

演奏表情付けシステムのためのユーザ主導型音楽構造分析

田中 駿 二^{†1} 橋田 光代^{†1} 片寄 晴弘^{†1}

本研究は、階層的フレーズ構造の分析に焦点を当て、ユーザとのインタラクションを通じて、ユーザが求めるフレーズ構造を直感的かつ効率的に抽出するシステムの実現を目的とする。従来の演奏表情付け研究において、演奏生成の前処理にあたる音楽構造分析処理の自動化は、システムが情緒豊かな演奏表現を自律的に行うための最重要課題である。しかしその一方で、人間主導による音楽構造分析のためのシステムという観点から、分析結果に対する効率的な修正手法や操作インタフェースのあり方に関する議論は後回しにされてきた。本稿では、音楽構造解釈における多義性に配慮しつつ、煩雑になりがちな分析作業を効率的に支援する方法論について述べる。

User-oriented Musical Structure Analysis for a Performance Rendering System

SHUNJI TANAKA,^{†1} MITSUYO HASHIDA^{†1}
and HARUHIRO KATAYOSE^{†1}

The purpose of our study is to develop a system for extracting the phrase structure intended by the user intuitively and efficiently through user-system interaction. Previous studies on performance rendering mainly dealt with automation of musical structure analysis because it was considered the most important issue to autonomously achieve expressive performances. They did not therefore address efficient methods or user interfaces for correcting analysis results although it is also an important issue from the viewpoint of user-oriented musical structure analysis. In this paper, we address a method for efficiently supporting musical structure analysis taking into account ambiguities in musical structure interpretation.

1. はじめに

演奏の表情付け (performance rendering) は、音楽情報処理の中でも中心的な研究課題のひとつである¹⁾。演奏の表情付けシステムのさきがけとしての取り組みは、1980年代の Frydén²⁾, Clynes³⁾ らの研究にさかのぼる。1990年以降は、GTTM⁴⁾ や IRM⁵⁾ などの認知的音楽理論の利用、学習システム^{6),7)} や事例ベース推論によるアプローチ⁸⁾ も見られるようになったのに加え、2002年度からはシステム生成演奏の聴き比べコンテスト Rencon^{*1} が開催されている⁹⁾。

これまでの表情付けシステムの多くは、演奏生成処理の自動化に焦点が当てられてきた。中でも、演奏生成の前処理にあたる音楽構造分析処理の自動化は、システムが情緒豊かな演奏表現を自律的に行うための最重要課題である。浜中らは、GTTM⁴⁾ を計算機実装用に拡張し、楽譜から階層的なフレーズ構造を自動的に獲得するシステムの開発に取り組んでいる¹⁰⁾⁻¹²⁾。しかしながら、楽曲に対する音楽構造や演奏表現には、部分的に複数の選択肢が存在する機会が多い。浜中らはそのような場合の構造解釈に対応させるものとして、ルール間の優先度パラメータを多数導入したものの、すべての場合に通用するような共通のパラメータが存在するわけではない。

ここで、表情付けシステムを用いて演奏デザインを行うユーザの立場について考えてみる必要がある。自動化による演奏表情付け、音楽構造分析研究が進む一方で、生成物に対する修正手法やユーザインタフェースがどのようにあるべきかという議論は後回しにされてきた。結果、演奏デザインの過程において、システムの出力した生成物に対し修正を施したい場合に煩雑な操作を必要とすることが少なくなかった。人間が使う上でそのような操作性は重要で、人間のデザイン活動を支援することができるシステムの開発が求められる。

本研究は、表情付けシステムにおける音楽構造分析に焦点を当て、ユーザとのインタラクションを通じて、ユーザが求めるフレーズ構造を直感的かつ効率的に抽出することが可能なシステムの実現を目的とする。以下、第2章で構造分析について検討し、第3章ではシステムの提案手法、第4章でアルゴリズムについて述べ、第5章で実際の分析事例を挙げて考察を行う。

^{†1} 関西学院大学理工学研究科
Kwansei Gakuin University

*1 <http://www.renconmusic.org/>



図1 フレーズ構造の多義性（モーツァルトのピアノソナタ K.331 第1楽章の冒頭4小節より）。
両者の違いを体感するには、最下層のフレーズ境界で息継ぎをしながら口ずさんでみるとよい。

