

## 事例ベースで生成された演奏表情を用いて 自然な演奏を行なう伴奏システム

西田深志, 鈴木泰山, 徳永健伸, 田中穂積  
東京工業大学 情報理工学研究科 計算工学専攻  
〒152-8552 東京都目黒区大岡山 2-12-1  
{nishida, taizan, take, tanaka}@cs.titech.ac.jp

あらまし

我々はこの数年「人間の演奏と協調して演奏する伴奏システム」の研究を進めている。従来の伴奏システムには演奏表情の付いた楽譜が必要であるが、それを人手で入力することは非常に困難な作業である。本論文では、演奏表情の付いた楽譜を自動生成できるように事例ベースの演奏表情生成手法をシステムに組み入れた。また、人間の演奏に対して複数のパートが自律した演奏を行なえるように、事例ベースの演奏表情生成手法および伴奏システムを拡張した。これにより、複数の独立したパートによる自然なアンサンブルが可能になり、パートの特徴やパート間のバランスを考慮した演奏表情を自動的に生成した。

キーワード 事例ベース 演奏表情自動生成 伴奏システム アンサンブル

## An accompaniment system using performance expression generated from an example-based approach

Fukashi Nishida, Taizan Suzuki, Takenobu Tokunaga, Hozumi Tanaka  
Department of Computer Science,  
Graduate School of Information Science and Engineering,  
Tokyo Institute of Technology  
2-12-1, Oookayama, Meguro, Tokyo, 152-8552  
{nishida, taizan, take, tanaka}@cs.titech.ac.jp

Abstract

We have been working on the development of an accompaniment system which melds in with human performance. However it is very difficult to input performance expression data as required by conventional accompaniment systems. First, we introduce an example-based approach for generating performance expressiveness as used by the accompaniment system. Second, we extend both the accompaniment system and the example-based approach for generating performance expression so that the system can perform plural parts independently. This enables both a more natural-sounding ensemble effect and the automatic generation of expression balancing the plural parts.

key words

example-based automatic generation of performance expressive accompaniment system ensemble

## 1 はじめに

音楽情報処理の一分野として伴奏システム [4, 6, 7] の研究が行われている。我々が数年に渡り研究している伴奏システムとは、演奏者と「協調して」伴奏を行うシステム [1, 3] である。伴奏システムが「協調する」とは、人間の独奏者にただ追従するだけでなく、音楽的状况によって独奏者に合わせるべきところは合わせ、伴奏者が自由に演奏するべきところは、自主的な演奏意志を持って、表情豊かに演奏することである。

伴奏システムは、自然で表情豊かな演奏を行うために、どのように伴奏を演奏するかという演奏意図を記した「演奏表情の付いた楽譜」を必要とするが、従来の伴奏システムではこの演奏表情を手で決定し入力していた。ここでの「演奏表情」とは、人間が楽曲を演奏する時に付けるテンポの揺らぎや音量の変化などを指す。「演奏表情」の与え方は人間の感性によるので、これを数値化するという作業は非常に困難であり、すべての音に演奏表情があることを考えると膨大な作業量になる。そのため、既存の伴奏システムでは数多くの曲のデータを作成することが難しかった。

また、従来の伴奏システムの多くは人間の演奏者に対して伴奏の演奏者が一人であるものを対象としていた。このようなシステムでは三重奏や四重奏の曲に対してそれぞれのパートが自主性を持つような伴奏を行なうことはできなかった。

本論文では、これらの従来の伴奏システムの問題点を解決し、自然なアンサンブル演奏を行えるようにするために以下の手法を提案する。

1. 事例ベースの演奏表情生成手法 [9, 10, 11] を伴奏システムに導入する  
事例ベースの演奏表情生成手法とは、計算機による自然な自動演奏を行うために、演奏事例から類似した旋律を検索し、検索された旋律から演奏表情を抽出し、演奏対象曲に対して適用する手法である。この手法を用いて、伴奏システムが自主的な演奏を行うために必要な「演奏表情」を自動生成する。
2. 伴奏システムの演奏者を複数にする  
パート毎に個別の演奏者を用意することで、それぞれのパートが独立して演奏を行えるように伴奏システムを拡張する。これにより、パート

ごとの自主性を考慮した自然なアンサンブル演奏が可能になる。

まず、第2章では堀内らの自主性を持つ伴奏システム [1, 3] について説明し、その問題点について触れる。第3章では、伴奏システムの入力となる演奏表情を生成する事例ベースの演奏表情生成手法について説明する。第4章でその解決法及び本研究で構築したシステムについて述べる。

