奏データを、伴奏システムにどう受け渡すかが、重要な研究ポイント。ソリストにとって、気持ちの良い演奏をお願いしたい。

A: 人間同士の演奏をうまくとれていないので、今後の課題として進めていきたい。

C: 堀内 精度の問題もあって、なかなか実現が難しいところもある。

### (11) 高密度キーによるデータ入力システムの開発

米谷昭彦, 新居里織, 戸莉吉孝 (名工大)  
記録: 荒木円博 (豊田中研)

携帯電話などのための文字入力装置の発表だった。装置は13mm×31mmの領域に6×3個のキースイッチを配置したもので、携帯電話の場合、裏面に取り付けられる。複数のキーを人差し指と中指で同時操作することによってかなやアルファベットが入力できる。

Q: 堀内 (千葉大) 装置を横に配置し、親指と人差し指を操作にあて、残りの指と親指の付け根で携帯電話を支える方が使いやすいのではないかと?

A: 今回提案した方式との比較評価を考えていきたい。

Q: 荒木 (豊田中研) 考えにくい想定だが、仮に長時間使うとなると手の腱に負担がかかるのではないかと?

A: 今後評価しようと思う。

### (12) ピアノ指導方法の差異が及ぼすピアノ学習への影響について

大島千佳, 西本一志, 小長谷明彦 (北陸先端大)  
記録: 西村憲 (会津大)

Q: 片寄 (和歌山大) この研究の目標は、ある特定の生徒にとって適切な指導法を見出すことにあるのか。それとも、一般的に良い指導法を見出すことにあるのか。

A: 各生徒の個性を引き出すために効果的な、一般的に良い指導法を見出すことが目的である。

Q: では、たくさんの事例を集めることが次のステップか?

A: その通りだ。

Q: 指導者と生徒の組によって、適切な指導法が大幅に異なるということはないのか?

A: たしかに、同じ指導法でも生徒によって効果が異なることは経験している。また、適切な指導法は、指導者と生徒間の相性にも大きく依存しているという見解を持っている。ただし、単に相性が良いからといって、指導の効果が大きいとは限らないようである。相性と適切な指導法との関連については、現在、その研究方法を模索中である。

Q: 後藤 (電総研) 指導者の演奏との差が小さいのに、生徒の自分の演奏に対する主観評価が低いときには、指導者の演奏に対する主観評価も同様に低いのか?

A: 調査していない。

C: velocity と duration (inter-onset interval) の方が用語として適当) だけをもとに演奏の近さを判定することが妥当かどうかを検証する目的で、指導者の演奏に対する主観評価も行った方がよい。

Q: 淵野 (北見工大) velocity の比較において、打鍵強度の個人差が問題とならないか?

A: 値を正規化しているので問題ない。

C: 田辺 (カクタソフト) velocity は音の大きさを表す物理量との対応に何ら根拠がなく、また楽器メーカー、音源、音色によってその対応はまちまちであるので、十分注意が必要である。特に、実験に使う velocity 検出つきピアノは常に同じものを使用するべきである。

C: 堀内 (千葉大) 正規化の際に、velocity の非線形性が悪影響を及ぼすことのないように注意が必要ではないか。

### (13) 音楽理解における音楽知の役割に関する一考察 — 音楽に内在する暗黙知の解明・J.S. バッハ「マタイ受難曲」における暗黙知の機能 —

村瀬喜代美, 小長谷明彦 (北陸先端大)  
発表: 西本一志 (北陸先端大)  
記録: 西村憲 (会津大)

C: 大野 (NTT) 被験者は真に知識の無い状態と言えるのか。実験に使用したジャンルの曲が、教会に関係するという程度の知識は最初から持っているのではないかと。

C: 多田 (ATR) 被験者に偏りがあるのではないかと。

Q: 平田 (NTT) わざと間違った解説を与えて同様の実験をし、結果を比較してみると良いのではないかと。1つの曲に対して異なる解説を与えて、その反応がどれも解説通りであった場合には、曲にはメッセージ性が無いと言えるのでは?

A: 面白い実験だと思う。ただし、間違った解説による影響が強すぎて本来のメッセージが打ち消されるということもあり得る。したがって、間違った解説を与えるのではなく、むしろ、異なるメッセージを込めた似たような曲を使い、結果を比較してみる方がより適切だと思う。

Q: 堀内 (千葉大) 曲自体より、その調性感や旋法が主に効いているのではないかと。

A: そうかも知れない。

Q: 間瀬 (NTT) チェックポイント以外のデータは、どうだったのか。

A: 測定データは残っている。アルファ波以外のデータについても、測定は行っているので、吟味する必要があると思う。

Q: 堀内 安静時との比較も必要ではないかと。

A: その通りだ。

Q: 片寄 (和歌山大) 心理学や脳波の専門家からの指導は受けているのか。

A: 間接的には指導を受けていると思うが、よくわからない。

Q: 堀内 今日の演奏と昔の演奏との違いを無視しているのではないかと。

A: 無視している。また、指揮者が音楽にメッセージを込めている可能性もある。すべての要因を考慮するのは困難であると思う。