

## LG-006 音楽聴取に関する一検討：グループと頂点の推定

## A Study of Musical Listening: Estimation of a Group and an Apex

橋田 光代<sup>†</sup> 片寄 晴弘<sup>‡</sup> 野池 賢二<sup>‡</sup> 保科 洋<sup>¶</sup> 河原 英紀<sup>†</sup>

Mitsuyo Hashida Haruhiro Katayose Kenzi Noike Hiroshi Hoshina Hideki Kawahara

## 1. はじめに

人間は音楽をどのように聴き、理解しているだろうか。このことは、音楽を提示する作曲家や演奏家らにとって最大級の関心事であるとともに、心のはたらきを研究する人間にとっても、重要、かつ、非常に興味深い研究対象である。

普段、耳にする演奏に対し、我々は、フレーズや楽曲構造を見いだすことが出来る。クラシックやポップスのソロ演奏では、指揮者あるいは演奏者が、自身が理解した楽曲構造をより明確化することを目的に演奏表現を与える。そこでの微妙な演奏表現が、聴取者にフレーズや楽曲構造を認知させる要因になっていると考えられる。一方、ミニマルミュージックに代表されるような単調・機械的な演奏に対しても、我々は、フレーズや楽曲構造を見出すことができる。つまり、シンボリックな音符の並び自体にグループを感じさせる要因が含まれていることになる。

本稿では、これら2つの事項に着目し、音楽聴取、特に、グループ認知の計算モデルについて検証を行っている。その際（人間のための）演奏解釈理論として定評のある、保科理論<sup>[1]</sup>を題材とし、聴取実験に基づくメタモデルの検討、理論の定式化を目指す。

以下、第2章では保科理論の概要を紹介し、第3章では、表情のついた演奏の聴取実験に基づいて、グループ認知モデルの考察を行う。第4章で、定式化の一つの事例として、グループ中の頂点推定アルゴリズムを示す。第5章で、音楽解釈システムへの応用と課題について検証する。

## 2. 保科理論

保科理論は、演奏者が音楽的なフレーズ表現を行うために、どのようなことを考え、実施していくべきかを実践的に解説した演奏解釈理論であり、直観的でわかりやすい理論として定評がある。音楽的なエネルギーの流れを、放物線運動や振り子運動に例えたシンプルな概念で説明している。

## 2.1 基本的な考え方

フレーズ表現の基本は、一般にも広く知られているように、フレーズの核となる頂点に向かってだんだん音を強め、頂点を越えた後はだんだん弱くさせることである。この表現を与える対象は、音楽的なまとまりを表すグループであり、独立したひとつのグループにはひとつの頂点<sup>||</sup>が存在する。保科理論では、フレーズとは、1つあるいは複数連なったグループに、図1に示すような階

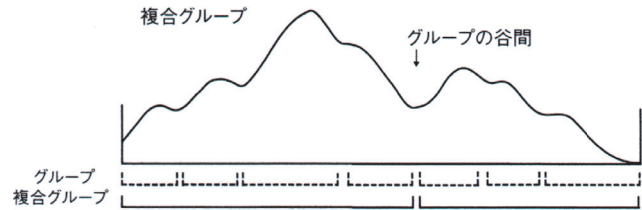


図1: 保科理論によるグループ表現

層的な演奏表現を与えられたものである。

このような表現を行うためにもっとも重要となるのが、楽譜からグループを見つけ、それに頂点を与えることである。どの音が頂点となるかは、楽曲分析を行い、音符の並びや和声、拍節、アーティキュレーションなどの組み合わせで決定する。分析の対象自体はIRM<sup>[2]</sup>やGTMM<sup>[3]</sup>をはじめとする従来研究と大きな違いはないが、保科は、頂点としてより強い条件を持つ音型の導出に重点を置いている。たとえば、倚音は頂点として強い条件を持つが、順次進行する音型と後続音が反行する音型とでは後者の方がさら強い条件を持つ。さらにそれが拍節的に強勢の位置であれば、ほぼ確実に頂点になる。

