



# SIGMUS 第5回 研究発表会 質疑記録

1994年2月4日 NTT 武蔵野研究センタ

(1) コード進行をベースにして作曲を支援するシステム  
内橋由佳, 平野健太郎, 伊丹誠, 伊藤紘二 (東京理科大)  
記録: 鈴木 (東京高専)

Q: 長嶋 (ASL) Sun から MIDI を出力する方法は?

A: インターフェースボードを自作している.

Q: コード中の基音以外のベース (分数コード) をサポートしているか?

A: 考えているが今のところはない.

Q: 1小節内の音符単位での分割はできるか?

A: できる.

Q: コードモーションの種類は?

A: 現在は C, F, G7 の 3 種類である.

Q: 次に移る隣のコードの動きだけを見ているのか?

A: そうである. 2 つ前は見ていない.

C: 小坂 (NTT) 今回はコード進行から見た作曲であったが, 実際使う立場から見るとメロディが先に浮かぶ場合が多いと思う.

Q: 高橋 コードネームに別の名前を付けられるか?

A: 違う名前も付けられる.

Q: 自作のコードでもモーションとして登録可か?

A: 登録可能である.

(2) 標準 MIDI ファイルからのメロディの自動抽出法  
鷲坂光一 (NTT)  
記録: 増井 (富士通研)

Q: 長嶋 (ASL) (メロディチャンネルの選択例として用いた) 224 曲の MIDI ファイルデータの提供ベンダ数は? それぞれメーカーごとに癖がある.

A: ローランドとカワイなど. 2, 3 社である.

Q: ピアノソロ曲は 1 チャンネルにデータが全部入っているのか?

A: 1 チャンネルに入っている.

Q: ソロ曲におけるメロディチャンネル選択結果の可否はどう判断するのか?

A: ソロ曲は 224 曲中には含まれない. 処理内容の説明としてピアノソロを例にあげた.

Q: SMF には最初からパート毎にヘッダにメロディやコードといった記述が含まれているのでは? だとするとメロディ抽出の意義は?

A: ヘッダにおけるテキスト部分にはパート情報は必ずしも含まれない. 用いた SMF データには含まれていなかった.

Q: 平賀 (図情大) 「1 つのチャンネルを選択」とは「全曲を通じて」の意味か?

A: 歌謡曲を用い, 歌詞の入っている声部が選択されたかどうかを判断した.

Q: 成功例では, 途中でメロディを担当している声部 (チャンネル) が替わったりしないのか?

A: しない.

Q: (メロディチャンネル選択における) 失敗の原因は?

A: 音域のずれの問題, 伴奏の派手さなどが考えられる.

Q: 声部の交替が原因になって失敗しているものがあるのでは? 一般のデータでは, 頻りに声部の交替があるのが普通ではないか?

A: そうではなかった.

Q: 1 つの声部だけがメロディを担当するというのは仮定として強過ぎるのではないか? 現実の演奏では何人もがメロディを担当することは多い. 現実の音楽データベースとしての利用を考える場合, 考慮が必要であろう.むしろ音楽データベースという観点からは, 検索側で「メロディ」をきっちり押えておく必要があるのではないか?

A: 仮定が違う. メロディは意味空間を構成する一つの要素であり「意味検索」のためのツールと使う. メロディの指定によって音楽を検索するために使うのではない.

Q: 検索の Key はメロディではないのか?

A: 「意味検索」の Key は必ずしもメロディではない. オーサリングシステムにおいて絵にあった音を選ぶ場合に「どういう絵か」「どういう音楽か」がそれぞれ定義できていれば「その絵にふさわしい音」としてマッチングを取れる可能性がある.

Q: 鈴木 (東京高専) 「似たようなメロディの曲を探したい」というのが研究目的なのか?

A: 部分集合である. 絵と音とのマッピングを行なうソフトがあるが, 例えば, あるシーンで, 絵と音のマッピング状況など, 絵が持つ性質にしたがってサーチが可能かもしれない.

Q: MIDI は音楽情報としては記号化されたものである. データベースとして考えると, もっと一般的な PCM データを扱えると応用が広がるのでは?

A: 同感であるが, 今回はまず MIDI データを利用した.

Q: 伊藤 (東京理科大) メロディ・コードの雰囲気と絵の雰囲気を関連されるやり方については?

A: 現在, マッピングについていいアイデアを探している.

Q: 自動化というより, 人間が介入するのがいいのか? 絵の評価と音の評価をつなぐ感性的なチャンネルというのは, 人間の感覚を介する以外にはないのか?

A: 例えば, 悲しい絵に悲しい曲を結び付けるのは感性情報のレベルであるが.

Q: そのカテゴライズは人間がやるしかないのか?

A: 例えば人間と機械で異なる観点が存在するとして, その判断基準が存在すれば, そのマッピングによっては整合性がとれるかもしれない.

Q: その規範は人間がきめるべきなのか?

A: 人間がきめるべきかどうかは... 判断が難しい.

Q: 片寄 (LIST) 感性レベルでのメディア変換を目指しているのか?

A: いや, なるべく感性には踏み込まないようにしている.

C: それで正解だと今は思う.

Q: 後藤 (早大) メロディの抽出に用いたパラメータはどう設定したのか? 重み設定は?