2. 楽曲構造分析

2.1 階層的なフレーズ構造と解釈の多義性

音楽構造として扱われる要素には、代表的なものとして、旋律（音の連続的な並び）、拍節（リズム）、和声などが挙げられる。各要素においては、ある時刻における音とそれに前後する音との間になんらかの関係を持ったまとまり（フレーズ）が形成され、複数のフレーズがさらに互いに関連付けられて、図1に示すような階層的なフレーズ構造が形成される。このフレーズ構造に対する解釈、表現、聴取者への伝達が音楽表現の主題といっても過言ではなく^{13),14)}、これまでに様々な楽曲分析においてフレーズ構造が注目されてきた¹⁵⁾⁻¹⁷⁾。

楽曲におけるフレーズ構造は、大局的には多くの人に共通のものとして捉えられるが、階層レベルが局所的になるにつれて、異なる構成を採り得る場合がある。図1で言うと、図上と図下では、四小節全体でひとまとまりという点では同じであるが、その内部（下位レベル）においては、図上は小節単位でのフレーズ構造であるのに対し、図下は、この楽譜においては音価の長い四分音符の後でフレーズを区切るように構成されている。

このように、同じ楽譜に対して異なる局所的フレーズ構造が構築できるのには次のような点に基づくと考えられる。

- (1) 演奏上の時系列イベントに対して得られる聴覚的なまとまり感（拍節表現による発音時の強弱差、各音符が発音されるタイミングの間隔のばらつきなど）
 - (2) 過去の演奏聴取体験から人間が経験的に得てきた音楽的知識や慣習、法則
- 構造分析においてこれらをどの程度重視すべきかは、最終的には分析者や音楽家の裁量に任

されており、このことが、計算機で一意的階層的フレーズ構造を同定することを困難にする大きな原因となっている。

2.2 GTTMに基づくフレーズ構造分析

階層的なフレーズ構造を形式的に記述する方法論として、Lerdahlらは音楽経験者の認知的直観を整理したうえで、生成文法に基づく構造分析手法（GTTM）を示している⁴⁾。GTTMでは、楽譜からフレーズ構造を得るための個々のフレーズを導くルール群が提唱されており、構造を記述するためのグルーピング構成規則ルール（Grouping Well-formedness Rules: GWFR）と、認知的直観を具体化した選好ルール（Grouping Preference Rules: GPR）に分類される^{*1}。これらをさらに拡張、定式化し、計算機による完全自律型の構造分析を目指したのが浜中らの研究^{11),12)}である。

浜中らは、前節で述べたような多義的なフレーズ構造解釈に対応させるものとして、ルール間に優先度パラメータを設けた¹¹⁾。しかしながら、扱う優先度パラメータの数が多く、構造分析においてどのルールが有効であるかを検証することが難しいという問題がある。その後¹²⁾において、パラメータを自動で取得する手法の提案が行われたが、多義性の問題もあり、必ずしも十分な精度が得られていない。

2.3 ユーザ主導のフレーズ構造分析支援

演奏の表情付けは一種のデザインである。多義性の解消やユーザの嗜好を反映するためには手作業が伴う。人間の手作業を含めた演奏デザイン全体での効率性やユーザ支援の在り方を考慮していくことが求められる。図1程度の短い入力楽譜ならばともかく、表情付けを施すのに階層的なフレーズ構造を入力しなければならないとなると、演奏時間が1～2分程度の小曲^{*2}でさえ、かなり煩雑な作業となる。

浜中らはGTTM Editorというシステムにおいて楽曲をタイムスパンツリーで表現し、視覚的に木構造の従属関係の操作と構造の分割を可能にしている。しかし、GTTMのルールをよく理解していないユーザの利用を想定した場合、優先度パラメータを操作してユーザが意図する演奏表現の構造に近づけることは難しい。デザイン支援の観点からは、ユーザが直感の赴くままにシステムを操作することができ、なおかつ煩雑な操作を行わずに最大の効果を上げることが求められる。ユーザにとっては、優先度パラメータを操作するより、フレーズ、つまり、グループの境界を直接操作（= Direct Manipulation）し、必要に応じ

*1 このほかにも、拍節構造を得るルール群、フレーズ構造と拍節構造に基づいて楽曲の木構造を記述する2種の簡易ルール群が提唱されている。

*2 ポピュラー音楽の楽曲は通常3～5分程度、クラシック音楽では独奏曲でも10分を超える作品が散在する。

て、優先度パラメータの利用法に關与するというインタフェースの方が使いやすい。このような考え方に基き、本研究では、以下のような構造分析のインタラクティブデザイン指針を採用することにした。