## 2.2 「悲愴」の解説

本章冒頭で述べたように、保科理論は実践的であり、文献<sup>[1]</sup>の中で楽曲の具体的な分析例および表現例を挙げている。この節では、ベートーベンのピアノソナタ「悲愴」第2楽章の保科の解説と演奏例について紹介する。

図2に、冒頭8小節の表現と文献<sup>[1]</sup>で所収されている演奏データを示す。この8小節には全部で6つのグループがあり、そのうち3~4小節目の2つは複合的なグループとしてまとめられ、大きく5つのグループに分かれる。

フレーズ表現は各声部ともに与えられるが、とりわけ、終始十六分音符が続く内声が全体の演奏表現の詳細を表出している。特に、5小節目（第2拍の旋律を除く）、6小節目のグループは、旋律、バスともに音符自体が四分音符しかなく、特にクレッシェンドを表現するためには内声が重要な役割を持つ。これは逆に言うと、ここでは内声のクレッシェンドを聴き取れなければ、このグループでのフレーズ表現、すなわち頂点（5小節目のB $\flat$ 、6小節目のA $\flat$ ）に到達するために必要なエネルギーを感じることができず、結果としてこのグループを認知することが難しくなることが予想される。

## 3. 聴取実験

ベートーベンのピアノソナタ「悲愴」第2楽章冒頭8小節を題材に、グループ聴取傾向に関する調査実験を

\*\*文献<sup>[1]</sup>では、この音を重心と呼んでいる。

<sup>†</sup>和歌山大学システム工学研究科

<sup>‡</sup>科学技術振興機構さきがけ研究 21「協調と制御」領域

<sup>¶</sup>関西学院大学心理学研究科、理工学部情報科学科

<sup>||</sup>兵庫教育大学

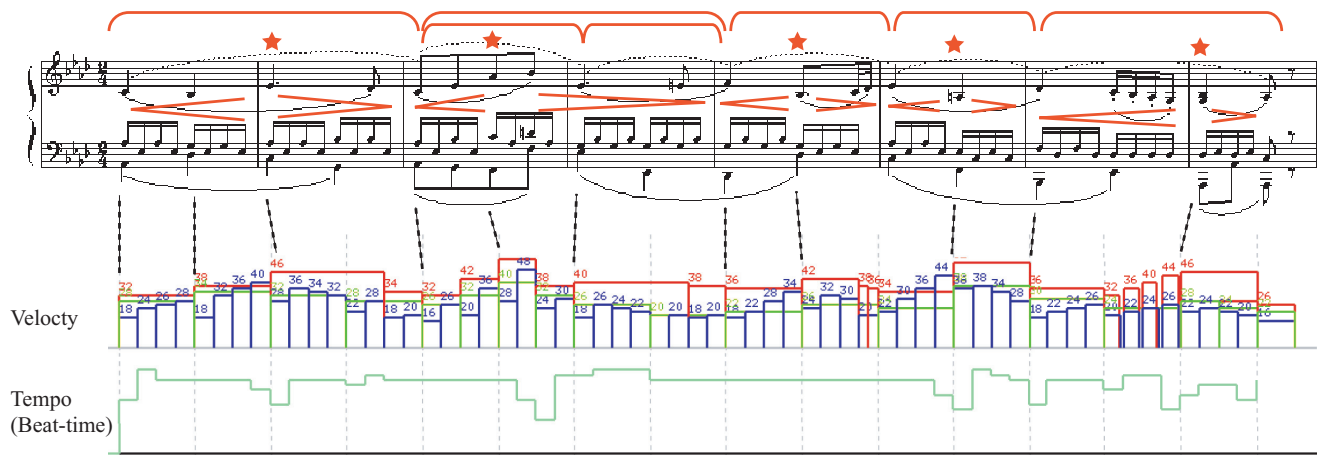


図 2: 「悲愴」演奏データとその表現意図 (文献 [1] 所収)

