

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

～ニューラルネットワークによる局所部分に於ける演奏特徴の抽出～

小田安彦 白川健一 村上豊 梶川嘉延 野村康雄

(関西大学工学部)

我々は、特定演奏者の意識的な特徴、及び無意識な特徴を演奏者情報として備え、システムにとって未知な曲に対しては、この演奏者情報を加味した演奏者に勝るとも劣らない演奏を実現する自動演奏システムの構築を進めてきた。本システム構築にあたり、従来のルール処理だけから構成される自動演奏システムに加え、汎用性の高いニューラルネットワークを併用した。これにより演奏者のルールだけでは補い得ない特徴をも抽出する。今回、このニューラルネットワークによる処理について報告する。

Construction of automatic piano playing system wearing performer's characteristic

～Extraction of performer's characteristic in local part by Neural Network～

Yasuhiko Oda, Shirakawa Kenichi, Yutaka Murakami, Yoshinobu Kajikawa and Yasuo Nomura

(Faculty of Engineering, Kansai University)

We have been constructing on automatic piano playing system that wears performer's intentional and unintentional characteristic, which can play as if he really played. This system uses both the rules and the neural network. The neural network can extract better the performer's characteristic in local part of performance than the rules only. Here we report extraction of the performer's characteristic by the Neural Network.

1・はじめに

ピアノ演奏に於いて、楽譜にはその演奏を決定する基本的な情報しか記されていない。つまり、その曲に対する強弱、緩急などの表情付けは演奏者にかなり委ねられているのである。そのため、演奏者は楽譜上の発想記号、楽曲構造等を独自に解釈し演奏する。またこの際、演奏者の感覚的、身体的なものからくる弾き癖などの局所部分に生じる無意識な揺らぎが加味される。そのため、実際の演奏にはその演奏者の個性が感じられ、また演奏者の感情が伝わってくる。逆に、楽譜通りの機械的な演奏というのは、聴いているものに何の感動も与えてはくれないものである。

我々はプロの演奏者より、上記のような演奏に於ける意識的、及び無意識な演奏特徴をその人の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、その演奏者独自が持つ個性が感じられる演奏を実現するものである。

我々は演奏者情報の抽出に当たってエキスパートシステム、及びニューラルネットワークを併用して用いる。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使用し、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、或いは、ルールだけでは補い得ないものに対してはニューラルネットワークを使用し抽出処理を行う。

このように両者を併用することにより互いの持つ問題点を補い合うことができるのである。つまり、ルール処理であるエキスパートシステムに於ける未知なものに対する対応の低さをニューラルネットワークのそれに対する

汎用性の高さがカバーするのである。

ニューラルネットワークを用いた同様の研究例として Neil Todd⁽²⁾等がある。しかし、我々のようにエキスパートシステム、ニューラルネットワークを併用して演奏者情報を抽出するシステムは他にみられない。

今回、このニューラルネットワークに於ける処理について報告を行う。

2・システムの構成

Fig.1 にシステムの構成図を示す。

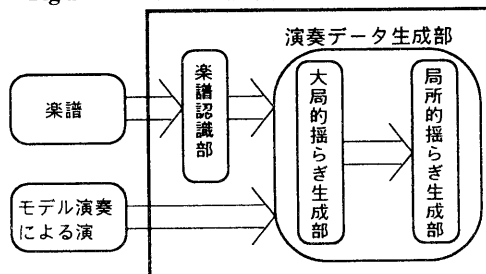


Fig. 1 Construction of System

システムは楽譜認識部⁽³⁾、大局的揺らぎデータ生成部⁽⁴⁾、局所的揺らぎデータ生成部⁽⁵⁾から構成され、大きく分けて2つの処理の流れを持つ。一方は特定演奏者のもつ演奏者情報の抽出、そして、他方は未知曲に対する演奏データの生成である。まず演奏者情報の抽出について述べる。

まず、楽譜認識部に於いてスキャナで楽譜を取り込み楽譜上のすべての情報を楽譜データに変換する。次に、大局的揺らぎデータ生成部に於いて、楽譜データ及び、その楽譜に対する演奏者による実際の演奏データから、演奏者が楽譜中の発想記号、楽曲構造等を、どのように解釈したかをエキスパートシステムにおけるルールの構築という形で抽出する。これにより演奏者の解釈部分だけからなる大局的揺らぎデータが生成される。次に、局所的揺らぎデータ生成部に於いて、楽譜データ、演奏データ、大局的揺らぎデータより、演奏者の持つルール

