

聴取者傾向を加味した GTTM グルーピング規則適用の 演奏表情パラメータへの拡張

野池 賢二^{t1} 橋田 光代^{t1,t3}
竹内 好宏^{t4} 片寄 晴弘^{t1,t2}

本稿では、GTTM であまり積極的には述べていない、Grouping Structure 選好ルール (GPR) の適用対象の、演奏表情パラメータへの拡張について述べる。その際、GTTM において明示的に述べられていないルール適用の閾値を、聴取者ごとに定めることにより、聴取者傾向を反映した「その人らしいグループ境界推定」を行うことを目指す。

Expansion of GTTM Grouping Rules for Involving Expression Parameters

KENZI NOIKE, MITSUYO HASHIDA, YOSHIHIRO TAKEUCHI
and HARUHIRO KATAYOSE

In this paper, we describe expansion of GTTM Grouping Preference Rules for involving performance expression parameters. In GTTM, the author is not described explicitly about threshold for rule applying. We adjust the threshold value to each listener to reflect listener's inclination, and aim to inferring grouping boundary according to his listening tendency.

1. はじめに

音楽構造を自動的に分析する機構は、音楽認識・理解システムにとどまらず、生成系のシステム、特に、予測・補完を伴うシステムにおいて、なくてはならない重要な役割を担う。構造分析機構の計算機上への実装可能性という観点で有望視されている音楽理論の一つに GTTM (A Generative Theory of Tonal Music) がある¹⁾。これは、Grouping Structure, Metrical Structure, Time-Span Reduction, Prolongational Reduction の 4 つのサブ構造分析理論から構成されており、そのうちの Grouping Structure と Metrical Structure の一部が、浜中らによって実装されている²⁾³⁾。GTTM の実装上の困難さのひとつに、ルールの優先順位の与え方がある。浜中らの研究では、ルールの適用を制御する調節可能なパラメータを導入することで、その解決に取り組んでおり、パラメータ調節

の結果、3 人の専門家による“正解”との F 値評価を向上させている。

GTTM は、基本的に楽譜上のシンボリックな情報から分析を行う理論であり、浜中らの実装は、それに則った実装となっている。しかし GTTM はその中で、演奏者による微細な演奏表情の違い (演奏表情パラメータ値の変化) によってグループ構造知覚が変化することにも、積極的ではないが触れており、この点の考慮を踏まえた定式化と実装はなされていない。

我々は、以前から演奏表情パラメータの操作によって構造知覚が変化することに注目しており、モーツァルトの K.331 を題材とした聴取実験を行っている⁴⁾⁵⁾。その結果、Henle 版、Peters 版の判別に用いるパラメータの判別閾値や優先度合いが、聴取者によって異なることを確認した。これは、構造認知において、シンボリックな情報を上回って、演奏表情パラメータの情報が必要な役割を果たすことを示しており、さらに、その効果が聴取者によって傾向が異なることを示している。

これらを踏まえ、本稿では、GTTM の Grouping Structure の選好ルール (以降 GPR と記す) の適用対象を、楽譜上のシンボリックな情報から演奏表情パ

^{t1} 科学技術振興機構さがけ研究 21
PRESTO, JST

^{t2} 関西学院大学理工学部
Kwansei Gakuin University

^{t3} 和歌山大学システム工学研究科
Wakayama University

^{t4} 京都府立須知高等学校
Kyoto Prefectural Syuuichi High School

ここでいう演奏パラメータの操作とは、MIDI レベルの発音制御を指しており、具体的には、打鍵時刻、離鍵時刻、打鍵強度の操作である。

ラメータへ拡張することで、より人間の認知に近い音楽グループ構造知覚分析を行うことについて述べる。

まず、2章でGPRの作用について吟味し、その適用対象の演奏表情パラメータへの拡張を定式化する。3章では、その定式化に則ったGPR適用の実装について述べる。4章では、この拡張が有効であることを示す初期的な実験結果を示し、最後に、今後の展望について述べる。

