

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 ～ 音符属性に着目した個人的解釈ルールの生成～

白川健一, 小田安彦, 熊谷俊行, 梶川嘉延, 野村康雄
関西大学工学部

我々は特定のモデル演奏者の演奏者情報(特徴)を持つピアノの自動演奏システムの構築を行ってきた。本システムは、システムに市販の楽譜を入力することにより、あたかもモデル演奏者が演奏したかのような演奏を行う。このような機能を実現するため、本システムではエキスパートシステムとニューラルネットワークを用いており、本稿ではエキスパートシステム部に於けるルール生成について報告する。

Construction of automatic piano playing system wearing performer's characteristic.

-Generation of the personal interpretation rules from note's attributes-

Ken-ichi Shirakawa, Yasuhiko Oda, Toshiyuki Kumagai
Yoshinobu Kajikawa and Yasuo Nomura
Faculty of engineering, Kansai University

We have been constructing an automatic piano playing system wearing a performer's characteristic. The system generates his performance data on a score as if he really played piano. In order to realize the system, expert system and neural network are used. This paper describes the rule-generating of the expert system.

1. はじめに

今日、電子楽器の発達により、音楽経験のない人でも計算機に演奏データを入力するだけで比較的簡単に一応の演奏及び、作曲ができるようになってきた。しかしながら、素人の人がいきなり音楽性の高い演奏を作ることは、やはり困難なことであろう。また、電子楽器用に様々なプロの演奏者による演奏データを納めたソフトも多く市販されているが、聴きたい曲が自分の好きな演奏者のものである例は少ないのではないだろうか。そこで我々は、楽器をピアノに限定し個々のプロの演奏者による実際の演奏データより、その人独自の演奏技術、及び演奏感性を演奏者情報として抽出しデータベースとして備えたピアノの自動演奏システムの構築を行ってきた⁽¹⁾。これは、システムに楽譜さえ入力すれば好みのプロの演奏者の演奏者情報を反映させた音楽性の高い演奏を実現するものである。すなわち、プロの演奏者がその楽譜を実際に演奏したものに、勝るとも劣らない演奏を出力するのである。

プロの演奏者による演奏データを楽譜と照らし合わせると、演奏者は必ずしも楽譜通りの演奏を行うのではなく、楽曲に対して演奏者独自の解釈、またその演奏者の感覚的、身体的なものからくる弾き癖などの無意識な揺らぎが加味され、個性的な演奏がなされていることが分かる。実際、楽譜通りの表情のない演奏というものは聴いている者に対してなら感動を与えず、上記のような良い意味での楽譜からの逸脱があってこそ、演奏者の個性を感じられ心地よく思われることが分かる。我々は演奏者情報の抽出にあたって演奏者の音楽的知識からくる独自の解釈部分、すなわちルール化が可能であるものに対してはエキスパートシステムを用い、次に感覚的、身体的なものからくる無意識なもの、すなわちルール化が不可能なものに対してはニューラルネットワークを用いて抽出処理を行う。

本稿では、エキスパートシステム内のルール生成についての報告を行う。同様の研究として、G. Widmerの研究⁽²⁾、片寄の研究⁽³⁾等がある。G. Widmerの研究ではLerdahlとJackendoffの理論、Narmourの理論をもとに音楽構造を解析し、音楽構造と実際の演奏から演奏ルールを抽出す

るものである。また、片寄らは「類型は同様に演奏される」⁽⁴⁾という仮定をもとに自己相関関数とNarmourの「暗意-実現のプロセス」の分析に於けるゲシュタルト的な制約を用いて学習を行い、特徴ある演奏データに沿って類型の演奏法を抽出している。これらの研究では、人工知能を用いた音楽情報処理へのアプローチであるが、我々はエキスパートシステムとニューラルネットワークを効果的に組み合わせたシステムを構築しており、この様な研究は他には見られない。

