

## 表情情報の操作を目的とした音楽の構造モデルに関する研究

鈴木泰山, 徳永健伸, 田中穂積  
東京工業大学 情報理工学研究所  
{taizan, take, tanaka}@cs.titech.ac.jp

本論文では、計算機で音楽の表情情報を処理することを目的とした楽曲の構造モデルの提案を行なう。我々はここ数年行なっている「人間と協調する伴奏システム」の研究の一貫としてリハーサルシステムの研究を進めている。このシステムは独奏者の演奏や独奏者との対話からその特徴や意図を抽出し、伴奏システムの演奏に反映させることを目的としている。リハーサルを行なうためには楽曲の構造の知識が必要であると考えられるが、現在のシステムは楽曲構造に関する情報を殆んど保持していないので有効なリハーサルが行えない。リハーサルシステムに構造情報として楽曲の構造モデルを導入することで独奏者の演奏の細かな特徴の抽出や要求に対する有効な対応が可能となる。

### A Computational Model of Musical Structure for the Manipulation of Musical Expression

Taizan Suzuki, Takenobu Tokunaga, Hozumi Tanaka  
Graduate School of Information Science and Engineering  
Tokyo Institute of Technology

This paper proposes a computational model of musical structure which is aimed at dealing with musical expression. We have been developing a "Rehearsal System" as a part of our "Cooperative Accompaniment System". The purpose of this system is to extract the characteristics and intention of a solist from performance and conversation, and to adjust the performance of the accompaniment system according to this information. Rehearsal requires some musical knowledge, making it limited in its existing form. The model enables our system to analyze a solist's performance effectively and it is helpful in coping with a solist's requests.

#### 1 はじめに

われわれはここ数年に渡って「人間と協調する演奏システム」に関する研究を行なっている [1]。その研究の一つとしてリハーサルシステムがあげられる [2]。人間が音楽を演奏する際には実際に楽譜に記載されている情報の他に、テンポや音の強さなどの演奏表情に関する情報を演奏者が音楽知識に基づいて付与し表情豊かな演奏を行なっている。この曲に付加された演奏表情に関する情報を演奏プランと呼ぶ。演奏プランは演奏者が音楽知識に基づいての主観的に付加している情報なので演奏者によってその与え方が異なることが多い。演奏者が単独で演奏する場合に問題は無いが、複数の演奏者が競演する場合に

はそれぞれの演奏者が異なる演奏プランを持っているとアンサンブルが不自然なものになってしまう。伴奏システムにおいても計算機と独奏者の演奏プランの相違は、アンサンブルを不自然なものにする大きな要因の一つであり、より自然なアンサンブルを行なうためには人間と伴奏システムの間でリハーサルにあたるような作業を行ない、お互いの演奏プランを同じものにしておく必要がある。このためのシステムがリハーサルシステムである。

リハーサルシステムでは、独奏者の演奏とシステムの演奏プランとのずれや、独奏者との対話によって得られたシステムへの要求などをもとに、システムの演奏プランの変更を行なう。

リハーサルを有効に行なうためには独奏者の演奏とシステムの伴奏とのずれを演奏の特徴として検出するだけでは不十分である。そのずれが楽曲のどの範囲に適用されるものであるのか、また楽曲のその部分に特有の特徴なのかそれとも繰り返しなど同様の構造を持つ部分にも適用できるもののかなどを考慮しなければならぬ。したがって、楽曲の構造に関する知識が必要になってくる。

そこで、本研究では、リハーサルシステムにおいて独奏者の演奏プランの抽出を助け、また独奏者との対話から得られた要求に対して柔軟に対応するための楽曲の構造情報として、表情情報の操作を目的とした構造モデルの提案を行なう。

まず、第2章で我々のリハーサルシステムの概要について説明し、第3章では現状のシステムとその問題点を示す。第4章ではその解決法として、表情情報の操作を目的とした構造モデルについて述べる。