- ユーザが楽曲の多義性（曖昧性）を決定する
- ユーザはフレーズ構造を直接操作する
- システムは、ユーザの決定した部分以外の補完分析を担当する。

3. システムデザイン

本研究では、少ない入力力でユーザの意図を反映させた楽曲全体の階層構造を獲得可能にすることを目的としてシステムデザインを行っている。図2に、本システム全体の流れを示す。

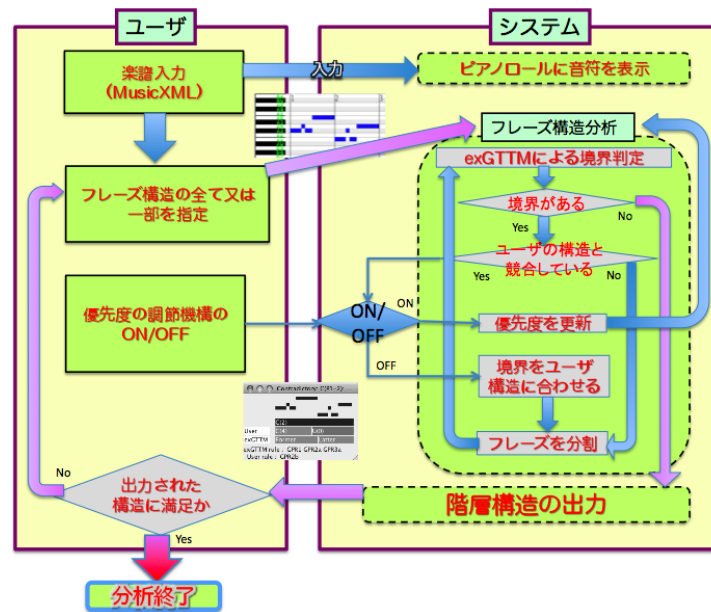


図2 システム概要図

3.1 分析過程

システムとユーザのインタラクションによって最適な階層構造を獲得する。インタラクションプロセスは以下の通りである。

step1 ユーザが、楽曲に対してフレーズ構造の全て又は一部を指定する

step2 指定されたフレーズに沿って、システムによる構造分析が行われ、その結果がフィードバックされる

step3 ユーザが納得する構造が得られるまで step1,2 を繰り返す

これまでのユーザ主導型演奏デザインシステム¹⁸⁾ではユーザがすべてのフレーズ構造を手入力する必要があった。本提案手法では、ユーザが特に指定したい部分のみのフレーズの境界を与え、それ以外のフレーズ境界および階層的フレーズ構造をシステムが実施するというシステムデザインを採用している。ユーザが指定した場所以外のフレーズ境界および階層的フレーズ構造は、浜中らの exGTTM を利用して分析が行われる。

システムは、ユーザが指定したフレーズ境界に対応した GPR のユーザプレファレンスを保持している。ユーザが指定しなかった部分のフレーズ境界の自動分析については、a) ユーザプレファレンス優先、b) ユーザの指定した同型モチーフのみに対してのユーザプレファレンス優先、c) exGTTM 優先、の3つのオプションが用意されている。

階層的なフレーズ構造の分析（縦方向）については、構築した階層構造がユーザの指定するフレーズ構造と exGTTM での基本 GPR とが反駁する可能性がある。この場合、常にユーザの指定したフレーズ境界を優先する形で分析がなされる。

以上の処理によってシステムが推定した構造に対して、ユーザが修正を施したい場合がある。その場合、ユーザは、当該の階層的フレーズ境界を直接修正することができ、階層構造もそれに合わせて実時間で再分析がなされる。あわせて、システム内部の GPR のユーザプレファレンスの更新もなされる。ユーザは、随時、フレーズ構造に対応した演奏プレイバックを聞きつつ作業が実施できる。この一連のインタラクションを繰り返すことによってユーザ主導の構造分析を実施する。

3.2 ユーザプレファレンス

ユーザプレファレンスは浜中らが GTTM の各選好ルール間の優先度を決定するために導入したパラメータを用いている（図3）。パラメータは(1) S_R : GPR2a, 2b, 3a, 3b, 3c, 3d, 4, 5, 6 の各ルールの強さを表すパラメータ ($R \in 2a, 2b, 3a, 3b, 3c, 3d, 4, 5, 6$)、(2) T_4 : GPR4 で、GPR2, 3 の効果が明白であるかどうかを決める閾値、(3) $\hat{\sigma}$: GPR5 で用いる平均をグループの中心とする正規分布の標準偏差、(4) W_m : GPR6 で、各音の発音時刻の類似

