

自動表情付システム MorPH

真栄城 哲也, 樋口 直史, 中口 孝雄
ATR 人間情報通信研究所 第6研究室
要旨

MorPH は, SMF で記述された楽譜通りのポップス系の楽曲に, 自動表情付を行うシステムであり, インターネットで公開している. 対応している楽器は, ベース, ギター, サクソフォンおよびドラムである. 1小節分の好みのドラムパターンをユーザーが指定すると, そのパターンに基づいて, 曲の部分毎に合ったドラムパターンを, 自動生成する. ユーザーの指定パターンが無ければ, 自動的にドラムパターンを生成する. 初心者用の Web でアクセス可能な MorPH light 1mg と, 細かい設定が可能な MorPH Rhapsody の2種類のクライアントシステムがある. 初期段階でのユーザーの反応は, 特にドラムに関して評価が高い.

MorPH — Musical Performance Humanizer

Tetsuya Maeshiro, Naofumi Higuchi and Takao Nakaguchi
Dept.6, ATR Human Information Processing Res. Labs.

Abstract

MorPH: Musical Performance Humanizer processes popular musics in SMF format generated precisely from musical scores, and it is open for Internet. Supported musical instruments are bass, guitar, saxophone, and drums. Given the preferred drum pattern of one measure from the user, MorPH generates automatically drum patterns that matches each part of music. Two types of the client system are available: MorPH light 1mg, for beginners, accessible through web browsers, and MorPH Rhapsody, which allows detailed specifications of parameters. Initial response from the users is affirmative, particularly the drums.

1 はじめに

楽曲の自動表情付は, クラシックやジャズなど様々なジャンルに対して, また, 単に音に強弱を付けるシステムから, 曲の解釈に基づいた表情付を行うシステムまで, 多々発表されている [1, 4, 5, 6, 7, 2, 8]. MorPH は, Musical Performance Humanizer の略で, 我々の自動作曲/編曲システムの研究の一環であり, 作曲/編曲システムの出力を, 人間の演奏らしく聞こえるように基本的な表情付を行う. 対象は, ドラムパートを含む楽曲, 特にポップス系の楽曲とし, 楽譜通りに記述された SMF (Standard MIDI File) データに, 音の強弱, 発音タイミング, ゴースト音の付加などの処理を行い, 表情付された SMF データを出力する. 現在対応している楽器は, ドラム, ベース, ギター, そしてサクソフォンである. また, MorPH の特徴は, 楽曲に合ったドラムパートの自動生成であり, ユーザーが指定する一小節分のドラムパターンを基に, ドラムパートを生成する.

MorPH の目標として, 特に, 実際に使えるシステムの構築が挙げられる. MorPH の使用方法として, 2通りが考えられる. 1つは, 手動で行うには細かすぎると同時に, 時間がかかる表情付の作業を自動化し, より高度な表情付に専念するというものである. 一方, 2つ目は, 楽譜通りに入力された MIDI データを, できるだけ簡単に表情付するものである.

MorPH は、自動編曲システムのサブシステムとして位置付けられているが、単独システムとして実現されており、インターネット上で公開している (<http://morph.hip.atr.co.jp/>) (図 1)。2種類のクライアントが用意されており、1つは Java で書かれたブラウザ上の初心者向けインターフェース MorPH light 1mg, そしてもう1つは表情付の細かい設定が可能な Windows 95/98 専用の MorPH Rhapsody である。いずれのクライアントも、ATR 内にある MorPH サーバーに表情付する SMF データと、表情付の設定パラメータを送り、表情付された SMF データを受け取る。一般公開の目的は、(1) 多数のユーザーからの評価を得ること、そして (2) 表情付けの好みのデータの収集と解析である。

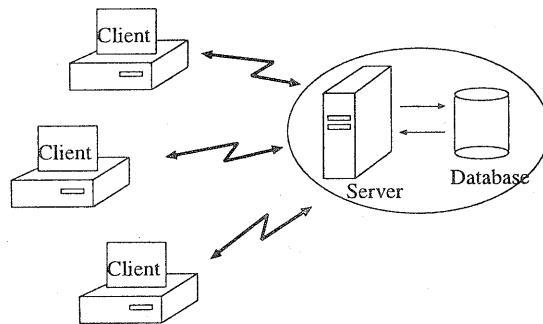


図 1: MorPH システムの構成図

第 2 節では、各楽器の表情付について述べ、第 3 節では、MorPH light 1mg および MorPH Rhapsody について述べる。第 4 節は考察および結論である。

なお、実装の詳細については、「自動表情付システム MorPH の実装」を参照されたい。また、実際の処理結果は、MorPH のホームページ、<http://morph.hip.atr.co.jp> を参照されたい。