## 2 リハーサルシステム

### 2.1 リハーサルシステムについて

我々が研究を進めている伴奏システムは、独奏者、すなわち人間の演奏におけるテンポの揺らぎに柔軟に対応し自然なアンサンブルを行なうことを目的としている。自然なアンサンブルを行なうためには楽譜に記載されている情報のほかに、テンポの変化や打鍵の強さなど楽譜からでは判断しにくい演奏の表情についての情報を付与することが必要である。そこで、我々のシステムでは独奏及び伴奏の楽譜とともに、伴奏の表情をどのようにつけるかという情報、即ち伴奏の「演奏プラン」を付与している。

演奏プランは、実際の楽譜などに基づいて、多くの人に共通していると思われるものを人手で与えているが、いずれにせよ細かい部分、特に具体的な数値の設定などには入力者の主観的な判断が多分に含まれている。したがって実際に演奏を行なってみるとシステムの演奏プランと独奏者の演奏プランが異なっているために演奏がうまく噛み合わなくなってしまうことが多い。

演奏プランの相違は、人間同士でアンサンブルを行なう場合にも頻繁に発生する。初対面の二人がアンサンブルを行なってお互いに納得のいく演奏ができることはまずないであろう。このような場合、演奏者どうしは何度かリハーサルを繰り返すことでお互い

の演奏プランの相違を認識し、よりよいアンサンブルを実現しようとしていると考えられる。計算機と人間の間でアンサンブルを行なう場合にも、リハーサルを行なってお互いの演奏プランをあらかじめ知るとともに、その相違を認識することが自然なアンサンブルの実現に有効であると考えられる。

Vercoe は独奏者とのアンサンブルを何度も行ない統計情報から独奏者固有の演奏の傾向を抽出するリハーサルのモデルを提案している [3]。しかしながら彼の方法にはいくつかの問題点がある。まず、統計情報を収集するためにアンサンブルを何度も行なわなければならないことがあげられる。あまりに多くの演奏を行なうことは独奏者に対してかなりの負担になるし、繰り返し演奏することによって独奏者の演奏自体が変化してしまうことも考えられる。また、独奏者の演奏プランを統計情報だけで抽出できるかどうか疑問である。計算機を相手にアンサンブルを行なうとなれば、独奏者は自然と計算機の演奏をある程度考慮して演奏してしまうので、独奏者の演奏プランが正確に伝わらないことも考えられる。

### 2.2 ツーフェイズリハーサル

我々はこれらの問題点を解決するリハーサルの手法としてツーフェイズリハーサルという方法を提案した。人間同士でリハーサルを行なう場合を考えてみると、演奏者はただ単にお互いの演奏に合わせて演奏して、相手の演奏の特徴を理解するだけではなく、その演奏によって得られた情報をもとに相手の演奏プランについて質問をしたり、また逆に相手に自分の演奏プランを伝えて何らかの要求をしたりすることでお互いの演奏プランをより近いものに行なうとしていると考えられる。そこで、リハーサルシステムに、従来の独奏者の演奏からその特徴を抽出する手続きの他に、独奏者との間で質問や要求を交換する手続きを導入し、より人間同士のものに近いリハーサルを実現したものがツーフェイズリハーサルである。ここでは、前者の手続きを「演奏フェイズ」、後者を「コミュニケーションフェイズ」と呼ぶ。

「演奏フェイズ」ではその時点でシステムが保持している演奏プランを利用して独奏者とのアンサンブルを行なう。それと同時に独奏者の演奏と独奏部分の楽譜などとの比較を行ない、システムが独奏者に期待している演奏との相違を独奏者の演奏の特徴として記録する。独奏者の希望によっては曲の途中

からアンサンブルを開始したり途中で中断したりすることも可能である。

「コミュニケーションフェイズ」では「演奏フェイズ」で得られた独奏者の演奏プランに基づいて、独奏者と対話を行なう。対話はシステムから独奏者への質問と独奏者からシステムへの要求の二種類からなる。独奏者への質問は、システムが抽出した独奏者の表情づけの特徴の確信度が低い場合に、独奏者にシステムの解釈が正しいか確認する目的で行なわれる。システムへの要求は主にテンポや音量について独奏者の希望をシステムに直接伝える手段として利用される。

