

演奏上での頂点とグループ境界の聴取モデルについて

片 寄 晴 弘^{†,††} 橋 田 光 代^{††} 野 池 賢 二^{††}

演奏者の多くは自分が理解した音楽構造を伝えるという目的をもって音楽演奏を行うが、その演奏を聴いた聴取者が認知する音楽構造は必ずしもユニークなものとはならない。我々は、聴取者の音楽経験やアテンションによって聴き方が変わるという前提に立ち、さまざまな聴き方の存在を統一的に説明する聴取モデルの構築を目指している。本論文では、演奏表現に関して十分考慮されたベートーベン悲愴ソナタ第2楽章を題材として、音楽に関して何かの経験を持つ被験者のグループ境界、グループ中での重要音知覚に関する典型パターンを整理する。その上で、音量輪郭の追跡処理とパートアテンション、さらにスキーマに基づく選択的注意によって頂点とグループ開始点の聞き分けを説明するモデルを提示する。

On Perception of Apices and Group Boundaries in Listening Music

HARUHIRO KATAYOSE, MITSUYO HASHIDA and KENZI NOIKE

A Music player performs music with a goal to convey the audience the structure that she/he conceives. However, the audiences do not always find the structure as the player intends. The authors are based on a standpoint that there is no right ways in perceiving music, and are trying to make a generalized model that explains various listening tendency. This paper, picking up well-elaborated "Piano Sonata No.8, pathetic, the 2nd. movement" composed by Beethoven, examines heard patterns of music boundaries and apices, of those who experienced in music. Then we propose a model that explains various listening tendencies regarding boundaries and apices, based on search for dynamic-contour, control of attention toward each part, and attentions to schemata.

1. はじめに

演奏表現における最も基本的な作業は、演奏者や指揮者が自らが理解した音楽構造を伝えるということである。規範となる音楽環境で訓練を受けてきた聴取者に限れば、演奏を聴いて演奏者の意図（想定した構造）が概ね理解できる。はからずとも専門家の閉じた世界の中では受容されるべき音楽構造としては共通の理解が存在し、一つの正解が存在することを前提に、理論立て、教育の展開がなされてきた。

楽譜からの音楽構造知覚¹⁾や自動採譜（表情を含む演奏の楽譜化²⁾は、音楽認知科学あるいは音楽情報科学の重要な研究テーマとして、今までにも積極的に取り組まれている。それらの研究の多くも、暗黙のうちに受容されるべき音楽構造には正解が存在するという観点に立って構築されてきた。聴取者の認知的な音

楽構造理解という視点を導入し、この分野での研究思想に大きな影響を与えている理論の一つに GTTM³⁾がある。GTTM は演奏パラメータが変わることによって音楽構造の知覚が変化することに触れているが、ここでも、聞き方の多様性については積極的な主張はなされていない。

モーツァルトのピアノソナタ K.331 を題材としたグループ境界の判別に関する先の調査⁴⁾で、我々は、同じ演奏であっても異なった音楽構造の理解を示す聴取者がいることを確認した。また、十分に音楽教育を受けた被験者でも、音楽を構成する要素に対してのアテンションの持ち方で聞き方の傾向が異なるということも確認し、以来、複数の聞き方の存在を許容した上で、それらを統一的に説明する聴取モデルの構築に取り組んでいる。音楽を聞き方の多様性という点でとらえるなら、数多くの興味深い事項が浮かび上がってくる。その中でも、本論文では、同じ演奏のある特定の音に対し、グループのスタート点ととらえる聞き方とフレーズの頂点ととらえる聞き方が存在するという現象について焦点を当てる。

[†] 関西学院大学

Kwansei Gakuin University

^{††} 科学技術振興機構さきがけ研究 21

PRESTO, JST

以下、音楽グループとフレーズ、その表現について楽典的に説明されている事項と我々の立場について述べる（第2章）。次に、演奏表現に関して十分考慮されたベートーベン悲愴ソナタ第2楽章を題材として、音楽に関して何かの経験を持つ被験者の頂点・グループ境界知覚のパターンを調査し、その典型パターンの整理を行う。さらに、アテンションのコントロールによって、グループ境界知覚が変化することを示す（第3章）。その結果に基づき、目立つ音の追跡処理とパートアテンションによって頂点とグループ開始点の聞き分けを説明するモデルを提案する（第4章）。

2. 音楽のグループとその表現

2.1 グループとフレーズ

伝統的な西洋音楽に限らず音楽には表現の単位となるまとまりが存在し、グループないしはフレーズと呼ばれている。作曲の際に言及されるモチーフとはグループの代表例である。フレーズとは、発声すなわち一息の範囲で表現されるまとまり感をより意識した用語であり、一般的にはこちらの呼び名で呼ばれることの方が多い。フレーズは物理的な一息の範囲を超え、1~2小節のものから、4小節あるいは8小節など広いレンジの演奏上のまとまった表現対象をさすことも少なくない。階層的なものとしての理解も定着しており、比較的大きな（長い）フレーズの構成要素として、サブフレーズという用語も普通に使われるようになっている。