表 1: 演奏 1) でグループ境界を感じた場所: 被験者 39 名. 下記に示した音の手に境界があったことを示す

場所	グループ境界	人数
4 小節目	E $\flat$	31
5 小節目	B $\flat$	4
6 小節目	A $\flat$	9
7 小節目	D $\flat$	20

表 2: 演奏 1) で感じた強調音

場所	強調音	人数
4 小節目	E $\flat$	3
5 小節目	B $\flat$	4
6 小節目	A $\flat$	36

行った. 大学生 39 人に, MIDI で作成した演奏を聞かせ, 自分が感じたグループ境界の位置と, 音楽的に強調されたと感じた音 (以降, 強調音と呼ぶ) について記入してもらった. 合わせて, 被験者には, 「悲愴」に関する知識や音楽経験, 好きな音楽ジャンル, 普段の音楽環境などに関するアンケート調査を実施した. 被験者には, あらかじめ, ごく基本的なフレーズ表現についてのみ解説を行い, 後述の場合を除いて楽曲や (1) の演奏意図についての解説は特に行わなかった.

演奏データには, (1) 文献 [1] に所収されたデータ, (2) (1) の内声を和音に置き換えて旋律と同時に発音させたもの, (3) 表情なしの 3 種類を用意した. (1) は, 文献の中で行われた楽曲分析に基づいて保科自身が演奏表情をつけたもので, すべての音符の Velocity と, 十六分音符単位でのテンポが示されている. 演奏順は, (2)(1)(3) とした.

聴取実験の結果を表 1, 2 に示す. 4 小節目から 6 小節目のグループについては, 演奏データが意図しているグループと異なる位置に境界を感じた被験者が半数を超える結果となった.

ここで, 6 小節目の A $\flat$  手前に着目する. 第 2.2 節でも述べたように, 保科は, この場所を, 内声のクレッシェンドがフレーズ表現として重要な役割を担う, つまり, グループとしては切れないと解説している. しかし, 演奏 (1) において, この場所を境界と感じた被験者が 9 人いた. この被験者らは, 内声に留意した聴き方をしていなかった可能性がある. そこで, フレーズは, メロディだけではなく内声によってもサポートされることを解説し, 同じ演奏データを聴いてもらった. すると, 9 人中 8 人がこの場所に境界を感じなくなった. この結果は, 声部への注意 (以降, アテンションコントロールと呼ぶ) が音楽聴取に大きな影響を与えていることを示している.

逆に, A $\flat$  手前をグループ切れ目と感じなかった被験者が 30 人いる. この 30 人に対し, 実験 (2) において, 同様に, A $\flat$  手前をグループ切れ目と感じたか否かを調べてみた. この結果, 5 人は, A $\flat$  手前をグループの切れ目と感じていることが確認された. 実験 (2) では, 内声でのフレーズ表現がなされていない. つまり, この 5 人は内声を含め, トータルな音楽表現を聴いている可能性が高い.

そこで, この 5 人と他の 15 人, さらに, 上記の 8 人の 3 つのグループに対し, 音楽経験, 音楽指向にどのような傾向があるかを調べてみたこの結果を, 図 3 に示す. ここで実施した評価は順位尺度に基づくものであり, 距離尺度の絶対性は, 必ずしも保証されるものではない. しかしながら, 「BGM として音楽を聴く」「静かに聴きたい」「一人で聴きたい」「集中して聴きたい」「期待感を持って聴く」などのプリファレンスに関し, トータルな音楽表現を聴いていると思われる 5 人と, それ以外の被験者の間で, 明確な差が確認された.