度と音高差の類似度のどちらを重視するかを決めるパラメータ (5) W_l : GPR6 で並列な区間の長さをどのくらい重視するかを決めるパラメータ (6) W_s : GPR6 で音符 i が並列な区間の始端あるいは終端のいずれになる方を重視するかを調節するパラメータ (7) T_{low} : 低レベルの境界であるかどうかを決める閾値の 7 種類 15 個である . 値域はそれぞれ 0~1, W_s のみ 0~0.1 となっている .

GPR1	局所的な境界が前後の遷移より強い
GPR2a(スラー/休符)	i 番目の音符の消音時刻と $i+1$ 番目の音符の発音時刻の時間間隔が前後の遷移より大きい
GPR2b(アタックポイント)	i 番目の音符の発音時刻と $i+1$ 番目の音符の発音時刻の時間間隔が前後の遷移より大きい
GPR3a(音高差)	i 番目と $i+1$ 番目の音符の音高差が前後の遷移より大きい
GPR3b(強弱)	i 番目と $i+1$ 番目の音符のダイナミクスの変化が前後の遷移より大きい
GPR3c(アーティキュレーション)	i 番目と $i+1$ 番目で楽譜上の音符の長さや実際に演奏された音の長さの比の変化があり、前後で変化がない
GPR3d(音価)	i 番目と $i+1$ 番目が異なった音価で、前後が同じ
GPR4	GPR2,3が集中する境界を優先する
GPR5	隣接するフレーズの長さが等しくなることを優先する
GPR6	音符が並列な区間の始端か終端になる

図 3 GTTM のグルーピング嗜好ルールを拡張した浜中らのルール群 (exGTTM)

ユーザがフレーズを指定して分析を行った際、そのフレーズと似たようなフレーズが楽曲中の他の場所に存在する場合は同じような構造で分割することが望ましい . ここでは (i) ユーザプレファレンスをユーザが指定したフレーズ構造に合わせて調整し、楽曲全体を分析することで対応する方法と、(ii) 指定したフレーズと近似しているフレーズを楽曲中から検索し同じ構造とする方法の二通りの中から任意の分析法を選択し、適用する . 数値調節や検索の作業はシステムによって行われるため、ユーザはこれらを意識することなく分析作業を行うことができる .

GTTM を理解しているユーザであれば、逆にユーザプレファレンスを操作して構造分析を行うことも可能である . 同じようなフレーズでも演奏表現を変えたい場合は先ほどの二通りの方法を使わずにあらかじめ設定されたユーザプレファレンスのみを用いて分析することが可能である .

3.3 ユーザへのフィードバック

システムはユーザに対して視覚的にフィードバックする機構を持つ . それぞれの構造に対して何故その箇所が分割されているかという情報を GTTM の嗜好ルールを用いて表示している . また、分析の際にユーザが指定したフレーズ構造と exGTTM の分析した境界とで競合が起きた場合、ユーザの指定したフレーズと exGTTM が出力するフレーズ構造とでどのような違いがあるかを表示する . システムの分析結果やユーザが選択した構造が嗜好する理由を示すことにより分析への理解を深めることができる .

4. アルゴリズム詳細

本システムは CrestMuseXML API ^{*1} を用いて開発を行っている . 構造分析のアルゴリズムには浜中らの exGTTM ⁴⁾ を用いている .

4.1 構造分析プロセス

システムが行う処理は以下の通りである .

- step1 楽曲全体をフレーズ構造のトップレベルとするフレーズを生成する
- step2 exGTTM に基づき、フレーズに嗜好ルール (GPR) を適用し、境界を求める
- step3 境界が存在する場合、その境界がユーザが指定したフレーズ構造の中に存在すれば step4 へ進む . 境界が存在しない場合、この一連の処理を抜ける
- step4 優先度調節機構の ON/OFF を確認し、
 - (1) ON であれば、優先度パラメータを更新し、step1 に戻る
 - (2) OFF であれば、境界をユーザ構造に合わせ、step5 へ進む
- step5 境界でフレーズを二分割し、子階層のフレーズを生成する
- step6 生成した子階層のフレーズそれぞれに対して step2~6 を繰り返す

4.2 競合箇所の処理

exGTTM で求めた境界がユーザが指定したフレーズ構造を分割する位置である場合、以下の 2 つの処理のどちらかを実行する . 2 つの処理はユーザによって選択される .

