

## 声部協調型音楽解釈モデルに基づく演奏の表情付け

橋田 光代†,†† 野池 賢二† 片寄 晴弘††††

演奏表情付けに関する近年の研究においては、複数声部を伴う楽曲の声部ごとの演奏表現に対する考慮が不十分であった。本研究では、楽曲の各声部に対するアテンション（注意）の移動に着目し、1) 各声部に独立した表情付けを行う、2) アテンションが移動する際、当該の音の音長を延ばす、3) 複数声部の発音時刻がグループ構造境界で一致するよう、オンセットタイミングのスケールリングを行うことを基本とする音楽解釈モデルに基づく演奏表情付け手法を提案する。

### Performance Rendering Based on“PopE”: A Musical Interpretation Model for Polyphrase Ensemble

MITSUYO HASHIDA,†,†† KENZI NOIKE† and HARUHIRO KATAYOSE††††

This paper proposes a musical interpretation model for Poly-phrase Ensemble: PopE. A polyphonic musical piece has to be rendered as it may exhibits independent expression of each voice, as well as considered synchronization. PopE, considering “attention transition” generates expressive performance based on the following procedure: 1) giving individual expressions to each part, 2) stretching lengths of notes which are located at the attention transition, 3) stretching (shortening) lengths of notes of parts attention is not given, as the parts length may fit to the attention part. The paper also describes how the middle part of Chopin’s “Fantasie-Impromptu” awarded NIME-Rencon was rendered using PopE.

#### 1. はじめに

演奏表情付けに関する研究 (Performance Rendering) は、音楽情報処理の中でも中心的な研究領域の一つである<sup>1)</sup>。2002年度からシステム生成演奏の聴き比べコンテスト (Rencon) も開催されており、情緒領域を扱うシステムの評価基盤の先導的な取り組みとしても注目されつつある<sup>2)</sup>。

演奏の表情付けシステムに関する研究は、1980年代の Sundberg<sup>3)</sup>、Clynes<sup>4)</sup>らの研究にさかのぼることができる。1990年以降は、GTTM<sup>5)</sup>やIRM<sup>6)</sup>などの認知的音楽理論の利用、学習システム<sup>7),8)</sup>や事例ベース推論<sup>9)</sup>などのアプローチも見られるようになった。

音楽構造解析については部分的な人間の介入を前提とするもの、フレーズを含め主旋律の表情付けに関してはかなりの表現力を持ったシステムが始め

ている。FIT-Rencon (2002年)において、竹内は専門家の立場から「人間の演奏者にひけをとらない演奏も出始めている」と指摘している<sup>10)</sup>。一方この段階では、音の縦軸についての表情が乏しく、また、副旋律の演奏表情を評価できるものはほとんどなかった。特にこの状況はショパンを題材に選んだエントリーに顕著であった。

IJCAI-Rencon (2003年)では、ショパンの作品を指定しての聴き比べとなった。この年の上位3位までは、音の縦軸の表情を扱ったものが選ばれた<sup>11)</sup>。しかし、これらのエントリーは、人間が演奏表現を実施し、そのプロセスを支援するというタイプのシステムに基づくものである。つまり、演奏生成モデルに基づいて自動的に演奏を生成するというタイプのものではなかった。

我々はメロディを含む複数旋律から構成される音楽に自然な演奏表情を与えるための演奏生成モデルの構築を目的として、声部協調型音楽解釈モデル (PopE) のデザインを進めてきた。そのモデルに基づきショパンの「幻想即興曲」から展開部32小節の演奏生成を行

† 科学技術振興機構さきかけ研究 21「協調と制御」領域 PRESTO/JST

†† 和歌山大学システム工学研究科  
Wakayama University

††† 関西学院大学理工学部情報科学科  
Kwansei Gakuin University

い、NIME-Rencon (2004年) に出品した<sup>☆</sup>。NIME-Rencon においては、試行的なチューリングテストとして人間の演奏を交えての聴き比べが行われたが、我々のエントリは、「演奏の人間らしさ」、「演奏の好み」の評点において、人間の演奏と比べての有意差は検出されず、NIME-Rencon での Rencon Award (第一位) に選ばれた<sup>12)</sup>。

本稿では、声部協調型音楽解釈モデル (PopE) のデザインについて述べるとともに、NIME-Rencon にエントリした「幻想即興曲」の生成状況について述べる。