## 2 自主性を持つ伴奏システム

### 2.1 伴奏システムの概観

堀内らの伴奏システムは、人間の独奏者の演奏に合わせてコンピュータが伴奏を行うものである [1, 3]。この伴奏システムの目的は全体の演奏として良い演奏を実現することにある。

一般に、独奏者は自分の演奏意志に従って自由に演奏し、伴奏者はそれに合わせて演奏するように考えられがちである。しかし実際には、独奏者も伴奏の演奏を聴きながら、その演奏と調和・協調するように演奏を動的に制御しており、独奏者と伴奏者は互いに影響を及ぼしながらアンサンブルの演奏を作っている。このため、自然な演奏を行うためには、伴奏システムはただ単に独奏者の演奏に追従して自分の演奏を変化させるだけでは不十分であり、音楽的状况に合わせて、伴奏システム自身も自主的に自分の表情豊かな演奏を行うべきである。

そこで、伴奏システムの自主的な意志に相当するものとして「演奏表情」を与える。これによって伴奏システムは表情豊かな演奏を行うことができる。また、伴奏システムが独奏者に対してどの程度自立して演奏するかを表す尺度として「独立度」を導入している。

伴奏システムの構成を図1に示す。伴奏システムの処理は以下の2つのモジュールに分かれる。

1. Listener:  
独奏者の演奏を入力とし、楽譜との比較からそれらの間の対応関係を調べる。
2. Performer:  
独奏者の演奏に合わせて、伴奏パートの演奏をスケジューリングし、伴奏の演奏を行う。

伴奏システムには、あらかじめ人間の演奏する独

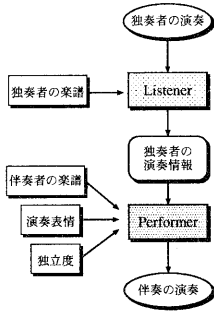


図 1: 伴奏システムの構成図

奏パートとコンピュータの演奏する伴奏パートの楽譜が与えられている。このシステムでは、Listener が独奏者の演奏を追跡し、その楽譜上の位置を認識する。Performer は、その独奏者の演奏に合わせて表情豊かな演奏を行う。自主的な演奏を行なうために伴奏システムは演奏表情にしたがって演奏する。しかし、伴奏システムが勝手に演奏すると独奏者の演奏との間にずれが生じてしまうので、「独立度」を用いて演奏表情を動的に修正しながら自然な演奏を行う。

### 2.1.1 独立度

独立度は、伴奏システムが独奏者の演奏に対してどの程度独立して演奏するかを表す尺度で、楽譜の位置毎に変化し、0~100の整数値をとるパラメータである。数値が高い程独奏者の演奏に対する独立の度合いが高くなる。独立度0と100は特別の意味を持ち、独立度100では独奏者の演奏に影響を受けること無く演奏し、一方、独立度0は独奏者と同時に演奏を行う。

### 2.2 従来の伴奏システムの問題点

従来の伴奏システムの問題点としては、まず、演奏表情を手で与えていることが挙げられる。演奏表情は個々の音符に対して与えなければならないが、これを人手で入力することは非常に困難である。また、どのように演奏表情を付けるかという問題もある。例えば、ピアノからフォルテまで大きくする時に、一つの音毎に大きくするのか、最後に一気に大

きくするのかといった違いを機械的に判断し、作成することは難しい。

次に、伴奏の演奏者がシステム一人であるものを対象としているために、伴奏システムは複数のパートが自主性をもった演奏をすることが出来ないことが挙げられる。従来の伴奏システムに複数の伴奏パートを演奏させる場合、それらを複合した単一のパートを演奏することになり、それぞれの伴奏パートの自主性を十分に考慮できない。

次の章では、前者の問題点を解決するために伴奏システムに導入する事例ベースの演奏表情生成手法について説明する。

## 3 事例ベースの演奏表情生成手法

### 3.1 演奏表情

人間が音楽を演奏する時には、ただ楽譜に忠実に演奏しているのではなく、テンポの揺れや音量の変化など様々な表情を付けて演奏する。このように、演奏中に演奏者によって付けられる様々な表情のことを演奏表情と呼ぶ。人間は演奏時には必ず何らかの演奏表情を付けるが、その演奏表情は、すべての演奏者に共通するものではなく、演奏者によって異なったものになる。