● 優先度パラメータの調節

優先度パラメータ S_R ($R \in 2a, 2b, 3a, 3b, 3c, 3d, 4, 5, 6$) を調節する . exGTTM の分析結果が嗜好するルールの優先度 (exGTTM preference rule) を 0.1 下げ、ユーザの指定したフレーズが嗜好するルールの優先度 (USER preference rule) を 0.1 上げて再

*1 <http://www.crestmuse.jp/cmxf/>

分析を行う。

$$S_R = \begin{cases} S_R - 0.1 & S_R \text{ exists only in exGTTM preference rules} \\ S_R + 0.1 & S_R \text{ exists only in USER preference rules} \\ S_R \pm 0 & \text{otherwise} \end{cases} \quad (1)$$

$R \in 2a, 2b, 3a, 3b, 3c, 3d, 4, 5, 6$

● 境界の再設定

exGTTM で求めた境界がユーザが指定したフレーズ構造を分割する位置である場合、ユーザグループを保ち、GPR5 に倣ってユーザグループの上位階層を分割する。ユーザグループの始点と終点のうち、上位階層の中心点に近い方を分割位置とする。上位構造 P(Parent) の分析結果がユーザ指定のフレーズ構造 U(User) を分割する時、分割位置 A の計算は以下の通りである。始端時刻を s(start)、終端時刻を e(end) とする。

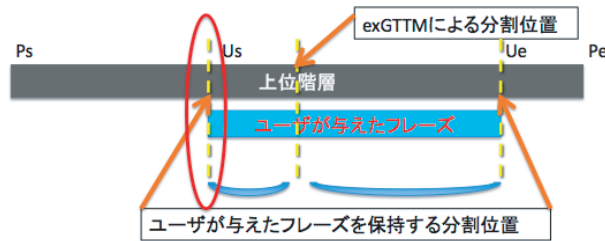


図 4 競合箇所の処理

$$A = \begin{cases} U_s & \left| \frac{Pe + Ps}{2} - U_s \right| \leq \left| \frac{Pe + Ps}{2} - U_e \right| \\ U_e & \text{otherwise} \end{cases} \quad (2)$$

優先度パラメータを動かさずに分析を行う場合には強制的にユーザの構造に合わせて分析を行う。分割点は GPR5 のルールに基づいて、ユーザのフレーズ構造の前後のうち、上位構造の中心点に近い位置で分割を行う。その際は競合を起こした場所を表示してユーザに知らせる。

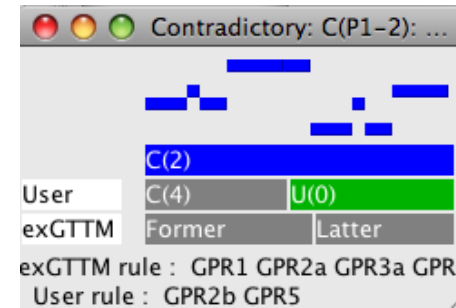


図 5 フレーズ構造の競合

5. システムの動作と考察

5.1 分析事例

モーツァルトのピアノソナタ K.331 第 1 楽章の冒頭 4 小節を対象に、システムを用いた構造分析を行った。

図 6 はユーザが楽曲全体に渡って空白がないようにフレーズを指定したものである。図中の緑色の部分がユーザが入力したフレーズ、それ以外の部分がシステムより生成された上位階層と下位階層のフレーズである。分析の過程において、ユーザのフレーズとの間に競合が生じ、システムがユーザのフレーズに合わせて修正を施した部分が青色で示されている。