我々のフレーズとサブフレーズ、グループに対する考え方は、基本的には上記の一般的な理解にしたがうものであるが、議論を行っていく上で、以下に示すような定義を与える。

- (1) グループとは比較的小さなフレーズあるいはサブフレーズそのものを指す。
- (2) フレーズ（サブフレーズ）の終了音の次の音が次のフレーズ（サブフレーズ）の開始音となる
- (3) フレーズ（サブフレーズ）の終了音の次のフレーズの開始音と一致することがある。

以上の考え方は、GTTM³⁾で定義されるグループ・イングストラクチャあるいはTimeSpan木で定義されるスパンとほぼ同等のものである。ただし、我々はグループやフレーズは、各パート（声部）毎に定義できるという前提にたっており、厳密には、GTTMの前提とは異なる。

2.2 音楽のグループとフレーズの導出

演奏者の行う音楽のグループやフレーズの分析は楽譜に従って実施されるものである。音楽的に十分な洞

察力を持った演奏者が、よく練られた楽譜中の繰り返しパターン、和声進行、スラーを代表とする演奏表意を読み取っていけば、サブフレーズを上位のフレーズレベルにまとめるかどうかを除いて、ほぼ一定の解釈に定まるといわれている³⁾⁵⁾⁶⁾。

GTTMは、ホモフォニーを中心とする伝統的西洋音楽に対して、グループやフレーズの知覚のベースとなるルールと、メタルールに関する整理を行ってきた。グループ構造、拍節構造をベースにCompositionの構造を表すTimeSpan簡約と緊張弛緩構造を表すProlongational簡約の導出を行っている。ツリー状の解析を行うのが大きな特徴である。

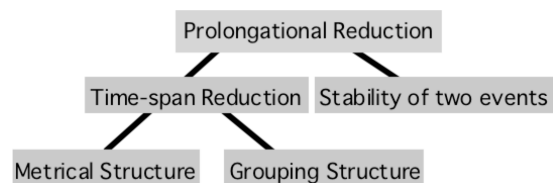


図1 GTTMの基本構造

IRMは、図2に示すように、メロディに対して進行の暗示と解決に関して、入れ子を許す形でアノテーションを行い、その意味解析をすることによって、音楽構造の理解を行うことを目指している。

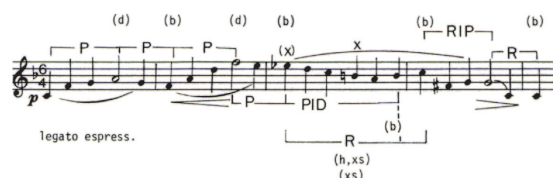


図2 IRMの例

保科は、演奏においてどの音にエネルギー（演奏上の力点）を置くかという視点にたち、基本的は倚音、フレーズ上での頂点ピッチが高くかつ長い音を中心にアナクルーズとデジナンスの関係をを用いたフレーズの把握と表現法の提示を目指している。人間が理解するという目的に立てばかなりわかりやすい理論である。

以上に示した三つの例に示すように、フレーズ解析の手がかりとなる原則および初期的な音楽グループの発見手段については既に提示されている。しかしながら、GTTMにおけるメタルールの扱い、IRMにおけるアノテーションのまとめあげ、保科の理論でのアナクルーズとデジナンス、フレーズバウンダリの発見など十分に定式化されていない部分が存在する。作品（楽譜）の中には、構造の解釈自体に多様性を許すも

のも少なからず存在している．よく練られた楽譜であれば，ほぼ一定の構造解釈に定まるとは述べたが，この部分のモデル化は大きな課題である．

2.3 音楽のグループ・フレーズの演奏表現

一旦，グループやフレーズの理解が確定すれば，演奏においては聴取者にそれが解るような形で表現されるべきである．図3にモーツァルト K.331 の二つの構造の例を示す．上が Peters 版楽譜によって支持される構造，下が Henle 版楽譜によって支持される構造である．

グループ・ゲシュタルトの観点からはグループとグループの間に間をあける³⁾，グループ開始の音にアクセントをつける（音量をあげる）などの表現が指摘されている⁷⁾⁸⁾．



図3 K.331 二つの版の存在

一方，フレージングに留意した事項としては，図4に示すような形で，頂点（フレーズ中で最も重要になる音）に向かって，音量と演奏速度が増加し，逆に，頂点を過ぎれば，それらは減少するというような表現が一般的なものとして理解されている．

