

4Z-5

ニューラルネットワークによる テンポに関する演奏者情報の抽出

○村上豊 熊谷俊行 梶川嘉延 野村康雄
関西大学工学部電子工学科

1. はじめに

ピアノの演奏において、演奏者は必ずしも楽譜通りには演奏せず、演奏者独自の特徴がみられる。これは、演奏者によるその楽曲の解釈、及び演奏者の感覚的・肉体的要因により生ずる無意識な揺らぎに依るものである。これら楽譜からのずれは、その演奏をより心地よい音楽性の高いものとするために必要不可欠なものであり、これがその演奏者の個性を形成する重要な要素となっている。

我々は、上記のような演奏者の音楽性の高い演奏を再現する方法として、この特徴を演奏者情報としてデータベースに備え、未知なる曲に対してこの演奏者情報を加味することにより、あたかも実際にその演奏者が演奏したかのような演奏を実現する自動演奏システムの構築を行ってきた⁽¹⁾。

この演奏者情報を抽出する手法として、従来はAI的な手法がその主流を占めていた⁽²⁾。しかし、このような手法は既存の曲に対する解析に於いては有効であるが、未知曲に対しては柔軟性を欠きあまり良い結果が得られていない。そこで、未知なものに対しても柔軟に対応できるニューラルネットワークにいち早く着眼し、これらを併用する方法を検討してきた⁽³⁾。以下に両者を併用し、演奏者情報の抽出を行う方法について述べる。

2. 演奏者情報

ピアノの演奏に於いて、演奏者独自の特徴が見られるパラメータとしては、鍵盤打鍵の強弱を表すヴェロシティと時間的な緩急変化を表すテンポが挙げられる。またテンポは、音符の出現するタイミングを表すステップタイムとその音符により鍵盤が押されていた時間を表すゲートタイムがある。ここで、これらのパラメータ値の揺れが演奏者情報となる。

今回は、特にテンポに関する演奏者情報を抽出する手法について述べる。

3. 演奏者情報の抽出

演奏者情報を抽出する手段として、ルールを用いた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する。しかし、演奏者情報にはさまざまな要因が複雑に組み合わせられており、演奏者情報の中で各々どの部分をエキスパートシステムにより抽出し、どの部分をニューラルネットワークで抽出すべきかは、一概には言えない。そこでまず、各々を個別に用いて演奏者情報の全てを抽出することを目指す。

テンポに関する演奏者情報を抽出する時のシステムの概念図を Fig. 1 に、未知曲に演奏者情報を加味した演奏データを生成する時のシステムの概念図を Fig. 2 に示す。

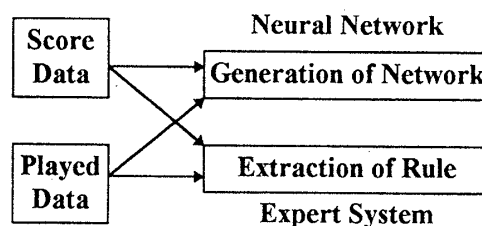


Fig. 1 Extracting System

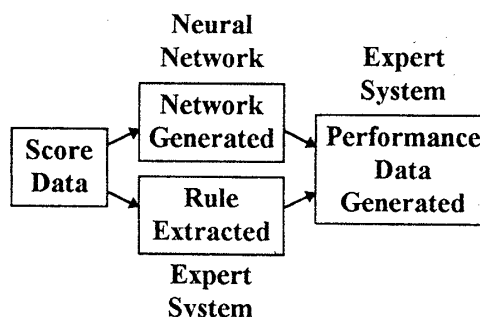


Fig. 2 Generating System

未知曲の認識再生時は、未知曲の楽譜データから、生成済みのルールとネットワークを用いて演奏データを生成する。そして、両者の演奏データの中でより演奏者情報を抽出していると考えられる部分をエキスパートシステムにより組み合わせ最終的な演奏データを出力する。

今回、これまでテンポについては行われてこなかった、ニューラルネットワークを用いた演奏者情報の抽出について述べる。

Extraction of a Performer's Characteristic on Tempo
by Neural Network

Yutaka Murakami, Toshiyuki Kumagai,
Yoshinobu Kajikawa, Yasuo Nomura

Department of Electronics, Faculty of Engineering,
Kansai University

3-3-35 Yamatecho, Suita, Osaka 564, Japan

4. ニューラルネットワークの構造

使用するニューラルネットワークは、4層構造の階層型であり、学習アルゴリズムとして誤差逆伝搬法を用いた。学習は音符単位で行い、入力データとしては学習する音符だけではなく、その前後の音符も同時に入力する。また、学習音符以前に演奏された音符に対しては、実際の演奏データも同時に入力する。これは、未知曲の認識再生時にそれまでの出力結果を入力するリカレント型にするためである。それでは、具体的な例を Fig. 3 に示す。