## 2. 音の縦軸の表情付けとその課題

演奏の表情付けに関する現在の課題の一つに、揃いすぎた音の縦軸に対する自然な表現がある。複数の演奏者によって構成されるアンサンブルにおいては、それぞれの演奏者が自身の音楽的表現を実施した上で、演奏者間のインタラクションを通して全体的なテンポの統制が行われていると考えられる。テンポの同期はアンサンブルにおける演奏上の重要な技術的課題であるが、一方で自身のパートが担う独自の旋律的表現を行うことも重要であり、そのバランスが演奏表現上の特質となる<sup>13)</sup>。このことは、単一ポリフォニー楽器、例えば、ピアノ用に作曲された楽曲においても考慮されるべきである。

音の縦軸の表情付けに留意した研究としては、Raphael の Orchestra in a Box<sup>14)</sup>、奥平、片寄らの iFP<sup>15)</sup> などの演奏支援システムがあげることができる。これらのシステムは、人間の既存演奏を、テンポ推移 (基準となる拍) と各音の微細なタイミングのズレとを分離したテンプレート (モデル) として利用することにより、拍内の微細な演奏表情を保存した状態での拍打型演奏生成インタフェースを実現している。Orchestra in a Box, iFP ともに、演奏技術はシステムが担うため、演奏者は自分のフレーズ表現そのものに集中できるという特長を有している。但し、楽曲毎に表情のついたテンプレート (モデル) を用意しなければならないという制約がある。

本研究では、単一曲のテンプレートではなく、アンサンブルやポリフォニー音楽 (多声部音楽) の自然な演奏表現を行うための一般的なモデルのデザインを対象とする。この目的に対しては、それぞれの声部に対し個別に演奏表現を施すという方法が考えられる。この考え方は直感的にも受け入れやすいものであるが、このままでは全体的には全く統制が取れない演奏

とになってしまう。つまり、多声部音楽の演奏表現においては、以下の同期処理に関する問題を解決する必要がある。

- 楽曲のどの時点でタイミングを合わせるのか
- 同一演奏区間内の演奏表現内容 (音楽的な主張) が声部によって異なる場合、それらをどのように調停するのか

## 3. 声部協調型音楽解釈モデル (PopE)

声部協調型音楽解釈モデル (PopE: Musical Interpretation Model for Polyphrase Ensemble) は、アテンション (3.1 節) の移動と、GTTM のグループ構造、延長簡約<sup>6)</sup> を利用することにより、前章で述べた問題を解決するとともに、アテンションの移動に係る演奏ルールを導入することによって構成される今までに無い演奏生成モデルである。以下、モデルの概要について整理する。

準備:

PopE では、以下に示す音楽構造に対して、演奏ルール (3.3 節) を適用することによって演奏を生成する。これらの内、一部 (グループ構造、頂点) は我々の別研究によって自動解析が可能である<sup>16)</sup>。将来的にはすべての構造の自動解析を行うことが課題になるが、現時点ではユーザが与えるようにしている。

[音楽のパート]

楽曲を複数のパート (旋律、副旋律、伴奏旋律等) として取り扱い、それぞれの流れを個別パートとして取り扱う。

[アテンション]

時間軸上において最も優勢なパートとして考えられる音列にアテンションを付与する。

[グループ構造]

各パートに対し、個別にグループ構造を与える。

[延長簡約]

各パートに対し、個別に延長簡約を付与する。

演奏の基本ルール:

表情付けの基本ルールについては、以下に示すような簡易なものを用いるものとする。

[グループ構造の表現]

各グループに対する、グループであることを明確化するための表現

[タメの表現]

アテンション移動時、および装飾音や連符等極端に短い音の引き延ばし

同期処理:

以下のような同期処理を実施する。

<sup>☆</sup> 当時はモデル名を WALTZ として出品した。

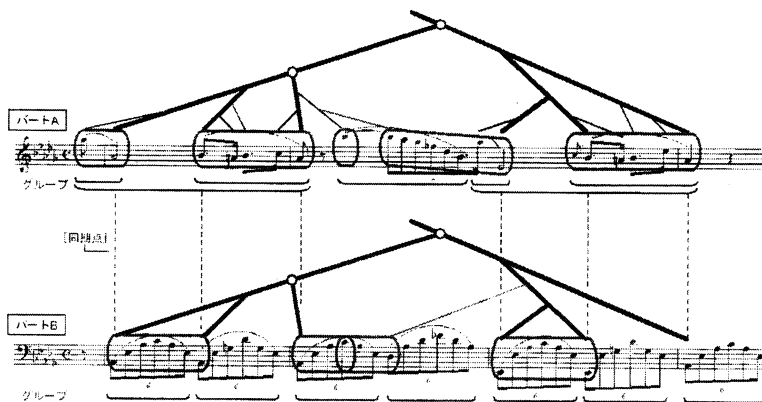


図1 同期点の算出

### [基本的なテンポ表現]

アテンションが付与されているグループのテンポ推移を楽曲全体のテンポ推移とする。

### [同期点の算定]

各声部に与えられた延長簡約を比較し、大局的な構造が一致する場所を同期点とし、その部分においては、同時に発音させるものとする。

### [非アテンションパートのテンポ表現のリマップ]

同期点に挟まれた区間において、アテンションのある声部を基準として、ルール適用後のIOI比を保ったままスケールリングする。

以下、同期処理および演奏の基本ルールについて説明を行う。

### 3.1 アテンション

アテンション（注意）とは、関係ない刺激を無視して活動の焦点を絞り込むことを指す。特に、視覚領域の認知心理学において研究が盛んであり、例えば、絵画を見た時、主観的には絵の全体を眺めたつもりでも、実際にはその情景のうち数箇所の小さい部分にしか注目していないということが知られている<sup>17)</sup>。

音楽においては休符や長音を含め常時主旋律に対してアテンションが向けられていると考えられてきた。事実、演奏の表情付けシステムはその観点で構成されることがほとんどであった。しかし、我々は同期処理を考えるにあたり、アテンションの移動が頻繁に行われていると考えることにした。基本的に時間的な空白区間を作らないという前提のもと、当該時刻において最も前面に聞こえてくる音の列に対し、アテンションがあるという情報を付与することとした。

### 3.2 同期処理

図1および図2に同期処理の様子を示す。PopEでは各パートでのグループ構造が必ずしも同期的に揃っ

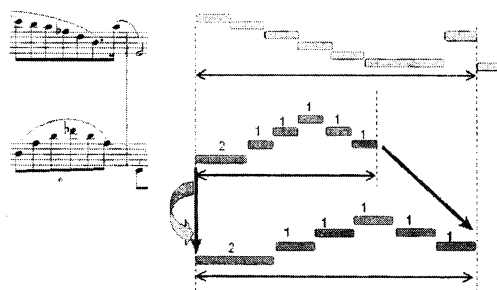


図2 同期処理におけるスケールリング

ている (homogeneous) とは限らないという前提のもとに、同期点の算定を行う。各グループの開始点もしくは終了点のいずれかがパート間で一致する拍節上の音符（の開始時刻）を同期点と考えるものとする。ただし、パート間で緊張-弛緩構造のレベルが異なる音符については、フレーズングの階層が異なるため、同期点としては取り扱いたくない。このような要請を満たすものとして、延長簡約において、枝の長さが同じ、かつ、拍節以上で一致する音符を同期点として選び出すものとする。延長簡約とは、時間スパン簡約と二つのイベント間の安定性から導出される構造である。図1において、ツリー構造として記した線が延長簡約である。2小節目の二つ目のアルペジオ（パートB）の開始音符 Db は同期点とならないことに注意されたい。

PopEでは主旋律や伴奏、内声等の区別を付けず、それぞれのパートグループに対して演奏ルールを適用することで、IOI（テンポ）に変化を与える（図2）。楽曲の全体的な時間進行は、アテンションを付与したパートの演奏ルール適用後のIOI時間進行（テンポ進行）が支配するという立場をとるため、アテンションの付与されていないサポート（図では、パートB）に



図 3 フレーズング

については、そのパートグループの前後での同期点を合わせることを目的として、当該 IOI の並びに対して、その比が保存されるような形でスケールングを行う。

### 3.3 演奏ルール

表現力を豊かにするという目的を持ってルールベースによる表情付けシステムを構築する場合、通常、できるだけ多くの演奏ルールを用意する。本研究では、モデルの骨格部分の妥当性を確認することを目的として、少数の本質的なルールセットを用いるようにしている。ここで取り上げるルールは、**グループ構造の表現**、**タメの表現**のかかわるもののみである。