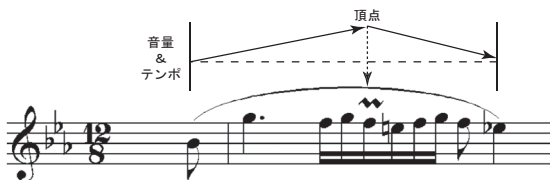


図4 フレーズ表現

これらの例が示すように，グループ表現を見た場合，グループ開始アクセントの表現といわゆるフレージングと呼ばれる二つのタイプが存在していることがわかる．この二つの典型的な表現例の仕方の存在とその聴取モデルにおける統合が本論文の主題となる．以下，ベートーベン悲愴ソナタ第2楽章による，グループ（フレーズ）境界の知覚に関する実験例を紹介する．

3. 聴取傾向の調査

3.1 調査の概要

本研究での一番目の課題は同一の演奏がどのように聴取されるかを調べることである．実験にあたっては，



図5 ベートーベン悲愴ソナタ第2楽章（冒頭8小節）

1) 十分音楽的に演奏データが練られたものであること，2) その演奏の生成にあたっての音楽的構造が明記されていること，が望まれる．この要請を満たす対象として，保科の文献⁶⁾に所収されているベートーベン作曲，悲愴ソナタ第2楽章冒頭（図5）を取り上げることにした．

図6にこの曲の演奏データを示す．このデータでは，各小節のスタートタイム（四分音符を1.0とする）毎の括弧内に音名，ペロシティ，持続時間を記述している．テンポの変化については，BEATTIMEという記述子の後に，四分音符あたりの占有時間を記述するようにしている．なお，ペダルに関してはOFF，ゲートタイムに関しては100%，7小節目のスラーとスタッカートがかかった部分については，40%に設定し，音源 GigaPiano を用いて波形データに変換したデータを使用した．音データについては <http://ist.ksc.kwansei.ac.jp/~katayose/mus52/> を参照されたい．

この実験に先立つ予備調査において，音楽体験の少ない被験者は，自身の聞き方を外在化することが困難なことが確認されていた．今回の実験では某かの音楽経験のある被験者：音大修士卒2名（一名はピアノ演奏科卒業，一名は音楽デザインコース卒業），音楽大学学生（情報音楽コース3年生）9名，エレクトーン経験者1名，バンド経験者4名を対象に聞き方の分類を試みた．

調査では，まず，被験者に，第2.1節で示したような音楽グループに関する定義と重要音に関する説明を行った．その上で，演奏を聴いた上で，スラー等の情報を外した楽譜に，1) 表現上でのグループ，2) グループ内での重要音，について書き込んでもらうとい