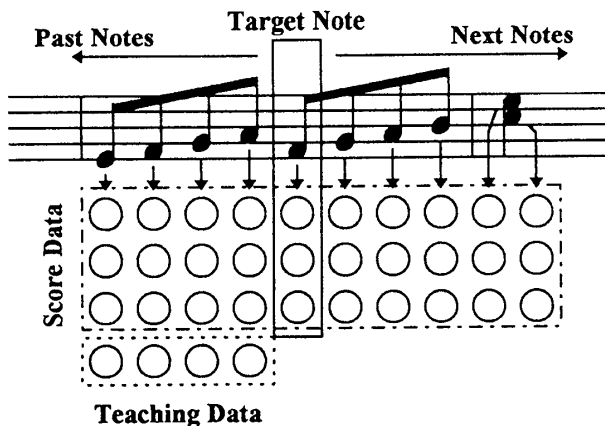


Fig. 3 Input Data

この例の一番右にある様な和音に対しては音高の高い順に演奏されると仮定する。もしここで、実際の演奏順序が異なる場合はステップタイムが負の値を採ることになる。

5. 実験結果

今回楽譜データとして、1音符あたり音高、理論的なステップタイム、理論的なゲートタイムの3ユニットを用いたもの、そして、スタッカート、スラーの存在を表すユニットを加えた5ユニットを用いたもので実験を行った。ここで理論的なというのは、曲全体を一定の速度で演奏したものととして算出した値のことである。

ステップタイムを学習した場合の結果を Fig. 4 に示す。グラフでは、分かり難いが、両者を聴き比べると若干1音符あたり5ユニットを使用したものがよい結果が得られた。ここで、両者のネットワークは、学習曲に対して同程度の認識再生結果がでるように学習させている。つまり、5ユニットを用いた方がより未知曲に対応できる、言い換えれば、より演奏者情報を抽出できたわけである。また一般的に、入力ユニット数が増加したことにより学習により時間がかかるのではないかと考えられる。しかし実際にはより少ない学習回数で所望の誤差まで減少するため、5ユニットを用いた方が学習時間が短くすむ。

また、ゲートタイムについてであるが、これはあまりよい結果が得られなかった。

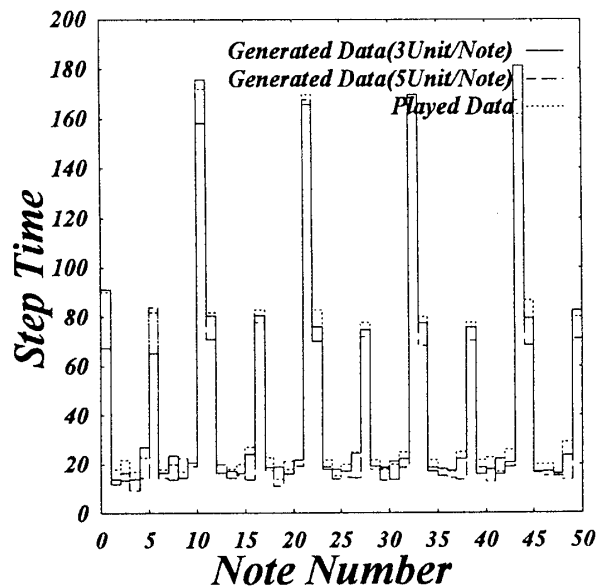


Fig. 4 Step Time

6. 考察

今回ゲートタイムについてあまりよい結果が得られなかったのは、ステップタイムの揺らぎがあまり大きいとリズムが狂いとて聴けないものになるのに対し、ゲートタイムはその音符に対してのみ効果を持ちその音が鳴り終わると他に影響を及ぼさない事から揺らぎが大きく成るためであると考えられる。

また、ステップタイムに関してもまだ聴いていて不自然に感じる場所が多々ある。これは、考慮している楽譜の範囲が狭く限られているため、曲全体の構成に関してなにも考慮に入られていないためであると考えられる。これを考慮するには、曲の起承転結を把握する必要や、だんだん遅くといった一概には数値化できない演奏記号を考慮する必要性からエキスパートシステムに頼らなければならないと考えられる。そこで、Fig. 1, Fig. 2 のニューラルネットワークの処理部分の前段に曲全体の流れに対する音符の位置付けを表すデータを生成するエキスパートシステムを加えることにより曲の起承転結のはっきりした演奏データを生成できると考えられる。今後は、このエキスパートシステムが具体的に生成するものとして、拍単位、または小節単位のステップタイムの検討を続ける。

【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) 小田 他：第49回情処全大 1-323 (1994)