A: パラメータ数は6つ。選択結果が外れた要因を調べて、フィードバックさせる。80曲を対象にパラメータを設定したが、すべてを正しく選択するには設定できない。ある曲で選択できるようにパラメータを変えると、今まで正しかった曲で選択できなくなることがある。

C: 長嶋(ASL) 抽出したメロディのマッチングを取るの簡単か? このデータベースを JASRAC に用意すれば、登録時のチェックに使えるのでは? アレンジに関わりなく、メロディの盗作チェックができそうに思われるが。

### (3) UNIX 上の MIDI シーケンサーとそのインタフェース 高田敏弘 (NTT) 記録: 増井 (富士通研)

Q: 片寄 (LIST) イベントの管理にはタイムスタンプを使っているのか?

A: MIsend というライブラリ関数があり、メッセージは一度カーネルにためられ、時間がくれば RS-232C ポートから送出される。MIDI 程度の精度なら最小限度まで十分管理できる。

Q: カーネルに一度に送っても OK か?

A: そうだ。メッセージを送る場合には、何秒間かデータ送出行なわなければ、ユーザ側のプロセスをサスペンドさせるようにも設定できる。

Q: 竹内 (NTT) どのくらいの精度なのか? RS-232C だから 1m sec 近くか?

C: 長嶋 (ASL) MIDI では 320  $\mu$  sec に 1 つ、.1 m sec で最大 3 発 (3 バイト)。

Q: 竹内 RS-232C ポートでそこまでの精度が出るのか?

C: 長嶋 SC-55mkII の場合は、(受けとったデータを) 内部で保留して発音を遅らせ、同時発音の精度を上げているらしい。

Q: 後藤 (早大) MIsend の呼び出しは、以前のイベントのタイムスタンプが来るまでブロックされるとあるか?

A: 設定で変更可能。たとえば 5 秒間はブロックしないように設定できる。

Q: ブロックされる場合は、ノートオンで和音が届くと同時に鳴らないのでは?

A: 完全にブロックすると駄目。ブロックされるまでの時間的余裕を設定することで、カーネル上のバッファ溢れを調節できる。

Q: MItimecreate, MIsend についてもそういう機能があるのか?

A: 用意されている。

Q: 森 (浜松ポリテク) Internet 上での広域 DB で Mosaic を利用する場合、Indigo を持っていないと受け手は MIDI データの再生ができないのか?

A: 現状ではそうだが、テンポがヨタってもいいなら、Sparc でも MIDI 再生ができる。IBM-PC, Mac でもクライアントがあれば可能だろう。

Q: midifile の解析ライブラリ: midifile 1.11 は SGI についているのか?

A: SGI 添付でなく PDS である。MIDI 関係の各アーカイブサイトにも置かれている。

Q: どのワークステーションでも利用可能なのか?

A: Sparc 版の場合は、midifile ライブラリと、自前のコマンド送出力部を用い、タイミングの制御はループで行なっている。

### (4) Conferences, Concerts, and Cocktail Parties: Besides Immersion Michael Cohen (会津大) 記録: 平賀 (図情大)

Q: 平賀 (図情大) 音の再生はヘッドホン、スピーカーのどちらによるのか?

A: Circumfering ヘッドホンによる。したがって [発表時のような] スピーカー用には optimize されていない。

Q: どのようなパラメータを操作しているのか。特に距離との関係は?

A: いろいろなパラメータがあるが、考慮していないものも多い。Distance attenuation は、基本的には gain attenuation だけで処理している。したがってドライで残響のない環境を作っていることになる。

Q: Lyon (慶大) このような技術を実現するための band width はいつ頃実用になるか?

A: ISDN は次世紀には本格化するとすると、あと 6 年ということになる。ついでに、音源について clustering のようなことも考えられる。画像では pixel 集合が 1 つの単体を表したり、逆に pixel もさらに細かく分割できると同様、音でも音やチャンネルの集合を 1 つの単位として扱うということだ。

Q: 小坂 (NTT) copy and paste でクローンを数多く作ったときの聞こえを論じたが、検証実験は行なったか?

A: 特にやってはいない。Precedence effect (Haas effect) の実現方法には、他にもいろいろある。Precedence effect 自体のシミュレーション、平均化処理、source sink の手動割り当てなど。音楽は多重度が高いので、こういった研究に向いている。

### (5) Sinusoidal model の特性分析と音合成への適用 小坂直敏 (NTT) 記録: 平賀 (図情大)

Q: 伊藤 (NTT) デモのサンプリング周波数は?

A: 楽音は 48kHz サンプリングだが、情報は 44.1kHz 分である。分析の窓は 2048 に固定。音声は 12kHz サンプリングだが、情報があるのは 5kHz くらいまで。

Q: 楽音は歪みが耳についた。サンプリングレートの利害が原因か?

A: そうだろう。

Q: 歪み量は?

A: まだ現実の信号については測定していない。

Q: MQ アルゴリズムの原論文とは位相項の近似式が違うのか?

A: やったのは結局論文通り。日本ではまだ試されていないようなのでやってみた。

C: 長嶋 (ASL) 音声、自然音については OK だが、楽器音は頭が歪んでいる。Sine 波 200Hz で 3Hz の偏差というのは「楽器」のピッチ設定精度としてはおもちゃ程度で誤差が大きく、せめて 9-10 bit の周波数設定精度がほしい。

A: 実用面ではその通りだろう。とりあえず今の範囲では検証が目的である。

### (6) Time Varying Extensions to the Moore Phase Vocoder Eric Lyon (慶大) 記録: 平賀 (図情大)

Q: 小坂 (NTT) 音楽家の意見は聞いてみたか?

A: 全般にインパクトはあるようだ。特に non-musician の反応が大きい。

Q: 音楽家への普及は考えているか?

A: 考えている。現在の 5 通りの processor を 20 通りくらいに増やし、公開するつもりである。使い方は各人によって変わるだろうから、選択幅は大きいほどいいだろう。