音源，スピーカーによって聴取傾向が微妙に変化する可能性は否めない

```

Header (
  GENERATOR "smf2note Ver. 0.40 2003-06-16(Mon)"
  BEATTIME 1875.00 4
  BEAT 2 / 4
)

= 1
1.00 BEATTIME 2142.857 4
1.00 ( C4 32 1.00 ) ( G#3 18 0.25 ) ( G#3 18 0.25 )
1.25 BEATTIME 1764.706 4
1.25 ( D#3 24 0.25 )
1.50 BEATTIME 1875.00 4
1.50 ( G#3 26 0.25 )
1.75 BEATTIME 2000.00 4
1.75 ( D#3 28 0.25 )
2.00 ( A#3 38 1.00 ) ( G3 18 0.25 ) ( C#3 34 1.00 )
2.25 ( D#3 32 0.25 )
2.50 ( G3 36 0.25 )
2.75 ( D#3 40 0.25 )
= 2
1.00 BEATTIME 2222.222 4
1.00 ( D#4 46 1.50 ) ( G#3 28 0.25 ) ( C3 32 1.00 )
1.25 BEATTIME 1875.00 4
1.25 ( D#3 36 0.25 )
1.50 ( G#3 34 0.25 )
1.75 ( D#3 32 0.25 )
2.00 BEATTIME 1935.484 4
2.00 ( A#3 22 0.25 ) ( G2 28 1.00 )
2.25 BEATTIME 1818.182 4
2.25 ( D#3 28 0.25 )
2.50 BEATTIME 1875.00 4
2.50 ( C#4 34 0.50 ) ( A#3 18 0.25 )
2.75 ( D#3 20 0.25 )
= 3
1.00 ( C4 32 0.50 ) ( G#3 16 0.25 ) ( G#2 26 0.50 )
1.25 ( D#3 26 0.25 )
1.50 ( D#4 42 0.50 ) ( A#3 20 0.25 ) ( G2 32 0.50 )
1.75 ( D#3 36 0.25 )
2.00 ( G#4 54 0.50 ) ( C4 28 0.25 ) ( F2 40 0.50 )
2.25 BEATTIME 2142.857 4
2.25 ( G#3 48 0.25 )
2.50 BEATTIME 2500.00 4
2.50 ( A#4 38 0.50 ) ( D4 24 0.25 ) ( F3 32 0.50 )
2.75 BEATTIME 1818.182 4
2.75 ( G#3 30 0.25 )
= 4
1.00 ( D#4 40 1.50 ) ( G3 18 0.25 ) ( D#3 26 1.00 )
1.25 BEATTIME 1764.706 4
1.25 ( A#3 26 0.25 )
1.50 ( G3 24 0.25 )
1.75 ( A#3 22 0.25 )
2.00 BEATTIME 1875.00 4
2.00 ( G3 20 0.25 ) ( D#2 20 1.00 )
2.25 ( A#3 20 0.25 )
2.50 ( E4 38 0.50 ) ( G3 18 0.25 )
2.75 ( A#3 20 0.25 )

= 5
1.00 ( F4 36 1.00 ) ( G3 18 0.25 ) ( C#2 22 1.00 )
1.25 ( A#3 22 0.25 )
1.50 ( G3 28 0.25 )
1.75 ( A#3 34 0.25 )
2.00 ( A#3 42 0.75 ) ( G3 24 0.25 ) ( C#3 26 1.00 )
2.25 ( D#3 32 0.25 )
2.50 ( G3 30 0.25 )
2.75 ( C4 38 0.125 ) ( D#3 20 0.25 )
2.875 ( C#4 36 0.125 )
= 6
1.00 ( D#4 34 1.00 ) ( G#3 22 0.25 ) ( C3 24 1.00 )
1.25 ( D#3 30 0.25 )
1.50 ( G#3 36 0.25 )
1.75 BEATTIME 2068.966 4
1.75 ( D#3 44 0.25 )
2.00 BEATTIME 2307.692 4
2.00 ( A3 52 1.00 ) ( D#3 36 0.25 ) ( F2 38 1.00 )
2.25 BEATTIME 1764.706 4
2.25 ( C3 38 0.25 )
2.50 BEATTIME 1818.182 4
2.50 ( D#3 34 0.25 )
2.75 BEATTIME 1875.00 4
2.75 ( C3 28 0.25 )
= 7
1.00 BEATTIME 2222.222 4
1.00 ( C#4 36 1.00 ) ( F3 18 0.25 ) ( A#1 30 1.00 )
1.25 BEATTIME 1875.00 4
1.25 ( C#3 22 0.25 )
1.50 ( F3 24 0.25 )
1.75 ( C#3 26 0.25 )
2.00 BEATTIME 2000.00 4
2.00 ( C4 32 0.099 ) ( C#3 20 0.167 ) ( D#2 24 1.00 )
2.25 BEATTIME 1818.182 4
2.25 ( A#3 36 0.099 ) ( C#3 22 0.083 )
2.50 ( G#3 40 0.099 ) ( C#3 24 0.167 )
2.75 BEATTIME 2307.692 4
2.75 ( G3 44 0.099 ) ( C#3 26 0.083 )
= 8
1.00 BEATTIME 2068.966 4
1.00 ( G3 28 1.00 ) ( A#3 46 1.00 ) ( C#3 22 0.25 )
1.00 ( G#1 28 0.50 )
1.25 BEATTIME 1935.484 4
1.25 ( D#3 24 0.25 )
1.50 ( C#3 22 0.25 ) ( G#2 24 0.50 )
1.75 BEATTIME 2142.857 4
1.75 ( D#3 20 0.25 )
2.00 BEATTIME 1875.00 4
2.00 ( G#3 26 0.50 ) ( C3 16 0.50 ) ( C3 16 0.50 )
END

```

図 6 ベートーベン悲愴ソナタ第 2 楽章（冒頭 8 小節）の演奏データ

う手続きを取った。グループの数については特に指定せず、グループ内での重要音については「演奏者が演奏上で最も重要な音と考えている音であり、音量やテンポ、アゴーギクの変化で表現されている。グループ内で通常一つ見つかるはずであり、エネルギーを持って聞こえることが多い」と説明を行った。なお、実験については本人が納得できるまで何回も聞いてもらうようにした。

3.2 グループの知覚実験結果と検討

図 7 に演奏データの実際の動きについて示す。は音量のピーク、縦線は「ため（間）」の入った部分、メロディ、内声のそれぞれに対し、その声部の上部にクレシェンド、ディミヌエンドを付記している。尚、の大小はピークの強さ、縦線の太さはための大きさに対応させている。

調査結果の代表例（二人以上に支持されたもの）を図 8 に示す。この図において重要音については印で

記している。2 小節目の印についてはグループ内の重要音ではあるが、音量は前の音より小さいとしたグループの指摘を表している。a) は保科が想定した構造である。b) はバンド経験者 3 名と音楽大学生 1 名に支持されたグループ境界という意味で、Pops Listener (P.L.) タイプと呼ぶことにした。c), d) については、それぞれ、3 人の音楽技術者志向の学生に支持されたグループ境界ということで Music Engineer (M.E.) タイプ、その他の音大生内 3 人に支持されたグループ境界ということで、Music School Student (M.S.) タイプと呼ぶことにした。e), f) は、筆者二名（音楽解釈システムの研究に従事し、演奏表現にコンシャスになっている）がどちらの聞き方も可能ということで最後まで判断に苦しんだグループ境界である。それぞれ、Expression Listener (E.L.) タイプ、Continuity Listener (C.L.) タイプと呼ぶことにした。

