

演奏ルールの抽出について

上符裕一[†] 青野裕司[†] 片寄晴弘^{††} 井口征士[†]

[†] 大阪大学基礎工学部システム工学科

^{††} (財) イメージ情報科学研究所

音楽演奏における表情付けの研究において楽譜と演奏表現を結びつける「演奏ルール」の抽出に関する研究が盛んになっている。最近ではAIやニューラルネットなどの手法を用いた研究がいくつか提案されている。我々は重回帰分析を用いた演奏ルールの抽出に関する手法を提案している。この手法は、速さや強さといった演奏表現情報をテンポ、ベロシティといったバリューに置き換え、それらから楽譜中の様々な情報を条件節として重回帰分析を用いて演奏ルールを抽出するものである。

本稿では、より抽象化したレベル（人間が理解しやすい表現パラメータ、例えばノリやタメなど）でのルールの抽出を行い、得られたルールについて考察を加える。

Extract Musical Performance Rules

Yuichi Uwabu[†] Yushi Aono[†] Haruhiro Katayose^{††} Seiji Inokuchi[†]

[†] Faculty of Engineering Science, Osaka University

^{††} Laboratories of Image Information Science and Technology

uwabu@inolab.sys.es.osaka-u.ac.jp

As for researches of musical performance generation, extraction of musical performance rules linking between score and performance becomes popular. Recently, some researches using methods of artificial intelligence or neural net have been proposed. We have been proposing extraction method of musical performance rules using multiple regression analysis. In this method, using expression information in performance (speed and dynamics) as values (tempo and velocity) and various informations in score as conditions, musical performance rules are extracted from them.

In this paper, we describe extraction of rules on more abstracted level (understandable expressive parameters for human), and discuss about its rules.

1. はじめに

近年、表情豊かな自動演奏を行うための演奏者の音楽解釈モデルが注目されつつある。演奏には、楽譜という作曲家の意図しているイメージを記号化した情報を、物理的な音響情報に変換する作業が必要となる。その手続としては、まず楽譜を読み、そこに与えられた意味を分析しする。そして、その意味に応じた表現を与え、演奏につなげていくことになる。前者は、楽曲分析にあたり、Lerdahl & Jackendoff^[1]や Narmour^[2]による理論などが注目されている。また、「演奏ルールの抽出」と呼ばれる研究はその後者、つまり分析されたものに表現を与えていく過程についての研究といえる。

演奏ルールの抽出に関する研究は、音楽情報科学の分野においていくつか行われている。小田、白川ら^[3]は、エキスパートシステムとニューラルネットを組み合わせたシステムを提案している。エキスパートシステム部では、大局的なルールを抽出している。また、感覚的、身体的なものからくる無意識的なものはルール化が難しいとしてニューラルネットを使用し、局所的なルールを付加している。Widmer^[4]は、人工知能的アプローチで音楽解釈システムの実現を行っている。用いられた手法は、説明に基づく学習 (Explanation-Based

Learning) を用いたもので、音楽知識から演奏ルールの学習を行おうというものである。あらかじめ与えられる音楽知識は、Lerdahl & Jackendoff や Narmour による理論から記述される音楽構造と楽譜から得られる音価などの表層構造である。

我々は、重回帰分析を用いた演奏ルール抽出の手法を提案している^[5]。この手法は、条件節となる説明変数と学習例となる目的変数の関係をバリュー情報で表すことができる点で演奏ルールの抽出に非常に有効な方法である。本稿では、目的変数となるバリュー情報をより人間のわかりやすい記述に抽象化し、得られた演奏ルールについての考察をおこなう。

2. システムの概要

システムの構成を図2に示す。ルールの抽出には、MIDIに代表される物理的なデータ記述でなく、拍を基準にしたもののように人間が理解しているのに近いと思われる記述を対象とする。そうすることのメリットとしては、人間へのフィードバックがしやすく音楽教育などに利用しやすいということがあげられる。そのようなデータ記述を演奏パラメータと呼ぶ。演奏パラメータ抽出部では、実演奏と楽譜を比べることで実演奏で行われた音楽の表現をわかりやすい形にして抽出する。