演奏表情は以下の3つに分類される。

- 音符と音符の時間間隔  
これは、テンポと呼ばれるもので曲の速さを表している。一般的に、テンポの値は楽譜の冒頭に記載されているが、曲全体を通して同じテンポで演奏すると、機械的で不自然なものになってしまう。人間の演奏ではテンポは常に変化し、その変化の与え方が演奏表情の一つの要素になっている。具体的にはルバート(リタルダンド)、アツェレランド、フェルマータといったものがある。
- 音の強さ  
音の強さとは、具体的にいうと鍵盤楽器では打鍵の強さ、管楽器では息の強さ、弦楽器では弓と弦の摩擦力や弦を指で弾く時の強さである。音の強さもまた、曲全体を通して同じ強さで演奏すると不自然であり、面白みの欠ける演奏になってしまう。

- 楽器の奏法

これには、楽器特有のものも含まれ種類も様々である。鍵盤楽器では打鍵から離健までの時間、弦楽器では弓を動かす速さや圧力の変化などが挙げられる。演奏の技法としては、スタッカートやレガート、ピブラートなどがある。また、弦楽器ではピッチカート等、演奏の方法自体が異なるものもある。

これらの演奏表情は、演奏者の解釈の違いや演奏者の演奏スタイルなど、演奏の行われる状況によって多様に与えられる。しかしながら、この演奏表情の与え方は演奏者の感性に強く依存するために、最適解といえるものが無く、またその評価も聴衆の感性によるために絶対的な評価は存在しない。

### 3.2 事例ベースの演奏表情生成

鈴木らによって提案された事例ベースの演奏表情生成 [9, 10, 11] は、人間の演奏を参考にして、未知の曲に対して演奏表情を生成する手法である。この手法の手順を図2に示す。具体的には、人間の演奏を収集した演奏データ集の中から、対象曲に類似した旋律や旋律断片などの演奏事例の検索を行う。そして、検索結果から旋律の類似度の高さや演奏状況の近さを考慮して、演奏事例の重要度を求め、それを基に演奏表情のパターンを抽出する。抽出されたパターンを対象曲に適用することで、演奏表情を生成する。

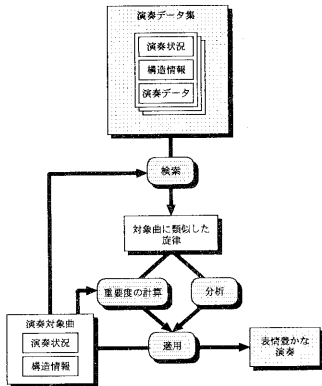


図 2: 事例ベースの演奏表情生成手法

### 3.3 事例の重要度

演奏表情を生成する際には、演奏データ集から検索された事例を分析し、そこに共通してみられる演奏表情のパターンを抽出する。このとき、全てのパターンを対等に扱うことには問題がある。演奏表情パターンの分析を行なう際には、それぞれの事例の有効性を評価し、その有効性の高い事例を優先的に考慮する必要がある。そこで、事例の重要度を旋律の類似度及び演奏状況の近さから求める。

#### 3.3.1 旋律の類似度

旋律の構成要素としてメロディとコードとリズムが挙げられる。類似していると感じる旋律ではその構成要素に多くの共通点が見られる。よってそれぞれの旋律の構成要素の特徴を比較し、共通点や相違点を調べることで旋律の類似度を客観的に評価することが可能である。メロディ、コード、リズムについて、それぞれ類似度を計算し、その結果を用いて旋律そのものの類似度を計算する。

#### 3.3.2 演奏状況

3.1節でも述べたように、演奏表情は多様なものであり、演奏者や演奏時の気分、演奏スタイルなど状況の変化により、その与え方が変化する。このように、演奏時における状況は演奏表情の与え方に大きく影響している。

演奏状況は以下に挙げる例のような要素を持ち、それぞれの要素が様々な素性を持つ。

- 演奏者  
素性としては、演奏者の経験や趣向、癖などがあり、演奏表情に大きく作用している。
- 演奏スタイル  
演奏スタイルの素性は、クラシックやロック、ジャズといった音楽のジャンルに相当する。
- 演奏時の気分  
演奏者がどのように演奏したかという要素で、素性としては「明るく」「優雅に」などが挙げられる。