以上の実験により, 音楽グループ識別は人によってかなりの幅があること, アテンションコントロールが聴き方に影響を与えること, 音楽グループの聴き方には, その人が普段どのような態度で音楽を聴いているか, が影響を与えていることが確かめられた. 紙面の制約上, ここでは割愛するが, グループ境界と強調音の識別には興

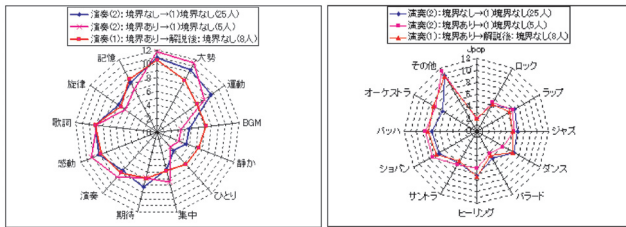


図 3: アンケート結果：軸は順位（内側が上位）を表す

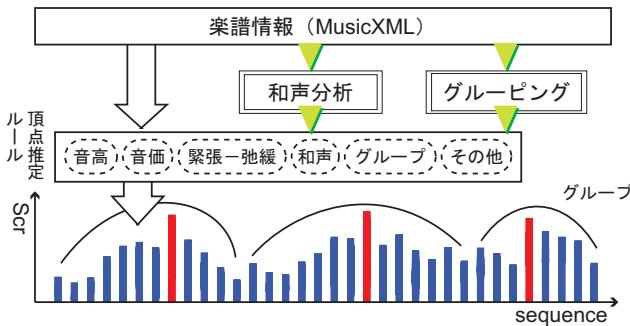


図 4: 頂点推定の流れ

味深い相関が見いだされている [4].

#### 4. 頂点推定

本稿では、演奏時にフレーズ頂点とする音を推定する手法を構築するために、保科理論の定式化を進める。定式化にあたって、1) 保科理論で不足する推定条件や定義を追加し（以降、頂点推定ルールと呼ぶ）、2) それらに音楽的内観に沿う形で一定の評価ポイントを設けた。

##### 4.1 手順

頂点推定のプロセスを図4に示す。入力する楽譜情報は Finale で入力編集した MusicXML 形式ファイルである。おもに旋律部分を対象とし、楽譜情報には音高と音価のほか、基本的な楽曲形態（調号、小節、拍子など）や、一般的に広く用いられる演奏指示記号（スラー、アクセント、テヌート、スタッカート）が含まれる。ただし、*f* や *mp*、クレッシェンドのように、強弱変化を直接的に指示する強弱記号は除外した<sup>†</sup>。推定に先立ち、本稿では、あらかじめ手作業で和声分析の結果とグループ構造を与えておく。

実装にはプロダクションシステムの形式を取る。あるグループ  $G_k$  が音符列  $N_1, N_2, \dots, N_x$  を構成するとき、頂点  $Apex_k$  は、 $G_k$  に含まれる 1 音  $N_n$  について次式を満たす。

$$Apex_k = \arg \max_{n=1}^x Scr(N_n)$$

このとき、 $N_1$  は  $G_k$  の開始音、 $N_x$  は終了音を表

<sup>†</sup>これらの記号は、保科理論では、頂点を導くための直接的な根拠としては扱わない。音型から選ばれた頂点との関係から、最終的にグループ全体をどのように演奏するかを決定する参考として利用する。