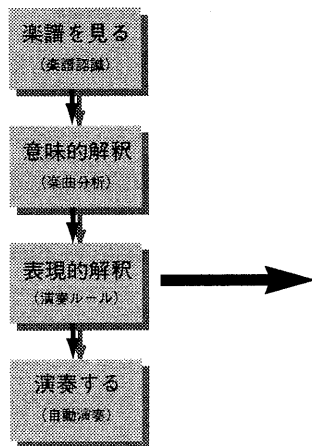


図1 楽譜から演奏への流れ

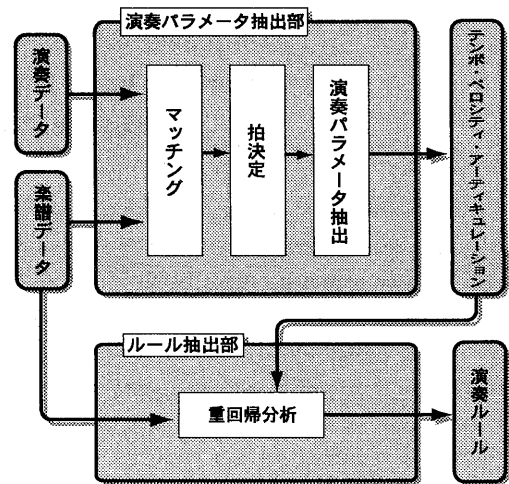


図2 演奏ルール抽出システムの概要

そして、演奏ルール抽出部では抽出された表現がどのような状況でどのように用いられているのかを統計的手法のひとつである重回帰分析を用いることでルールとして抽出する。

以後、各部の詳細な説明を行う。

3. 演奏パラメータの抽出

演奏パラメータとして今まで扱われてきたものは、テンポ（アゴギグ）やベロシティ（デュナミク）だけであった。しかし、演奏者はテンポとベロシティだけで演奏を制御し、どの音も楽譜上の音価どおりにレガートで演奏しているわけではない。ノリを出すために音長を変化させたり、タメを作るために発音を遅らせたりしている。そのため、今回は従来のパラメータに加えにアーティキュレーション情報を導入する。アーティキュレーションの定義は辞書によれば、「発音を明確にすること」または「1フレーズ内の旋律をより小さな範囲に区切り、それにある形を与えること」だが、ここでは音長(Onset-Offset Interval)の変化及び、発音位置の揺らぎ(Onset Deviation)と定義する。実演奏から、以上のようなテンポ、ベロシティ、アーティキュレーション情報を取り出すためには図2のように3つのプロセスが必要になる。

3.1 入力データ

入力されるデータは、楽譜に記されているとおりにシーケンサに打ち込み、Standard MIDI File (SMF) 形式ファイルにしたものを楽譜データとする。また、電子ピアノやキーボード等のMIDI楽器による演奏をシーケンサで記録しSMF形式ファイルにしたものを演奏データとする。

3.2 マッチング

演奏データと楽譜データをマッチングすることにより、演奏データ上の音の小節位置や拍位置がわかるようになる。マッチングは、単旋律だけでなく和音を含むものも可能である。

和音の場合は、楽譜上では同時になる音が、実際の演奏では図3のように音が時間的に分散される。そのため、和音の間だけ先読みをして図4の

ように順番の入れ違いに対応している。

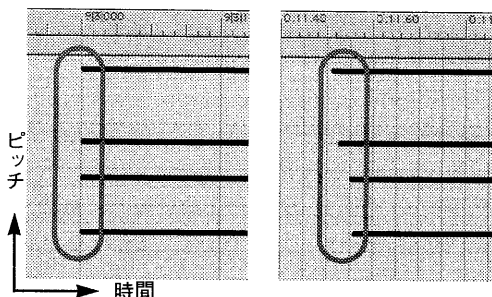


図3 楽譜データ (左) と演奏データ (右)