「演奏フェイズ」と「コミュニケーションフェイズ」の組合せを「ステージ」と呼び、一連のリハーサルではこのステージが何度か繰り返される。

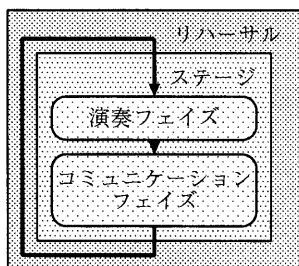


図 1: ツーフェイズリハーサル

### 3 現状のシステムとその問題点

#### 3.1 現在のリハーサルシステム

現在のリハーサルシステムは「自主性を考慮した伴奏システム」の機能を拡張したシステムとして Macintosh 上に実装されている [4]。この「伴奏システム」では伴奏部分の演奏プランを楽譜情報の一部としてあらかじめ与えている。楽譜情報はテンポの変更や音符の演奏など(以下これらをイベントと呼ぶ)とそのイベントが発生する拍との組合せからなる逐次的なデータ [1] として記述されており、リハーサルシステムでは、独奏者の表情づけの特徴や独奏者からの様々な要求をもとに、このデータ上の演奏プランに対し修正や変更を行なう。

演奏フェイズでは、独奏者の演奏のイベントを楽譜情報と同じような逐次的な情報の羅列として記録

する。その際に、楽譜情報に記述されたイベントの間隔と比較し、独奏者の演奏での拍の間隔が楽譜情報とどの程度ずれているかについて計算する。

コミュニケーションフェイズではこれらの情報を基に独奏者の演奏プランを抽出する。われわれの伴奏システムでは、独立度 [1] という概念を利用することによって計算機にある程度の自主性を持たせているので、局所的なテンポの揺らぎの影響を受けて伴奏が不自然になることは少ない。従って、局所的なテンポの揺らぎを細かく抽出する必要はないと考えている。逆に、大局的なテンポの変化が独奏者の演奏プランと異なる場合は不自然な演奏になりやすい。例えば、間奏など独奏者の演奏が休みになり伴奏システムが自主的に伴奏の演奏を行う場面に変化する場合、独奏者がそれまで演奏していたテンポから計算機の演奏プランにあるテンポに急激に変化し、演奏自体が不自然になることがよくある。そこで、このシステムでは主に大局的なテンポの変化に関する独奏者の演奏プランを抽出し、これにもとづいて楽譜情報にあるテンポを変更し、次の演奏に反映させている。また、独奏者が全体のテンポや音の強さなどを変更するように要求した場合には、その範囲とどの程度変化させるかを入力させ、それに基づいて楽譜情報に記述されている演奏プランを変更する。

#### 3.2 現在のシステムの問題点

奥井はこのリハーサルシステムを利用することにより独奏者の演奏プランに近い伴奏を行なえることを示した [4]。しかしながら、人間同士のリハーサルに比べると現在のリハーサルシステムは不十分な点が多い。

まず問題点としてあげられるのが、楽曲内の繰り返しや類似した構造に対する演奏プランの取り扱いである。通常の楽曲の場合、一つの楽曲の中に同様の構造を持つ部分が何度も出てくることが多い。これらの部分に対しては同一の演奏プランが適用されることが予想される。勿論、類似した部分に異なる演奏プランが適用されることもあるが、演奏プランが根本的に異なることは稀であり、基本となる演奏プランは共通のものである場合が多い。したがって独奏者が何らかの要求を行なったときに、その曲に出現する類似した部分に対しても暗に同様の変更を期待していることが多いと考えられる。

しかしながら現在のリハーサルシステムでは、楽

曲の繰り返しや構造的な類似に関する情報を持っていないため、ある部分に対する独奏者の要求を曲中の類似した部分に適用することができない。よって、独奏者は繰り返しなどによって類似した部分が登場する度に同様の要求を行わなければならない。主題やフレーズなど比較的大きな部分に対する要求であっても少なくとも数回程度は同様の要求を繰り返す必要があるだろう。それがもし、「一拍目をもっと強く」のような小節単位の要求であったとすると、曲の小節数と同じ回数だけ要求を行わなければならない、とても現実的とはいえない。