2 表情付

表情付けは、個々の楽器もしくはパートを個別に行うのではなく、全パートを統一的にを行い、全体として統一のとれた表情付を行うことを目標としている。

2.1 入力 SMF データ

MorPH は、フラット、もしくはジャストな SMF データのみを入力データとして受け付ける。これは、入力がフラットなデータでなければ、MIDI データから実際の楽曲が推測できないからである。ここで、フラットな SMF データとは、音の長さおよび発音時間が、MIDI ファイルの分解能の分数比もしくは整数比であるデータを指す。例えば、分解能が 480 の場合、4 分音符の長さは 480、16 分音符は 120、そして、小節の始めの音符の発音時間 t は、 $t \% 1920 = 0$ を満たさなければならない。なお、% は、余りを意味する。

2.2 サクソフォン

サクソフォンの処理では、ノート・ベロシティおよびエクスプレッションの設定を行う。エクスプレッションの値は、実際に人間がウインドコントローラを演奏したデータを用いている。処

理はフレーズ単位で行い、フレーズの始め、中間、終り、そして各音の長さとその次の音の音程から、発音のベロシティとエクスプレッションのカーブを生成する。

2.3 ギター

ギターの処理は、(1) ソロ、(2) コード、そして(3) アルペジオの3種類に分けられる。(1)のソロの部分は、演奏の個人差が非常に大きいため、基本的と考えられる部分のみ、チョーキングやハンマリングの処理を行っている。(2)のコードに関しては、コードと弾く弦の対応表をシステム内に持っており、ストロークの処理と、各弦で発音される音のベロシティの生成を行う。システム内に存在しないコードが見付かった場合、動的計画法[3]を用いて弦へのマッピングを行い、それも失敗すれば、入力コードの構成音を単純に順次弦に対応付ける。また、コードと弦の対応表を用いて、入力SMFデータのギターのパートに単純にコードの構成音から、実際にギターで演奏される音を自動生成できる。例えば、「ドミソ」を「ドソミソド」のようにである。アップおよびダウンストロークの選択は、自動であり、1拍を8分音符で表した場合、8分音符に重なるコードはダウン、それ以外はアップとしている。そして、(3)のアルペジオは、(2)のコード同様、まず音を弦にマッピングしてから、音を伸ばす処理を行っている。弦へのマッピングを行うことで、より精度の高い表情付けが可能である。

2.4 ベース

ベースの表情付けは、ハンマリング、ダウンスライド、ビブラート、そしてゴースト音の付加である。

2.5 ドラム

MorPHの1つの特徴は、ドラムパートの自動生成である。ドラムパートの自動生成には、楽曲のリズムパターンの推測が不可欠だが、本研究では、ベースが楽曲のリズムを担うとの認識に基づいている。MorPHは、一拍分のスネアドラムとキックドラムの全ての組合せを内部に持っており、似ているパターン、そしてリズムが細かい/粗いパターン同士の関係のように、構造化している。さらに、フィルインのパターンも同様であり、類似/細かい/粗いパターンを関連付けている。

ゴースト音として、16部音符および32部音符のパターンが追加される。これらのパターンは、実際に使われる、もしくは使えるパターンのみをシステム内に持っている。

また、出力SMFデータに加えられるドラムパターンの生成は、ユーザの指定する1小節分のドラムパターンと、楽曲から推測されるドラムパターンから、両者の共通項を取り、相違点を段階的に付加することで、ユーザのパターンから楽曲のパターンへ段階的に変化するパターンの生成が可能になる。

MorPHが生成するハイハットの強弱と発音時間は、人間の演奏データに基づいており、1小節単位で持っている。より細かい単位、例えば1拍の場合、全体的なアクセントの統一感が失われてしまうため、小節単位としている。アクセントは、強弱および発音時間の組合せのため、両者をセットで持ってあり、このデータを用いて、次に述べる全パートの発音時間を生成している。

2.6 全パート

MorPHの特徴は、楽曲に含まれる全パートを、統一的に表情付けすることである。その1つとして、曲のリズムのアクセントを担うハイハットの生成されたデータを用いて、全パートの発

音時間を生成していることが挙げられる。

3 MorPH のクライアント

MorPH には、初心者用の “MorPH light 1mg” および細かい設定が可能な “MorPH Rhapsody” の2種類のクライアントが用意されており、インターネットで公開している (<http://morph.hip.atr.co.jp>)。ホームページには、アンケート記入フォームも用意しており、MorPH の評価に有効だと考えられる。また、対象ユーザーの可能な限り増やすために、再生用の音源は機種を特定せず、GM (General MIDI) 対応の音源としている。

