

Title	演奏家の経験的知識を用いた音楽構造分析
Author(s)	守山, 算哉
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/617
Rights	
Description	Supervisor:小長谷 明彦, 知識科学研究科, 修士

修 士 論 文

演奏家の経験的知識を用いた音楽構造分析

指導教官 小長谷明彦

北陸先端科学技術大学院大学
知識科学研究科知識システム基礎学専攻

守山 算哉

2000 年 3 月

要 旨

音楽構造分析は認知心理学的な視点を活かした研究が多い。しかし認知心理学的な視点に立つ研究は計算機での実装が難しい。計算幾何学的な視点と認知心理学的な視点、それに音楽的に妥当な視点を取り入れて統合モデルを生成する事が求められている。

そのような視点に立つ先行研究をふまえて、多声音楽の構造分析を試みる。本研究はその先駆けとなる声部分離に注目している。多声音楽は複数の声部が独立に旋律を並列に奏で、かつそれぞれが緻密に組み合わさっている。その構造はバロック様式を思い起こすような巧妙さで、この構造にある種の価値がある事は疑いない。音楽構造の研究が難しい理由の一つに音楽の何をどう聴いてどう感じているかがわかりにくく、また客観的に評価できないことがあげられる。音楽活動は人間の生活と強く結びつき、まるで普遍的な物がないように思われる。また、音楽は個人の興味のおもむくまま聴くことに反対する必要はなく、個人の評価を否定する事はできない。そんな音楽ではあるが、ある種の音楽には特別な価値が存在すると考えられる。

クラシカル音楽では、その構造に価値を見出し、それを知らない物でも十分に感情を揺すぶられるが、それを知っている物には、感情を越えた世界を知る権利が与えられる。これは何世紀にもわたる聴き手による自然選択に裏付けられている。我々は対象をこの種の音楽に絞り、その深層を探るための構造分析モデルを作るために、文献調査を行い、それをもとに多声音楽の声部分離を試みる。多声音楽ではそれぞれの声部の存在を認知する事がそれ以降の音楽解釈に意味がある。である。聴き手の振るまいに似せるために、経験による事前学習による知識ベース、音のゲシュタルト的認知、記憶機構を検討しモデルを作成した。次にそのモデルを使って J.S.Bach の作品を分析し検討した。この検討によると提案モデルによって認識可能な特徴と認識不可能な特徴がある事がわかった。認識不可能な特徴はどのようにして克服すべきかの検討を行い、来世紀に向けた課題を提言した。

目次

1	はじめに	1
1.1	本研究にいたるまでの背景	1
1.2	研究対象に蔓延る問題	1
1.3	音楽構造分析と多声音楽	2
1.4	本論文の目的	4
1.5	本論文の構成	5
2	音楽の聴き方	6
2.1	聴き手の信念・感情	6
2.2	クラシカル音楽の価値	6
2.3	作曲家・演奏家・聴き手	7
2.4	構造という考え方	8
2.5	情動という考え方	9
2.6	ゆらぎという考え方	9
2.7	聴き手の音楽的能力の習得について	10
2.8	外的要因と内的要因	10
2.8.1	経験	11
2.9	構造に価値を見出す聴き方	12
2.9.1	経験は生得的に習得できるのか？	12
3	音楽構造と分析	14
3.1	先行研究	14
3.1.1	音楽構造分析の先行研究	14
3.1.2	認知心理学的アプローチと情報工学的アプローチと音楽学的アプローチ	15
3.2	音楽構造を研究するという事 — 3つの仕事 —	16
3.3	時間という物	16

3.3.1	現在はどこか？	17
3.4	複数の視点を持つ音楽構造	18
3.5	多声音楽に注視する	19
3.6	J.S.Bach と音楽構造	19
3.6.1	Johann Sebastian Bach (J.S.Bach) の生い立ち	19
3.6.2	対象とする作品、平均律クラヴィーア曲集第2巻	19
3.6.3	対象とする楽譜	20
3.6.4	音楽家の感覚	20
4	声部分離モデルの構築	22
4.1	モデルにする項目としない項目	23
4.1.1	モデル化すべき物	23
4.1.2	モデル化しない物	25
4.2	旋律のパタン	26
4.2.1	パタンの分類	26
4.2.2	声部を構成するパタンの類似性	27
4.2.3	聴き手の声部分離の仕方	28
4.3	モデルの動き	28
4.4	パラメータの考察	28
5	声部分離実験	29
5.1	実験条件	29
5.1.1	MIDI データ	29
5.1.2	パタンの作成	29
5.1.3	パタンの発見	30
5.1.4	生成される物	31
5.1.5	パタンの長さに関する予備実験	31
5.2	実験結果	31
5.2.1	声部分離の結果	31
5.2.2	正解率	32
5.3	考察	36
5.3.1	楽曲考察の表	36

5.3.2	提案モデルと近傍モデルの比較	37
5.3.3	表の検討	37
6	まとめ	40
6.1	今後の課題	40
6.1.1	モデルの改良について	40
6.1.2	声部分離を越えたモデルの拡張について	44
6.2	結論	45
A	実験環境	48
B	用語説明	49

第 1 章

はじめに

本研究は多声音楽に含まれる独立した声部が並行に存在する構造を計算機上で把握させようと試みる音楽情報科学分野の音楽構造分析に関わる研究である。本研究の成果は自動演奏、演奏支援などの領域に適用出来る。本論文では声部分離プロセスのモデル化を検討し、J.S.Bach の平均律クラヴィーア曲集第 2 巻で実験をした。本研究をはじめるとあたり音楽における知とは何かという根元的な興味を充足する事を目指す。音楽における知を体系化する事は音楽を計算機で表現可能にする上でも効率的な音楽教育を考える上でも真に有意義である。

1.1 本研究にいたるまでの背景

演奏家は演奏家としての能力を身につけるために形式化されていない知識を経験的に習得する。音楽的知識の習得は多くの演奏家の目標であり、習得を促すのは指導者の役割であるが、どういう経過で音楽的知識が習得されているかはまだわからない所が多い。演奏家は演奏時に音楽を解釈しているが、その時に音楽構造が重要な役割を担うと言われている。

この音楽構造の把握を形式的に行えれば、音楽教育、計算機による自動演奏・自動伴奏・自動作曲システムにも役立つ。これらのシステムは複合的な成果で楽譜認識、音楽構造分析、演奏表現生成等の基礎技術が組み合わさって実現される。音楽構造の形式化によって、より高度に知的な社会文化が育成され、ゆとりある社会の実現が期待できる。

1.2 研究対象に蔓延る問題

音楽構造分析の研究が難しいのには大きく 2 つの理由がある。

1. 聴き手の聴き方が一意ではないこと。

2. 音楽の構造は複数の視点を持つこと。

1つ目の問題の詳細については第2章「音楽の聴き方」で述べる。一般に音楽の聴き方は一意だと思われているが、実際には様々な聴き方があり、お互いに聴き方に関しては意志疎通は難しい。音楽構造に関しては理解する事を試みる聴き手と興味を示さない聴き手がいる。しかし聴き方に全くルールがない訳ではない。我々はその中から特定の聴き方に対象を絞る。後に説明するが、ここでは作品毎の価値はその作品中に内在すると仮定して、作品の構造に音楽的価値を見出す聴き方に対象を限定する。これを前提に2つ目の問題に取り組む。

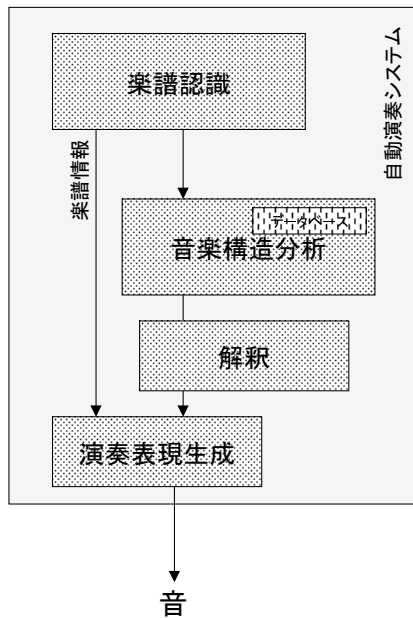
2つ目の問題の詳細については第3章「音楽構造と分析」で述べる。音楽構造はプリズムのように見る視点によって変化する。これは人は一回の認知では有限の認知能力しか持たない事に関わっている。この問題は音楽を現実社会から切り離し、楽譜に表記されている情報で閉じたシステムと考える場合でも生じる。

1.3 音楽構造分析と多声音楽

これまでの先行研究での音楽構造分析の位置づけを図1.1に示す。図1.1は音楽構造分析の研究成果が実際に音楽情報科学分野でどのように活かされるかを示している。計算機による自動演奏システム・自動伴奏システム、もしくは自動作曲システムは、楽譜の認識部、楽曲の分析部、楽曲の解釈部、演奏再現部の複数の部分から構成される。音楽構造と演奏表現には関係性がある事は実証的にわかっているので、演奏再現を行うために構造の解釈が必要になる。音楽構造分析部は演奏家の経験的知識を知識データベースとして持つと言える。

音楽構造分析の概観を図1.2に示す。図1.2は音楽構造分析を行う場合に複数の要因が緻密に絡んで構成されている事を示している。要因は図に示した以外にも数え上げればいくらでも出てくる。これらの要因は相互に深い関連性がある。よって、これらの要因の一つだけを取り上げて楽曲の全てを語る事は不可能である。従って、楽曲から取り出したい視点を考えてから、必要な面だけを取り出す方法を適宜検討する必要がある。楽曲から取り出したい視点は楽曲毎に違っている。その視点は楽曲の本質であり、本質は楽曲毎に異なるからである。図1.2右下に表している「J.S.Bachの音楽構造分析に向けて」が本研究で取り出す視点を反映すると思われる要因である。「心」は心理学的アプローチ、「音」は音楽的アプローチ、「情」は情報工学的アプローチが必要な事を意味する。それぞれの説明は3.1.2で行う。

音楽構造を形式的に明らかにする取り組みには情報工学的知識や認知心理学的知識、当然音楽学的知識、音楽的経験が求められる。本研究では多声音楽の構造分析に取り組む。多声音楽の構



自動演奏システムの内部に音楽構造分析が組み込まれる。自動演奏システムは

- ・楽譜・演奏の認識
- ・音楽構造分析・解釈
- ・演奏表現生成

からなる。音楽構造分析によってある特徴を持つ作品群のデータベースが構築されればその作品群にまとまる新たな作品を生成する作曲システムが構築出来ると言われている。

図 1.1: 構造分析の位置

造分析はその音楽の分かりにくい事もあり、音楽学的分析はあれども、情報工学的な分析ほとんど知られていない。多声構造は聴き手が声部を明確に分離し、かつ融合する等の特徴があり、そこには音楽的価値が多く含まれると考えられる。

多声音楽の特徴を図 1.3 に示す。多声音楽はある主題や応答が組み合わせあって楽曲が構成される。本研究では主題や応答は情報工学的にパターンと見なす。パターンはある決まった順序で出現したり、平行に出現したり、巧みに変異をしたり（ミューテーション）、少しずつ重なったりする。さらに重なって2声と同時に鳴り響く場合に、その2声はばらばらに聞こえるのではなく、さらにまとまって新たな声部が認識出来る場合もある。本研究で目指す所はこのような構造を計算機が自動的に把握する事である。

音楽構造の理解には経験が必要と仮定する。多声構造の理解にはまず独立した声部を事前に把握する事が前提であると仮定し、これをモデル化し実験を行う。モデル化には認知的な特徴を中心に含むが、生得的な特徴だけでなく、転回構造のような物はイディオムの知識（形式知）として事前に理解しているとし、それがなくては理解ができないと仮定する。音楽作品には色々な