もう一つ、独奏者の演奏の特徴抽出能力が非常に貧弱である点があげられる。現在のシステムでは、一定以上の長さの範囲で、楽譜上のテンポと独奏者のテンポの差異が段々と変化する、あるいは常にはほぼ一定の差がある場合に、それを独奏者の演奏プランとして解釈している。しかしながら、長い範囲にわたって計算機の演奏プランと独奏者の演奏プランのテンポに差がある場合には、計算機が独奏者に質問したり自主的に修正を行ったりするまでもなく、最初のステージのコミュニケーションフェイズで独奏者からの要求としてあげられる。しかも、それ以外の細かい変化を抽出する能力がないので、第二ステージ以降は独奏者がシステムに対し一方的に要求を出すだけになることが多い。

よいリハーサルは独奏者と伴奏者が対等な立場でお互いに意見を述べあうことで成立すると考えられる。したがって現状のように独奏者がシステムに対して殆んど一方的に要求を行なうだけのシステムは、リハーサルシステムと呼ぶには不十分である。

## 4 音楽構造のモデル化

### 4.1 音楽知識の必要性

我々の伴奏システムでは、計算機に表情情報や演奏プランを生成させるのは様々な音楽知識を要する非常に困難な作業であるという考えから、計算機には楽曲の構造情報などの音楽知識を殆んど与えず、演奏プランなどその生成に音楽的知識を必要とする情報は全て人間があらかじめ生成して計算機に与えるというアプローチをとっていた。従って、このような伴奏システムの上に成り立っている現在のリハーサルシステムは構造情報も含め音楽的知識を殆んど保持していない。

一方でリハーサルという作業はかなり音楽的知識が必要な作業である。特にコミュニケーションフェイズでは、言葉など直接的には曲と関係ない情報として与えられる独奏者の要求を音楽の表情情報と結び付けて演奏プランに反映させたり、独奏者の演奏データから独奏者の演奏の特徴や演奏プランを抽出するなど、通常人間が行なう際にもある程度の音楽知識を必要とする作業を行わなければならない。また、演奏フェイズにおいても楽譜情報と比較してテンポの変化や拍のずれを検出するだけでは不十分であり、さらにそれらの情報と曲の構造の関係などを考慮し、それぞれの特徴が楽曲のどの範囲に対するものかなど、音楽的な意味をある程度解釈する必要があると考えられ、そのためには楽曲の構造に関する知識が不可欠である。このようなことから、楽曲の構造情報などの知識情報を持たずにリハーサルを行なうこと自体に無理があると思われる。

### 4.2 構造モデルと表情情報の操作

リハーサルは、独奏者の演奏からその演奏表情の特徴を抽出し伴奏部分の表情情報を修正するという点で表情情報の学習、生成に近いタスクであると考えられる。一方で3.2節で述べたような問題点を解決するためには、楽曲のそれぞれの構造上の境界や、それぞれの部分同士の構造の類似性など、対象曲の構造に関する情報が必要である。

表情情報に関する知識の自動獲得や楽曲の構造情報の記述についての研究は音楽情報処理の分野の中でも歴史が古く、数多く発表されている[5-12]。それぞれの研究では楽曲の構造や表情情報を取り扱うための様々なモデルが提案されているが、いずれもリハーサルシステムに適しているとはいえない。

そこでリハーサルにおいて独奏者の演奏の特徴や演奏プランの抽出を補助し、また独奏者から出される要求を構造的に取り扱うための、楽曲の構造に関する情報として、表情情報の操作を目的とした構造モデルを提案する。

#### 4.2.1 構造モデル

このモデルでは、一つの楽曲は階層的な手続きの実行として表現される。それぞれの手続きは以下の情報を持っている。

- 手続きの名前

- 手続きに適用されるパラメータ
- 実行する手続きに関する情報
  - 実行する手続きの名前
  - 実行を開始する小節と拍
  - 手続きに渡す引数
  - 実行の条件

定義は次のような書式で行なう。

```
elm <手続き名> (<手続きパラメータ>) begin
  (<開始小節>, <開始拍>):
    <手続き名>(<引数>) "<条件>"
  ...
end
```