2. システム構成

本システムをFig.1に示す。

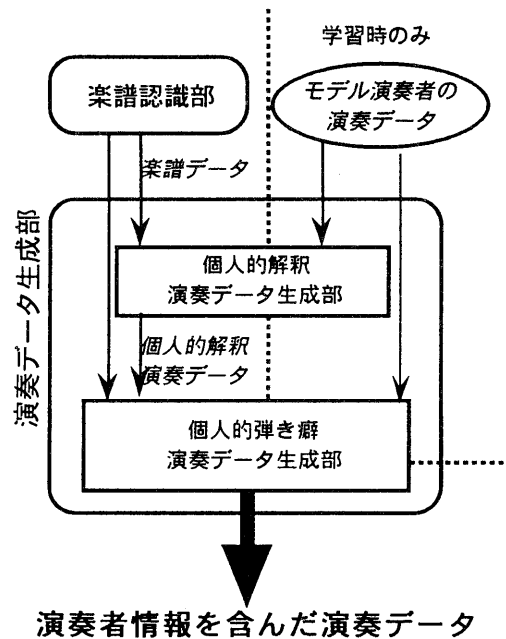


Fig.1 Flow diagram of performance data generation.

本システムは大きく2つの部分に分かれている。楽譜認識部では、イメージスキャナによって市販の楽譜を画像データ化し、ニューラルネットワークを用いて画像データから演奏のための情報の抽出を行っている。また、情報の認識時において五線はノイズとなる。そこで、ニューラルネットワークを用い五線の消去を行った上で認識を行っている⁽⁵⁾。

本システムのもう一つの部分が、演奏データ生成部である。演奏データ生成部では、楽譜認

識部で取り出された楽譜データを元に、演奏データを生成する。演奏データ生成部は更に2つに分かれており、1つが個人的解釈演奏データ生成部、残る1つが個人的弾き癖演奏データ生成部である。個人的解釈演奏データ生成部では、楽譜データをもとに音楽知識を用いて演奏者が行う独自の解釈をエキスパートシステムで生成する。また、個人的弾き癖演奏データ生成部では、楽譜データと個人的解釈演奏データを参照し演奏家の感覚的、身体的な無意識な演奏をニューラルネットワークを用いて生成する⁽⁶⁾ ⁽⁷⁾。また、生成された演奏データはMIDI楽器に入力され実際の演奏となる。今後は、個人的解釈演奏データ生成部について述べる。

3. ルールの自動生成

現在提案されているルール生成法の中には、音楽知識を取り入れたルールを用いた研究が多く見られる。しかし我々は音楽活動に対し情報処理的な立場からアプローチを行い、その限界を確かめた上で、音楽知識を用いたルールに取り組む考えである。

本節では、従来我々が行ってきたルールの自動生成法⁽⁸⁾について説明を行う。本研究では、メロディーラインに着目し、かつ演奏パラメータのうちヴェロシティ（音量）のみを扱っている。ルールで扱う情報は楽譜データから容易に得られるものである。例をTable1に挙げる。

Table 1 Attributes.

強弱記号, 速度記号, コード, 繰り返し拍子, 曲中に於ける位置

ここで、演奏者が演奏するときにはヴェロシティを変化させるが、その変化の要因のうち最も重要な情報は楽譜上の強弱記号であると考えられる。つまり、演奏者は強弱記号に従いヴェロシティに大局的な変化を持たせ、更にその他の情報を用い局所的なヴェロシティ変化を加えると考えられる。そこで、Table1で挙げた情報のうち、強弱記号のみを扱うルールを生成し、生成されたルールを用いて大局的なヴェロシティデータを自動生成する。次に、強弱記号以外の情報を扱うルールを自動生成し、それらを用いて局所的なヴェロシティデータを生成す

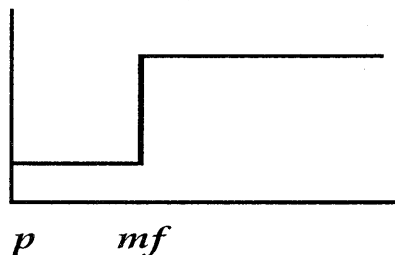
る。そしてこの2つのヴェロシティデータを加え合わせ最終的なヴェロシティデータとする。