## 4 自然なアンサンブルを行なう伴奏システム

### 4.1 システムの目的

2.2 節で挙げた問題点を解決するために、本論文では以下の2つのことを行なった。まず、演奏表情を自動生成するために、事例ベースの演奏表情生成手法を導入した。これにより、構造情報付きの楽譜を入力するだけで楽譜に演奏表情を付けることが出来るようになり、入力にかかる労力が減少した。また、独立度も簡単なヒューリスティックで自動的に生成した。

次に、アンサンブルのテンポ及び独立度の概念を拡張し、伴奏システムの演奏者は一人だけでなく、システムが演奏するパート数の演奏者が演奏できるようにシステムを修正した。これによってシステムが複数のパートの自主性を考慮することが出来るようになり、それぞれを独立したパートとして扱えるようになった。尚、伴奏パートが一つである時には従来の伴奏システムと同等である。

### 4.2 システムの概要

本システムの構成を図3に示す。

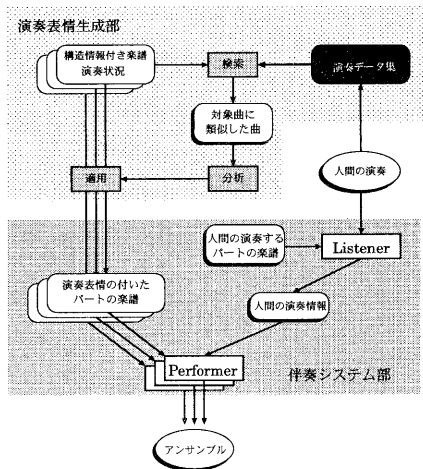


図3: システム構成図

システムにはあらかじめ、すべてのパートの構造

情報付き楽譜を与える。システムは事前に、それぞれのパートについて、事例ベースの演奏表情生成手法を用いて、演奏表情付きの楽譜を生成する。演奏時には、自動生成された楽譜を元に、人間の独奏者とアンサンブル演奏を行なう。独奏者の演奏はシステムに入力され、システムはこの演奏に合わせてリアルタイムで伴奏パートの演奏を出力する。独奏者の演奏と伴奏システムの演奏は共に MIDI 信号である。

### 4.3 演奏表情生成部

演奏表情生成部では、入力を各パートの構造情報付き楽譜とし、3章で紹介した事例ベースの演奏表情生成手法を用いて、テンポ、音の強さなどの演奏表情が付けられた楽譜を出力する。

人間がアンサンブルの演奏を行なう時は、一人で演奏する時とは異なり、他のパートの演奏を聴きながら、パート間のバランスを考え、それらの影響を受けて表情を変化させて演奏をする。しかしながら、事例ベースの演奏表情生成手法では、単一パートで構成された曲を対象としていた。そのため、この手法をそのまま利用して各パートの演奏表情を生成すると、アンサンブル演奏を行なう時の演奏表情とは異なる、そのパートを単独で演奏した時の演奏表情を生成してしまう。その演奏表情を使ったアンサンブルは、各パートが勝手に演奏している不自然な演奏になるだろう。

そこで、ソロとして演奏したのか、アンサンブルで演奏したのか、また、アンサンブルを演奏した場合はどのパートを演奏したのかという情報を演奏データ集にある個々の事例の演奏状況に記述する。それぞれのパートの演奏表情を生成する時は、同じパートの演奏が事例の中に存在すれば、それが優先して検索される。

これにより、そのパートの特徴やパート間のバランスを考慮した演奏表情が生成できるようになる。

### 4.4 伴奏システム部

伴奏システムが複数のパートの演奏を同時に行なうために、次のことを行なった。まず、Performerを伴奏のパート数用意し、それぞれのPerformerが一つのパートを担当し独自に演奏できるようにした。

具体的には、マルチスレッドを用い、パート毎にスレッドを割り当てることでパート数の Performer を作成し、演奏を行なえるようにした。伴奏の演奏者が複数になったことに伴い独立度を拡張するとともに、ヒューリスティックを利用して独立度を自動的に生成できるようにした。また、自動生成によって付けられた演奏表情のテンポのとらえ方の解釈を新たにした。

#### 4.4.1 独立度の拡張

従来、独立度は「伴奏システムが独奏者の演奏に対して、どの程度独立して演奏するかを表す尺度」として与えられるが、これを拡張して「それぞれのパートがアンサンブルのテンポに対して、どの程度独立して演奏するかを表す尺度」とし、人間のパートも含むすべてのパートに与えるものとした。独立度はその与え方のルールを考慮し、以下のような簡単なヒューリスティックで与えた。