#### グループ構造の表現ルール：

演奏家は自身の理解した音楽的構造を明確化するという意図をもって演奏表現を行う<sup>18)</sup>。その中でもグループ構造の表現は根幹的なものであるが、代表的な表現手法としては、以下の二つが知られている。

- グループ開始音に対するアクセント付与<sup>18)~20)</sup>
- フレーズング（テンポ、音量の山型表現<sup>6)</sup>）

グループ開始音に対するアクセント付与に関しては、グループ開始音の音量および音長を、初期値に対する倍率を与えるという形での演奏ルールを用意した。

フレーズングについては、グループ中の一音を頂点音（延長簡約の二重弛緩点によって与える）とし、グループに含まれる音の音量および局所テンポを、グループ開始音から頂点音にかけて線形増加させ、頂点音からグループ終了音にかけて線形減少させるという形でのルールの定式化を行った（図 3）。ここでの増加率（と減少率）が演奏ルールのパラメータである。なお、延長簡約を与えることによりグループは階層構造として記述される。階層毎にグループ表現ルールを適用するものとする。

#### タメの表現：

PopE でのもう一つの表現に関するルールは以下のタメに関するものである。

#### 装飾音、連符等の極端に短い音

演奏表現において、装飾音や極端に短い音（連符等）に対しては、音長をのばして局所的にテンポ

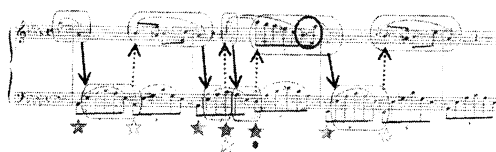


図 4 ルール説明図

を遅くしたり、音量を強めたりすることで、この部分の演奏を強調することが多い。対応する規則として、これらの音の音量および音長の初期値に対する倍率を与える。

#### アテンション移動

アテンションが別の声部に移動する際、アテンションが移動した音（図 4 中の☆マーク参照）に対してタメを与えるものとして音長比率を与える。パラメータについては、グループの階層レベルやつながりごとに設定するものとしている。このルールの設定の意味については第 5 章にて検討を実施する。

## 4. ショパン「幻想即興曲」での実施例

この章では、以上で説明したモデルに基づき生成した NIME-Rencon へのエントリ作品について説明を行う。

### 4.1 演奏生成のステップ

演奏生成対象としては、従来の表情付けシステムが取り扱ってこなかったショパンの「幻想即興曲」展開部から 32 小節とした。この部分は、2つの声部 A、B があり、全声部の発音時刻箇所 204 に対し、両声部が一致するのは 37 箇所（約 18%）である。3 拍子系（六連符）の声部 B に対して、声部 A は、2 拍子系の音型や、直前のタイからつながる複雑な七連符が含まれる。この部分の演奏においては、声部毎の適切な表現と声部間の連携が必須である。モデルの機能を確認するには適切な題材である。この対象に対し、以下のステップによって演奏を生成した。

#### step1 楽譜入力

Finale を用いた楽譜入力。この際、Finale の自由記述機能を利用し、構造情報、アテンション情報を付与。その情報を MusicXML ファイルとして出力。