音楽構造分析の概略

音楽構造分析の色々

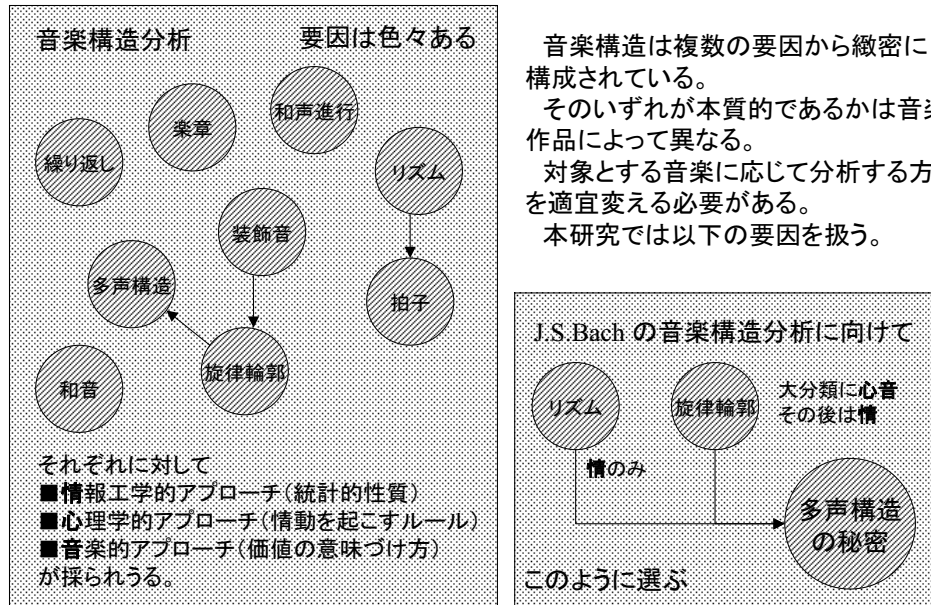


図 1.2: 音楽構造分析の色々

要素があり不変的なモデルを構築するのは困難なので特定の楽曲に対象を絞って検討する。

正しい音楽構造は人間が与えることとする。ただし、何が正しい音楽構造かについては、現状では著者の主観が大きく客観的な評価が難しいため本論文では詳しくは論じない。

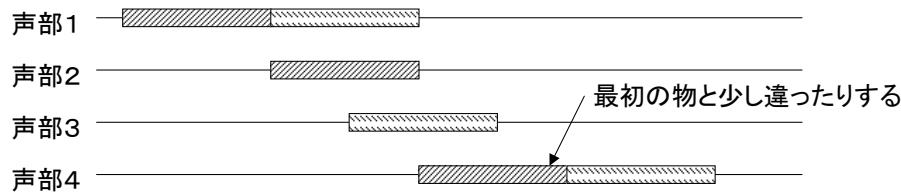
1.4 本論文の目的

本研究の目的は多声音楽の声部を分離させる事にある。分離させる事で計算機が音楽構造を理解する事に一步近づける。

ポリフォニー音楽は多声構造をとるが人が音楽の本質を理解する場合にこの多声構造の理解が不可欠と考える。多声構造には多くのルールがあるが多くの例外もある。計算機上で多声音楽の声部分離を試みる事で音楽構造を分析し音楽の美的価値付けの客観的指標を得る事を目指す。

多声音楽に取り組むの図

声部の関係に潜む価値は何か



多声音楽では上記のように、ある主題や応答(ここではまとめてパターンと言う)が組み合わさっている。パターンは順番に出るだけでなく、巧みに変異したり、ずれて同時に重なったりする。さらに、同時に提示されている部分は2声が鳴り響くにも関わらず、その2声が融合されて別の旋律が聞こえるような工夫もある。本研究の目指す所はこのような構造を計算機が把握する事である。

図 1.3: 多声音楽に取り組む

1.5 本論文の構成

本論文の構成は以下の通りである。はじめに、第2章で個人によって異なる「音楽の聴き方」について述べる。第3章で聴き方は人によって異なる点をふまえて、その中の特定の音楽の聴き方に範囲を絞り(形式)その中での「音楽構造と音楽構造分析」について述べる。ここでは特定の音楽に関わる基礎知識(作曲家の説明・主張、多声音楽の説明)も述べる。第4章でその音楽作品に有効と思われるモデルを提案する。第5章でそれを計算機上で実装し音楽作品を分析、実験結果を示し考察を行う。第6章で今後の課題を述べまとめとする。付録として実験環境と用語説明と使用した楽譜と分析結果のデータを図示した物を掲載する。

第 2 章

音楽の聴き方

この章では、音楽の聴き方には色々な方法があり、全ての人が同じ聴き方をしていないという意見を検討する。

大きく分けると聴き手は二手にわけられる。熱心な素人の音楽愛好家や音楽家のグループとそれ以外のグループで、彼らの目的は異なる。前者にとっての音楽は生活の中心または一部であり、音楽活動や経験が人生の価値を高める活動そのものである。一方、後者に属する人々にとっては音楽は必要不可欠なものではなく、贅沢なものともみなされる。音楽を評価しない人々は、音楽が束の間の楽しみ以外、何の意味も持たないと考えている。つまり必要なものではなく、害のない道楽であると考えている。さらに音楽愛好家の中でも色々ある。

音楽に対する考え方の相違がある。音楽の価値は色々な側面からつける事ができるため、全てを対等に評価する事は一般には難しい。

2.1 聴き手の信念・感情

同じ作品を聴く場合でも聴き手の信念により曲の価値はかわり、聴き手におこる生理的反応もかわる。また同じ聴き手であっても、聴取時にどのような感情下であるかによっても、反応はかわる。[22]

2.2 クラシカル音楽の価値

クラシカル音楽はバロックであれば 1680 年から以降数百年間、多くの人に支えられて発展してきた。クラシカル音楽の魅力は多いが、最大の特徴はその豊かな響き、複雑であり、尚かつ統一感の感じられる構造にあると言える。これは古代バロック建築を一目見て美しいと思うのに似

音楽の聴き方について

みんな同じように聴いているのか？

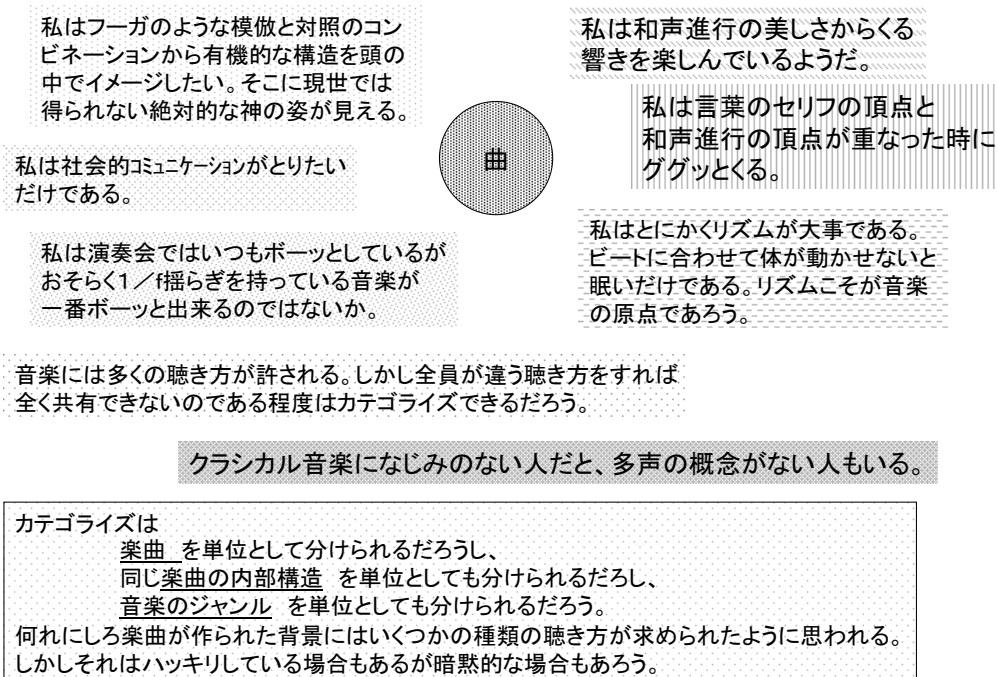


図 2.1: 音楽の聴き方

ている。クラシカル音楽には後に述べる構造という考え方と情動という考え方のどちらも含まれる。だが伝統的な音楽研究をみてもわかるように、情動よりも音楽作品の大局的構造、和声などの表現手法に価値があると考えられる。

2.3 作曲家・演奏家・聴き手

三者の関係は図 2.2 で示すように、作曲家の意図を演奏家が運び聴き手に伝えると考えることができる。だが、このような関係と割りきってしまう事ができない可能性もある。なぜならば、聴き手が作曲家と同じ意図を理解していなくても音楽の本質に問題はないとすればいいからである [20]。

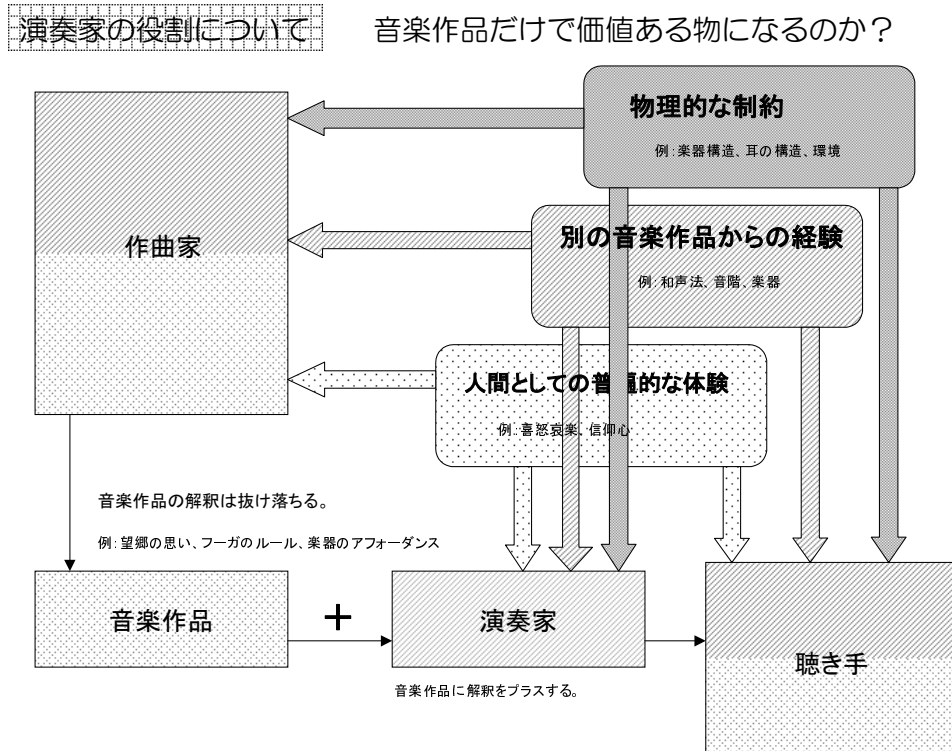


図 2.2: 作曲家・演奏家・聴き手

2.4 構造という考え方

構造に興味を持つ考え方は演奏家でも音楽認知研究者でも共通している。これは言い換えるとシステム論的音楽観と言える。音楽を聴くという行為を注意深く観察すれば音楽の価値は音と音との関連性によって意味づけられる事がわかる。つまり音は単体では意味を持ち得ない。

次に音と音との関連性は同一レベルだけでなく階層的な振る舞いを持つ。この考え方によって小節のような概念や、楽章のような概念が存在する。この事は、古くからソナタ形式が階層構造を持つ事と言われている事からも明らかである。Jackendoff らの [3] 代表的理論でも階層モデルが提案されている。

システム論的音楽観では音と音との関連性を決めるのはシステムである。システムはどうなっているかわからないが和声理論、旋律理論等はシステムの一部である。システムの入力は通常は作曲家が神から受け取ったメッセージであったり、古典派以降ならば靈感であったりする。また鳥の声や電線のうなり DNA の場合もありうる。入力作曲家の外にある場合、作曲家はシステ

ムを作る行為が作曲行為と言える。

出力は伝統的な方法では楽譜であるが、インタラクティブな作品であればリアルタイムに計算機から出力される音響情報の場合もある。つまり、作曲行為において一番重要とみなされている所は、音の関係性を記述するシステム作りであり、それが紛れもなく音楽構造であると言える。聴き手は演奏はしなくとも音と音との関連づけに関しては無意識的にしろ絶えず能動的な作業をする。これが音楽構造を理解する作業と言える。