- 全体的に低めに設定する  
伴奏システムは複数のパートを演奏するので、独奏者と個々の伴奏パートのずれが大きくならないようにする。
- 短い音符は高めにする  
短い音符でそのパートが細かい動きをしている時は、メロディや弦楽器の刻み等のように、アンサンブル全体のテンポに大きく影響を与えていると考えられるために独立度を高めに設定する。
- 長い音は低めにする  
長い音を演奏している時は、テンポにはあまり影響を及ぼさないと考えられる。また、そのパートはアンサンブルに合わせようとするために低めに設定する。
- 長い休符や長い音の後は独立度を 0 にする  
長い休符や長い音の間はそのパートはテンポに影響を与えることはなく、次に音を出す時はアンサンブルに完全に合わせる。

#### 4.4.2 演奏表情の解釈

従来の伴奏システムでは自動生成された演奏表情のテンポに忠実に演奏していた。しかし、独奏者の

演奏したいテンポと伴奏システムの楽譜に付けられた演奏表情のテンポが食い違う時に、独奏者が長い休みや長い音になると、その間にシステムは急激にテンポを戻そうとするために自然ではなくなっていた。

例えば、図4のように独奏者の演奏しているテンポが 130 で、楽譜に付けられたテンポが 100、アンサンブルのテンポが 120 とする。次に独奏者が休みに入ると、そのことで伴奏システムの独立度が高くなり、伴奏システムだけがアンサンブルのテンポを作る状態になる。その時、伴奏システムの楽譜のテンポが 110 に変わると従来の解釈では伴奏システムは、アンサンブルのテンポが 110 になるように急激にテンポを落していく。しかし、楽譜に付けられたテンポは 100 から 110 に上がっているのにここでテンポを落すことは伴奏システムの演奏意図とは異なると考えられる。

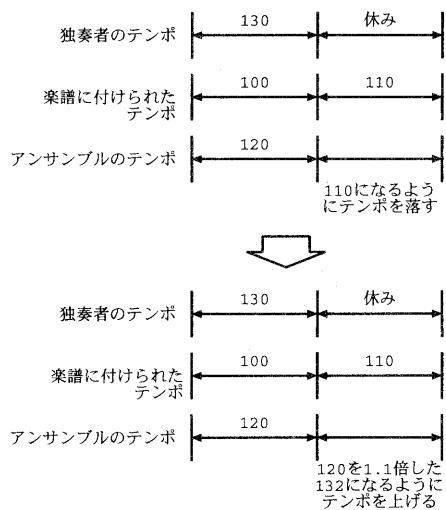


図 4: 演奏表情の新たな解釈

そこで、楽譜に付けられた演奏表情の新たな解釈として、次に演奏する音のテンポを式 1 によって求める。

$$nexttempo_{system} = localtempo_{ensemble} \times \frac{nexttempo_{system\_score}}{localtempo_{system\_score}} \quad (1)$$

$localtempo_{ensemble}$  は、アンサンブルの局所的なテ

ンポで、 $nexttempo_{system\_score}$  は、次に演奏する音のシステムの持つ楽譜に記載されているテンポである。 $localtempo_{system\_score}$  は、 $localtempo_{ensemble}$  と同じ範囲のテンポを平均したもので、 $nexttempo_{system}$  は、この式で最終的に求まる伴奏システムが次に演奏する音のテンポである。

式1では、伴奏システムの演奏意図を楽譜に付けられた演奏表情のこれまでのテンポと次に演奏するテンポの比で表している。先ほどの例では、前のテンポが100で次に演奏するテンポが110なので、伴奏システムはテンポをアンサンプルのテンポが132になるようにテンポを上げて演奏を行なう。

#### 4.5 リハーサル効果

伴奏システムにおいてはリハーサルの研究も数多くなされている [2, 5, 8]。伴奏システムにおけるリハーサルとは、本番の演奏でより良い演奏ができるように、独奏者の癖や演奏意図を学習することである。

本システムでは演奏表情の自動生成を事例ベースで行なっている。そのため、伴奏システムを使って行なった独奏者の演奏を新たな事例として、演奏表情生成部の演奏データ集に加え、再度演奏表情の生成を行なうことで、リハーサルと同様の効果が得られると考えられる。