#### step2 パラメータ設定

演奏生成にかかわる各種パラメータを別途テキストファイルとして用意。

#### step3 レンダリング

上記 2 種類のファイルを読み込み、演奏ルールを

\* P. Casals は Rainbow の法則と呼んでいる

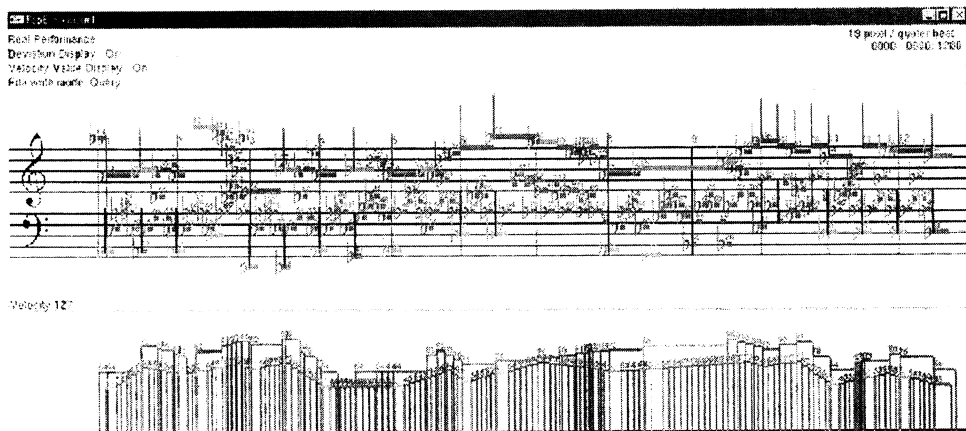


図 5 生成例図

表 1 演奏パラメータ

パラメータ	音量	音長
初期値 (演奏表情なし)	パートA: 64 パートB: 48	BPM: 112
グループ開始 アクセント	パートA: 130% パートB: 150%	100% (変化させず)
フレーズ表現 (頂点の高さ)	階層 1 (小): 125% 階層 2 (大): 150%	100% (変化させず)
タメ	装飾音 極端に短い音	100% (変化させず)
	アテンション の移動	パートA: 200% パートB: 100% (該当なし) パートA+B: 200% パートB+A: 125%

適用し、SMF 形式で演奏を出力。

#### step4 後処理

ルールに基づくペダリング<sup>☆</sup>と、手作業による二ヶ所の修正(図 4 中赤○印 B $\flat$  音の音長を 150%、および A $\flat$  音の音長を 250% 延長)の実施。

後処理の手作業については、試聴の結果短く聞こえたため、修正を加えたものである。NIME-Rencon においてはこの修正を行ったものと行わなかったものの両者を出品したが、そこでは、両者に有意な差があるとは判断されなかった<sup>12)</sup>。ここでの修正の意味については、第 5 章でも取り上げる。

#### 4.2 演奏ルールのパラメータの設定

演奏ルールについてはできるだけシンプルなものにするという目標のもと、アナリシスパイセンシスの結果、表 1 に示すパラメータを採用した。この表に示

すように、表情付けのための演奏ルールは、グループ開始アクセントにかかわる音量表現(2 レベル)、フレージングにかかわる音量表現(2 レベル)、装飾音・極端にかかわる音長表現(1 レベル)、アテンション移動時の音長表現(2 レベル)の、計 7 つに限定した。演奏生成例を図 5 に示す。この演奏は、<http://www.m-use.net/research/PopE/> から試聴が可能である。

### 5. 検 討

この節では、PopE における特徴的な演奏ルール「タメ」を中心とした検討を行う。

PopE では、もともとテンポを支配するパート設定を目的としてアテンションパートを導入した。アシュケナージ、中村紘子らの演奏を分析する過程において、我々は、装飾音符や連符に加え、アテンションパートの移動部分にも「タメ」が存在していることに気づき、その結果「タメ」に関する演奏ルールの構成に至った。

装飾音符や連符については解釈上の表現に加え、早いパッセージであるが故に演奏の際の運動負荷が大きいと思われる。卓越した演奏者にとってこのことが技術的な制約になることは考えにくい。一般的な演奏者にとっては負荷となるはずである。また、それを受容する鑑賞者もそのストレスを共有する可能性があり、それゆえ、楽譜上の音価より長めに演奏することが一般化したという仮説を持つことができる。

一方、アテンションパートの移動部分の「タメ」についてはどうだろうか？ 言葉や音の途中に白色ノイズ等が挿入された場合、音の長さは、もとの音の長さより、短く感じられるという<sup>22)</sup>。津崎らは、対象の変化を知覚する際の知覚コストが時間経過感覚をマスクするという可能性を指摘している。パート間でのアテ

☆ ペダリングに関する主要な技法<sup>21)</sup>のうちレガート・ペダルのみを採用し、一律に、伴奏(パート B)の各グループ先頭音の発音時間内で、踏み替え(コントロール・チェンジ 64:sustain の値を n から 127 まで線形増加させる)を行った。

ンションの移動にもなんらかの知覚コストが伴うと考えれば、そのコストを代償するために「タメ」が必要であるという解釈が成り立つ。

次に、手作業による二ヶ所の修正(図4中赤○印B♭音およびA♭音の音長の延長)の意味について考える。NIME-Renconのエントリにおいては、この箇所はアテンションの移動が行われない部分として取り扱った。しかし、この部分では、7度および11度の跳躍が見られる。大きな跳躍においては同一パート中でもアテンションの移動が起こると考えることにすれば、手作業による修正部分も、統一的なモデル(アテンションの移動によるタメの挿入)による説明が可能である。