3.1 MorPH light 1mg

MorPH light 1mg は、できるだけ単純な操作で MorPH を用いた自動表情付を行えるようにしたクライアントシステムであり、Java で書かれている (図 2)。従って、Netscape や Internet Explorer から実行できる。基本的な操作は、処理する MIDI ファイルの指定と、1小節分のドラムパターンの設定のみである。元の MIDI ファイルにドラムパートが存在しても、MorPH が生成したドラムパートと自動的に入れ換えられるように設定されており、ハイハットやドラムのゴースト音の細かさなどは、始めから設定されており、変更できない。また、ドラムのフィルインの位置も固定である。

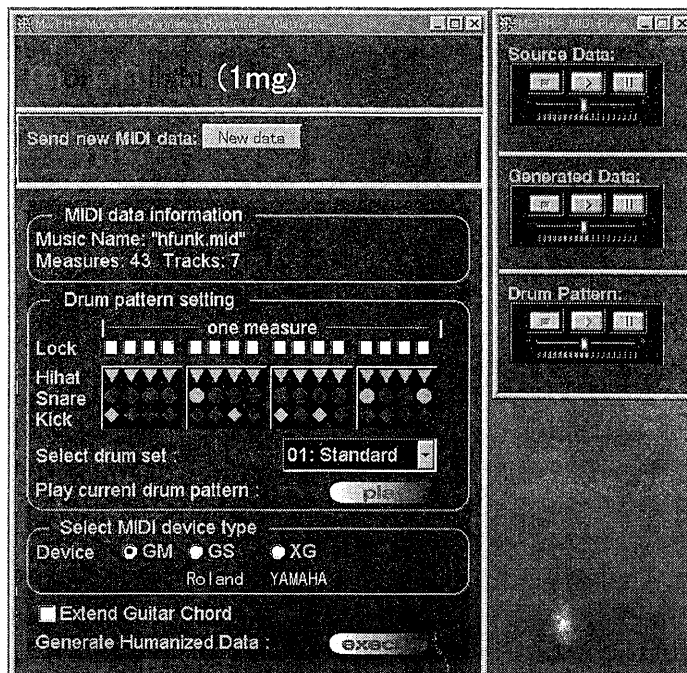


図 2: MorPH light 1mg の画面

3.2 MorPH Rhapsody

MorPH Rhapsody は、Windows 95/98 専用のクライアントであり、MorPH light 1mg とは異なり、表情付けのパラメータの設定が可能である (図 3)。クライアント用のインターフェースのみを備えており、表情付けの処理は ATR 内のサーバが行うため、実行時にはネットワークに接続している必要がある。ドラムの表情付に関しては、フィルインの位置も含めて指定可能であり、曲の部分毎に別々の設定を行える。