2.5 情動という考え方

情動理論は最近の流行なので理論的研究、その理論の応用例ともに数多くの研究が見受けられる [7] [19][23]。情動理論は非常に簡単な調性音楽で研究されているが、殆どはホモフォニー音楽を対象にしている。これはポリフォニー音楽が現在はずでに廃れてバロック以前の遺産となっているからと考えられる。しかしバロック以降の音楽にもポリフォニー的な特徴は多く含まれる。ホモフォニー音楽は巧みな和声を使った色彩感あふれる所に特徴があるが、それは感情に対応していて、結局情動理論に結びつくと考えられる。一方ポリフォニー音楽は聴き手は神（洗練された人）としてかれ、直接感情に訴える技法をしているとはいえない。これは情動理論に合致しないと思われる。

2.6 ゆらぎという考え方

音楽に $1/f$ ゆらぎがあるという研究報告がある [21] [14] [31] [16]。その理論の基盤は未だ定かではないが、揺らぎの要素があることは経験的事実として知られている。要はそのゆらぎがどの要素に含まれていて何のためにあるかである。 $1/f$ ゆらぎの存在を主張する研究者は $1/f$ ゆらぎの存在場所について非常に曖昧である。ある時は演奏のそれぞれの音を鳴らすタイミングの正確なリズムに対する誤差であり、ある時は旋律を構成する音列の音程差である。演奏タイミングにかかわる部分にゆらぎの要素が含まれる事は大いに考えられる。

経験的事実を裏付ける可能性のある根拠の一つに人の生体に揺らぎの要素がある事があげられる。生体のゆらぎのために人が演奏をするときには揺れぐ可能性がある。特に人間の呼吸などの影響は大きく受ける可能性がある。呼吸のような生体リズムは音楽の本質的な部分と考える事もある。

しかし旋律を構成する音列に $1/f$ ゆらぎが含まれると考える事はそれほど自明ではない。なぜならば $1/f$ ゆらぎは階層的な構造を持っていなければならず、音程差という離散的な数値をしか

も音と音の間隔で求める場合、膨大なデータを処理しなければ $1/f$ ゆらぎとわからないからである。おそらく音楽の構造に含まれるゆらぎは単一な方向に階層的な $1/f$ ゆらぎではなく、むしろ階層的ではなくとも複数の要素が関係する揺らぎが存在すると思われる。

2.7 聴き手の音楽的能力の習得について

義務教育における音楽教育、音楽家になるための音楽的教育のどちらも副次的にではあれど音楽構造を理解する事を目的にしている [28]。だが音楽は高度な音楽的教育を受けていない聴き手にも楽しむ事ができる。なぜならば聴き手は音楽から任意の視点で自分なりの価値を引き出す事が可能だからである。強いリズム感を基本として身体との動機によって価値を見出す事もできるし、流行歌では社会的コミュニケーションの道具として音楽を価値づける。これらはどれも間違っている訳ではない。聴き手がそれぞれの楽曲にある適切な価値を見出す事が大切である。当然レベルの差はあるが、どのような聴き手でも自分に合わせた価値を見いだせる所が音楽の素晴らしい所である。

これを前提として、さらに考えると、しかし作曲家が楽曲の構造に価値を置いたならば、楽曲の構造に価値を求める聴き方をしなければ聴き手の意図は見えても作曲家の意図は見えてこない。弦楽四重奏にロックバンドなみの迫力を求めるのは筋違いという物である。やはりクラシカル音楽を聴くときにはクラシカル音楽のどのような点が面白く楽しみが見いだせるかを意識的にしろ暗黙的にしろ知って聴くべきである。ストーリーによれば、彼は音楽が万人にとって有意義な事を認めた上で、さらに「確かに感覚的な快楽は音楽体験の一部ではあるが、音楽家や素人の音楽愛好家たちは、偉大な音楽がそうした楽しみ以上のものをもたらすことを知っているのである。」と言っている [9]。我々の興味はこの『楽しみ以上のもの』が何かという点であるといえよう。本研究ではそれはおそらく構造の理解にあるという立場に立つ。

2.8 外的要因と内的要因

図 2.3 と図 2.4 のように音楽作品は楽譜情報だけで成り立っているわけではなく、それを理解するだけの沢山のバックグラウンドが必要な事がわかる。それらの多くはひとまとめにして経験という事が出来る。

音楽構造分析の問題点 どうやって外的要因と分離するか。

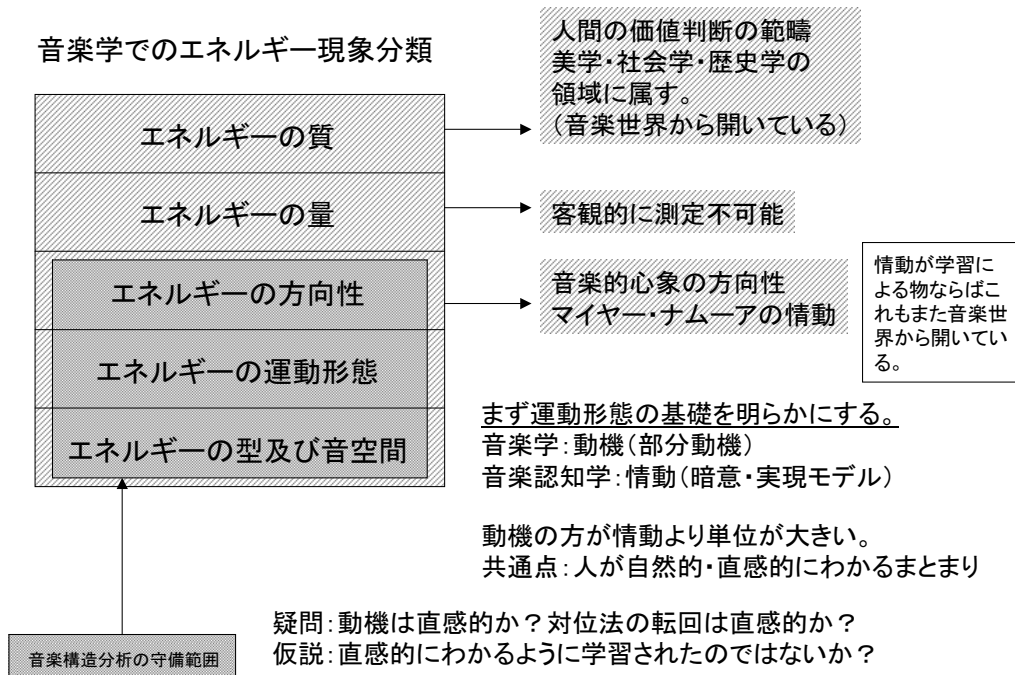


図 2.3: 外的要因と内的要因

2.8.1 経験

経験には

1. 音楽の内部の音の組み合わせ情報を学習する事
2. 共通の感情体験や社会的立場により共通の音楽現象に興味を持つこと

が考えられる。これによって新規に音楽を聴いた時にどのように解釈するかがかわってくる。
さらに現在の状態として

1. 聴いている時にどのような気持ちか
2. 聴いている時はどのような社会的環境にいるか

によっても解釈がかわってくる。

音楽経験について

価値観は経験によって身につくのか？

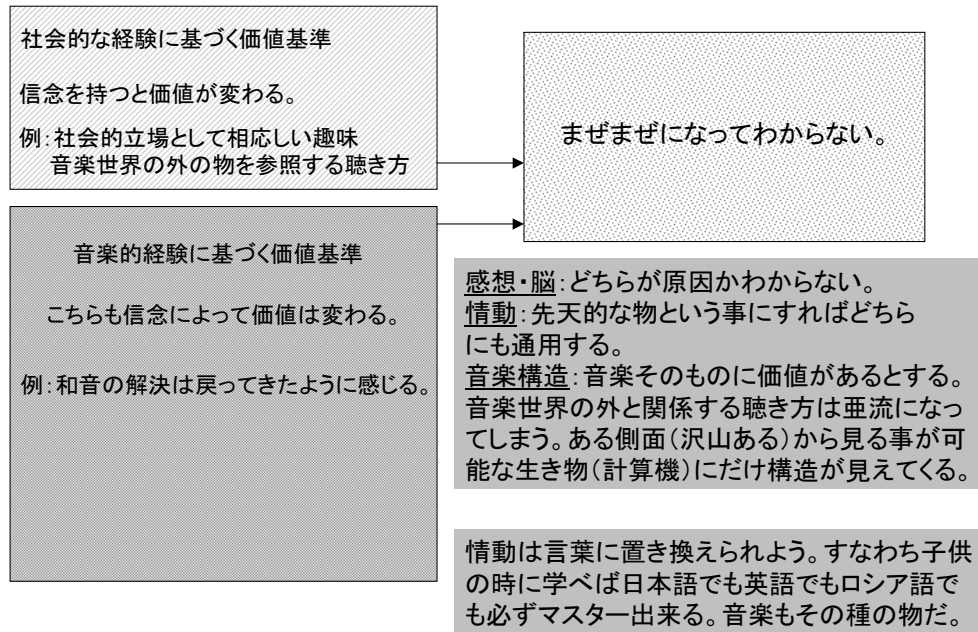


図 2.4: 音楽経験

2.9 構造に価値を見出す聴き方

古来作曲家が構造に注意を払って作曲を行っていたので聴き手がそれにのっとして解釈を行うのは間違っていない。これに反する事は可能だがこれはいわばゲームのためのルールと同じである。現代音楽には、抽象画と同じように聴き手の解釈にゆだねられる幅が大きい作品がある [13]

2.9.1 経験は生得的に習得できるのか？

言語はどのような子供でも自然に学び習得する事が出来る。言語を生得的に持っているのではないが言語を学ぶという事を生得的に持っているのである。従って日本人の子供でもアメリカで育てれば知能的障害がない限り英語をマスター出来る。

音楽にもそのような要素があると考える人がいる。鈴木は誰でも教育すれば演奏家になれる、という信念のもとにヴァイオリンによる才能教育運動を起こした [29]。才能教育の面でも音楽構造の面でも共通する事は、感受性だけで演奏や聴取を行うのではなく、用意されている形式的知

識を先に与える事で効果的に音楽を理解する能力をつける事である。実際問題、クラシカル音楽の高次の構造の存在を多くの人が習得していると考えるのは難しい。対位法の魅力、高度な和声の効果の違い、を普通の人々が習得しているとは言い難い。しかし低水準な和声構造や論理構造は誰でも習得できる。おそらく日常的に音楽にふれることで習得したのだろう。しかしそれ以上を経験のみで習得出来る人は少ない。(例外として一部の黒人ミュージシャン、JAZZ ミュージシャンには結構経験的に音楽を習得した人を見うけられる。これらの音楽の特性として学習すべき規則性が少なく演奏スタイルについても属人性が強いことなどがあげられる)

例えば対位法には転回という技法がある。これは旋律の高低が逆になるが、これは楽譜という二次元空間の理解、音楽作品上でのある構造レベルでの遊びというような見方を習得してからでなければ理解しがたい。従って経験的な知識だけではなく、形式的な知識を必要になってくる。

第 3 章

音楽構造と分析

聴き手は楽曲構造を理解する場合にどのようにして理解しているのだろうか？音楽は時間とともに作られる。ある時点毎に構造を作っているのだろうか？作っているならばどうやって作っているのだろうか？我々はこの地点より過去の知識だけでなく先の知識もわかることを仮定する。経験的に理解するということは今聞いている音楽作品と同じ音楽作品を以前に聴いた事があるとか、同じ作曲家のつまり似た形式の作品を聴いた事があるという事である。つまり先が読める。

殆どの音楽構造分析研究は和声構造やリズム構造に注意が向いている。これは現代のポップスやロック他多くのジャンルでポリフォニー的な形式の影響が少なく、和音を一つの音として感じ取る音楽の聴き方や旋律があったとしても主旋律の休止中に対旋律として使う等、同時性が少ない作品ばかりのために、興味が薄い事が考えられる。しかし、クラシカル音楽にはこのポリフォニーの効果は作品の価値を考える上では欠かせない一要因になっていることは間違いない。本研究の新規性はこの多声音楽に注目している点にある。