以上の検討事項は現段階では仮説の域を超えない。今後は、Clynesのアプローチ<sup>4)</sup>のように、PopEの演奏家毎の特徴記述再現能力を評価するとともに、知覚コストの代償に関する文献考証を進めていく必要があると考えている。

## 6. まとめ

本論文では、声部協調型音楽解釈モデル“PopE”の構成について紹介した。アテンションの移動という視点を導入することによって、複数声部を持つ音楽の自然な表現が可能となった。また、4種類、7パラメータの簡単な演奏ルールのみで生成した演奏が、他システムの生成演奏に勝ることが聞き比べによって実証された。演奏ルールの数が少ないということは、このモデルが音楽表現の本質を捉えている可能性を指し示すものである。今後は、本モデルの記述能力についての検証を行うものとして、「幻想即興曲」に対する演奏者毎の違いがどの程度記述できるのかに関する実証を進めていく予定である。現段階では、楽曲構造、アテンションを人間が与えるという形での実装を行っているが、その自動化に取り組む必要がある。アテンションの抽出に対してはマルチエージェントモデルの導入を検討している。その上で、様々な楽曲に対するレンダリング事例を増やしていく予定である。

**謝辞** 本研究は、科学技術振興機構さきかけ研究21「協調と制御」領域研究として実施されました。

## 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) 平賀, 平田, 片寄: 蓮根, 目指せ世界一のピアニスト, 情報処理, Vol. 43, No. 2, pp. 136-141 (2002).
- 3) Fryden, L., Sundberg and J.: Performance Rules for Melodies. Origin, Functions, Purposes, *International Computer Music Conference (ICMC) Proc.*, ICMA, pp. 221-225 (1984).
- 4) Clynes, M.: A Composing program incorporating Microstructure, *International Computer Music Conference (ICMC) Proc.*, ICMA, pp. 225-232 (1984).
- 5) Lerdahl and Jackendoff: *A Generative Theory of Tonal Music*, MIT Press (1983).
- 6) Narmour, E.: *The Analysis And Cognition Of Basic Melodic Structures*, the University of Chicago Press (1977).
- 7) Widmer, G. and Tobudic, A.: Playing Mozart by Analogy: Learning Phrase-level Timing and Dynamics Strategies, *International Conference on Auditory Display (ICAD) Proc.*, pp. 28-35 (2002).
- 8) 片寄晴弘: 音楽感性情報処理に関する研究, 博士論文, 大阪大学基礎工学部 (1991).
- 9) 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 (1998).
- 10) 橋田, 野池, 平賀, 平田, 片寄: FIT 2002 RENCON Workshop —報告と課題—, 情報処理学会研究報告音楽情報科学 2002-MUS-48, pp. 35-99 (2002).
- 11) 片寄, 平田, 平賀: IJCAI-RENCON の報告と課題, 情報処理学会研究報告音楽情報科学 2002-MUS-52 (2003).
- 12) 野池, 橋田, 平田, 片寄, 平賀: NIME04 RENCON 開催報告と次回への課題, 情報処理学会研究報告音楽情報科学 2005-MUS-59 (2005).
- 13) W. ビストン: 管弦楽法, 音楽之友社 (1967).
- 14) Raphael, C.: Orchestra in a Box: A System for Real-Time Musical Accompaniment, *Proceedings of International Joint Conference of Artificial Intelligent (IJCAI)* (2003).
- 15) 片寄晴弘, 奥平啓太, 橋田光代: 演奏表情テンプレートを利用したピアノ演奏システム: sfp, 情報処理学会論文誌, pp. 2728-2736 (2003).
- 16) 橋田光代, 片寄晴弘, 野池賢二, 保科洋, 河原英紀: 音楽聴取に関する一検討: グループと頂点の推定, *FIT2004* (2004).
- 17) R. カーター: 脳と意識の地形図, 原書房 (2003).
- 18) 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).
- 19) 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).
- 20) 竹内好宏: 認知的視点による演奏解釈の研究, 修士論文, 兵庫教育大学大学院 (1994).
- 21) 菊池有恒: 演奏のための楽典, 音楽之友社 (1996).
- 22) 津崎実, 加藤宏明: 持続時間収縮錯覚と時間間隔による文脈の関係について, 情報処理学会研究報告音楽情報科学 2002-MUS-57-28 (2004).