図 8 には、当該のグルーピングを支持する根拠になりうる情報として、拍節構造支持（小節線でグループ境界）: Metrical Structure Consistency(MC)、フレージング支持（グループの中ほどに）: Phrasing Consistency(PC)、グループ開始アクセント支持（グループの最初の音が）: Group Starting Accent Consistency(SC) の記号を付記している。

図 8 と図 7 での演奏上重要な音に対する聞こえ方と実際の音量の動きとはほぼ一致している。異なっている部分は、メロディ 4 小節目の E 音、5 小節目の C 音に重要音を感じている被験者が多いということである。いずれも前の音と比べて実際の音量は小さい。3 小節目と同じ、付点四分音符 + 八分音符の組み合わせということといえば、2 小節目が該当する。2 小節目の音量 (velocity レベル) 差が 12 であるのに対し、3 小節目では 2 である。また、5 小節目の C 音の前の音との音量差は 4 である。定性的な解釈の域を超えないが、保科の指摘するように音の減衰曲線を越えた音が目立って聞こえる、その背景として、付点付きの音符（の支配する時間の長さ）があると考えれば、一応の説明が可能である。

図 8 においては、それぞれの聴取者の聞き方は、おおまかには、上にいくほどポップスに慣れた聞き方、下に行くほど伝統的な西洋音楽の表現に慣れた聞き方と考えられる。

b) P.L. タイプと c) M.E. タイプは、7 小節目のグルーピングの判断を除けば、ほぼ同様の聞き方をしている。P.L. タイプのリスナーは 7 小節目と 8 小節目

この説明に異論を持つ音楽経験者も存在しうる。今回の被験者の中では上記の説明に関しては同意は取れていた。

後で被験者に確認したところ、より大きく聞こえたと答えた。

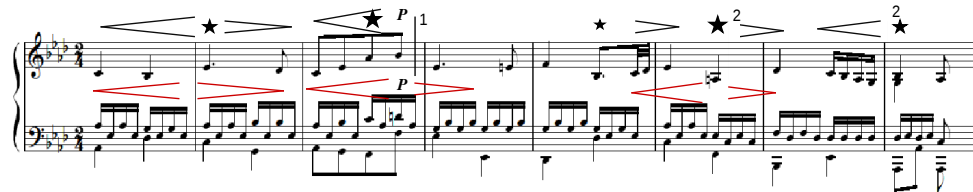


図 7 実験につかわれた演奏データの実際の動き（演奏データの視覚的解釈）



図 8 「悲愴」聴取結果

で境界を感じており、拍節構造を重視していると言える。それ以上に、重要と感じた音をグループの最後の音としていることに注目したい。ポップス、特にビート系の音楽では、グループの最後の音にアクセントが与えられることがある。伝統的な西洋音楽の表現ではあまりないことであるが、スキーマの一つになっていると考えられる。

d) M.S. タイプ、e) E.C. タイプは、重要な音に対する認識はほぼ同じである。これに対してグループのとらえ方は大きく異なっている。M.S. タイプは、グループ開始アクセント重視型、E.C. タイプは、フレージング重視の聞き方をしていることがわかる。

E.C.、C.L. タイプは筆者らのグルーピングである。この実験を行う際は、極力自身の聞こえを優先した。

結果として保科の提示したグループに似た傾向になっている。演奏解釈システムの研究を行っている関係上、通常より、フレージングに対してコンシャスになっている可能性は高い、5、6 小節目を複合グループととらえるなら、C.L. タイプは、ほぼ、演奏作成者の意図をほぼその通りにとらえているといえよう。

3.3 アテンションコントロールによる聞き方の変化

最近の研究によって、耳をそばだてるという行為によって当該の周波数領域での知覚感覚が向上するということが知られている⁹⁾。このようなアテンションコントロールによる聴覚の変化は、高次の音楽認知によっては、より顕著であると予想される。前節ではスキーマに基づく説明を試みたが、本節ではアテンションコントロールの影響を考究する。

図 7 のデータの動きからは、6 小節目のマークをグループの中央ととらえるタイプは 6 小節目の内声に留意している、逆に、P.L. タイプ、M.S. タイプの被験者は、内声にあまり留意せずにメロディ中心で聞いているという仮説が成り立つ。そこで、P.L. タイプ、M.S. タイプの被験者、さらに、図 8 には記載していない被験者 M.K.（エレクトーン歴 12 年）に対して、内声に留意して全体を聞くよう要請し、再び、前節の実験を実施した。同時に、内声に留意して聞いていると思われる被験者 Y.K.（ピアニスト）に対して、内声に対する意識を減らして聴取を行うよう要請して実験を行った。この様子を図 9 に示す。