3.1 先行研究

音楽情報処理はここ数年多くの人に注目されている。計算機で音楽を扱う事は自動演奏や自動作編曲、及び音楽活動の支援に役立つ。それぞれの研究は互いに深く関わり合っているがその中でも特に本研究に関係すると思われる項目について言及する。

3.1.1 音楽構造分析の先行研究

音楽構造分析で注目される対象をいくつかあげる。

1. 楽音の群化問題
2. パラメータの問題
3. 根拠情報の不足問題
4. 単調性・非線形性の問題
5. フレーズ表現と音楽的グルーピング問題

その実装例として、解釈による学習・適応、コーパスによる構造解析リアルタイム音列予測がされる。

3.1.2 認知心理学的アプローチと情報工学的アプローチと音楽学的アプローチ

曲の構造を理解するモデル作りには大きく分けて認知心理学的アプローチと情報工学的アプローチと音楽学的アプローチの3者がある。

認知心理学的アプローチ 認知心理学的アプローチは音楽の認知過程を科学的に探求しながら音楽を理解しようとするアプローチである [11]。代表的な認知心理学的アプローチによる音楽構造の解析理論には、Meyer の「暗意-実現モデル」の流れをくむ Narmour のクロージャー解析理論 [7] と、Schenker の音楽分析理論の流れをくむ Lerdahl と Jackendoff による生成音楽理論 [3] がある。認知心理学的アプローチは「聴き手の聴取プロセスに則しているか」が鍵になる。これらの理論はかなり精緻に組み立てられているが、理論に例外が全くない訳ではなく、いざ例外が起こった場合には、人の主観によって決める必要があるなど、計算機での実装には問題点がある。

音楽学的アプローチ 音楽学的アプローチは演奏家や聴き手の理解のためにされる。他のアプローチに比べて、経験的であり、計算機での実装を考慮していないので、モデル化しにくい。音楽学的アプローチは最終的な評価の場である、演奏の場で作られた物なので、他のアプローチに比べて正当であると認められる。欠点は、計算機での実装を考慮せず、実装の部分は演奏家の暗黙知に頼っているために、客観的な評価が出来ない。演奏家の暗黙知には楽曲の暗黙知の他にも、音楽全般に関わるものが含まれるため、音楽学からの成果をそのままモデル化する事は非常に難しい。

情報工学的アプローチ 情報工学的アプローチの利点は他のアプローチに比べてより根元的な理論があり、計算機で普遍的に実装できる点である。情報工学的アプローチでは、現象として曲が理解されている訳ではなく、工学的に似ているかどうかを数値で評価するのが基本

である。近年、計算機資源のスピードや記憶容量が発達した事で、大量のデータを用いて問題解決を行うシステムが研究されている。音楽構造の履歴をためておいたコーパスを用いて、曖昧性の問題を解消する計算機モデルを構築する事が目指されている。この分野の流れは、他のアプローチと組み合わせたルールベースから事例ベースへと移行している。現象的に理解しているように見えればよしとするのかどうかはまだ結論は出ていない。

我々は高度な音楽理解のためには現象的な理解、つまり情報工学的アプローチだけでは限界があると考えている。認知心理学的アプローチと音楽学的アプローチの成果を上手く組み合わせ、必要な所にだけ情報工学的アプローチを適用するのが好ましいと考える。

3.2 音楽構造を研究するという事 — 3つの仕事 —

我々は音楽を聴くとき、単に個々の音を聞き流しているわけではなく、そこに様々な音楽構造を知覚している。音楽構造のモデルを作成するには、そのような構造を表現するための知識表現の枠組みを与え、それを操作するための処理機構を定式化し、実現する事が必要となる。音楽の意味は複数の音同士が継時的・同時的に組み合わせられた構造に見出される。その意味は単語のようにあらかじめ規定されておらず、曲の文脈に依存して動的に生じる。音楽構造の認知モデル作成は

1. 音楽的構造を知識表現の問題として定式化する事
2. その構造を構成する要因についての理論を与える事
3. それをアルゴリズム化してプロセス・モデルを作成する事

の3点からなる。本論文では2番目と3番目を横断的に設定し、その上で、実際の実装をおこなう。1番目は直感的に記述したレベルに止まり定式化までは至っていない。

3.3 時間という物

音楽は時間芸術である。決して戻ることの出来ない時間を相手にしているので音楽の振興は一方通行のように感じられるが、時空間を再構成していると考えれば一方通行ではなくなるだろう。時間に沿って飛び込んでくる音を全て記憶すれば後に完全な構造が得られる。しかしおそらく飛び込んでくる音を全て記憶する事は不可能である。何か重要な部分を選び出しそれを人の内部に作り上げたデータベースに照会しながら音楽的解決を行っているだろう。そのときに注目す

判断する地点

現在はどこか？



幅の中の処理は

- 音を1個ずつすぐに分離するのか？ (Serial Single-Choice Model)
- 聴いていく途中で選択枝が複数ある場合に
その地点の決定は保留して先を読むのか？
(Serial Indeterministic Model, Parallel Multiple-Analysis Model)
- この範囲を聞き終わってから分離するのか？

図 3.1: 現在の幅

るのは暗意・実現モデルに代表されるある種の期待である。人はルールによって音の期待をするだろう。未来は未確定である。我々の形成行為はその行為による何らかの結果を見込んでいる。しかも結果はたまたまあるのではなく、未来における結果を目的として捉えている。つまり形成行為は必然的に未来に向けられ、その未来は確定していない。ここでは何か時間を計る単位を用いて時間を作り上げていると思われる。

3.3.1 現在はどこか？

過去・現在・未来という定義はいささか曖昧である。なぜならば人の認知できる現在は音一つ分以上の時間幅を持っているかもしれないからである。

そのように考えるとその中での処理方法がいくつか考えられる。[6] では Serial Single-Choice Model, Serial Indeterministic Model, Parallel Multiple-Analysis Model がある。

Serial Single-Choice Model では一度に一つの解釈を選んで計算する。

Serial Indeterministic Model では一つのピースに対して一度に一つの解釈を選んで計算する。しかし、解釈が分岐するような場面ではその場では決定せずに、進行していくに連れて証拠を集め単一の解釈を割り出す。この処理のアイディアは Marcus(1980) and Fraizer and Foder(1978) の言語パーサに基づいている。このパーサはローカルな部分に対して予備解釈を計算し、全体を通して正しい解釈が見つかるまでは各部分を結合しない。このモデルは厄介なバックトラックを避けることが出来る。しかし決定に際しては、可能な候補群すべてを評価しなければならない。また、音楽的経験から考えられる方法とも適さない。現実には予備解釈できる数はそう多くはないはずである。

Parallel Multiple-Analysis Model は分岐点にさしかかったらプロセスが複数に分岐してそれぞれが解釈を追うというものである。そしてもっともらしさの閾値が一定の値より低くなったものは捨てていく。こうすることにより、作品の終わりまで到達した解釈は作品全体の最も最もらしい解釈となる。このモデルも音楽的経験に合致しない点に問題がある。このモデルによれば5つの解釈の可能性があれば5つの処理を平行し行うが我々は通常このように感じていないと思われる。

おそらく我々は実際には意識しては Serial Single-Choice Model 方式の処理しか出来ない。ただし無意識中では様々な解釈が平行して存在できる。ただそれを感知していないだけである。Jackendoff はこのように考えている。

ここでは Serial Single-Choice Model の方式が有力に思われるが実際には Serial Indeterministic Model の要素が含まれる事も許されると考えられる。なぜならば実際に解釈を施す場合には解釈済みの情報の他に解釈はしていなくとも与えられた複数時点の情報を持っているからである。

3.4 複数の視点を持つ音楽構造

音楽は時間（音価、リズム）、音量（音の強さ）、音色、グルーブ感（ゆらぎ）、高次の構造等が複雑に絡み合っている。そしてその絡まりかたは音楽によって異なる。そのため普遍的なモデルが作りにくい。複数の視点をニューラルネットで自由に選んで模倣モデルを作る研究例がある。さらに感情の表現に構造の把握が重要な事も多くの研究者は認めている。

複数の視点が作品中に存在している。無意識中では複数の解釈が存在し、最も顕著な解釈だけが意識されるという主張がある。Bever, Garrett, and Hurtig(1973)。これは Jackendoff の立場にも近い。複数の視点を持つという考え方は [2] にも見受けられる。

3.5 多声音楽に注視する

音楽の構造には色々なものがあるが、我々は多声構造を持つ多声音楽に注視した。なぜならば、多声構造という特性が複数の視点を持つ考え方に結びつく可能性を持つ事と、多声音楽にこそ構造の価値が埋め込まれていると思ったからである。多声音楽（ポリフォニー音楽）は音楽史上でも、ホモフォニー音楽に先駆けて登場した。その形式は比較的古くからある。

3.6 J.S.Bach と音楽構造

3.6.1 Johann Sebastian Bach (J.S.Bach) の生い立ち

J.S.Bach はドイツの音楽一族の一人で、1685.3.21 アイゼナハに生まれ 1750.7.28 ライプツィヒに没したの作曲家・オルガン奏者である。幼少からルター神学の基礎を学んでおり、後年の彼を理解するうえで重要である。J.S. バッハが活動した 17 世紀末から 18 世紀前半にかけての時代は、音楽史上後期バロックという様式概念の中に包括され、バッハはその頂点を成す完成者と考えられる。バッハとともにバロック時代は終わり、古典派へ通じる新しい流れにはいる。しかし、バッハの生前から既にこのような新しい音楽の動きは後進的なドイツでも浸透していた。16 世紀以来北ドイツの音楽文化を支配したルター正統派信仰に支えられた市民性と、フランスを中心におこった啓蒙的宮廷文化に支えられた新しい音楽観が融合したところに、ドイツにおける後期バロックの特殊性が存在している。生前のバッハは、オルガンやクラヴィアの名手として、また対位法技術者の大家として高く評価された。一方、当時の若い世代には伝統的で重厚な対位法を駆使したバッハの作風に批判的で、作曲家としてのバッハはヘンデル、テレマン、グラウプナー、グラウン兄弟よりも低く評価されがちであった。生前のバッハ批判が物語るように、音楽の様式がバロックから古典派へと移り変わる時代の転換期にあって、バッハは古風な、旧様式の代表とみなされたため、死後は急速に忘れられた。

3.6.2 対象とする作品、平均律クラヴィーア曲集第 2 巻

平均律クラヴィーア曲集第 2 巻は 1742 年にまとめられた。1742 年にまとめられたものの、ある曲はずいぶん前にあったものをそのまま転用した。またある曲は、そうした以前に書かれた作品をさらに改作して組み入れたとか、さらにこのために新たに書き下ろす等のため、具体的に何年に作曲されたかは特定できない。したがって、ある時期のバッハのスタイルというように統一されたものは見られない。比較的若いスタイルの曲もあるし、晩年のスタイルを示しているもの

も見られる。とはいってもフーガの技法のような抽象的な世界にまではいってないと考えられている。ヘルマン・ケラーによれば第2巻に散見されるフーガの特徴として、ホモフォニックな間奏部の挿入によるフーガ形式の弛緩が見られると指摘されている。この作品はクラヴィーアを意図して作曲した物とされている。その結果、2声の動きを主体とする等のクラヴィーア的な構成が見受けられるとも言える。それがすなわち弛緩型と思われる。声楽曲への編曲を考える場合、発声上の都合から困難な所が多くバッハはそれを意図したとは思えない。クラヴィーアが音の進行について制約があることは極度の跳躍と極度の早さをのぞけば殆どない。早さについてはいくらかの指針がある。

3.6.3 対象とする楽譜

本論文では平均律クラヴィーア曲集の中でもロンドン原点版を用いる。その理由は

1. 信頼のおける MIDI データが入手出来るため。
2. 作品構造がアルトニコル版よりも簡単で分析しやすいため。
3. バッハの初期の作曲スタイルを先に知ることが有意義と思われるため。