2. Grouping Structure 選好ルール

2.1 GPRの内容とその作用

GPRは、大きな分類で7種あり、その内容を簡単に述べると次のようになる。

- GPR1, alternative form
非常に小さなグルーピングの回避。
- GPR2 (Proximity) 時間的な近さ
連続する4音 n1, n2, n3, n4 に関するルール
 - GPR2a (Slur/Rest) スラー/休符(間)
n2の終わりからn3の始まりまでの時間間隔がn1の終わりからn2の始まりまでの時間間隔、および、n3の終わりからn4の始まりまでの時間間隔よりも長い場合、そこはグループ境界として聴かれる。
 - GPR2b (Attack-Point) 発音開始時刻
n2の始まりからn3の始まりまでの時間間隔がn1の始まりからn2の始まりまでの時間間隔、および、n3の始まりからn4の始まりまでの時間間隔よりも長い場合、そこはグループ境界として聴かれる。
- GPR3 (Change) 変化
連続する4音 n1, n2, n3, n4 に関するルール
 - GPR3a (Register) 音高変化
n2, n3間の音高変化がn1, n2間の音高変化、および、n3, n4間の音高変化よりも大きい場合、そこはグループ境界として聴かれる。
 - GPR3b (Dynamics) 強弱変化
n2, n3間で強弱変化があり、n1, n2間、および、n3, n4間では強弱変化がない場合、そこはグループ境界として聴かれる。
 - GPR3c (Articulation) アーティキュレーション変化
n2, n3間でアーティキュレーションの変化があり、n1, n2間、および、n3, n4間ではアーティキュレーションの変化がない場合、そこはグループ境界として聴かれる。
 - GPR3d (Length) 音長の変化

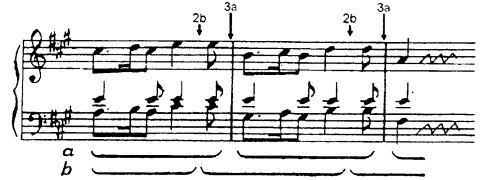


図1 K.331の二つのグルーピング候補

Fig. 1 Two candidate of musical grouping for K.331

n2, n3間で音長変化があり、n1, n2間、および、n3, n4間では音長変化がない場合、そこはグループ境界として聴かれる。

- GPR4 (Intensification)
GPR2, GPR3で挙げられる境界候補の効果が比較的明白である場合、より大きなレベルでもグループ境界となりうる。
- GPR5 (Symmetry)
グループを二つのサブグループに分割するとき、等分割となるグルーピングを選好する。
- GPR6 (Parallelism)
並列性のあるグルーピングを選好する。
- GPR7 (Time-Span and Prolongational Stability)
Time-span reductionとProlongational reductionの双方、あるいは、どちらかの結果がより安定的であるグルーピングを選好する。

GPRの内容とその作用を吟味すると、グループ境界候補を具体的に挙げるルールは、GPR2, GPR3である。残りのGPRは、これらによって仮定されたグループ境界候補の尤もらしさを補強、あるいは、取捨選択し、複数のグルーピング候補からの選好を行うルールと言える。したがって、本稿では、グループ境界候補挙げと演奏表情パラメータとの間に直接的な関わりがあるGPR2, GPR3を中心に考察する。

2.2 GPRの適用と演奏表情パラメータ

GPRの適用のされ方が、演奏表情パラメータ値の違いによって影響を受けることについては、GTTMの3.5節、「The Performer's Influence on Preferred Hearing」で、1節を割いて述べられている。ここでは、この節の要旨を簡単に紹介する。

図1は、モーツァルト作曲のピアノソナタK.331の冒頭部分である。この図は、二つのグルーピング候補a, bが対立していることを示している。シンボリックな情報だけを見た場合、グルーピングaを支持するGPRは、GPR3a(Register), GPR4(Intensification), GPR6(Parallelism)であり、音高変化による境界候