だけでは補い得ない無意識なものを局所的揺らぎデータとしてニューラルネットワークの学習という形で抽出する。以上の流れにより、演奏者の意識的な揺らぎ、無意識な揺らぎを大局的揺らぎデータ、局所的揺らぎデータとして抽出、学習するのである。

次に、演奏者情報抽出後の演奏データの生成について述べる。システムに未知な曲に対する楽譜を入力してやる。楽譜は先程と同様に楽譜認識部に於いて楽譜データに変換される、次に大局的揺らぎデータ生成部に於いて、この楽譜データに構築した演奏者の解釈ルールを適合させ、その未知曲に対する大局的揺らぎデータを生成する。局所的揺らぎデータ生成部に於いて、これを楽譜データとともにニューラルネットワークに入力し演奏者の無意識な揺らぎを加味する。以上の流れにより演奏者の意識的な揺らぎ、及び無意識な揺らぎを加味した、すなわち演奏者情報を加味した演奏データを生成する。

3・局所的揺らぎの抽出

ここでは、ニューラルネットワークによる、ルールだけでは補い得ない演奏者の無意識な揺らぎの抽出について述べる。前述したように抽出処理はニューラルネットワークによる学習という形で行われる。Fig.2に実際に用いるニューラルネットワークの構造を示す。

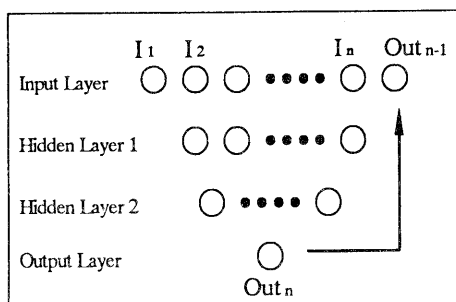


Fig. 2 Construction of Neural Network

我々は図に示すような4層からなる階層型ネットワークを用いる。一般にニューラルネットワークに於いて階層が薄いと入力データに非常に依存した出力がなされ学習曲に対しては有力であるが未知なものに対する対応性が低下する傾向がある。また、階層が厚ければ未知なものに対する対応性は向上するが、学習が困難になる傾向がある。そこで適当な階層として4層のものを用いている。そして学習手段としては強力な学習能力を有するバックプロパゲーション法を使用する。また学習単位は音符単位で行う。これは1つ1つの音符が楽譜上において最小の単位となっており演奏者の小節にまたがった揺らぎの抽出に都合が良いからである。

次に、入力データ、教師データについて述べる。前節で述べたように入力データとして楽譜データ及び、大局的揺らぎデータを、そして教師データとして演奏者の実際の演奏データを用いる。楽譜データは音高、音価、音長、個々の音符に加味される記号、複数の音符に加味される演奏記号などである。

今、学習しようとする音符を現在の音符、学習音符に対して直前の音符数個分を過去の音符、直後の音符数個分を未来の音符と定義する。

現在の音符に於ける学習に際して、過去の音符数個分に対する楽譜データ、大局的揺らぎデータ、実際の演奏データ、そして未来の音符数個分に対する楽譜データ、大局的揺らぎデータを入力データとして挿入する。これは、演奏者の実際に演奏を行う際の「楽譜の先読み」、「直前までの演奏の影響」等を考慮に入れるためである。

Fig.3 にペロシティ（鍵盤打鍵の強弱変化）に関する、システムによる未知曲に対する演奏データの再生結果を示す。なお演奏者情報抽出のための学習曲として「ブルグミュラー25の練

習曲」より5曲を抜粋，プロの演奏者として荒憲一氏の演奏データを使用した．システムに対する未知曲として同練習曲より「小さな嘆き」という曲を使用した．

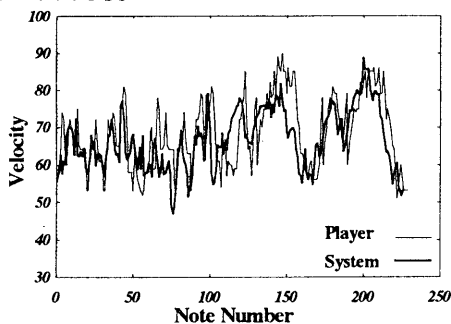


Fig. 3 Output of System

Fig. 3 に於いて，演奏者情報の抽出度，すなわち演奏者情報の反映の評価を相関値で表すと0.595であった．

4・多様なネットワークの使用

ここでは，単一のネットワークからなるシステムの問題点と多様なネットワークからなるシステムについて述べる．楽譜中には演奏の表情を決定する様々な演奏記号，発想記号が配置され，それらは曲中に於いて重要な役割を果たしている．例としてスタッカート記号，休符記号を挙げると，前者はおもに曲中の盛り上がり配置されたり，後者はどちらかという曲中の段落区切れ，すなわち盛り下がり配置されるなどその役割は異なれど曲中に於いて非常に重要なものである．そのため演奏者はその記号近傍に於いて特別な揺らぎをもつと考えられる．