その中のフーガを用いる。その理由はプレリュードに比べてフーガの方がそれぞれのフーガで形式に統一感があるため。である。さらに計算の都合上、途中で拍子が変わる物は除外してある。拍子が変わる作品については作品を大きくわける仕組みを備えたモデルを作れば対応できる。

3.6.4 音楽家の感覚

音楽家がバッハをわかるプロセスについて興味深い話がある [12]。

- 小林：長いことバッハを勉強して試行錯誤を重ねているうちに、あるとき、自分にとってのバッハはこれだ、と啓示のように頭に浮かんでくる。そういうものだと思いますね。
- 伊達：そうなんです。これは広い幅での音楽体験とその結果生まれる洞察力といってもよい。さらに、バッハを習得する順番として、
- 伊達：とにかく第1段階として、提示部ではテーマの性格、間奏部では特に調性を含めて和音的感觉、そして3番目に、基本的に流れの止まることのないフーガという形の中での音楽的区切りなどの認識が必要なわけでしょう。

過去のピアノ指導者もフーガというものをよくつかんでいなくて、とにかくテーマが大事だからと、それを強調しておけばなんとかなるという教え方が支配していた事もある。フーガの理解は、才能のある人であれば、フーガの構造など（明示的に）知らなくとも暗黙的に理解してうまくもっていていた。しかし、高度な才能のない人（直感で理解できる才能であって演奏の才能とは少し異なる）と計算機には何かしらの形でこの構造を教える必要がある。パターンを前もって網羅するという事は暗譜が出来ている状態に近いと思われる。その状態から声部を分離する。これは構造理解への第一歩である。

第 4 章

声部分離モデルの構築

多声音楽のためのモデルは始めに並行する旋律の存在を理解する事が必要になる。そこで本研究では多声音楽の声部分離をモデル化する事を試みる。声部を認識する時に注目しているのは旋律と思われる。多声音楽の特徴は和声よりも対位法的な観点にたった旋律に一番の特徴があるからである。先行研究で単旋律の構造分析に使われていたモデルをもとに検討する。本研究では聴き手を模範にモデルを作りたいので作成するモデルには認知的な側面が含まれる。

声部を分離するために各々の声部に存在する旋律を構成する音の推移パターンを手がかりにする。推移パタンの長さは比較的短く短期記憶と思われる範囲に収めるのが好ましい。推移パタンの長さは他に、事前に与えるデータの数が少ないためにそれに見合う長さにする。パターンを用いた先行研究には [5] [4] [27] がある。

パターンを使った手法に限らず音楽認知モデルには達成すべき条件がある。モデルが音楽構造の抽象化表現が可能な事。構造単位とその関係、特に階層関係が表現でき階層的であること。構造の構成要素が分離的にモジュールとして扱え、繰り返し構造が識別できること。しかし本研究の目的は声部分離なので、階層関係の表現は必要条件から外す。音楽は記憶するために高度に構造化・抽象化されている。よって音楽の高次の構造は「知識表現」つまり記憶の形式と密接に関わる。それは多声音楽でも同様である。ここで多声音楽を認識する手がかりを考える。複数の声部が同時に提示されても各々の声部を認知し感じ取る事を可能にするのは、あらかじめ注目する旋律パターンを聴き手が持ち合わせていて、選択するパターンが決まっているからだと思われる。これは聴き手が旋律を並列的に処理している事を意味する。パターンを使った手法に必要な問題を以下に掲げる。

- 視点
- 事前の声部の把握

- イディオムのパタンの把握
- パタンの反復
- 反復発見までにグループ認識
- ゲシュタルト的グループ化 (近接性、類似性、連続性)
- イディオムのパタン
- 音の協和性、音階組織や和声構成による音楽固有の性質（物理的、心理的だけでなく文化的背景に依存）

以下、本研究でモデル化すべき問題と保留すべき問題を検討する。

4.1 モデルにする項目としない項目

本研究の目的は声部分離が最終目標にあるのではない。声部分離の後の高次の構造把握、それによる高次の表現までも見据えたモデルが必要とされる。そのためモデル化は慎重に行いたい。

4.1.1 モデル化すべき物

視点

視点と言っても視点は沢山あるために全ての視点を把握する事は無理である。だが、提案モデルではどのような視点が大切かについて、はっきりとした意志を持って計画されている。モデル化で重要な事は音楽的に意味のあるモデルを作る事と考えているからである。今までにも述べたが、まず第一に旋律に注目する事、そしてその中でも輪郭に注目する事である。提案モデルは旋律に関してモデル化した物なので全てにこの視点が活かされている。

事前の声部把握

声部分離には事前に各々の声部を独立して認識している事が考えられる。ゆえに独立した声部の学習が効果的でありうる。これは事前にすでに分離されたデータからパターンを取得する事で模倣する。実際の演奏家であれば、声部を一つずつ弾いてみて学習する事や楽譜を見て視覚的に声部関係を把握する事に値する。



図 4.1: 転回

イディオムのパターン

今回の対象はフーガなので計算機はフーガ形式を理解出来なければいけない。人がフーガ形式をどのように理解しているか考えて見るのが役に立つ。フーガの形式の理解は人間に生得的に備わっていると考えるのは難しい。フーガがどのような物が先に知っている必要があるだろう。

例えばフーガには転回という技法がある。これは、従来からバロック音楽やその他のクラシカル音楽に含まれる転回構造になれ親しんできた物には容易に把握出来るが、親しんでいない物には理解できるだろうか？転回構造が生得的に把握できるかどうかは疑問がもたれる。つまり、楽譜という手段を用いた図として、もしくは高低が把握できた人に、これが上下逆になっています、と教示する事による、聴取以外の方法によって習得された可能性がある。もしそうであれば、フーガを分析するモデルにこの転回の仕組みを意図的に教える事は人の認知プロセスになんら反していない。これは経験的に習得される事はされるが事前に形式知としての理解が必要であろう。

パタンの反復

パターンは反復されなければ意味がない。反復されなければパターンをパターンと見なし得ない。よって生成するパタンの数は反復が可能になるように、そして聴き手が反復とわかるように作らなければならない。提案モデルでは以上の事には検討しているが、分離実験後にどの部分でどのような反復が存在したかという長期記憶に関わるようなより上位の階層については言及していない。

反復発見までにグループ認識

大きな物を把握する場合には記憶構造の面から言っても階層構造を作らなければならない。そして階層構造を作るためには、いくつかの物をグループ化して認識するという、物をまとめる作業を行わなければならない。提案モデルでは単旋律の時間的に近傍で短期記憶で扱うと思われる範囲だけをパタンとして扱っている。それ以上のより大きな構造のグループ化については提案モデルでは言及していない。

ゲシュタルト的グループ化 (近接性、類似性、連続性)

物の理解にはゲシュタルト的グループ化が存在するという考え方がある [8]。問題はゲシュタルト的な部分をどこに見出すかである。提案モデルでは音高差を用いてゲシュタルト的グループ化がおこると仮定している。提案モデルでは扱っていないが、リズムにゲシュタルト的グループ化がおこるとも考えられるし、単なる音高差ではなく、音高の調性的な距離でゲシュタルト的グループ化がおこるとも考えられる。

音の協和性、音階組織や和声構成による音楽固有の性質

提案モデルでは直接的にこれらの問題について触れてはいない。作品の特徴としてはフーガのルールをイディオムのパタンとして埋め込むのみである。だが、12 音楽の音高の並びが周波数の大小で並ばないように、提案モデルでも暗黙的に D は C より高く E は D より高い等の性質を埋め込んでいる。

4.1.2 モデル化しない物

拍節情報を用いない

音楽作品を聴く時には拍節情報、つまり小節は聞こえない。これは作品の構造がわかってから、後に表出してくる物である。むしろ音楽構造を見極める場合に拍節構造が重要な役割を果たしている事は間違いない。だが実際に鳴っている音楽から楽曲構造を分析する事をふまえたため、提案モデルでは拍節情報は用いない。拍節構造はともするとそれ一辺倒になりがちで、事実、演奏家でも拍節構造をベースに旋律解釈をしてしまう場合がある。

人は確かにわずかの音が耳に入るだけで拍節情報をイメージ出来るがそれは未来までの情報を一度頭の中で整理しているから浮かぶためで、未来の情報、つまり経験なしには出てこない。拍



図 4.2: 拍節構造

節情報は音高・リズム情報から副次的に生み出され、実際には経験としての音高・リズム情報から拍子を先行して感じ取ると考えられる。

リズム

提案モデルではリズムを直接的に使っていない。リズムはパターンを作る単位が固定な事で暗黙的に含まれてはいる。リズムに関するモデル化は [5] で述べられている。後にリズムのパターンも独立並行してモデル化したものと比較する事も考えられる。

4.2 旋律のパタン

声部を分離する時にはおそらく聴き手が旋律を認識している。つまり音がならされた瞬間の楽音情報以外に、過去の情報を使っていると考えられる。旋律は音と音の関係から紡ぎ出されるからである。多くの研究は和声的な特徴や単旋律を扱っている。並列に旋律が存在する多声音楽を扱っている物はまだほとんどない。本研究ではこのパタンを多声音楽の理解に使う。旋律パタンをどのくらいの大きさで認識しているかは検討しなければいけない。

4.2.1 パタンの分類

本研究ではパタンの分類に音の進行型を使っている。ある時点で鳴らされた音が直前に鳴らされた音より高いのか低いのかという分類をしている。一般的な分類法には調性音楽に用いられる十二音の音程からその差を使うことが考えられる。しかしこの十二音の空間が線形であるという

保証はないしむしろそうでないと考えられる。これはバロック音楽をさらにさかのぼった時に音の進行形こそ意味があったと考えられたためである。さらに今回ひきだそうとしているのは旋律分離による音楽構造であるので、和声理論に関わってくる調性構造はなるべく排除して考えたためである。本研究では旋律の輪郭をボタンに使用することにする [27]。

4.2.2 声部を構成するボタンの類似性

どのような旋律が似ていると見なすかは難しい。似ていると感じるかどうかは先に述べた聴き手の信念により大きく変わってくるからである。認知する分解能が低い人にとっては多くのボタンが似ていると感じるだろう。過去に行われた研究では [25] に類似性についての考察があり、[27] に旋律ボタンの分類について述べられている。本研究では旋律の輪郭をもとに分類を行う。この方法は CCARH の Themefinder でも採用されている。その中でも輪郭を up step, up leap, down step, down leap, same, それに difference, has rest の 7 つに分類する。このアプローチの利点は調性情報を取り扱う事である。これから対象とする作品にも調性はあるが、調性構造を踏まえた理論である Narmour 理論に対して本研究のアプローチは調性を出来るだけ無視しようとしている。このような取り組みは [5] の研究にも見られる。さらに、このアプローチは類似性比較にも影響を及ぼす。このアプローチは性質の類似性比較により近いと考える。輪郭の進行の方向こそが旋律の基本要因と考える。

類似性の問題は、どのような要因のもとで類似性が成立するかと、認知システムの中でどのように処理され、どういう役割を果たしているかという問題とに分けられる。前者について実際の曲を調べた研究によると、類似性のボタンはかなり多用であるとされている [4]。平行移動や部分的变化、音程、リズム等の特徴的な部分を抽出・強調するものなど様々である。簡約も類似性の一種の評価である。しかしこれらはなかなか人間の直感的な認識のレベルまでは到達できない。認知システム上での役割を考えると類似性の存在を比較するのは難しくなる。例えば 2 つの音の音程差が小さいものが関連があり、離れている物は関連が低いというような線形的な振る舞いは持たないだろう。

リズムとの関係

旋律の類似性は音の進行形だけにあるのではない。拡大・縮小の技法は時間軸上に拡大したり縮小したりする事によって表現する。同じ音の進行形であっても異なる旋律に感じとる例は数多くある。本実験ではボタンを構成する tick を指定してあるので暗黙的にリズムは表現出来る。し