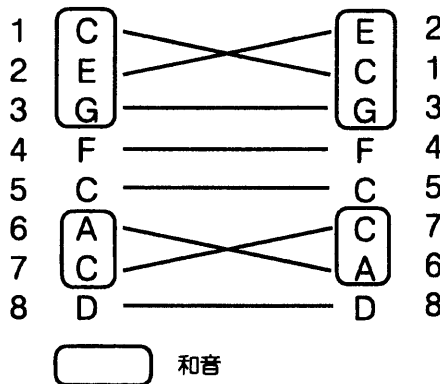


図4 マッチング

3.3 拍位置の決定

拍位置を決定することはテンポやアーティキュレーション情報を抽出する上で必要である。ここでは2つの方法をとっている。

1つは、拍感がある特定の音によって引き起こされるという前提のもとに拍位置を決定する方法である。まず、楽譜上で拍位置にあるもので、単音しかない場合はその音の時刻を拍位置とする。また、和音の場合は左手（伴奏パート）でベロシティの一番大きい音の時刻を拍位置とする。しかし、拍位置に音がない場合や伴奏パートが意図的に拍位置よりずらされるような演奏が行われる場合があるので、伴奏が拍を促すような曲に適用することが望ましい。

もう1つは、演奏者が自然と感じる拍位置をタッピングにより入力してもらう方法である。タッピングされた位置が常に拍位置とされるため、1つ目の方法で述べた欠点は解消される。

3.4 演奏パラメータの抽出

以上のようなプロセスにより、演奏パラメータを算出できる。

(1) テンポ

従来の研究ではテンポという発音時刻の間隔 (Inter-Onset Interval) から求めるものが多いが、(伴奏パートから、または入力された) 拍位置が決まっているので、以下のように求められる。

$$T_i = \frac{60}{b_{i+1} - b_i} \quad (1)$$

T_i : i 拍目のテンポ [BPM]

b_i : i 拍目の実演奏での拍位置の時間 [s]

(2) ベロシティ

マッチングをとることで、楽譜上の各音符が実際の演奏でどのぐらいのベロシティで弾かれたのか分かる。ベロシティは、0 ~ 127 の値をとる。

(3) 拍位置での揺らぎ

決定された拍位置よりどのぐらい前に出ているか、遅れているかを求める。

$$A_{ij}^T = \frac{p_{ij} - b_i}{b_{i+1} - b_i} \quad (2)$$

A_{ij}^T : i 拍目の j 番目の音のズレ [拍]

p_{ij} : i 拍目の j 番目の実演奏での音の発音時刻 [s]

(4) 拍内での揺らぎ

拍位置と拍位置の間で実際に弾かれる位置からどのぐらいずれているかを内挿して求める。

$$A_{ij}^T = \frac{p_{ij} - b_i}{b_{i+1} - b_i} - (s_{ij} - i) \quad (3)$$

s_{ij} : i 拍目 j 番目の音の楽譜上での拍数 [拍]

(5) 音長の変化

実際の音価よりどのぐらい短く、または長く弾かれたのかを求める。

$$A_{ij}^L = \frac{L_{ij}^P}{b_{i+1} - b_i} - L_{ij}^S \quad (4)$$

A_{ij}^L : i 拍目の j 番目の音の音長のズレ [拍]

L_{ij}^P : i 拍目の j 番目の実演奏での音長 [s]

L_{ij}^S : i 拍目の j 番目の楽譜上での音長 [拍]

以上より求められた演奏パラメータの抽出例を図5に示す。

BeatNumber	Tempo	Note	Velo	NoteOn	NoteOff
2.000000	77.293247	G#3	48	0.000000	-0.416974
3.000000	67.351992	C#1	35	0.000000	-0.665595
3.000000	67.351992	E4	65	-0.019293	-0.334405
4.000000	88.381732	C#3	18	-0.002110	-0.702532
4.000000	88.381732	G#2	22	-0.008439	-0.637131
4.000000	88.381732	E3	32	0.000000	-0.687764
5.000000	96.084725	E3	36	0.000000	-0.853211
5.000000	96.084725	C#3	27	0.009174	-0.816514
5.000000	96.084725	G#2	31	0.000000	-0.651376
6.000000	88.010377	D#1	41	0.000000	-0.466387
6.000000	88.010377	G3	29	0.012605	0.075630
6.000000	88.010377	D#4	47	0.004202	0.088235
7.000000	81.822151	C#3	39	0.000000	-0.562500
7.000000	81.822151	A#2	15	0.000000	-0.589844
7.000000	81.822151	D#2	37	0.017578	-0.625000
8.000000	70.055086	C#3	39	0.000000	-0.625418
8.000000	70.055086	A#2	24	0.003344	-0.578595