P.L. タイプ 4 人の内、3 人は内声への留意によって、4 小節目一拍目を前のグループの終了音であり、かつ 6 小節目のマークを頂点とした長いフレーズの開始音として認識するようになった。一人はよくわからなくなったと答えた。前節では、ポップス特有のスキーマを聞き方の差の要因の一つとする考察を行ったが、内声の演奏表現への留意の度合いの低さがその背景になっていると考えられる。

M.S. タイプ 3 人の内の 2 人は、後半の二つのフレーズをまとめて聞くようになった、内一人は 8 小節目一拍目はさほど大きく聞こえなくなったと答えた。二名とも内声への留意で聞き方が変わることに驚き



図 9 アテンションコントロールによる聞き方の変化．上：コントロール後，下：コントロール前

を感じていた．もう一人の被験者は，最初のグループ開始音と変わらないと答えた．

M.K. は聞き方が大きく変わった．その中で，6 小節目をフレーズととらえるようになった．メロディに留意するように支持を受けた Y.K. は 6 小節目 2 拍目をグループ開始点として聞くようになった．以上，9 人中 7 人がアテンションコントロールによってグループ境界の識別が変わった．変化しないと判断した聴取者は 1 人しかいなかった．

以上を整理すると，1) 聞き方には，グループ開始アクセント重視型，フレージング重視型，拍節構造重視型などが存在する．2) アテンションを変えることによって聞き方が変わる．3) 内声への留意によって聴取傾向が変わる（フレージング重視型が増えることが多い），ということになる．また，仮説の域をこえないが，4) 伝統的西洋音楽の演奏経験あるいは聴取経験が多いほどメロディパート以外への留意が高い．5) 伝統的西洋音楽の経験者は，重要音をグループ終了音と感じにくい，ということが言えそうである．

次の章では，これまでの考察に基づき，1) パート毎の重要音，クレッシェンド，ディミヌエンド抽出処理（音量輪郭の追跡処理）2) パートへのアテンションコントロール，3) グループ開始アクセント重視型，フレージング重視型，拍節構造重視型等のスキーマへ

の選択的注意に基づく聴取モデルの提案を行う．

4. 重要音とフレーズの頂点に関する聴取モデル

この節では，まず，パートへのアテンションを考慮した，聴取者が感じる重要音を抽出する処理について述べる．続いて，その処理の結果から得られる重要音がグループ開始音か，フレーズ表現上の頂点音なのかを理解する処理について述べる．

処理内容は，実験による観測をもってその定義を行うべきであるが，精緻なモデル化を行う前に，モデルそのものの妥当性の確認が必要であると考え，ここでは近似的な定義を用いた．

4.1 聴取者の傾向を反映した重要音の抽出

重要音の抽出処理は，聴覚のマス킹現象のうちの同時マス킹と順向性マス킹に近い概念をモデル化したものである．本論文では，多声部からなる演奏を声部ごとに扱いやすくするために，便宜的に演奏情報をイベント単位で扱う．ただし，聴覚をモデリングするために，次のことを考慮する．

- 時間経過にしたがった音量の減衰
- 打鍵時音量にしたがった音高軸方向への影響

これらを考慮した上で，演奏情報を [時間-音高-音量] 空間にマッピングする．

マッピングするにあたって，ひとつの音イベントの音量を表す関数 $EP()$ を定義する． $EP()$ は，時間経過にしたがった音量の減衰を表す関数 $L()$ と音高軸方向への影響量を表す関数 $PA()$ から構成され，アテンションの強さを表す関数 $A()$ によりバイアスがかかるものとする．我々が用いた $EP()$ の定義を次に示す．

$$EP(v, li, len, t, pd) = L(li, len, t) \cdot li \cdot PA(pd) \cdot A(v, t)$$

$$L(li, len, t) = t \cdot li / (2 \cdot len)$$

$$PA(pd) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(pd-\mu)^2}{2\sigma^2}}$$

$$A(v, t) = \text{constant value defined by every } v.$$

li: 打鍵時の音量

len: 持続時間

t: 打鍵時刻からの経過時間

pd: イベント音高からの音高距離

v: 声部

$L()$, $PA()$, $A()$ の定義内容の詳細は紙面の都合で割愛するが，これらによってイベントの初期音量とイベント長にしたがったマス킹量に，アテンションの強さを加味して扱うことができる．