かし、パタンの類似性には影響しないので、拡大・縮小は意識的に処理を施す。

4.2.3 聴き手の声部分離の仕方

聴き手は声部分離をするときに出現パターンを全て暗記するわけではないだろう。しばらくのちに同時進行、斜進行、反進行のパターンを習得すると考える。

4.3 モデルの動き

モデルは時間の進行とともにデータが入力されると徐々に処理を行う。同時に入力されたデータは並列的に処理を行う。つまり

1. 処理は音楽の流れにそって処理される。
2. 曖昧性を持っている。
3. 小さな枠組み。音程情報しか使っていない。

4.4 パラメータの考察

声部数は MIDI データを前処理で分析する時に調べ、その声部分の声部を処理する。作品によっては和声的な表現の部分で声部とは言えないが音数が増える場合がある。その場合は多い音は切り捨てる。

第 5 章

声部分離実験

本実験ではモデルがポリフォニー音楽作品の声部を分離出来るかを確かめる。前章で述べたモデルと比較するために用意した、音程差で近傍に結びつくモデルと、それを全体としての最小値で移動するモデルの 3 つを用意した。

5.1 実験条件

音楽作品のデーターには MIDI データを用いる。MIDI データは Finale で作成された物をいったん Note on 情報と Note off 情報だけにまとめた XML 形式に変換する。その後 JDK1.2SDK1.3rc 上の Java Sound API 上に格納する。MIDI データは譜面上の長さになるように編集してある。従って通常の演奏の弾き方に比べて多少濁って聞こえる。

5.1.1 MIDI データ

入力には MIDI DATA を用いる。これは富田氏が Finale を用いて J. S. Bach's Das Wohltemperierte Clavier II: A Critical Commentary, vol.2 (Leeds, 1995) において作成した楽譜を使用したデータである。このデータには様々な情報が含まれているが Note on 情報と Note off 情報のみを用いる。

5.1.2 パタンの作成

パタンは予め用意された MIDI データからトラック情報をもとに既に分離されたデータを用いる。パタン作成には旋律パタンの輪郭の情報と、音高と音長の情報を、ともに考慮する。分析には音高、音長のデータのみを用いる。データを単位長 (tick) で分割する。単位長は作品毎に適当



図 5.1: 音高差の分類

な数値を指定する。今回用いた作品では全て 16 分音符を単位とした。音高差は以下のように分類する。

1. "up a step" : 3 半音以内の上行
2. "up a leap" : 4 半音以上 9 半音以内の上行
3. "up a big leap" : 10 半音以上 12 半音以内の上行
4. "down a step" : 3 半音以内の下行
5. "down a leap" : 4 半音以上 9 半音以内の下行
6. "down a big leap" : 10 半音以上 12 半音以内の下行
7. "repetition" : 同音反復

音長が単位長より長い場合には同音反復とみなす。なぜならば、これによりリズムパターンを表現出来るからである。これをそれぞれの作品毎に全声部で作品のはじめから終わりまで調べ出現頻度を数える。パタンの長さは 2 から指定された長さまでの全てについてそれぞれ求める。最大の長さは実験によって検討する（今回の作品では 18 を上限とした）。パタンは作品毎に毎回生成する。なぜならば作品固有の特徴が音楽を強く支配していると考えられるからである。

5.1.3 パタンの発見

ある地点でのパタンの予測はそれ以前に存在しているパート数 × 新たに出現した音の数の組み合わせ分の可能性がある。それらのパタンの出現可能性を調べる。その後、新たに出現した音を検討しパートの追加処理を行う。パタンは長い物から探す。パタンの発見は二段階にわけて行われる。はじめに過去に完全に一致したパタンを探し、その後、音程差が近い物を結びつける。これはパタンが存在しない時にその後の処理を継続するためである。

5.1.4 生成される物

実験によって声部分離の結果が正解と対応付けられて出力される。出力は時間軸にそって、声部毎に学習時に与えた音程、声部分離の結果の音程を対応づけて表す。

5.1.5 パタンの長さに関する予備実験

実験のポイントはパタンをいかに作るかである。音楽を聴く時の処理は右脳で聴くと考えられている。右脳で聴くという事は「ひらめき」がメモリベースになるという事である。「ひらめき」は記憶の連想である。楽曲を記憶するためには短期記憶から長期記憶を形成しなければいけない。その典型的な方法は繰り返す事である。量が少なく簡単な処理ならば、記号として扱ってもよいが、音楽は記号処理のように簡単な処理ではない。左脳が記号処理を行うのに対して、右脳はメタ記号的処理を行っている。このメタ記号はすなわち、縮小パタン＋付加情報と言える[10]。この縮小パタンは全ての可能パタンからも、実際に出現したパタンからも、適度に縮小されている必要がある。

音楽作品中からパタンを生成する事でどれくらいの種類のパタンが存在するかがわかる。全ての可能パタンと生成されたパタン、それに縮小パタンと呼べるだけの、圧縮が出来ているかを概観で検討した。その予備実験の結果、例えば、パタンが 16 の場合にはフーガⅠでは作られたパタン数が 1724 に対し、その中のパタンの種類は 1310 となり、それぞれのパタンの出現頻度数は 1 ないし 2 となる。よって縮小パタンとしての機能を持たないと考えられる。まとめると、パタンが 16 以上の場合にはパタンとしての機能を持っていないと言える。

5.2 実験結果

5.2.1 声部分離の結果

実験結果を付録に付けた。結果の図は楽曲毎にわけられ横軸が時間軸、縦軸が音程に対応している。横軸は基本単位で区切られ、この 3 曲の場合はいずれも 16 分音符が基本単位になる。便宜上基本単位 8 個毎に区切り線を入れてある。縦軸は B と C の境目に横線を入れてあり、鍵盤楽器の黒鍵に対応する物には薄灰色の色を付けてある。これらは視覚的に理解しやすいようにつけた物であって、これらの情報を分析には利用していない。声部はそれぞれ黄色、緑、青、赤の色で色づけされている。特定の時間の特定の音程のマスが上下にわかれているが、上側が分離を試みた結果で、下側が正解である。わかりやすいように下側の色は上側の色に比べて多少明るく

してある。

5.2.2 正解率

成績を客観的に把握するために正解率を割り出した。正解率は直前の分離と同じ分離が出来ている割合で求められる。なぜならば正解通りの分離ではなく2つの旋律が完全に入れ替わっていてもよく、また途中から大きく2つが入れ替わる事もあるが、それは入れ替わり地点を間違えただけで、その後の分離は上手くいっているとみなせるためである。

表 5.1: 正解率 (間違い数)

	提案モデル	近傍モデル
フーガ I	94.2%(72)	96.0%(75)
フーガ II	92.6%(108)	94.8%(84)
フーガ V	97.0%(89)	97.8%(68)

フーガ I では提案モデルは 94.2%, 間違いの数は 72 個あった。近傍モデルは 96.0%, 間違いの数は 75 個あった。

フーガ II では提案モデルは 92.6%, 間違いの数は 108 個あった。近傍モデルは 94.8%, 間違いの数は 84 個あった。

フーガ V では提案モデルは 97.0%, 間違いの数は 89 個あった。近傍モデルは 97.8%, 間違いは 68 個あった。

この数値を見るだけでは、近傍モデルのほうが成績がよく好ましいように思われる。しかし、これは全体として見た場合の数値であって、どの部分をどう間違ったかはわからない。ゆえにこの数値のみを持って優劣を論じる事だけでは不十分である。そこで次に作品に旧来からされている演奏学的解釈 [12] をふまえて検討を行う。情報工学的に困難な箇所と、音楽的に困難な箇所が異なる可能性があるからである。

表 5.2: フーガ I

位置 [tick]	モデル名	声部	間違え方
22:7	T	ST	(a)(s)
24:1	K	ST	(a)
25:1	K	SAT	(A)
25:3	T	AT	(r)(y)
27:1	K	A	(B)
27:7	T	SA	(z)
29:1	T	SA	(n)(s)
29:7	T	SA	(n)(s)
30:3	T	SA	(m)
30:5	T	SA	(n)(s)(m)
32:5	T	SA	(n)(y)
35:1	K	SA	(D)
35:1	T	SA	(y)(a)
53:1	K	TS	(B)
70:3	K	SA	(E)
72:1	K	SA	(a)
72:1	T	SA	(a)
73:7	T	ST	(a)(s)
76:3	K	SA	(E)
78:1	T	SA	(y)
80-	T	SA	(b)

表 5.3: フーガ II

位置 [小節/2]	モデル名	声部	間違え方	位置 [小節/2]	モデル名	声部	間違え方
9:1	T	SA	(z)	35:7	T	AT	(n)
10:2	T	AT	(a)(n)	36:3	T	AT	(n)
10:3	T	AT	(a)(n)(m)	37:7	K	AT	(E)
12:3	TK	AT	(n)(y)	41:3	T	AT	(z)
16:3	K	SA	(E)	41:7	T	AT	(m)(z)
20:7	T	AT	(n)(z)	46:1	K	SA	(E)
24:3	K	BS	(B)	49:7	K	SA	(e)
27:5	T	AB	(n)(y)	51:5	T	AT	(z)
31:1	K	SA	(B)				
32:3	K	SA	(B)				
32:5	T	SA	(g)(z)				
32:7	T	SA	(m)(z)				
33:3	T	SA	(m)(n)				
33:5	K	SA	(D)				

表 5.4: フーガ V

位置 (小節/ 2)	モデル名	声部	間違え方
10:5	T	AT	(z)
12:3	T	AT	(n)
14	T	TB	(u) BEDCE
15:1	T	TB	(m)
16:1	T	BT	(n)
36:5	T	TB	(r)(n)
37:5	T	ST	(m)(z)
38:5	T	ATB	(m)(z)
45:3	T	SA	(g)
53:1	T	TB	(s)(z)
54:1	T	TB	(m)
56:3	T	ATB	(n)(z)
59:3	T	TB	(z)
60:1-3	T	TB	(m)(n)(u)(z)
66	T	A	(g)
67:5	T	TB	(m)
75:5	T	TB	(u)(z)
87-89	T	AT	(y)
93:1	T	AT	(n)

5.3 考察

5.3.1 楽曲考察の表

不正解部分に対する表の見方を説明する。位置は小節を単位とする場合と小節/2 を単位とする場合がある。A:B という形式になっていて、A は前記の単位を、B はその単位の中でいくつ目の基礎単位かを表している。モデル名は本研究で用いた提案モデルでの間違いか、近傍モデルの間違いかを表す。提案モデルは T で、近傍モデルは K で示される。間違い方は以下のリストにある記号に対応して表される。

- (A):近傍モデルで2個の音が同時になくなるため
- (B):近傍モデルで旋律が途中で切れる、または出現するため
- (D):近傍モデルで旋律の音域が重なるため
- (E):近傍モデルで距離が同じ選択肢があるため
- (a):学習時には音の頭で連続するか、音の途中も含めて連続するかを調べていないため。
- (b):和声的進行になり単旋律としての機能を失いかけているため
- (g):動機発見の失敗
- (m):直前に間違ったため
- (n):自然・音階的
- (p):普遍的知識がないため
- (r):休符があるため
- (s):リズムの変化・パタンの不足
- (u):融合を示唆・一つの旋律になる・新しい旋律になる
- (y):大局的構造を知らないため
- (z):正確にはわからない

5.3.2 提案モデルと近傍モデルの比較

パタンによる分離結果をパタンを使用しない分離方法と比較する事で検討する。パタンを使用しない分離方法には音程差が近い物が声部としてまとまる事を優先するルールに選択時の全ての音を視野に入れ、全体として最小の音程差になるルール（近傍モデル）を用いる。

フーガⅠでは1小節目から5小節目の頭のE音までがテーマで、5小節目から応答と対旋律が始まるが、テーマの冒頭と応答の冒頭はミューテーションを起こしているが音型は似ていて、パタンとしての形は同じになる。近傍モデルは結果的にかなり成績がよいが、重要な主題に使われる動機の冒頭を見失う場合がある。例えば、72小節目から始まる動機を発見できていない。一方、提案モデルは27小節目からのパタン以外はほぼ動機を発見できている。72小節目から始まる動機が発見できる事も特徴である。すなわち音程的に込み入っている部分でも、動機に注目していれば聞こえてくる事が表現出来ているといえる点で提案モデルの優位性を示せる。