補を、モチーフの並列性が支えている。グルーピング b を支持する GPR は、GPR1、GPR2a(Slur/Rest) GPR2b(Attack-Point)、GPR5(Symmetry) であり、主に、四分音符とそれに続く八分音符との発音開始時刻間隔による境界候補を、等分割による対称性が支えている。

GTMM 3.5 節によれば、グルーピングの解釈が複数あるとき、演奏者は、聴取者にどのグルーピングで聴かletたいのかによって、演奏表情を決めている、とある。図 1 の K.331 の例では、聴取者にグルーピング a を知覚させる場合、四分音符の発音終わりを八分音符まで延ばし、八分音符の発音時間を短くし、その音量を相対的に小さくする。グルーピング b を知覚させる場合、四分音符の発音時間を短くし、その後少しの間を入れ、八分音符の発音終わりを延ばす。これは、GPR2、GPR3 による境界候補を、演奏表情パラメータ値によって強調していることになる。すなわち、シンボリックな情報だけでは対立するグルーピング候補を、演奏表情パラメータ値によってどちらかを選択しやすくさせている。また、この GPR2、GPR3 の境界候補の強調は、GPR4 における「比較的明白な」の「明白さの度合い」とも関連し、GPR4 の適用可否の判断材料ともなり得る。

この演奏表情による強調は、竹内による実証的な実験結果⁶⁾⁷⁾とも合致しており、GPR 適用対象を演奏表情パラメータへ拡張することは、GTMM による音楽構造分析をより確かなものにするために、必要かつ不可欠であると考えられる。

2.3 適用範囲の演奏表情パラメータへの拡張

GPR2、GPR3 の適用対象を演奏表情パラメータへ拡張するために、ここでは、それらを演奏表情パラメータを用いて定義し、定式化する。

GPR2、GPR3 は、本来、「境界候補となり得るかどうか」を挙げる二値ルールであるが、本稿の定式化では、「境界になりやすさの度合い」を挙げる連続値ルールであるとし、実際にそこを境界とするかどうかは、それぞれのルールに設けられた調節可能な閾値による判定と、残りの GPR による選好に任せるものとする。この閾値設定を調節することにより、聴取者によって異なる選好の傾向を反映させることや、GTMM と同様の厳密な大小関係比較を条件として設定することも可能となる。これらのルールで得られる値が負値をとる場合、その絶対値は、「境界になりにくさ」を表しているとも言える。

以降、適用対象を演奏表情パラメータへ拡張した GPR2、GPR3 の定義を順に述べる。拡張した GPR に

は、それぞれにサフィックス p を付けて区別することとし、拡張した GPR を含む、GPR 全体を GPRp と呼ぶことにする。

- GPR2ap (Slur/Rest)

n2 を i 音目とし、n2 の終わりから n3 の始まりまでの時間間隔を $rest_i$ 、n2 の始まりから n3 の始まりまでの時間間隔を ioi_i と表記することにする。このとき、GPR2ap による n2、n3 間の「境界になりやすさ」 D_i^{GPR2ap} を次の式で定義する。

$$D_i^{GPR2ap} = \frac{(rest_i - rest_{i-1}) + (rest_i - rest_{i+1})}{ioi_i}$$

- GPR2bp (Attack-Point)

n2 を i 音目とし、n2 の始まりから n3 の始まりまでの時間間隔を ioi_i 、n2 の音価を $lenrest_i$ と表記することにする。 $lenrest_i$ は、n2 の直後に休符が続く場合、それを加算した値であるものとする。このとき、GPR2bp による n2、n3 間の「境界になりやすさ」 D_i^{GPR2bp} を次の式で定義する。

$$D_i^{GPR2bp} = \frac{(ioi_i - ioi_{i-1}) + (ioi_i - ioi_{i+1})}{lenrest_i}$$

- GPR3ap (Register)

n2 を i 音目とし、n2 から n3 への音程を reg_i と表記することにする。このとき、GPR3ap による n2、n3 間の「境界になりやすさ」 D_i^{GPR3ap} を次の式で定義する。

$$D_i^{GPR3ap} = (reg_i - reg_{i-1}) + (reg_i - reg_{i+1})$$

- GPR3bp (Dynamics)