○強弱記号を扱うルール

それでは、強弱記号を扱うルールの生成法を説明する。我々は強弱記号をそのヴェロシティに対する働きの違いから2つに分類している。

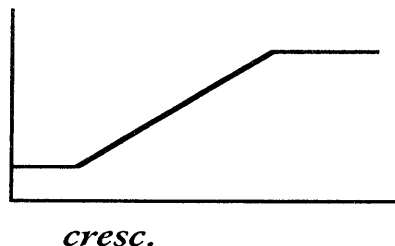
1. ラージ記号

f, mf, mp, p 等、記号が出現したときヴェロシティが階段状に変化するもの。



2. ミドル記号

cresc. <, decresc. 等、記号の有効範囲内でヴェロシティが継続的に変化し続けるもの。



この様に、強弱記号を2つに分類したためそれらを扱うルールも2種類必要となった。

- ① **if (p → mf) then +7.8**
- ② **if (mp → cresc. → f) then +1.2**

①はラージ記号に変化があったとき適合するルールであり、このルールが適合するとヴェロシティは+7.8変化する。②はミドル記号が存在するとき適合するルールであり、*cresc.*の有効範囲内でヴェロシティが+1.2づつ変化する。次に、それぞれのルールの具体的な生成法を説明する。

①のルールの場合、LHSには2つのラージ

記号の関係を設定する。また、ヴェロシティの変化はラージ記号が変化した位置で階段状に変化すると考えられるが、モデル演奏者の演奏データを調べるとある範囲内である傾きを持って変化していることが分かる。そこで、我々はモデル演奏者の演奏をラージ記号が変化する前後一定の範囲内で最小二乗法を用い直線近似する。この近似直線の縦軸方向の幅はラージ記号が変化したときのヴェロシティ変化量を表すので、その終点と始点の差をRHSの値とする。

次に、②のミドル記号を扱うルールの生成法を説明する。この場合LHSには、1つのミドル記号とその前後のラージ記号とする。ミドル記号の有効範囲内では、ある傾きを持ってヴェロシティが変化すると考えられる。そこで、ミドル記号の有効範囲内で一次近似を行いその傾きをRHSの値とする。

○強弱記号以外を扱うルール

強弱記号以外の情報を扱うルールの生成法を説明する。一般に、演奏者が演奏を行う際に見られる演奏の揺らぎは楽譜から得られる継続的な情報や、情報の変化によって起こると考えられる。そこで、我々は「情報は音符単位で取り出される」という仮定の下に音符の属性（特徴、情報）に基づくルール生成を行っている。

ここで生成するルールは、強弱記号以外の属性により変化するヴェロシティデータを生成する。そのため、ルール生成のもととなるヴェロシティデータには強弱記号の影響があってはならない。そこで、モデル演奏者の演奏から強弱記号を扱うルール生成で述べた近似直線をつなぎ合わせたものを差し引いたものをルール生成のためのデータとする。

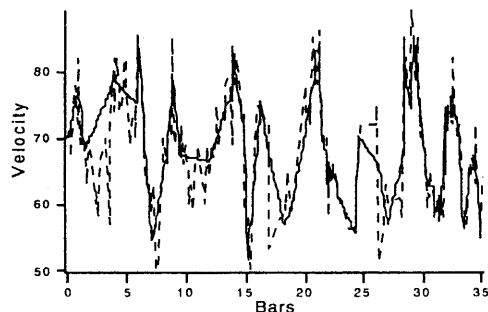


Fig.2 Played data and the approximation data.

Fig.2に於いて破線はモデル演奏者の演奏であり、実線は近似直線をつなぎ合わせたものである。つまり、ここで生成するルールは細実線から太実線を差し引いたデータをもとに生成されることになる。

ある楽譜に含まれる音符について、その属性とモデル演奏者の演奏から近似直線を差し引いたデータをもとに着目している音符とその直前の音符とのヴェロシティ差が記入されたTable3の様な表を考える。

Table 3 Attribute table.

音符	属性	ヴェロシティ差
A	A ₇ ,sf	1.2
B	C→F	2.5
C	G,riten.	-0.7
D	C→F	2.3

この様な表から、ルールのLHSを属性に、またヴェロシティ差をRHSとしたルールを考える。ところが、この様な方法を用いると同一のLHSを持ちながら異なるRHSを持つルールが生成されてしまう。そこで、同一のLHSを持つルールのRHSを平均し1つのルールにすることによりこの問題点を解決した。その結果、Table3から抽出されたルールは次のようになる。

If (A₇, sf) then 1.2
If (G, riten.) then -0.7
If (C → F) then 2.4

以上述べてきたルールの生成法をシステムの学習のために用意した数曲（学習曲）に対してコンピュータ上で自動的に行っている。

4. ルールの分割

我々が用いる学習曲はピアノ演奏のために作曲された音楽であるが、その表情は多種多様である。しかし、これまで述べてきたルール生成法ではルール生成の段階でLHSが同一の場合にはRHSを平均するという処理が含まれるため、生成されたルールは平均的なものになってしまう。このようなルールでは演奏データを生成する際、豊かな表情を持つ演奏データを生成