図5 演奏パラメータの抽出例

4. 演奏ルールの抽出

重回帰分析とは、説明変数を式(5)のように一次結合で表すとき、目的変数との誤差の自乗和が最小になるような結合係数を求めることである。重回帰分析を用いた研究は、山内ら^[6]のものがある。しかし、我々の手法では各条件節の有無を説明変数において「0」「1」で表すことで、どの条件がどのぐらい対象となる演奏パラメータ (目的変数) を変化させるのかを、結合係数において演奏ルールとして求めることができる。

$$\begin{bmatrix} x_{11} & x_{21} & \cdots & x_{m1} \\ x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (5)$$

$$\epsilon^T \epsilon \rightarrow \min$$

演奏パラメータは各条件における解釈だけでなく、それらの相互作用の影響もある。我々は、このような状態を表現するために説明変数間のAND処理を行い新たな説明変数をつけ加える処理を導入した。これにより、同じ説明変数が与えられている場所で、違う表現がされるような場合に対応することができる。

また、楽譜に書かれている情報は演奏に必要な最低限の情報であり、実際、演奏家は演奏を行う前に図1で示したように楽曲構造の分析のような

意味的な解釈を行うであろう。演奏ルールを扱う際には、そのように解釈された結果も条件節として扱う。



図6 構造解析されたワルツの楽譜

5. 実験

ショパンの「ワルツ作品64の2」を対象に、演奏ルール抽出の実験を行った。

まず、演奏データと楽譜データから演奏パラメータを抽出し、それらを目的変数とする。次に、図6のように構造解析された楽譜から、楽譜記号、構造情報などを説明変数として用意した。それらを表1にまとめる。

楽譜記号は、楽譜に明示的に書かれているものであり、ここではスラー、クレッシェンド、デクレッシェンドの3つを取り上げた。

構造情報は、音楽教諭で指揮者でもある専門家に構造解析を依頼したものである。なお、ここで用いた構造解析理論は保科理論に基づいている。保科理論ではグループにおける緊張と緩和の構造を持ち、その転換点（もっとも緊張度の高い点）を演奏において重要性の高い音符として、重心と呼んでいる。保科は明確な重心の求め方を示していないが、Lerdahl & Jackendoffの提唱するプロンゲーションナリダクションの解析手法を用いることでかなりの部分のアルゴリズム化が可能である^[7]。ここでは、説明変数として重心とグループの開始点とグループの緊張部（アナクルーズ）

とグループの弛緩部（デジナンス）の4つを取り上げた。また、拍節構造として1拍目、2拍目、3拍目にある音に説明変数を与えた。

視覚の認知構造に基づくゲシュタルト理論では、近接性、類似性、連続性、方向性などが基本的なものがあげられている。音楽にもゲシュタルト的なまとまりがあることは音楽心理学などの分野で検証されてきているが、ここでは音高の方向性に関して取り扱う。音高が前の音に比べ1度以上変化しているのものを上昇系、下降系に区別した^[8]。また、4度離れているものに関しては跳躍アクセント^[9]として別の説明変数を与える。

以上の中から重心とグループの開始点は、局所説明変数としてONとOFFの2つを与えた。これらの説明変数をもとに、重回帰分析を行うことで演奏ルールを抽出した（表1）。テンポ、ペロシティは差分値である。発音の揺らぎ、音長に関しては+が前に-が後ろにずれることを示す。

6. 考察

音楽学において、さまざまな演奏表現法が存在する。その中でも、フレージングに関する定義や方法はいくつもの見解があるが、竹内は認知的な視点からフレージングを「グループあるいはフレーズを聞き手がよりの確に認識できるように演奏表現を行うことである」と定義し、基本原則としてのフレージングルールを以下のようにあげている^[10]。