本来の GPR3b では、n1、n2 間、n3、n4 間で強弱変化がない場合にしか適用し得なかったが、GPR3bp では、本来の GPR3b を少し変更し、n2、n3 間の強弱変化量と n1、n2 間、n3、n4 間の強弱変化量との差を n2、n3 間の「境界のなりやすさ」として反映できるように定式化する。これにより n1、n2 間、n3、n4 間で強弱変化があっても、n2、n3 間の境界のなりやすさを評価できる。

定式化のために、n2 を i 音目とし、n2 から n3 への強弱変化量を dyn_i と表記することにする。このときの GPR3bp による n2、n3 間の「境界になりやすさ」 D_i^{GPR3bp} を次の式で定義する。

$$D_i^{GPR3bp} = (dyn_i - dyn_{i-1}) + (dyn_i - dyn_{i+1})$$

- GPR3cp (Articulation)

GTMM において、“Articulation” という用語の指す内容の明確な定義はないが、文献 1) の中での

適用例から推測すると、GPR3c 自身が演奏表情パラメータを評価をするルールであると考えられる。したがって、本稿での GPR3cp は、GPR3c そのものとする。

定式化のために、 n_2 を i 音目とし、 n_i の“Articulation”を $arti_i$ と表記することにする。このとき、GPR3cp による n_2, n_3 間の「境界になりやすさ」 D_i^{GPR3cp} を次の式で定義する。

$$D_i^{GPR3cp} = \begin{cases} 1 & arti_i \neq arti_{i+1} \text{ かつ} \\ & arti_{i-1} = arti_i \text{ かつ} \\ & arti_{i+1} = arti_{i+2} \\ 0 & \text{else} \end{cases}$$

● GPR3dp (Length)

本来の GPR3d では、 n_1, n_2 間、 n_3, n_4 間で音長変化がない場合にしか適用し得なかったが、GPR3dp では、GPR3d を少し変更し、 n_1, n_2 群の音長を表す量と n_3, n_4 群の音長を表す量との差を n_2, n_3 間の「境界のなりやすさ」として反映できるように定式化する。これにより n_1, n_2 間、 n_3, n_4 間で音長変化があっても、 n_2, n_3 間の境界のなりやすさを評価できる。

定式化のために、 n_2 を i 音目とし、 n_i の音長を len_i と表記することにする。このとき、GPR3dp による n_2, n_3 間の「境界のなりやすさ」 D_i^{GPR3dp} を次の式で定義する。

$$D_i^{GPR3dp} = |((len_{i-1} + len_i) / 2) - ((len_{i+1} + len_{i+2}) / 2)|$$

3. GPRp 適用ツールの実装

GPRp の作用を確認するために、GPRp を実際に適用できるツール IGp を作成した。現在、GPR2p、GPR3p が実装済みであり、それらの“その場での”適用が可能である。

3.1 IGp の入出力

IGp の入力データには、楽譜上のシンボリックな情報に加えて、演奏表情パラメータ値が必須であるため、MusicXML(4R)⁸⁾、または、note 形式⁹⁾ を入力情報として用いる。

MusicXML(4R) は、MusicXML の演奏表情パラメータ記述を積極的に活用し、MIDI を用いた演奏のために、ヴェロシティ値を正確に記述できるように微細な変更を施した MusicXML 仕様である。MusicXML(4R) は、Rencon¹⁰⁾ で配布する学習・評価用データセットの標準フォーマットとして策定が進んだファイルフォーマットであり、IGp の入力情報として