フーガⅡはフーガⅠに比べて細かいので分析結果も難しくなっている。提案モデルと近傍モデルはともに、14小節以降の拡大主題を発見したというよりは、音程が近いために旋律として認識しているだけのようと思われる。18小節目からは拡大反行主題が隠れているが、これも提案モデルでは分離は出来ているものの理解していると言えるかどうかは、詳細な検討が必要である。だが声部分離は比較的成功するので、今後、この結果から知見が得られる事に期待出来る。

フーガⅤでは主題を構成する動機のパタンがいたるところに出てくる。この動機を見つけられたかどうかを検討する。近傍モデルでは9-10小節目、11-12小節目では動機のパタンを見つけられないのに対して、提案モデルでは見つける事が出来る。特に88-90小節目のソプラノ、アルト、及びテノール、バスの音程が混じり合う所で提案モデルは動機を発見する事が出来る。

これらを総合的に解釈すると、提案モデルは聴くべきテーマの分離に成功しやすい事がわかる。提案モデルは分離して聴いていると思われない和声的な動きの箇所は誤りやすい事がわかる。

5.3.3 表の検討

提案モデルによる声部分離には上手く分離できる部分と出来ない部分がある。以下に、それぞれについて検討する。間違い方を(a)から(z)に分けたが、それをさらに大別すると以下のようなになる。

1. 動機が発見できない(g)
2. 音価のパタンが十分理解出来ない(a)(r)
3. 和声的進行の部分で旋律が見つけれられない(b)

4. パタンが不足、パタンの変化に追従できない (s)
5. よくわからない。聴いた場合も難しい (z)
6. 別の旋律を感じ取る (n)(u)
7. 直前で間違っただけに、予測不可能な状態に陥る (m)
8. 大きな構造をわからない (y)
9. 普遍的知識がない (p)

パタンによる提案モデルで声部分離で可能な事と不可能な事とその理由を調べまとめた。それは以下の通りになった。

1. 主題の冒頭動機を発見出来る。

動機は曲を性格づけるもので類型が頻繁に出現する。そのためパタンとして最も多く登録され、最も優先的に処理されるからである。聴き手も基本的にははじめにこの動機を感じ取るようである。動機が含まれていれば、音域的に交差しているような声部でも分離が可能であった。これは小節を作らずに連続的に処理していることも関係あると考えられる。

2. イディオムのパタンと見なした拡大・縮小・転回が見つけれられない。

本研究では主題、応答などの拡大・縮小・転回はイディオムのパタンとして記憶していると考えたが、これらのイディオムのパタンを考慮した物と考慮しなかった物との比較では十分な有効性は得られなかった。これはこれらのパタンがイディオムのパタンとしては考えられていない事を意味するという可能性もある。これについては考察の項で述べる。

3. 主題の変化に追従できない

主題が自由に変形されてくると、旋律の輪郭は同じでもリズムが変化するために対応が出来ないように思う。これはパタンが不足する事にも関係する。

4. 声部間で同じ音が重なる場合におかしな現象が生じる。

これはモデルが事前に曲を学習する時にパタンに音価情報を含まれていないため、パタンとの照合がうまくいかないのが原因であると考えられる。

5. 明らかに性質の違う声部が反転する。16分音符が続く声部と、4分音符が続く声部が反転する。このように、性質が異なる声部を分離出来ない場合がある。これは音価を知らない事に対する弊害と思われる。

6. 声部に休符が混じると、パタンを見失う。休符を楽音と同一のパタンとして扱う事に問題がある可能性がある。

7. 別の旋律を感じ取る大きな視点で見れば間違っているといえるものの部分的に見ればそこま

でおかしな選択でない場合があり得る。

全体としてわかること

これらの事から以下の事がわかる。

1. パタンの分類は up step, down step, up leap, down leap, same のように 7 段階で分類しても確かな判別が可能であった。これは聴き手の旋律認識がどのように行われているかについて興味深い。
2. 大局的構造が理解できない。モデルの間違い方からモデルが大局的構造を理解できない事がわかる。これはモデルを作成する上で検討していなかった事なので仕方がない。この問題を解決するには音価を使う方法や他いくつか考えられる。
3. 不自然な間違い方をする。極度の跳躍を同一の声部と認識したりする。バロック音楽を聴いている時に通常極度の跳躍は同一の声部と認識しない。局所解に陥っている。これは割り当てが難しかった音を強引に割り当ててしまうために生じられると思われる。
4. 異なる声部が混ざって新しい旋律を作る。聴き手が本来存在しない声部を作り出してしまふ。この部分に、多声音楽の価値がある可能性があるので、このような間違いは興味深い。
5. 学習データとしての知識が足りない。主題の変化や普遍的知識にも関係するが、作品内だけのデータでは知識が不足する。
6. リズムの形による旋律を認識できない。主題の変化を理解できない。音価を学習していないために、16分音符が連続する旋律と2分音符が連続する音符があった場合、それぞれはそれぞれで居続けようとする等の特徴をつかめていない。
7. パタンを構成している旋律がどう始まり、どう終わるのかをわかっていない。
8. 声部をわけて学習する事に意味があることがわかる。

第 6 章

まとめ

本論文では音楽の本質を探る音楽構造分析研究を調査した後に、多声音楽の多声構造の理解モデルを提案し実験を行った。あらかじめ旋律のパターンを調べる事で声部分離が可能な事が実験によって示唆された。本研究は提案モデルをより人の振る舞いに合わせた。さらに、分離のプロセスを見ることで声部間の関係がかいま見られた。旋律を意識する場合、声部が融合して旋律が浮き出てくる可能性が示唆された。これは声部の分離は紙一重に行われており、視点が少しでもかわると違う旋律が浮き出てくる事に思える。今後の課題として、モデルの改良と他の楽曲への適用があげられる。

6.1 今後の課題

6.1.1 モデルの改良について

今後いくつかの面でモデルを改良する事が有意義と考えられる。声部分離だけに関わる事、声部を意識した音楽構造理解に関わる事、両方に関わる事がある。以下に考えられる点をリストアップし、詳細に検討する。

1. パタンの探し方
2. 声部を分離する事
3. 旋律
4. 声部間の関係・融合
5. 注目
6. 調性構造 (パタンの距離)
7. 調性構造 (派旋律の入り切り)

8. 時間単位 (決定を下すまでの時間)
9. 音楽作品に共通の事前知識
10. 曖昧性の表現
11. 可能世界
12. 大きな構造
13. 文化的背景の影響
14. 知識表現の定式化

パタンの探し方

本研究ではパタンの検索時に長いパタンから探したが、長いパタンは同じパタンの数が少ない。その結果局所的な短いパタンばかり使われる可能性がある。どの程度のバランスで検索すればよいのか、パタンの最終的決定はパタンの照合結果のどの部分を持って行うのか等わからない所は多い。

声部を分離する事

今まで声部は分離する物としてその正当性についての議論は十分にしていなかったが、声部を分離する事については今一度検討する必要がある。なぜならば声部は旋律だけでなくそれ以外の要素がふんだんに織り込まれて構成されているからである。つまり旋律優位の考え方は保持するとしても他の要素は無視できない。特に和声的進行が強い所では声部分離はそもそも必要ない。和声的進行の場合和音という物自体が一つの物体として機能するからである。この事実は、今後のモデルの方向性に大きく影響する。この「ただ分離するだけではない」という所に多声音楽の魅力があるからである。

計算機の間違えかた部分は沢山ある。しかしこれはかなり熟練した人でも難しく、ましてや素人にはどこまで理解が可能かはなはだ疑問である。我々は今後は分離だけではなく声部の融合や、分離して認識する事が一般にどこまでされていて、どの程度重要かについてもあらためて検討をしていく必要がある。

旋律

多声音楽において旋律が重要な事は言うまでもないが、声部が常に旋律と言えない場合もある。重要な視点が旋律以外にある事もあり得る。それは例えば和声効果とのかねあいである。ま

ず第一にバロック時代という時代背景ではいかに J.S.Bach がポリフォニーを重んじる伝統的なスタイルをとったとしても、調性構造を用い和声効果を狙った作風は十分見受けられるので、調性構造を無視する事に限界がある。

さらに、前述の融合とも絡む問題だが、分離という試みについて、学習段階で声部を分離して構造習得をはかる事は音楽学習法からも理解できるが、学習済みの聴き手は分離をせずに逆に、複数声部の自由な融合をすることで音楽に新たな価値を見出しているように感じ取れる。その時には融合から導き出される新たな旋律輪郭のようなものも出てくるだろう。この問題は最も頭が痛い。

声部間の関係・融合

現状モデルから声部の融合現象を見つけるとすれば、基本パターンが他のパターンに先行して発見される事を利用して、本来存在しない声部を導き出す事が考えられる。そのためには、例えば斜進行の効果が使われる時に FUGUE V の 13 小節目で Alto の B に引き続き AG#F#E とつながる事を認識させたい。

聴き手にとって認識が難しい箇所がある。例えば音程の連続して聴こうとする習性によってソプラノとアルトが逆転する可能性がある。このような習性はモデルにも実現させることが理想である。

注目

注視点については検討の余地がある。本研究では主題と応答の動機に主要な注目がいくと見なしたが、このような認識の仕方の方向付けを示唆する助言は音楽指導の分野でまだある。このような助言を探し出し、モデルが重要と見なす点をあらためて調べるべきである。

調性構造 (パタンの距離)

本研究のモデルは先にも述べた通り調性とリズムを出来るだけ排除するようにした。しかし当然ながらこれらの要素は無視出来ない物である。相互の関連づけの方法についても検討がするとよい。

その足がかりとして調性構造を考慮した音高距離を求める多次元モデルが考えられる。例えば、転回パターンは調性構造とリズムを考慮してパタンの距離空間を作ると非常に近くなるように思われる。現在のモデルではこのような空間を表現しきれていない。音高差の距離を求める方法

として Shepard の 2 重螺旋モデルをもとに多次元空間モデルを作れる。これをもとに旋律のパタン分類すると調性を踏まえたモデルが構築出来る。これは先のモデルに調性構造を結びつけ、依然としてパタンを中心に分析する枠組みの範疇を脱しない。提案モデルを使うと、デコボコのパタンがデコボコであり続けるというような現象の表現はデコボコに見えるが音楽空間のある軸では平らでありゲシュタルト的な連続性を持っていてそれが人の認知パタンに合うといえよう。

イディオムのパタンの代表例としてあげた転回構造は、調性構造を考慮した上述のモデルを用いると、とても近い関係に存在する可能性がある。フーガ形式に見受けられるミューテーションはこの調性構造を考慮したモデル上での近傍に行くための変化と見なすのが相応しく思う。

これは複数の視点があるうちの視点の方向を変更することを意味する。

調性構造 (旋律の入り切り)

大きな構造を認識する場合には旋律の入り切りがわかる事が望ましい。パタンの出現の仕方から求める事も出来るが、調性を意識できれば、より容易に入り切りを理解できる可能性がある。

時間単位 (決定を下すまでの時間)

ある音をある声部と決めるのに時間がかかる事が考えられる。それは認識して処理するまでに時間がかかり、その間にまだ特定していない音がいくつかあらわれる事になる。つまり「記憶に幅」があるといえる。この幅を持たせないと強引に決定してしまうなどの問題が生じる。パタンについては記憶との関係からさらに検討しなければならない。本研究で述べているパタンは短期記憶と結びついていると思われる。音楽を聴いている時の現在はおそらく単音ではないので、どこまでが同時に処理していい時間なのかを考慮するべきであろう。これによって声部分離が未知の状態の音符が生じる。これは言い換えると先読みが出来るかともいえる。また、主題や応答の冒頭は構造を把握する上で特に重要だが、それを構成する始めの数音を聴いただけでは、その前に音がない場合もあり、パタンを発見できない。仮にあったとしてもそれが主題の発見に決定的ではなかろう。主題といったん思い始めたらその期待の解決のためにある程度の時間待つ事が許されると思われる。それは追加的なモデルで対応する。

音楽作品に共通の事前知識

モデルは声部分離をする音楽作品に限って事前に声部を学習したが、これには他の音楽作品と共通するメタ知識が存在する事が考えられる。パタンが固有の音楽作品に内在するものか、もし