図 7 ではユーザがフレーズを一カ所だけ指定し、優先度調節機構を OFF にした上で、シ

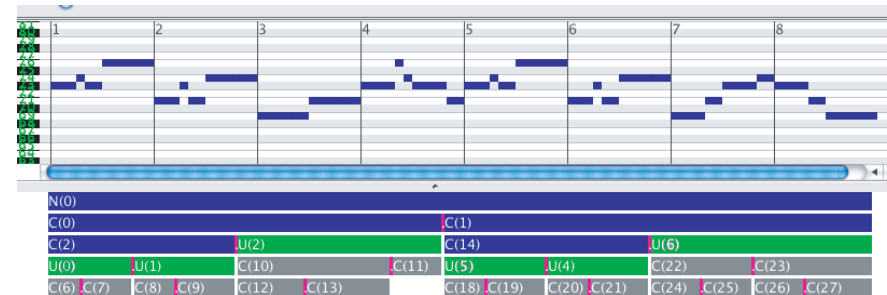


図 6 全てのフレーズを選択して分析

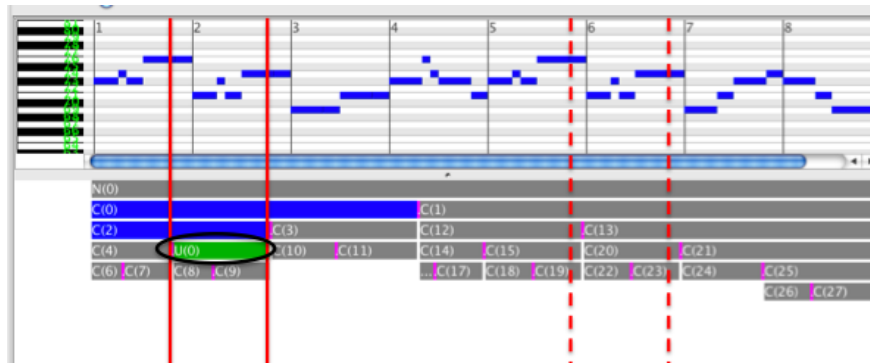


図 7 一カ所のみを指定して分析（優先度調節機構 OFF）

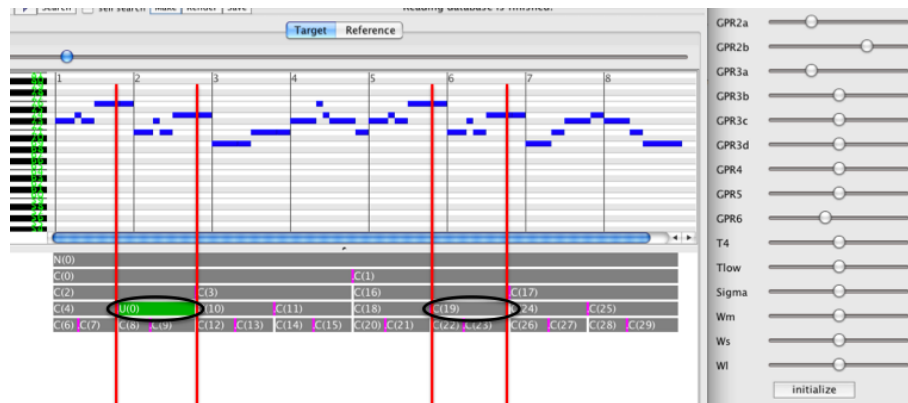


図 8 一カ所のみを指定して分析（優先度調節機構 ON）

システムに残りの部分を分析させた例である。優先度パラメータは同じ値に初期化した。システムによって階層構造は出力されたものの、この譜例で通常多くの人が判断するであろう4小節目と5小節目の間が区切れていなかったり、ユーザのフレーズと同形の旋律である部分（赤い点線内）がユーザの入力とは違う形で出力されている。

この状態で優先度調節機構を ON にし、システムに再分析をさせたのが図 8 である。ここではユーザのフレーズが成立するような優先度パラメータが計算され、それに基づいた階層構造が出力されている。図 7 ではユーザのフレーズの上の階層で競合を起こしていたが、

図 8 ではそれが解消された。また、ユーザのフレーズと同じ音形の部分も同様の構造が出力されるようになった（図 8 赤線内の黒丸印）。

5.2 検 討

第一章で述べたように、従来の音楽構造分析や演奏表情付けシステムは自動化を主眼において研究が進められてきた。浜中らは認知的音楽理論（GTTM）を導入し、構造分析の完全自動化を目指している。Kagurame Phase-II では MusicXML からスラーと連符の情報を抽出し階層構造を形成している。しかしながら第 2 章で述べたように、音楽構造は多義性があり、簡単に実現するものではない。中途半端な処理は人間が使うツールとしては使いにくく¹⁹⁾、ユーザの操作性とシステムの自動処理をバランスよく用いること（directability）がシステムデザインには重要である²⁰⁾。