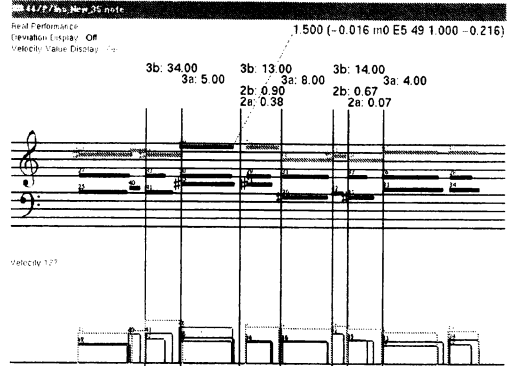


図 2 IGp による GPRp 適用
Fig. 2 GPRp applying by IGp

適した形式である。

note 形式は、豊田らの deviation データベース¹¹⁾、奥平らの iFP¹²⁾ で用いられている演奏表情情報の記述形式で、演奏情報を、楽譜情報とそれからのずれ情報の組で記述すること、演奏テンポを、多層の観点から記述可能であることが特徴である。note 形式は、微細な演奏表情パラメータ値を記述できるが、基本的に、スラーや f などのシンボリックな情報を記述しない。したがって、IGp の入力情報として過不足なく情報を与えるために、これらの情報はメモ情報として記述しておく。

IGp は、入力データに対して GPRp を適用し、その結果として得られるグループ境界候補を、音データのピアノロール表示の上に、「境界のなりやすさ」の値とともに表示する。IGp は、GPRp の適用を“その場で行うため、ピアノロール表示上の音データを時間軸方向に伸縮したり、音量を変化させたりすると、GPRp の適用結果が即座に反映され、適用のされ方の変化をその場で確認できる。IGp による GPRp 適用の様子を図 2 に示す。

IGp は、GPRp の適用結果を、浜中らの提案する GroupingXML²⁾³⁾ で出力することもできる。これにより、グループ構造分析結果を他のシステムで参照したり、データベースとして蓄えたりすることも可能となる。

3.2 GPRp 適用に必要なパラメータの算出方法

GPRp を適用するためには、入力データから GPRp 適用に必要なパラメータを算出する必要がある。IGp の実装では、各パラメータを次のようにして算出した。

● rest. ioi

GPR2ap, GPR2bp で必要な rest.ioi は、演奏表情パラメータ値から求める。各音の発音時刻、消

音時刻から算出する。

- *reg*
GPR3ap で必要な *reg* は、隣接する音の半音階での距離を用いる。
- *dyn*
GPR3bp で必要な *dyn* は、演奏表情パラメータ値からわかる、隣接する音の MIDI ヴェロシティ値の変化量を用いる。
- *arti*
GPR3cp で必要な *arti* は、各音に付随する演奏アーティキュレーションに関わる記号の情報を、便宜的に数値化して用いる。実際の演奏を評価するルールに与えるパラメータとしては粗い近似ではあるが、文献 1) で例示される適用例は、これでカバーできる。現在の実装では、スラーとスタッカートを *arti* として扱うことができる。
- *len*
GPR3dp で必要な *len* は、各音の音価を四分音符あたりを 1.0 とする時間尺度に換算した値を用いる。
- *lenrest*
GPR2bp で必要な *lenrest* は、各音の音価を四分音符あたりを 1.0 とする時間尺度に換算した値を用いる。ただし、*len* と異なり、直後に休符が続く場合は、その音価を加算する。

4. 評価実験

GPR2, GPR3 の適用対象を演奏表情パラメータに拡張したことによる効果を評価するために、初期的な評価実験を行った。

評価の題材として、我々が過去の研究において得ている、二人の被験者に対する K.331 の二つの版の識別空間調査結果⁴⁾⁵⁾ と、調査時に使用した、演奏表情パラメータ値の異なる 66 個の演奏データを用いた。これらを用いて、IGp によるグループ境界推定結果との再現率を求め、GPRp の効果を評価する。