することができない。そこで我々は、ルール生成のもととなる曲を表情別に分類し、分類されたそれぞれの学習曲グループからルールを生成する手法を提案する。つまりこの手法により、華やかな曲用のルール、寂しげな曲用のルール等が生成されることになる。

表情別に学習曲を分類するために学習曲に対しSD法と因子分析を行う。今回、SD法で用いた指標のうちいくつかを次に挙げる。

堅い	⇔	柔らかい
明るい	⇔	暗い
力強い	⇔	弱々しい
動的	⇔	静的

この様な指標を用い、学習曲をグループに分ける。実際に、学習曲25曲についてSD法、因子分析を行った結果をFig.3に示す。

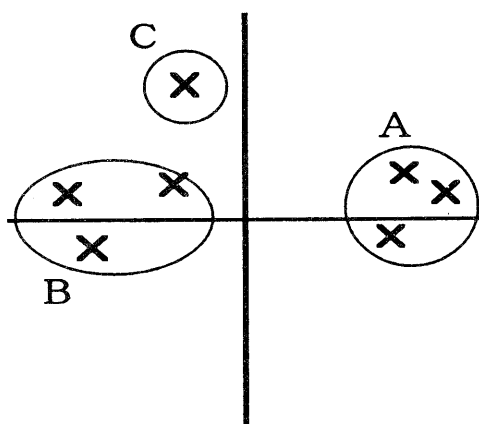


Fig.3 Result of factor analysis.

この結果から、学習曲は3グループに分類できることが分かる。そしてそれぞれのグループに対して3節で述べたルール生成法を用いルールを生成する。

5. 最適ルールグループの選択

4. 節で述べた手法により、表情別のルールが生成されたが、演奏データを生成する際にはそれらのルールグループの中から最適なルールを選択する必要がある。演奏データを生成しようとする曲（未知曲）に対してもSD法、因子分析が行えれば良いが、SD法を行うためには

視聴実験が必要であるので無理である。従って、分割されたルールグループと楽譜データから得られる情報とを結びつけるため手法が必要となる。この楽譜から得られる情報を我々は印象決定要素と呼び、それらをTable4に示す。

Table4 Decision factor of impression.

音高平均, 音高分散, メトロノーム記号, 調性, 単位時間当たりの音符数

ここに挙げた印象決定要素は楽譜データから容易に得られるものである。そして我々は、印象決定要素とルールグループを結びつけるための手法を2つ提案する。

○印象決定要素の分布を用いた選択方法

始めに、分割されたルールグループごとに印象決定要素を求める。Fig.4の様に各グループについて求められた印象決定要素をまとめる。

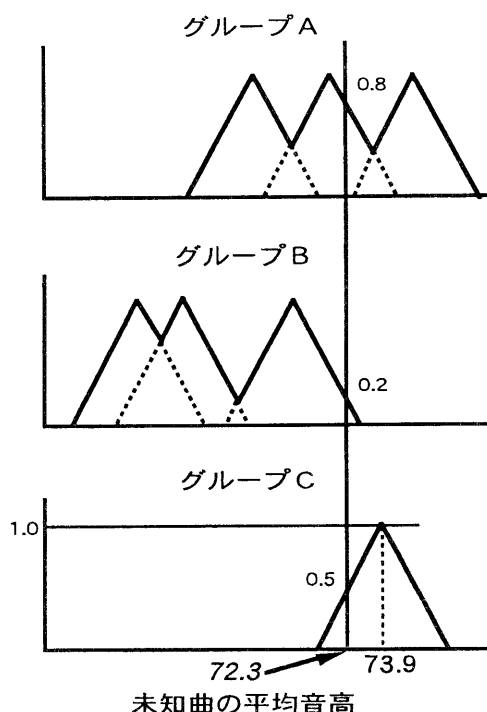


Fig.4 Spread of decision factor of impression.