聴取者が考える重要音は，単純に音量が大きいだけ

Y.K. 本人は自身の演奏では，6 小節目 2 拍目 A 音を小さい音で表現している．拍節を意識していると思われる．

ではなく、マスキングを踏まえた上での音量的に目立つ音であると考えられる。したがって、 $EP()$ をマッピングする際に、 $EP()$ の重なりについて差分をとったマップを作成し、そのマップ上で比較的高い山になる箇所を探すことで、重要音を抽出することができる。

さて、前節の実験により、アテンションコントロールのない状態での P.L. タイプ、M.E. タイプが考える重要音は、演奏者の考える重要音とは、ずれていることがわかった (5 小節目 C 音)。これは、聴取者のアテンションが主旋律に強く向いているからであると考え、主旋律の $A()$ を他の旋律の 2 倍にして差分マップを作成した (図 10)。図中の 6 個の * が P.L. タイプ、M.E. タイプが考える重要音である。これによると、マップからわかる重要音と、P.L. タイプ、M.E. タイプの考える重要音とが一致しており、本処理によって、聴取者の聴取傾向を反映した重要音の抽出ができることが確認された。

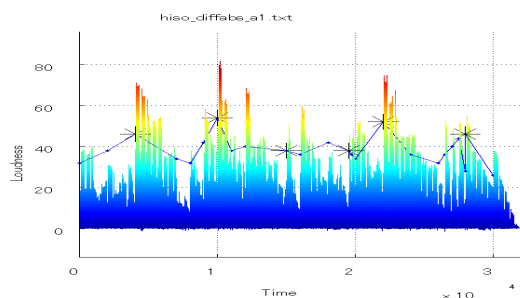


図 10 主旋律のアテンションを強めた「悲愴」差分マップ (時間-音量 平面)

4.2 グループ開始音と頂点音の弁別

前節の処理だけでは、抽出された重要音が、聴取者にとってのグループ開始音なのか、フレーズ表現上の頂点音なのかということの説明はできない。

重要音を頂点音と理解することに対して、我々は、聴取上の認知の観点から、次の二つの要因を仮定した。

ひとつは、聴取者のもつスキーマの作用である。GTTM などでも用いられているゲシュタルトによる説明では、重要音は頂点音にはなりえない。これが頂点音と理解される背景には、重要音と認識した音をフレーズの頂点音と理解するためのスキーマをその聴取者が持ち合わせ、それが働いているのではないかと考えた。そのスキーマの内容は、聴取実験後の聴取者へのインタビューや、演奏経験者の内省を踏まえると、フレーズ表現におけるクレッシェンド、ディミヌエンドへの意識の有無のようである。

もうひとつの要因は、フレーズ表現に関するスキ-

マが働くかどうかである。内声へのアテンションコントロールのない状態で、P.L. タイプが重要音を頂点音と理解することが少ない (あるいは、できない) のは、クレッシェンドの表現が知覚できなかったからであると考えられる。悲愴第 2 楽章冒頭中のクレッシェンドの表現は、内声へのアテンションが弱いと知覚できない。

この仮定に基づき、内声の $A()$ を他の旋律よりも大きくして作成した差分マップを図 11 に示す。右から二つの目の * が、図 7 の 6 小節目 2 拍目の * に対応する。これを見ると、* を中心として山型になっていることがわかる。これは、クレッシェンドとディミヌエンドによるフレーズ表現に対応しており、内声へのアテンションを強めることによってフレーズ表現が知覚できる状態になったことを示している。

ただし、この状態の差分マップから * をフレーズの頂点音と理解するためには、山型をフレーズ表現として理解するスキーマ処理が必要である。

我々はこれを、差分マップ M に、聴取者タイプ lt によって異なるスキーマ処理 $S()$ を施すことによってグループ境界 GB を得るという手続きでモデル化した。 $S()$ に含まれる処理には、たとえば、フレーズ表現重視スキーマ処理、グループ開始アクセント重視スキーマ処理、拍節構造重視スキーマ処理などが、聴取者タイプ lt によって異なる程度で組み込まれる。伝統的西洋音楽の演奏・聴取経験豊富な聴取者の $S()$ には、上で述べてきたフレーズ表現理解のスキーマ処理が含まれることになり、図 11 の右から二つ目の * は、フレーズの頂点音であると認識される。P.L. タイプの $S()$ には、そのスキーマ処理の働きが弱い、含まれないことになり、図 11 の * は、頂点音とは認識されにくくなる。

このスキーマ処理 $S()$ を施すという処理モデルにより、重要音を頂点音と理解する認知過程についても感覚的に理解しやすい形で説明できよう。

5. 検 討

我々は、演奏表現と聴取傾向の観察に基づいて、音楽グループ境界を識別するモデルの構築を進めてきた。冒頭でも述べたように認知的な視点を導入した音楽構造解析理論は他にも存在するが、楽譜および音符として表現される情報が中心的な解析対象である。演奏パラメータレベルで、グループ境界認知 (グループ開始アクセント重視型、フレーズ重視型、拍節構造重視型等) をとらえようとするアプローチはほとんどない。