評価実験に先立って、GPR2p, GPR3p から得られる「境界になりやすさ」の値から、グループ境界であるかどうかを判定する閾値を定める必要がある。本評価実験では、聴取者によって異なる聴取傾向を反映するために、この閾値算出を被験者別に行った。K.331 では、図 3 に示す楽譜上の位置が、演奏表情パラメータ値の違いによって、グループ境界として聴かれたり、聴かれなかったりし、それは聴取者によって異なる。これを利用し、グループ境界として聴かれるときと、聴かれないときの各 GPRp の値を調査した。本調査は、

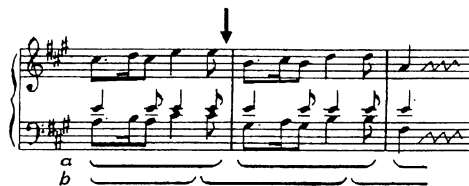


図 3 K.331 を利用した閾値算出点
Fig.3 Calculate point of GPR threshold

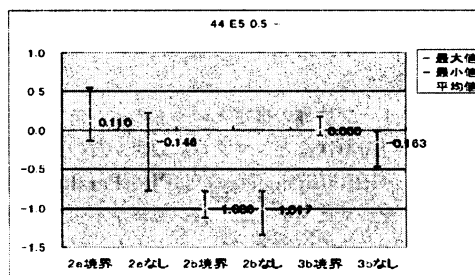


図 4 被験者 A の GPRp 値の分布
Fig.4 Distribution of each GPRp values (subject A)

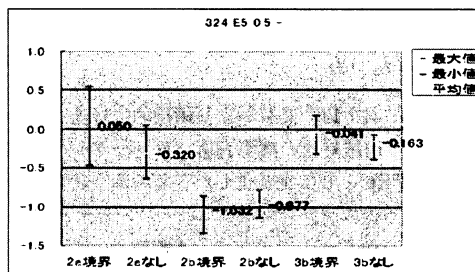


図 5 被験者 B の GPRp 値の分布
Fig.5 Distribution of each GPRp values (subject B)

演奏表情だけが異なる同一音列に対する調査となるので、聴取者によって閾値が異なる GPRp は、GPR2ap, GPR2bp, GPR3bp だけとなる。

調査結果を、被験者ごとに、図 4, 図 5 に示す。

図 4, 図 5 から、図 3 に示す位置をグループ境界と認知するときの GPRp の値の分布が、二人の被験者で異なることがわかる。また、値の分布の重なり具合から、GPR2b だけでは、このグループ境界を判別することができないことが読み取れる。

この調査結果から、GPRp 閾値を静的に定めることは難しいことがわかるが、本評価実験では、暫定的に、各 GPRp 値の分布の平均値を用いることとした。

この閾値を用いて、K.331 の二つの版の識別空間調査時に用いた演奏データ 66 個のグループ境界を IGp で推定し、その再現率を調べた。その結果、被験者 A

が75.0%、被験者Bが61.8%であった。被験者Bの再現率が少々低めではあるが、これにより、GPRのGPRpへの拡張によって、シンボリックな情報だけでは行えない聴取者ごとの傾向を反映した推定が、6割から7割半程度行えることが確認できた。

推定に失敗しているところを調べると、そこは、局所的な情報からグループ境界候補を挙げるGPR2p、GPR3pだけでは解決できない箇所であった。また、図4、図5において、GPR2p、GPR3pの値の分布が重なっていることからわかるように、これは本質的に、GPR2p、GPR3pによる閾値処理だけでは解決できない問題である。

本来、グループ境界の推定というものは、モチーフの並列性や等分割性など、マクロな観点からの評価が必要であり、GPR5(Symmetry)、GPR6(Parallelism)などの選好を加えた、トップダウン処理が不可欠である。図3に示す位置が境界であるかどうかの推定において、GPR1による隣り合う境界候補からの取捨選択、GPR5、GPR6による後続する境界候補の考慮によって解決が見込まれることから、残りのGPRpの実装とその評価が、今後の急務である。