Fig.4は平均音高について調べたもので、グループA, B, Cにそれぞれ3曲, 3曲, 1曲が含まれた場合の図を表している。平均音高の

値を3角形の頂点とし、重なった部分は上側の三角形を選択している。この時、未知曲の平均音高が72.3だったとするとFig.3にある縦の直線上がその値を表し、各グループとの交点がグループAからそれぞれ0.8,0.2,0.5となるのもっとも値の大きいルールグループAが演奏データ生成時に選択される。また、他に印象決定要素を考える場合には全ての印象決定要素についてグループ毎の和を求め、その和がもっとも大きかったグループを選択することにする。更に、各印象決定要素は印象を決定する度合いが異なると考えられる。そこで、グループ毎に和を求めるときに、それぞれの印象決定要素に重みの設定が必要と考えている。

○ニューラルネットワークを使った手法

ここで用いる、ニューラルネットワークの入力層の数は印象決定要素の数とし、出力層の数は分割されたルールグループの数とする。印象決定要素がネットワークに入力されると、出力層にあるユニットの内1つだけが発火する。出力層の各ユニットはルールグループを表しており、演奏データ精製時には発火したユニットが表すルールグループを選択することになる。

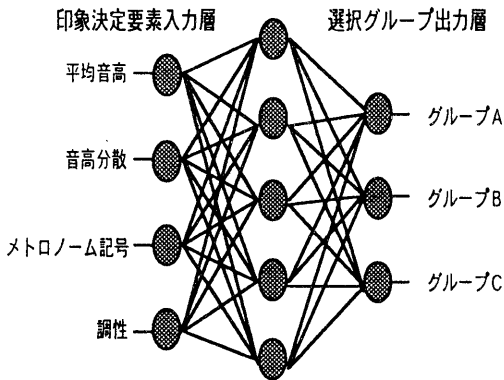
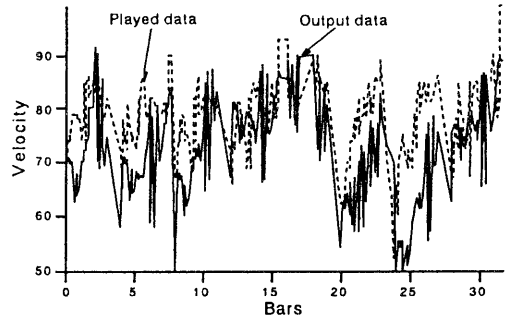


Fig.4 Neural network.

6. 実験結果

今回の実験ではブルグミュラー作曲「25のやさしい練習曲」を用い、モデル演奏者はプロのピアニストである荒健一氏とした。次に示す結果は、3節で述べたルールを用い「進歩」を未知曲として個人的解釈演奏データ生成部で処理を行った結果である。



この結果を見ると、数カ所でモデル演奏者の演奏と異なる部分があることが分かる。しかし、大局的な演奏の揺らぎはモデル演奏者の演奏に近い出力が得られている。この出力例は、個人的解釈演奏データ生成部からの出力であるのでモデル演奏者の局所的な揺らぎは捉えられていない。このデータが更に個人的引き癖演奏データ生成部を経ることにより更にモデル演奏者の演奏に近くなると考えられる。

誤差が生じたその他の原因として、4、5節で述べた手法が実現できなかったこと。また、楽曲構造などに関するルールがシステムに取り入れられていないためと考えられる。

7. まとめ

今回我々は、演奏者の特徴を模倣したピアノ演奏を行うシステムを開発を行うにあたり、エキスパートシステムとニューラルネットワークを用いた。この点が、我々のシステムの特徴であるが、どこまでをエキスパートシステムで行い、どこからをニューラルネットワークで行うかが大きな問題となっている。今後は、更に効果的な、エキスパートシステムとニューラルネットワークの組み合わせを考えていく必要があると考えられる。更に今後、音楽理論に基づいた新たな属性をルールに取り入れていく考えである。

【参考文献】

- 1)宮田 他：平4 春季音講論集2-7-4(平4.03)
- 2)G.Widmer：Proc. ICMC, pp.268-275(1993)
- 3)片寄 他：Proc.ICPR, pp780-784(1990)
- 4)片寄 他：平6 情処学研究会報告94-MUS-7-3
- 5)松本 他：平5 秋季音講論集1-6-5(平5.10)
- 6)小田 他：平6 秋季音講論集2-9-12(平6.10-11)
- 7)小田 他：第50回情処全大、1-373(1995)
- 8)白川 他：第49回情処全大、1-323(1994)