従来のリハーサルでは、演奏曲に特化された独奏者の癖や演奏意図を学習していた。そのため、演奏曲が変わる度にリハーサルをして学習し直さなければならなかった。一方、本システムのリハーサル効果では、同一の曲がなくても、独奏者の演奏が事例に含まれていれば独奏者の演奏表情の付け方そのものをある程度考慮した演奏表情がつけられるという点で大きく異なる。

また、伴奏システムを使って行なった独奏者の演奏が事例に含まれている場合、演奏者が同じであることや同じ曲を演奏していることから、演奏状況の類似性が高くなる。旋律の類似性においても、伴奏パートと独奏パートでは曲の構造がほぼ同じであることや、独奏パートとは別の場所に伴奏パートに同じような旋律が出現することなど、テンポに大きい影響を与えている場所に高い評価が得られ、リハーサルと同じような効果がある。

さらに、同じ演奏者が演奏した事例が演奏データ

集に含まれている場合には、演奏しようとしている曲が演奏データ中に含まれていなくても似ている旋律が曲の中に存在すれば高く評価され、演奏者の演奏表情の付け方を抽出し適用することができるので、ある程度の効果があると考えられる。

#### 4.6 実装

上で述べたシステムを SGI O2 Workstation 上に実装した。MIDI 楽器として YAMAHA Clavinova CLP-760 を用いた。

演奏データ集に含まれない新たな曲の1パートを独奏者に演奏してもらい、それに合わせて伴奏を演奏させた。その結果フレーズの始まりや終りでは自然なアンサンプルができていた。一方、不自然なところも多かった。その原因としては、マルチスレッドを用いたことで、伴奏パートと独奏パートが縦の線を揃えるべきところの演奏が若干遅れるようになったことが考えられる。

また、その演奏データを事例に加え、演奏表情を再度生成した後、同じ人に演奏してもらった。その結果良く似た旋律や旋律断片についてはリハーサルの効果が確認された。

### 5 まとめ

本論文では、従来の伴奏システムの問題点を指摘し、それらを解決するために以下の2つのことを行なった。

まず、伴奏システムに事例ベースの演奏表情生成手法を導入した。演奏状況としてアンサンプルの演奏かどうか、アンサンプル演奏ならどのパートを演奏したかという情報を導入し、これまで人手で入力していた演奏表情を自動生成できるようにした。また、独奏者の演奏を事例に含め、演奏表情を再度生成させることによって、伴奏システムにおけるリハーサルの効果が得られることも確認した。

次に、伴奏システムが複数の伴奏パートを独立したパートとして演奏できるように拡張することで、それぞれのパートの自主性を考慮したアンサンプル演奏が可能になった。特に、伴奏システムの入力である演奏表情をテンポの比として解釈することで、独奏者が休みになった時に起こる急激なテンポの変化による不自然さを取り除くことができた。

## 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] 藤井敦, 堀内靖雄, 田中穂積: “複数の人間と協調して演奏するシステム”, 人工知能学会 第7回 全国大会論文集,pp.567-570,1993.
- [4] Berry Vercoe: “THE SYNTHETIC PERFORMER IN THE CONTEXT OF LIVE PERFORMANCE”, Proc. of the 1984 ICMC, pp.199-200, 1984.
- [5] Berry Vercoe, Miller Puckette: “Synthetic Rehearsal: Training the Synthetic Performer”, Proc. of the 1985 ICMC, pp.275-278, 1985.
- [6] Roger B.Dannenberg: “An On-Line Algorithm for Real-Time Accompaniment”, Proc. of the 1984 ICMC, pp.193-198, 1984.
- [7] Lorin Grubb, Roger B.Dannenberg: “Automating Ensemble Performance”, Proc. of the 1994 ICMC, pp.63-69, 1994.
- [8] 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 情報理工学研究科, 修士論文,1996.
- [9] 鈴木泰山: “コーパスに基づく演奏表情の自動生成と演奏表情コーパスの設計に関する研究”, 東京工業大学 情報理工学研究科, 修士論文,1997.
- [10] 鈴木泰山, 徳永健伸, 田中穂積, “事例ベースの演奏表情生成手法に関する研究”, 情報処理学会研究報告, 96-MUS-21, pp. 7-12, 1997.
- [11] 鈴木泰山, 宮本朋範, 西田深志, 徳永健伸, 田中穂積, “Kagurame Phase-I —事例ベースの演奏表情生成システム—”, 情報処理学会研究報告, 98-MUS-24, pp. 61-68, 1998.