5. まとめ

本稿では、GTTMでもあまり積極的に述べていない、GPR適用対象の演奏表情パラメータへの拡張を行い、GPRpとして定式化した。その際、聴取者によって異なる聴取傾向を反映するひとつの方策として、GPRpを、二値ルールではなく、連続値ルールとして定義した。また、GPR2p、GPR3pを、実際に“その場で”適用できるツールIGpを作成し、GPRpの初期的な評価を行った。

IGpは、GPRpの適用ツールとしてだけでなく、演奏表情のビジュアルエディタとしての利用価値もあり、今後は、音楽構造分析機構をもった演奏表情エディタとしての発展も視野に入れていきたい。

GTTMの実装は、その実現によって受けられる恩恵が期待され、いままでにもいくつかの取り組みがあったが、ルールの優先順位や閾値の与え方、ボトムアップ処理とトップダウン処理を適切に組み合わせることなど、理論の中で明記されていない部分の実現の困難さによって、4つのサブ構造分析理論の完全な自動化はいまだになされていない。本稿で述べたIGpにおけるGPRpの実装も、現段階では、GPRのごく一部を、より現実的に即した形で実装したに過ぎない。GTTMを完全に実装するという事は、それ自体が意義ある研究テーマではあるが、我々は、既存の理論の完全実装

という観点にとらわれずに、自動的な音楽構造分析機構の実現というテーマに、柔軟な思考で取り組んでいきたい。

謝辞 本研究は、科学技術振興機構さがけ研究21「協調と制御」領域の研究テーマとして実施されました。

参考文献

- 1) Lerdahl and Jackendoff: *A Generative Theory of Tonal Music*, MIT Press (1983).
- 2) 浜中雅俊, 平田圭二, 東条敏: GTTM グルーピング構造分析の実装: ルールを制御するパラメータの導入, 情報処理学会研究報告音楽情報科学 2004-MUS-55, No. pp. 1 - 8, pp. pp. 1 - 8 (2004).
- 3) 浜中雅俊, 平田圭二, 東条敏: GTTM に基づく楽曲構造分析の実装: グルーピング構造と拍節構造の獲得, 情報処理学会研究報告音楽情報科学 2004-MUS-56, No. 56, p. 56 (2004).
- 4) 野池賢二, 片寄晴弘, 竹内好宏: 演奏からの音楽グループ構造の抽出 - K.331 を例として, 情報処理学会研究報告音楽情報科学 2002-MUS-47-19 (2002).
- 5) 野池賢二, 橋田光代, 片寄晴弘: 音楽グループ境界識別空間調査ツール WebMorton, 情報処理学会研究報告音楽情報科学 2003-MUS49-5 (2003).
- 6) 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).
- 7) Takeuchi, Y.: Performance Variables for Grouping Structure in two editions of the theme of K. 331, *Proc. ICAD-Rencon*, pp. 47-50 (2002).
- 8) Hirata, K., Noike, K. and Katayose, H.: Proposal for a Performance Data Format, *In Working Notes of IJCAI-03 Workshop on methods for automatic music performance and their applications in a public rendering contest*, pp. pp. 65 - 69 (2003).
- 9) NOTE format Specification:
<http://ist.ksc.kwansei.ac.jp/katayose/Download/Document/>.
- 10) Rencon web site:
<http://shouchan.ei.tuat.ac.jp/~rencon/>.
- 11) 豊田健一, 片寄晴弘, 野池賢二: 音楽解釈研究のための演奏 deviation データベースの作成, 情報処理学会研究報告音楽情報科学 2004-MUS-51, pp. pp. 65 - 70 (2003).
- 12) 奥平啓太, 片寄晴弘, 橋田光代: 音楽演奏インタフェース iFP - 演奏表情のリアルタイム操作とビジュアライゼーション-, 情報処理学会研究報告音楽情報科学 2003-MUS-51 (2003).
- 13) D. Temperley: *The Cognition of Basic Musical Structures*, MIT Press (2004).