表 3: 頂点推定ルールと評価ポイント

分類	Rule	Nnが各Ruleにおいて以下に該当するときの評価ポイント(w)			
		先行音	対象音	後続音	
音価	隣接する2音の第1音		後続音より短い	0	1
			" 長い	1	0
	同一音価が連続する音群の第1音			1	
	隣接する2音の第1音		後続音より低い	0	1
			" 高い	1	0
音高	隣接する2音の第1音		後続音より低い	0	1
			" 高い	1	0
	同一音価が連続する音群の第2音以降			発音順/音符数	
	進行到達音 (隣接する4音の第3音)	1) 上行-上行-下行	0	1	0
		2) 下行-上行-下行	2	1	0
旋律的 緊張-弛緩	倚音	音価が後続音より長い		3	
		" 後続音と同じ		2	
		" 後続音より短い		1	
		後続音が反行する		2	
		順次進行の途中		1	
	掛留音	後続音へ跳躍進行する		2	
		" 順次振興する		1	
		予備音が休符		1	
		予備音がタイあり	3	0	
		" なし	2	1	
	持続音	予備音がタイあり	2	0	
		" なし	1	0	
	刺繍音	上行して下行		2	
		下行して上行		1	
	経過音	上行形		2	
下行形			1		
	その他変化音 (臨時記号)		1		
音積 (和声)	和音	I		0	
		I6		1	
		I46		3	
		II		6	
		IV		1	
		V		2	
	借用和音	V7		3	
		VI		1	
		変化音である		1	
		" ではない		2	
先行和音から変化	カデンツ	V(7)-I	4	0	
	(隣接する2音の第1音)	V(7)-I6	2	3	
		V(7)-VI	2	1	
	音域	先行音より広がる		1	
	" 狭まる		-1		
	旋律密度	重音		1	
グループ	開始音			1	
	終了音			-1	
	最長音	グループにバウンド分割を含む		1	
		" 含まない		2	
	最高音	グループの平均音価より長く、0.25秒以上		2	
	その他		1		
	グループ中最大の上行跳躍音	グループの平均音価より長く、0.25秒以上	0	2	
	その他		2	0	
その他	アーティキュレーション	アクセント		1	
		テヌート		1	
		非装飾音		1	
	スラー (演奏時間1秒程度以下)	スタッカート		1	
		開始音		1	
	終了音		-1		
休符	強拍にあり、後続音が弱拍である		0	1	
	弱拍にあり、後続音が強拍である	1		0	

す。Scr( $N_n$ ) は、次式のように、 $m$  個の頂点推定ルール Rule( $N_n$ ) を適用した結果の合計を表す。

$$Scr(N_n) = \sum_j^m w_j * Rule_j(N_n)$$

頂点推定ルール Rule $_j$ ( $N_n$ ) は、 $N_n$  が頂点の根拠となる条件を満たすかどうかを判定する個々の関数である。

$$Rule_j(N_n) = \begin{cases} 1 & N_n \text{ is matched to Rule}_j \\ 0 & \text{else} \end{cases}$$

表3に示すように、頂点推定ルールは大きく6つに分類される。ドミナント・モーションなど連続する音の表現にも対応できるよう、各Ruleでは、該当する音の先行音や後続音に対しても評価が与えられる。

$w_j$  は Rule $_j$  の評価ポイントを表す重み係数である。ここでは、特に楽曲様式を意識しない一般的な演奏を想定

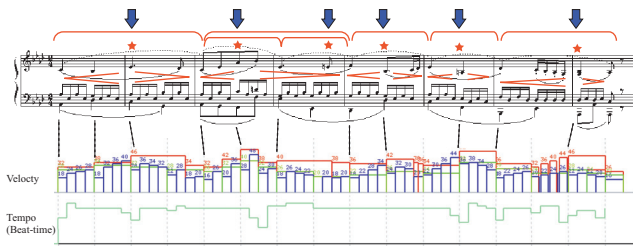


図 5: 演奏データの表現意図と推定された頂点

し、保科理論および数曲の譜例による予備調査を通して、理想に近い頂点推定ができる組み合わせを設定する。その際、各ポイントの数値が、音楽的な意味、あるいは演奏者の一般的な内省と照らし合わせても不自然でないよう配慮した。