図 2: 手続きの表記法

手続きの名前は一意に手続きが判別できるものであれば何でも構わないが、その手続きが曲の中で占める役割など音楽的に意味のある名称をつけることが望ましい。

手続きに適用されるパラメータは文字列のリストとして与えられる。内容は表情情報に関するものや楽曲の構造に関するものが中心であるが、特に情報の内容に規定はなく、無意味な文字列が含まれていてもかまわない。

実行する手続きに関する情報はその手続きから直接実行される手続きが書かれている。この部分は呼び出される手続きの数と同じ数だけある。まず実行する手続きの名前と、その手続きを何拍目に実行するかという情報が書かれている。実行を開始する拍は親手続きの先頭からの相対的な拍数である。実行の条件の部分ではその手続きを実行する場合の条件が書かれている。この条件は主に親手続きに渡された引数(文字列)リストに対するパターンマッチとして書かれ、繰り返しの最後の部分が変化する場合など、基本的な構造は同じであるが細かい部分がことなる場合に利用される。手続きに渡す引数は、手続き全体に適用されるパラメータと同様に文字列のリストとして与えられる。このリストの中には、呼び出される手続きに渡されるパラメータが与えられる。

最も基本的な手続きとして、Play という組み込み手続きが用意されている。この手続きは引数に渡された音を演奏するという手続きである。

図 3 はこのモデルを用いて「蛙の歌」を記述した例の一部である。手続き「KaerunoUta」が曲全体に

```
elm KaerunoUta () begin
  (1,1):Phrase1(motif1, ド 4) ""
  (3,1):Phrase1(motif2, ミ 4) ""
  (5,1):Phrase2(motif3, ド 4) ""
  (7,1):Phrase1(motif4, ド 4,
    スタッカート) "" end
```

```
elm Phrase1 (だんだん強く) begin
  (1,1):Up(once) "motif1"
  (1,1):Up(once) "motif2"
  (1,1):Up(twice) "motif4"
  (2,1):Down() "" end
```

```
elm Phrase2 () begin
  (1,1):Play(1) ""
  (1,3):Play(1) ""
  (2,1):Play(1) ""
  (2,3):Play(1) "" end
```

⋮

図 3: 楽曲構造の記述の例の一部:「蛙の歌」

相当し、「Phrase1」は「ドレミファミレド」に代表されるフレーズ、また「Phrase2」は「ドドドド」のフレーズである。三種類のフレーズが「Phrasel」のバリエーションとして記述されている。

#### 4.2.2 構造モデルによる演奏

楽曲の演奏を行なう際には曲全体を表す手続きを実行することになる。子手続きには引数として「親手続きの引数」「親手続きのパラメータ」「子手続きの引数」の順に文字列のリストを結合したものが与えられる。それぞれの手続きでは引数のリストを順に一つ一つ調べることで、曲全体に占める位置やその部分の表情の与え方を抽出することができる。図 3 で示した曲の演奏を行なった際の、手続きの実行と引数の関係は図 4 のようになる。

このモデルは、手続きが子手続きを実行する際に、子手続きを新たな手続きとして生成し分岐する。このようにすることで複数の手続きの並列動作や遅延動作が記述できるため輪唱などの記述が可能である。また多重奏全体やより大規模な交響曲への適用も容易である。

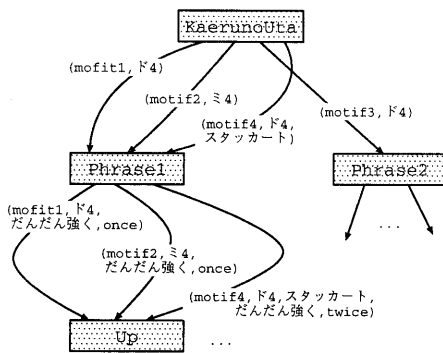


図 4: 楽曲の演奏の例

#### 4.2.3 演奏プランの記述と抽出

演奏プランはそれぞれの手続きへの引数や、手続きのパラメータとして文字列で与えられる。手続きのパラメータとして与えられる演奏プランは、その手続きが実行される状況に関わらず常に適用されるので、ここには「一拍目を強く」など、同様の構造の部分に一般的に適用される演奏プランを記述する。一方、手続きの引数として記述された演奏プランはそこで呼び出された手続きにしか適用されないので、「繰り返しの二回目は一回めよりやや強く」など、楽曲の細かな表情付けや大局的な変化を記述するのに適している。