これらの個別な特徴を非常に反映した曲を単一のネットワークにより同時に学習した場合，それぞれの特徴を十分に得難く，それぞれの特徴を相殺させてしまう恐れがある．また，ニューラルネットワークの学習に於いて，いわゆるローカルミニマムに陥りやすく学習に多大な時間を要することが起こりうる．そこで，

楽譜データより学習曲を休符記号の影響の強いグループ，スタッカート記号の影響の強いグループ，両者とは逆にそれぞれの影響の少ないグループという具合に分類しグループ毎にそれぞれの演奏者情報を学習抽出しシステム内に個別のネットワークを生成する (Fig.4)．これによりそれぞれのネットワークに各記号による影響を受けた演奏者の局所的な特徴を相殺することなく十分に備えることが出来る．

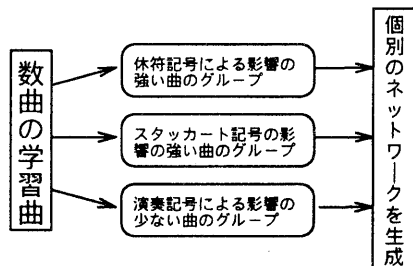


Fig. 4 Group Network

次に，この個別なネットワークを使用したときの未知曲の再生について述べる．まず未知曲に対する楽譜認識後，その楽譜データより状態ファイルを作成する．これはそれぞれの音符についてそれより過去（直前の音符），未来（直後の音符）数個分の内に休符記号，或いはスタッカート記号等がどの程度含まれているかをすべての音符について記したものである．再生時に於いてシステムは音符毎の処理時にこの状態ファイルを参照しその音符に対して適切なネットワークを使用して出力する．つまり状態ファイルは演奏に於いて各記号の有効範囲を決定する目安であり，言い換えるとシステム内の個別ネットワークの切り替えの目安となっている．

Fig.5 に多様なネットワークからなるシステムによる未知曲の再生結果を示す．なお学習使用曲として同練習曲より11曲を抜粋し，それぞれ休符用，スタッカート用，ノーマル用の3つのグループに分類して使用した．モデル演奏

者は先程と同様である。

また、状態ファイルに於ける有効範囲は処理音符を含め過去に4音符、未来に6音符である。

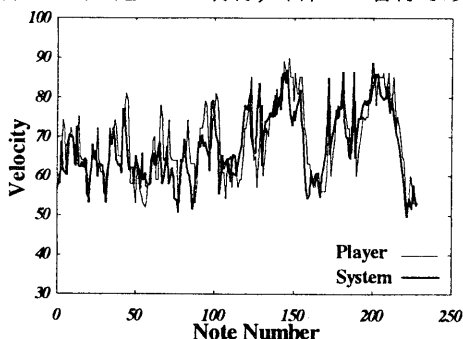


Fig. 5 Output by Several Networks

先ほどと同様に演奏者情報の抽出評価を相関値で表すと0.689であった。多様なネットワークを使用することにより、システムの出力がより演奏者に近づいたことがわかる。

今回は、非常に簡単なモデルとして休符、スタッカートだけについてネットワークを分けた。しかし更に着目する特徴を増やし、各ネットワークを有効に組み合わせ使用することにより向上があると思われる。

5・緩急変化に関する局所的揺らぎの抽出

これまで演奏時に於けるペロシティに関して演奏者情報の抽出を行ってきたが、ここでは演奏時の緩急変化に関する抽出について述べる。計算機上でピアノ演奏を行う際、演奏の緩急を表すパラメータとしてステップタイム、ゲートタイムがある。これらはそれぞれ、ある打鍵から次の打鍵までの時間、鍵盤を打鍵し続ける時間である。これらはペロシティと異なり楽譜中の音符からある程度の大きさが決定されるものである。すなわち、或る程度の基準値が楽譜から得られるという事である。そこで我々はステップタイム、ゲートタイムに関する演奏者の局所的な揺らぎ、特徴をこの基準値からの