表1 説明変数とルール抽出結果

		テンポ [BPM]	ペロシティ	発音の揺らぎ [拍]	音長 [拍]		
楽譜記号	スラー	-7.8	0.5	0.049	-0.230		
	クレッシェンド	-3.9	11.6	-0.045	-0.020		
	デクレッシェンド	36.2	-5.0	0.079	-0.177		
構造情報	保科理論	グループの開始点	ON	17.3	5.4	-0.022	-0.324
			OFF	16.3	4.1	-0.005	-0.077
		アナクルーズ	15.0	8.5	-0.041	0.177	
		デジナンス	-25.5	4.4	0.011	0.200	
	重心	ON	17.3	13.2	-0.078	0.061	
		OFF	16.3	-9.6	-0.001	-0.147	
	拍節構造	1拍目	-19.7	-10.0	0.005	0.149	
		2拍目	1.4	-8.0	0.007	0.217	
		3拍目	-16.8	-8.9	0.005	0.100	
	ゲシュタルト	上昇系	-6.8	2.6	0.027	-0.186	
下降系		-3.9	2.8	0.022	-0.083		
跳躍		-0.9	-8.7	0.027	0.005		

- (1) フレーズやグループの間に、音楽の流れの切れ目を用いる。
- (2) フレーズの頂点やグループの重心に向かって（アナクルーズ）、音量は次第に増加し、それ以後（デジナンス）は次第に減少する。
- (3) フレーズの頂点やグループの重心に向かって（アナクルーズ）、テンポは次第に加速し、それ以後（デジナンス）は次第に減速する。
- (4) フレーズやグループの開始部から次第に加速されるが、フレーズの頂点やグループの重心の直前でいったん減速し、頂点や重心で再び加速する。以後はデジナンスとして次第に減速する。これは演奏者が「タメ」と呼ぶ表現である。
- (5) フレーズやグループの開始は強勢によって明確に表現される。
- (6) フレーズに当たっては、特に拍節的な強拍部を演奏アクセントによって強調しない。

(3)や(5)については、明確にその傾向を示したルールが抽出されている。(6)についても、ワルツにおける強拍である1拍目がベロシティを下げるようにはたらいっている。しかし、(2)のデジナンスの部分では、フレーズルールとは逆のルールを抽出している。原因としては、グループ間の階層性を考えて説明変数を与えていないこと、相互作用の影響が各条件節間から抜け切れていないことがあげられる。しかし、むやみに説明変数やAND処理の回数を増やすことはアドホックなルールしか抽出できない可能性がある。また、アナクルーズ、デジナンスといった楽曲構造の演奏者による解釈の違いが演奏表現に大きく影響し、演奏ルールの抽出に大きな問題をもたらしているだろう。

7. おわりに

本稿では演奏表現をテンポ、ベロシティ、アー

ティキュレーションという演奏パラメータによって抽象化し、演奏ルールの抽出を行う手法を提案した。また、ルールの有効性を確かめられれば、音楽学で得られている知見の検証や音楽教育への応用が可能であろう。

演奏者の音楽解釈モデルを構築するという点では、楽曲構造の分析を行うのに人間よる処理が必要であり、環境や文化や個人差などの影響を大きく受ける。現在、これらの事項をオプションパラメータとして扱っているモデルの構築の検討を始めている。

参考文献

- [1] Lerdahl and Jackendoff : *A Generative Theory of Tonal Music*, MIT Press (1983)
- [2] Narmour : *The Analysis and Cognition of Basic Melodic Structures*, The University of Chicago Press (1990)
- [3] 白川, 小田他 : 演奏者情報を加味したピアノの自動演奏システムの構築, 情報処理学会研究報告, 95-MUS-10, pp.1-12 (1995)
- [4] Widmer : Understanding and Learning Musical Expression, *Proc. ICMC*, pp.268-275 (1993)
- [5] 青野他 : 重回帰分析を用いた演奏ルールの抽出, 情報処理学会研究報告, 95-MUS-11, pp.1-6 (1995)
- [6] 山内他 : GAを用いたピアノ演奏の局所テンポ解析, 情報処理学会研究報告, 95-MUS-14, pp.1-6 (1996)
- [7] 片寄, 竹内 : 演奏解釈の音楽理論とその応用について, 情報処理学会研究報告, 94-MUS-7, pp.15-22 (1994)
- [8] Conklin and Witten : Multiple Viewpoint System for Music Prediction, *Journal of New Music Research*, Vol.24, pp.51-73 (1995)
- [9] 村尾 : クロージャーの客観的測定に基づく構造音の抽出について, 音楽情報科学研究会夏のシンポジウム予稿集, pp67-72 (1992)
- [10] 竹内 : 認知的視点による演奏解釈の研究, 兵庫教育大学大学院修士学位論文 (1994)