例えば図 3 の場合「だんだん速く」は Phrase1 に一般的に適用されるので Phrase1 に特徴的な演奏プランであるが、「スタッカート」は Phrase1 を最後に演奏する時のみ適用される演奏プランで Phrase1 に特徴的な演奏プランではない。

独奏者の演奏から演奏プランを抽出する際には、独奏者の演奏とシステムの演奏プランとのずれをそれぞれの手続きのパラメータに追加する。何度か同じ手続きが呼び出されると、それらのパラメータに共通な部分と、呼び出される度にことなる部分が現れてくるので、共通な部分をその手続きのパラメータとして残し、呼び出す度にことなる部分は親手続きの引数として記述する。また全ての子手続きに共通に見られるパラメータがある場合には、そのパラメータを子手続きのパラメータから削除し自分のパラメータに追加する。この作業を全ての手続きに対して再帰的に行なうことで、独奏者の演奏プランとその適用範囲が抽出される。

## 5 おわりに

本稿では現在のリハーサルシステムに楽曲の構造に関する情報を持たせるためのモデルとして、表情情報の操作を目的とした音楽の構造モデルの提案を行なった。このモデルを利用すれば楽曲の構造して独奏者の演奏から効率的に演奏プランを抽出することが可能となり、また独奏者の要求を的確にシステムの演奏プランへ反映させることが可能となる。

今後は、まずこのモデルをリハーサルシステムに実装し評価を行なう予定である。また、構造モデルの楽曲の逐次データからの自動抽出の可能性についても検討を行なう。さらに初期の演奏プランの自動生成や、リハーサルによって得られた演奏者の演奏プランの特徴の学習や一般化などを行なう予定である。

## 参考文献

- [1] 堀内靖雄, 田中穂積: “自主性を持つ伴奏システム”, 人工知能学会論文誌, Vol.10, No.1, pp.72-79, 1995.
- [2] 堀内靖雄, 奥井学, 鈴木泰山, 田中穂積: “伴奏システムのリハーサル: ~人間とコンピュータとのアンサンブル(2)~”, インタラクティブシステムとソフトウェア II, pp.115-123, 近代科学社, 1994.
- [3] Berry Vercoe, Miller Puckette: “Synthetic Rehearsal: Training the Synthetic Performer”, Proc. of the 1985 ICMC, pp.275-278, 1985.
- [4] 奥井学: “自動伴奏システムにおけるリハーサル”, 東京工業大学 情報理工学研究所, 修士論文, 1996.
- [5] 田口友康: “自動演奏と演奏モデル: 演奏の芸術性を目指して”, bit 別冊, コンピュータと音楽, 共立出版, pp.42-53, 1987.
- [6] 野池賢二, 西岡大祐, 高田正之, 小谷善行: “曲の構造情報から表情付けを行なう自動演奏”, 情報処理学会第 44 年全国大会, 6U-3, 1992.
- [7] Neil P. McA. Todd: “Computational modeling of musical expression”, Proc. of the 1990 ICMC, pp.373-375, 1990.
- [8] 彌富あかね, 五十嵐滋: “音楽構造分析を用いたピアノ演奏の表情付け”, 情報処理学会第 50 年全国大会, 7S-6, 1995.
- [9] Peter Desain, Siebe de Vos: “Autocorrelation and the study of musical expression”, Proc. of the 1990 ICMC, pp.357-360, 1990.
- [10] Henkjan Honing, “An environment for analysing, modeling, and generating expression in music”, Proc. of the 1990 ICMC, pp.364-367, 1990.
- [11] 片寄晴弘: “音楽情報処理に関する研究”, 大阪大学 基礎工学部, 博士論文, 1991.
- [12] Gerhard Widmer: “Understanding and learning musical expression”, Proc. of the 1993 ICMC, pp.268-275, 1993.
- [13] 宮川和: “音楽におけるフレーズ間の類似性の評価及び相互関係の構造化に関する研究”, 東京工業大学 情報工学科, 卒業論文, 1996.