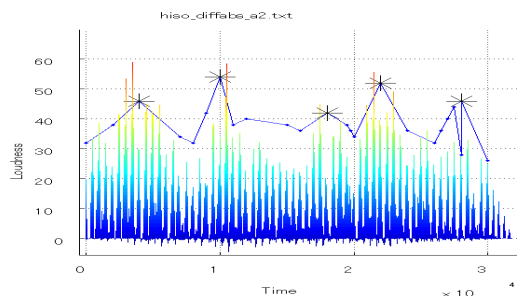


図 11 内声へのアテンションを強めた「悲愴」差分マッピング (時間-音量 平面)

GTTM や IRM は、その理論の提案者の音楽的な内省によって理論が構築されている。本研究では、より客観的な記述を目指し、簡単ではあるが、認知に関する実験を実施し、その上で計算モデルの構築を目指した。ここで実施した知覚実験は内省をサポートするためのものであり、音楽心理学の観点から見れば十分に計画されたものとは言えない。あくまでも興味の中心は計算モデルの作成であり、心理実験に深入りするのは避けることになるだろうが、今後、被験者に対する背景知識の説明の仕方を再検討し、アーティキュレーションが聞き方に及ぼす影響を調査していきたい。

今回、計算モデルに関しては、便宜的に、MIDI イベントデータを、時間周波数表現に展開して、音量輪郭の追跡処理を実装した。我々のモデルは本質的には、音響信号を意識したグループ識別モデルである。今後は、音響信号を対象として、パート追跡、重要音、クレッシェンド、ディミヌエンド抽出に取り組んでいきたい。また、メロディや内声のトラッキング処理に連動したパートに対する動的アテンションコントロール、などを実装していく必要がある。

今まで開発されてきた音楽解釈システムや作・編曲システムのほとんどは、オープンループ系のシステムとして開発されてきた。システムの生成物の選び出しと評価は使用者に任されており支援システムの域を超えるものではない¹⁰⁾。我々は、本論文で述べてきたような聴取モデルを音楽解釈システムや作・編曲システムに組み込み、フィードバック系のシステムとして構成することで、システムの自律性を高めていきたいと考えている。具体的には、聴覚モデルがパラメータの自動チューニング、さらには、生成系で使われるルールのマクロ化などに寄与すると考えている。

6. ま と め

本論文では、ベートーベン悲愴ソナタ第 2 楽章をとりあげ、音楽経験を持った被験者を対象としてグルー

プ知覚にかかわる聴取タイプの分類を行った。その結果に基づき、1) パート毎の重要音、クレッシェンド、ディミヌエンド抽出、2) パートへのアテンションコントロール、3) グループ開始アクセント重視型・フレージング重視型、拍節構造重視型等のスキーマへの選択的注意を基礎とするグループ聴取モデルの初期的な提案を行った。

今後の課題としては、グループ判別の再調査および調査法自体の再検討、メロディや内声のトラッキング処理に連動したパートに対する動的アテンションコントロール、音響信号を対象としての処理モデルの形成、パラメータの自動構成法、脳神経科学に根差した実験とモデル化など、取り組んでいくべきことは山積している。聴取モデルに関する研究は、音楽解釈システムや作・編曲システムなど応用システムを構築する上でも非常に重要である。一歩ずつ、取り組んでいきたい。

謝 辞

本研究は、独立行政法人科学技術振興機構戦略的創造研究推進事業さきがけタイプ「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) Spevak, C., Thom, B. and Hothker, K.: Evaluating Melodic Segmentation, *proc. ICMAI, LNAI*, pp. 168–182 (2002).
- 2) 片寄晴弘: マルチメディア情報学 4 巻「文字と音の情報処理」、岩波書店、音楽情報処理, pp. 163–220 (2000).
- 3) Lerdahl and Jackendoff: *A Generative Theory of Tonal Music*, MIT Press (1983).
- 4) 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).
- 5) Narmour, E.: *The Analysis And Cognition Of Basic Melodic Structures*, the University of Chicago Press (1977).
- 6) 保科洋: 生きた音楽表現へのアプローチ—エネルギー試行に基づく演奏解釈法, 音楽之友社 (1998).
- 7) Cooper, G. and Meyer, L.: *The Rhythmic Structure of Music*, the University of Chicago Press (1960).
- 8) Takeuchi, Y.: Performance Variables for Grouping Structure in two editions of the theme of K. 331, *Proc. ICAD-Rencon*, pp. 47–50 (2002).
- 9) James O. Pickles (著), 堀川順生, 矢島幸雄 (共訳): 聴覚生理学, 二瓶社 (1995).
- 10) 片寄晴弘: デザイン情報学入門, 日本規格協会, デザインと情報処理, pp. 161–183 (2000).