差という形で抽出学習する。現時点に於いて緩急に関する大局的揺らぎデータは未だ生成出来ていない。そのため演奏者の意識的なものからなる大局的揺らぎデータの代わりに、楽譜から得られる基準値を挿入する。すなわち楽譜上の基準値に演奏者の局所的な逸脱を加味する形になる。

Fig.6, Fig.7 に各パラメータについて演奏者の実際の演奏データと演奏者情報抽出後のシステムによる未知曲に対する再生データとの比較グラフを示す。なお、楽譜上の基準値からの差の形で示す。学習曲は同練習曲より7曲を抜粋して使用した。また、システム再生用の未知曲は同練習曲の11曲目「せきれい」を使用した。

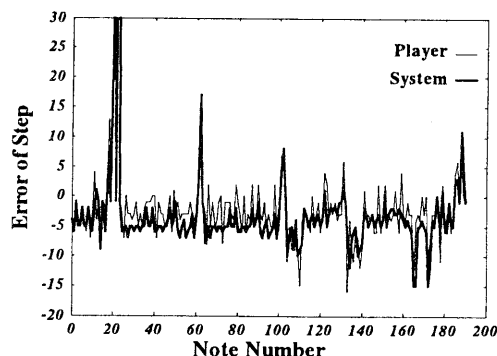


Fig. 6 Error of Step

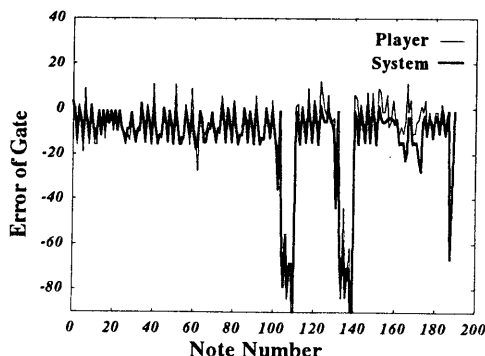


Fig. 7 Error of Gate

演奏者による演奏データとシステムによる演奏データとの相関値はステップタイムに於いては0.9771, ゲートタイムに於いては0.9663であった。

一曲の楽譜中に緩急に関する演奏記号, 発想記号はタイ, スラー等を除くと非常に少ない。

そのためベロシティに関する演奏者情報の抽出の時のように記号に関しての多様なネットワークの生成は困難である。やはり前述したように, 楽譜中の記号等だけではなく旋律パターンや楽曲中の起承転結等のような, より高次の音楽的要素毎にネットワークを生成する必要性がある。

7・まとめ

演奏者の揺らぎをルールによる処理だけで十分に抽出することは非常に困難なことである。仮に完全なルールを構築できたとしても非常に多くのデータベースが必要となり, また, 演奏者別にルールの構築を行うには多大な時間を有することになるであろう。そこで我々は, ルールだけでは補い得ない揺らぎに対してニューラルネットワークにより対応した。

このように, ルール処理であるエキスパートシステムとニューラルネットワークを併用してやることにより演奏者情報の抽出を行い, また未知曲に対して演奏者情報を十分に反映させた演奏が実現できた。

我々は演奏者情報を意識的なものによる大局的な揺らぎ及び, それ以外の無意識なものによる局所的な揺らぎとに分けて解釈を行ってきた。しかし現時点での大きな問題点として, どこまでをその区切りとするか判然としていないのである。ニューラルネットワークによる抽出処理は実際には局所的な揺らぎを“狭く深く”抽出させるのが望ましいのだが, 現在のシステムではニューラルネットワークに依存される部分が大きいと“広く浅く”なる恐れが

高い。そこで例として, 処理の流れを「ルール→ニューロ→ルール」或いは, 「ルール→ニューロ1(広く浅い)→ニューロ2(狭く深い)」のようにニューラルネットワークに対する準備範囲を狭める等, 今後さらに検討する。

【関連文献】

- 1) 宮田他: 平4 春季音講論集 2-7-4 (平 4.03)
- 2) Neil Todd: Journal of New Music Reserch (1994)
- 3) 松本他: 平5 秋季音講論集 1-6-5 (平 5.10)
- 4) 白川他: 第49 回情処全大 1-323 (1994)
- 5) 小田他: 平6 秋季音高論集 2-9-12 (平 6.10)