くは特定の音楽形式というそれよりも大きな枠組みに属すパターンなのか、という観点で分類し、必要に応じて利用する事が必要である。

曖昧性について (わからない音)

現在のモデルでは声部がわからない音は存在しないが、一時的または永久的に声部がわからない音を作る。そうすると曖昧なものを扱う仕組みが出来る。

選択されなかった物について (楽曲外から来る経験的パターン知識)

可能世界 [26] が選択されたパターンに影響するならば、最善のパタンを探すだけではなく、その近傍の情報も利用する必要があるだろう。

6.1.2 声部分離を越えたモデルの拡張について

より大きな構造について

声部分離が途中で入れ替わる点は大局的な構造を見ていないために起こる。今回のモデルでは最長パターン長を越える構造については理解出来ない場合がある。なぜならばパターンを小節に関係なく適応しているが、曲を聴く過程で拍節・小節構造が浮き彫りになってきて、部分的にパタンの類似性の評価方法に問題が生じると思われるからである。問題は、大きな形式による構造、例えば ABA 形式があるとすると、A の中と B の中のパターンは学習出来ても、A から B に向つる部分のパターンは頻度だけで求めるのに無理があるという点である。さらに細かい動きをするパターンは細かい動きをするパターンで居続けるのか、というような問題も存在する。それには小節の区切りを調べる事が必要であるが、[15] の研究が参考になる。パターンには本来の小節の開始点からか、中程からか等の違いによってパタンの差がある可能性がある。その違いの部分に大きな構造の変化点が多くある可能性がある。この部分は学習データが不足するので難しい。

パターンを構成する単位からその上位階層を構成するパターンを誘導したり、下位階層を演繹する事が考えられる [5]。ただし多声音楽音楽の構造把握であれば、主題、応答といった単位が一番大きく、それが声部間でどう構成されるかを把握する階層を作るのが好ましいだろう。

パターンを利用して分離を行うことは内在的にパタンの反復的提示がされる事を意味する。反復的提示は音楽の主要な性質で、広く議論がなされている。反復的提示が、どこでどのように見つけたかの情報を使うことが考えられる。

文化的背景の影響

第1章で述べたように、音楽には社会的立場や感情が影響を及ぼす。それが構造の特にパターン認識に及ぼす影響はどのような物であろうか。それは調べてみなければわからない。推測の域を脱しないが構造の作り方には影響を及ぼさなくとも、構造を把握出来るか否かに影響を及ぼすのではないか。

情報に関わる事：知識表現枠組みの定式化について

音楽構造を知識表現の問題として定式化する必要がある。本論文では簡単なモデルしか用いていないので、例えば [24] の FRM (Functional Representation of Music) のような定式化された知識表現形式を利用する事が考えられる。これは音高差や時間の距離、それによって作り出されるパターンが単純に線形でない事から、表層表現を作る枠組みといえる。しかし、FRM でも多声音楽に関しては FRM は本質的に 1 次元的な枠組みであり、フーガのように独立した声部が並行して進むような、多声音楽の表現には適さないとしている。各声部を個別に表す横方向の表現と、全声部を表す縦の表現を空間に表現する多次元的な表現の枠組みが必要である。

ある程度熟練した演奏家と分離の仕方と比較し、どのような点に弱いかわかる事が考えられる。パタンの類似性についての考察はさらに深める必要がある。例えば、被験者に数小節からなるある部分の声部の 1 つを聞き取ってもらいそれが出来るか、間違えるならばどこで間違えるか、という現象を記録し比較する。[27] に論じられているようにパタンの分類方法については多くの研究成果がある。

今後、是非注目する事は、調性構造を使うことで転回パターンを近距離に置く事が出来る事、長さをパターン化する事、合成を発見する事と言える。

6.2 結論

本研究によって声部分離にパターン構造が有効に働く事が立証できた。音楽作品内のパターンだけで素人と同等レベルの声部分離は可能な事がわかった。しかしより大きな構造や入り組んだ構造になると分離はうまくいかない。これはある程度熟練した演奏家にも判別が困難な場合がある。この理由はおそらく多声構造の理解には多くの音楽的経験が必要であり、声部分離は少しずつ行われる事、声部分離が完全でなくとも良いことを暗示していると考えられる。また、声部分離の結果からは単一の曲の情報からだけでの分離は不完全があることが示唆された。多声構造の理解に声部分離が必要条件でない事が示唆された。今後はこの部分に声部分離・声部融合が絡むのか

調べたい。音楽作品を聴く以前に多くの経験的知識が求められる結果、一般性や大衆性に疎くなるため、情動による感情とは別の立体的構造が構築されていくと考えられる。

謝辞

本研究を行うにあたり、ご指導ご鞭撻いただいた北陸先端科学技術大学院大学知識科学研究科小長谷明彦教授に深く感謝いたします。また、MIDI データの提供などのサポートをしていただいた富田庸さんに深く感謝いたします。同研究科藤波努助教授には有益なご助言をいただき深く感謝いたします。さらに、同研究科梅本勝博助教授には副テーマで熱心にご指導いただき、特に論文執筆の方法論について大変勉強させていただいた事に厚くお礼申し上げます。そして常日頃よりお世話になりました当研究室の皆様にも心より感謝いたします。

付 録 A

実験環境

- Windows98
- Finale 98
- Cakework Home Studio
- Java Development Kit 2 SDK 1.3 RC , Java Advanced Imaging API 1.0.2
- Roland SC-88VL

付 録 B

用語説明

音楽用語には専門家以外の者にはまぎらわしく誤解をするおそれがある用語がある。そこで誤解を招くおそれのある用語についてはここで説明をしている。

- 旋律 (melody, Melodie)

旋律とは色々な楽音がひとつのはっきりとわかる音楽的語句を形作って、ある目的のものを表現するように並べられたものである。連続的に響いているので楽音は線の形となる。線は直線であったり曲線であったりするが幾つかの特徴が見受けられる。ある方向へ順次的に進行した後に、その反対方向へ跳躍する等である。このような振る舞いは作曲家がある目的を表現するために行ったからなのだが、完全に自由な環境で旋律が作られるわけではない。旋律曲線を作る上で影響を与えている要因が幾つかある。

言葉の影響 言葉とそのつながりの文章とのもっている力を最上の条件で発揮させようとする場合

音階の影響 調性スキーマに従って作曲しなければいけなくなるために音階の制約を強く受ける。そのためかなり限定された音楽になる。クラシック音楽はこの調性構造をいかに自由な物にして表現豊かにするかに躍起になっていた。それにより拡張された和音、十二音音階や無調音楽が生まれた。

対位上の影響 対旋律が存在する場合にも多くの制約を受ける。

和声による影響 機能的和声法を土台とした時代の作品は、いかにも自由に旋律が動いているようでも実はその和声の中でしか動いていないことが多い。分散和音が旋律になる場合もある。

リズムによる影響

楽器による影響 楽器には演奏可能な音域、演奏可能スピード、発音数、音色の特性から来る最良の表現技法、などがあってその制限を越えることは出来ない。

- 多声音楽 (ポリフォニー)

複数の声部を持つ音楽のことをいう。ただしこの声部とはたんなる音ではなく、いわゆる実声部すなわち独立した意味をもつ声部のことをいう。和声音楽様式は複数の音で作る和音によって音楽の流れを作っているが、和音の構成音はそれぞれが独立したものでなく、複合してひとつの意味を表現しているものなので、多声とはいわない。技法としては対位法 (contrapunctus) という形で行われているものが代表的であり、楽曲形式としてはフーガがもっとも完成されたものといわれる。作曲の中では和声様式と対位法様式とは対立する物ではなく互いに補う物であり、バッハは両者を高度に調和させた音楽を作った。

- 対位法

対位法とは和音連結を対象とする基本的和声法と根本的に対照するものであり、主として、音の流れの横の線のための書式を扱う。よって音楽の本質的な在り方という点で、対位法を認めることは当然のことといえる。対位法の技法は、必然的に模倣という特殊な手法の基盤となる。対位法の主要な要素は、二つの声部を結合しつつ異なる旋律を同時に歌わすということである。

参考文献

- [1] Rita Aiello, et al. 音楽の認知心理学. 誠信書房, 1998.
- [2] Ian H. Witten Darrel Conklin. *Multiple Viewpoint Systems for Music Prediction*. Journal of New Music Research, Vol.24, N.1, pp.51-73, 1995.
- [3] Ray Jackendoff Fred Lerdahl. *A Generative Theory of Tonal Music*. The MIT Press, 1996.
- [4] Yuzuru Hiraga. *A Computational Model of Music Cognition Based on Interacting Primitive Agents, Proceedings of ICMC 1993*. International Computer Music Association, 1993.
- [5] Yuzuru Hiraga. *Structural Recognition of Music by Pattern Matching, Proceedings of ICMC 1997*. International Computer Music Association, 1997.
- [6] Mari Riess Jones and Susan Holleran (Eds.). *Cognitive Bases of Musical Communication*. American Psychological Association, 1992.
- [7] Eugene Narmour. *The analysis and cognition of basic melodic structures*. the university of chicago press, 1990.
- [8] W. ケーラー. ゲシタルト心理学入門. 東京大学出版会, 1971.
- [9] アンソニー・ストー. 音楽する精神 — 人はなぜ音楽を聴くのか? —. 白揚社, 1994.
- [10] 阿江忠. 階層構造の脳型コンピュータ. <http://www.aial.hiroshima-u.ac.jp/BC.html>, 1997 - 1998.
- [11] 阿部純一, 星野悦子. 音楽の認知心理学的研究について, 第 28 巻. 心理学評論, 1985.
- [12] 伊達純, 小林仁. バッハ 平均律の研究 2. 音楽之友社, 1993.
- [13] 岡本太郎. 今日の芸術 — 時代を創造するものは誰か —. 光文社, 1954.

- [14] 後藤慶一. ゆらぎの科学 3 — ゆらぎと音楽のかかわり —. 森北出版株式会社, 1993.
- [15] 寺井浩司. 類似フレーズに着目したピアノ演奏の自動採譜. 北陸先端科学技術大学院大学修士論文, 1998.
- [16] 寺本英, 広田良吾, 武者利光, 山口昌哉. 無限・カオス・ゆらぎ [物理と数学のはざまから]. 培風館, 1985.
- [17] 秋山好一, 松田稔, 中野稔. 日本の歌謡曲旋律の音高・音長推移パターンのスペクトル分析. 音楽音響研究会, 1991.
- [18] 諸井三郎. 音楽構造の研究 -中巻-. 音楽之友社, 1991.
- [19] 上符裕一, 青野裕司, 片寄晴弘, 井口征士. 演奏ルールの抽出について. 音楽情報科学 15-14, 1996.
- [20] 前川陽郁. 音楽と美的体験. 勁草書房, 1995.
- [21] 大橋力. ゆらぎの科学 1 — 音楽と音の非定常性 —. 森北出版株式会社, 1991.
- [22] 谷口高士. 音楽と感情. 北大路書房, 1998.
- [23] 長嶋洋一, 橋本周司, 平賀譲, 平田圭二. bit 別冊 コンピュータと音楽の世界 基礎からフロンティアまで. 共立出版, 1998.
- [24] 波多野誼余夫 (編). 音楽と認知. 東京大学出版会, 1987.
- [25] 平賀譲. 音楽認知研究の諸問題. 音楽情報科学 6-3, 1994.
- [26] 矢向正人. 楽曲分析と可能世界. 『美学』第 48 巻 3 号 (191 号) pp.1-11, 1997.
- [27] 矢向正人, 土屋景一, 荒木敏規. 旋律パターンの分類 -類似性判断と分析例-. 音楽情報科学 16-5, 1996.
- [28] 鈴木寛. SML の音楽家教育. <http://www.art.hyogo-u.ac.jp/hrsuzuki/Jindex1.html>, 1995 - 1998.
- [29] 鈴木慎一. 愛に生きる. 講談社, 1966.
- [30] 浅香淳. 新音楽事典. 音楽之友社, 昭和 52 年.
- [31] 諫早俊夫. ゆらぎの科学 3 — 音楽は聴手をどこまでだませるか —. 森北出版株式会社, 1993.