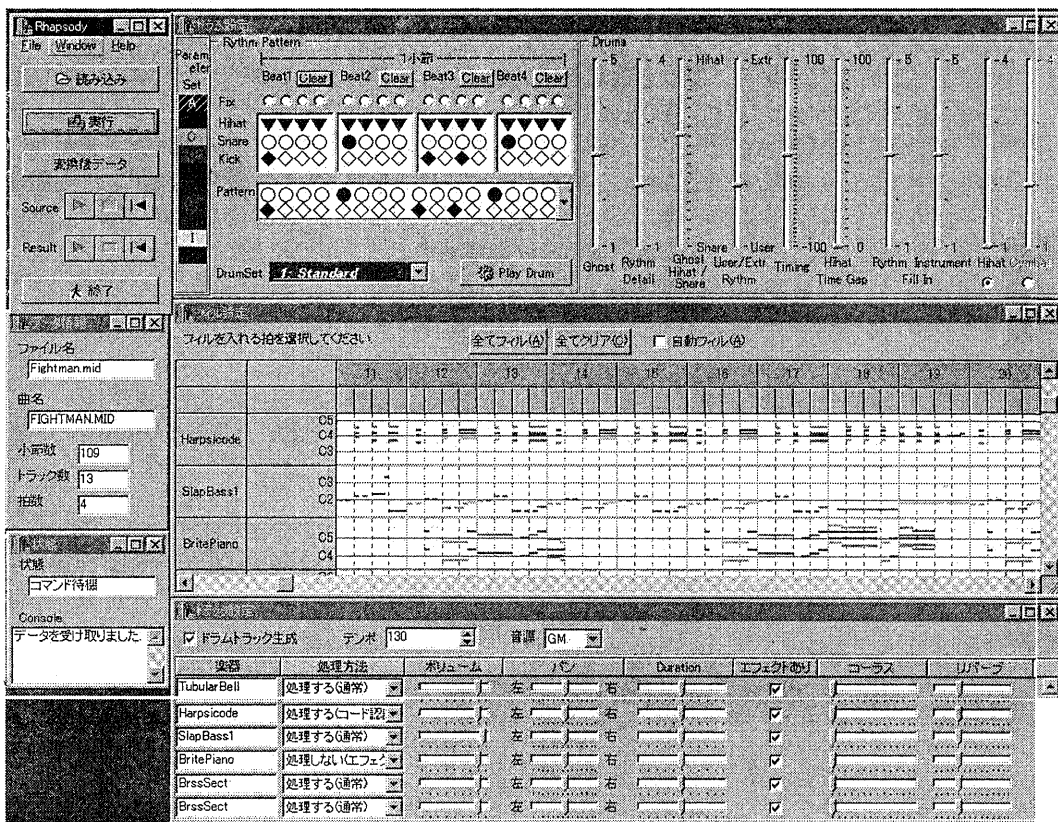


図 3: MorPH Rhapsody の画面

4 考察および結論

MorPH は、基本的な表情付けを自動的に行うシステムであり、現段階では、ギターソロのように、演奏の個性が強く出る部分の表情付けは、自動化していない。現在、クライアントは 4/4 拍子の楽曲にしか対応していないが、他の拍子の対応は容易である。

システムの評価方法として、ユーザーのアクセス数やアンケートを通しての反応が考えられる。既に得られている内部評価として、生成されるドラムパートの細かさが挙げられる。現段階までのユーザーの感想からは、MorPH の目標であった「実際に使えるシステム」に関して、いい評価が得られ、特にドラムに関しての評価が高かった。

MorPH の特徴は、高度で非単調なドラムパートを自動的に生成する点である。ドラムパートは、単純な方法を用いてユーザーと楽曲の両ドラムパターンの特徴を備えるパターンを生成するにも関わらず、曲を通して同じユーザーのドラムパターンではなく、部分毎に曲に合ったパターンを生成するため、聞いた感じが自然であるとの評価を内部的なテストで得ている。また、ユーザーのドラムパターンが与えられない場合には、推測するパターンに基づいてドラムパートを自動的に生成する。

表情付けの手法として、ルールベースにするか、それともケースベースにするかの選択肢がある。MorPH では、ルールベースにしているが、それは以下の理由によるものである。MorPH は、楽器毎の表情付けだけではなく、全体として統一感のある表情付けを目標としている。全体としての表情付けを行う場合、曲の解釈が不可欠となり、このことは、ケースベースの実装に不可欠な類似度の定義が非常に難しいことを意味する。

現在の MorPH の表情付けは、MorPH Rhapsody を用いたドラムの表情付けを除けば、我々が設定したルールに従っているため、表情付けの処理で生成される「仮想演奏家」は 1 種類しかない。そのため、ルールを一般ユーザーが編集できるように MorPH を拡張中であり、ルールの収集によって、表情付けもしくは演奏の好みを解析することが可能となり、クラシファイアシテムを用いた学習/進化機構を組み込む予定である。

参考文献

- [1] 青野 裕司, 片寄 晴弘, 井口 征士 (1997) “重回帰分析を用いた演奏表現法の抽出”, 情報処理学会 **38**:1473-1481.
- [2] T. Hoshishiba, S. Horiguchi and I. Fujinaga (1996) “Study of expression and individuality in music performance using normative data derived from MIDI recordings of piano music”, *Proc. ICMPC*, 465-470.
- [3] S. E. Dreyfus and A. M. Law (1977) *The Art and Theory of Dynamic Programming*, Academic Press, New York.
- [4] A. Friberg and J. Sundberg (1986) “A Lisp environment for creating and applying rules for musical performance”, *Proc. ICMC*, 1-3.
- [5] 福岡 俊之, 片寄 晴弘, 井口 征士 (1990) “音楽解釈システム MIS における自動演奏生成過程について”, 情報処理学会 全国大会, 1585-1586.
- [6] 片寄 晴弘, 福岡 俊之, 井口 征士 (1990) “音楽解釈システム MIS における演奏ルールの抽出について”, 情報処理学会 全国大会, 1587-1588.
- [7] 平賀 瑠美 (1998) “演奏の表情付け”, コンピュータと音楽の世界, 共立出版.
- [8] G. Widmer (1993) “Understanding and learning musical expression”, *Proc. ICMC*, 102-108.
- [9] G. Widmer (1994) “Learning expression at multiple structural levels”, *Proc. ICMC*, 102-108.