#### 4.2 アルゴリズムの評価

ベートーヴェンのピアノソナタ「悲愴」第2楽章を題材に、図2で保科が自身で行った分析を正解と仮定し、人間の分析と同様の頂点を推定できるか確認した。

図5に上部の矢印は推定された頂点、\*印は演奏者が意図した頂点保科が分析によって与えた頂点(\*印)と一致することが確認された。

### 5. 考察

我々は、グルーピングには、生得的な知覚基準に加え、音楽的環境によって聴取者の中に後天的に獲得されるスキーマが大きく関係していると考えられる。

今回特に注目したのは、6小節目の内声に対するアテンションの持ち方である。文献[1]が意図したグループでの演奏表現を聴き取るかどうかは、聴取者が内声を聴くという意識を持つかどうかによる。演奏(1)においてA<sub>1</sub>手前でグループが切れると聴いた被験者8人は、この部分の解説を受け、新しい聴き方を学んだことで、グループが切れないと感じるようになったと考えられる。また、演奏(2)ではグループが切れると答えたが、演奏(1)で切れないと答えた5人は、内声を聴くという習慣がすでに身につけている者である可能性が高い。日常的に内声へのアテンションが習慣づけられているのは、主に西洋クラシック音楽の系統に慣れた者が多いと考えられるが、今回の聴取実験では、それを明示的に示唆するような結果は得られなかった。代わりに、普段どのような態度で音楽に接しているかという点に有意な差が出た。被験者数を増やし、よりはっきりと差が現れれば興味深い知見となるであろう。

表3で示した頂点推定ルールは、保科が指摘した頂点の条件を整理したものであると同時に、どのような音が、音楽的に強い緊張を持っているかを示したものである。これらの音や音型を追っていく過程は、GTTMのプロロンゲーション・リダクションにおけるツリー記述の過程にも共通するところがある。

我々は、得られた知見の検証を目的を兼ねて、ここで述べたモデルを計算機による演奏の表情付けに応用している。具体的には、定式化した保科理論とプロロンゲ-

ーション・リダクション、さらに、アテンションコントロールを考慮した演奏生成モデルを構築した。このモデルを用い、ショパンの「幻想即興曲」の中間部に対して、次のような手順で表情付けを行った。(1)メロディと伴奏に対し、それぞれにプロロンゲーション・ツリーを記述する。GTTMでは、ホモフォニーに対してひとつのツリーを記述するが、ここでは、伴奏パートにGTTMどおりのツリーを与え、メロディパートには、単旋律としてのトナリティを重視したツリーを与える。これは、第4章での頂点推定のプロセスにおいて、和声に関する頂点推定ルール(I~VI)は単旋律から読み取れるものを適用したものに相当する。(2)発音時間を軸にして、各ツリーの幹が一致する場所で音の縦の並びが揃うように定めた上で、ツリーの緊張-弛緩関係に合わせて各声部に一定の強弱・テンポ変化を与える。この演奏を、2004年6月に浜松で行われたNIME04 Rencon Workshop\*での演奏聴き比べコンテストにエントリーしたところ、一般投票で「演奏の自然らしさ」という観点で高い評価を得、規定部門での優勝作品となった。

### 6. まとめ

本稿では、人間のための演奏解釈理論として定評のある、保科理論を取り上げ、ベートーヴェンのピアノソナタ「悲愴」第2楽章を題材とした聴取実験、ならびに、音楽グループ中の頂点推定アルゴリズムを示した。聴取実験では、音楽グループ識別は人によってかなりの幅があること、アテンションコントロールが聴き方に影響を与えること、音楽グループの聴き方には、その人が普段どのような態度で音楽を聴いているか、が影響を与えていることが確認された。また、頂点推定アルゴリズムの実装と合わせ、情緒豊かな演奏生成モデルのプロトタイプを作成した。そのモデルを用いて作成した演奏は、本年度実施されたNIME Renconにおいて最も高い評価を受けた。音楽聴取に関する計算モデルの研究は、シーズとしても大きな可能性を持っているといえよう。

我々の研究は始まったばかりである。今後、さらなる被験者実験の実施と評価、計算モデルの構築と応用に取り組んでいきたい。

### 参考文献

- [1] 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).
- [2] Narmour and E.: *The Analysis And Cognition Of Basic Melodic Structures*, the University of Chicago Press (1977).
- [3] Lerdahl and Jackendoff: *A Generative Theory of Tonal Music*, MIT Press (1983).
- [4] 片寄晴弘, 橋田光代, 野池賢二: 演奏上での頂点とグループ境界の聴取モデルについて, 情報処理学会研究報告音楽情報科学 2003-MUS-52 (2003).

\*演奏表情付けに関する評価研究の一環として、2002年よりシステムで生成した演奏の聴き比べコンテストを行い、一般参加者の投票で順位を決めている。http://shouchan.ei.tuat.ac.jp/~rencon/