これまでのユーザ主導型演奏デザインシステムではユーザがすべてのフレーズ構造を手入力する必要があった。本研究ではユーザがフレーズの一部を指定してシステムが残りを補完する手法を提案している。その際、ユーザの入力に合わせた処理の選択やパラメータの調整を自動的に行うことでユーザとのインタラクションを図っている。

6. おわりに

本稿ではユーザ主導で楽曲の構造分析を行う手法を述べた。ユーザが主導的にフレーズ構造を入力する機構と浜中らの exGTTM に基づく自動分析を組み合わせ、優先度パラメータの調節や構造を直接修正する機構を導入した。分析の結果、少ない入力でユーザが嗜好する多様な階層的フレーズ構造を出力できることを確認した。今後はユーザプリファレンスについて有用性の検討を行う。

参 考 文 献

- 1) 片寄晴弘：音楽生成と AI，人工知能学会誌，Vol.19, No.1, pp.21-28 (2004).
- 2) Frydén, L. and Sundberg, J.: Performance Rules for Melodies. Origin, Functions, Purposes, *International Computer Music Conference (ICMC) Proc.*, ICMA, pp. 221-225 (1984).
- 3) Clynes, M.: Secretes of Life in Music, *Proc. of International Computer Music Conference (ICMC)*, pp.225-232 (1984).
- 4) Lerdahl and Jackendoff: *A Generative Theory of Tonal Music*, MIT Press (1983).
- 5) Narmour, E.: *The Analysis And Cognition Of Basic Melodic Structures*, the University of Chicago Press (1977).
- 6) Widmer, G.: Learning Expressive Performance: The Structure-Level Approach,

Journal of New Music Research, Vol.25, No.2, pp.179–205 (1996).

- 7) 石川 修, 片寄晴弘, 井口征士: 重回帰分析のイタレーションによる演奏ルールの抽出と解析, 情報処理学会論文誌, Vol.43, No.2, pp.268–276 (2002).
- 8) Arcos, J., de Mantaras, R. and Serra, X.: SaxEx: A Case-Based Reasoning System for Generating Expressive Musical Performances, *Journal of New Music Research*, Vol.27, No.3, pp.194–210 (1998).
- 9) 平賀瑠美, 平田圭二, 片寄晴弘: 蓮根: めざせ世界一のピアニスト, 情報処理, Vol.43, No.2, pp.136–141 (2002).
- 10) Hamanaka, M., Hirata, K. and Tojo, S.: Implementing "A Generative Theory of Tonal Music", *Journal of New Music Research*, Vol.35, No.4, pp.249–277 (2006). An earlier version of this paper won the Journal of New Music Distinguished Paper Award at the International Computer Music Conference, 2005.
- 11) 浜中雅俊, 平田圭二, 東条 敏: 音楽理論 GTTM に基づくグルーピング構造獲得システム, 情報処理学会論文誌, Vol.48, No.1, pp.284–299 (2007).
- 12) 浜中雅俊, 平田圭二, 東条 敏: タイムスパン木獲得システムの完全自動化, 情報処理学会研究報告 2007-MUS-71, pp.93–98 (2007).
- 13) 小澤征爾, 堤 剛, 前橋汀子, 安田謙一郎, 山崎伸子 (編): 斉藤秀雄講義録, 白水社 (1999).
- 14) 保科 洋: 生きた音楽表現へのアプローチ: エネルギー思考に基づく演奏解釈法, 音楽之友社 (1998).
- 15) Reti, R.: ベートーヴェンピアノ・ソナタの構築と分析, 音楽之友社 (2003).
- 16) Badura-Skoda, E. and Badura-Skoda, P.: モーツァルト演奏法と解釈, 音楽之友社 (1963).
- 17) 池辺晋一郎: バッハの音符たち 池辺晋一郎の「新バッハ考」, 音楽之友社 (2000).
- 18) 橋田光代, 長田典子, 河原英紀, 片寄晴弘: 複数旋律音楽に対する演奏表情付けモデルの構築, 情報処理学会論文誌, Vol.48, No.1, pp.248–257 (2007).
- 19) D. ノーマン: 未来のモノのデザイン, 新曜社 (2008).
- 20) 片寄晴弘, 橋田光代: 生成系音楽支援システムの Directability 視点からの考察, 情報処理学会研究報告 2007-MUS-071, Vol.2007, No.81, pp.99–104 (2007).