

ニューラルネットワークを用いた 演奏者が持つ大局的揺らぎに関する検討

波多正臣 岡西正 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

ピアノを演奏する場合、演奏者は楽譜上から得られる情報をもとにして行うだけでなく、自身の意識的、無意識的な揺らぎを加味して奏でる。その結果、演奏する曲自体に演奏者独自の個性が備わり音楽性に富んだすばらしいものとなる。

我々はこれまで、演奏者が持つ特徴をニューラルネットワークによって抽出し、それを備えた自動演奏システムの構築を研究してきた^{(1),(2)}。本稿では、この自動演奏システムを構築するにあたって、演奏者が持つ大局的揺らぎについて検討しニューラルネットワークを用いて演奏者の特徴を学習・認識させた結果について述べる。なお、この実験に際しメロディライン上の音符のみを対象とし、ヴェロシティ（鍵盤打鍵速度）に限定し進めていく。

2 演奏者の特徴と大局的揺らぎ

ピアノを演奏する場合、演奏者独自の特徴は主にヴェロシティとテンポによって表現される。ヴェロシティは鍵盤打鍵速度、すなわち音量でありテンポは時間的な変化で表される。今回はヴェロシティについて検討しているののでこれについて述べると、演奏者が同じ曲を複数回演奏したときの演奏揺らぎというものは Fig.1 に示すように、ある大きな揺らぎに支配されていることが見てとれる⁽³⁾。しかも、これは演奏者が違えばそれぞれに独自の異なった揺らぎが存在するのである。つまり、この大きな揺らぎ、すなわち大局的揺らぎに多分の演奏者の特徴が含まれていると考えられる。このような揺らぎを抽出するには、複数回演奏したデータを平均すればよいのだが、実際にシステムを組む上のことを考慮すれば、演奏者が何回も同じ曲を演奏したデータを扱うといっ

たことはまずない。そこで大局的揺らぎの抽出法として、Fig.2 で示されるようにある一つの演奏波形を滑らかにすることによって取り出すものとする。

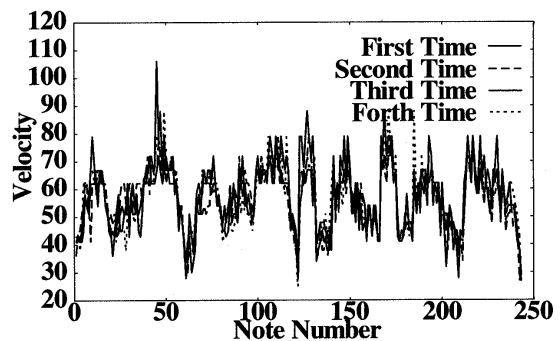


Fig.1: An example of fluctuations of a model player's performance to the same tune on the piano.

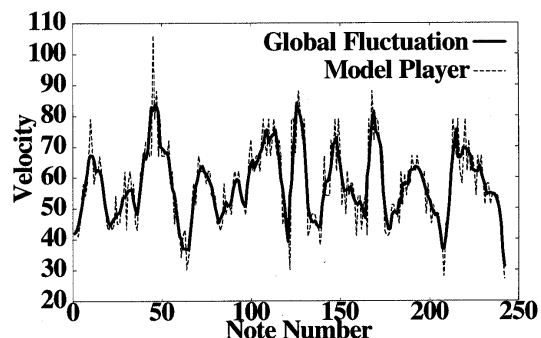


Fig.2: Method of extracting the global fluctuation from one played data.

3 システム構成

3.1 フレーズ処理

これまでのニューラルネットワークによる学習は音符単位で行ってきた。しかしながら大局的な流れを考えたとき、それが音符間の繋がりというよりはあるまとまった音符の塊、つまりフレーズと言う単位で構成されていると考えられる。そこで、大局的揺らぎデータをフレーズ単位で処理し生成する。また、フレーズ単位と音符単位で処理する場合のデータ数を比較すれば、前者はそれを著しく削減できることがわかる。つまり、ニュー

A Study on Global Fluctuations of a Player
by Neural Network.

Masaomi Hata, Tadashi Okanishi,

Yoshinobu Kajikawa, Yasuo Nomura

Department of Electronics, Faculty of Engineering,
Kansai University

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

ーラルネットワークに要する学習時間を大幅に短縮できるのである。

実際にどのようにして演奏データをフレーズ単位で処理するのだが、Fig.3 に示すようにあるフレーズ区間内のヴェロシティをレベルとスロープに分解する。ここでいうレベルとはフレーズ区間内におけるヴェロシティの平均値であり、スロープはそれらを直線近似した傾きの値である。

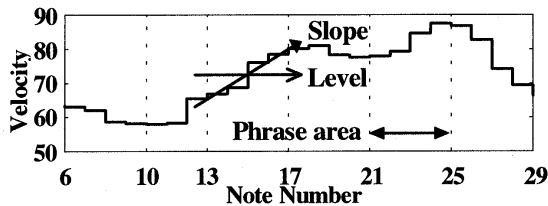


Fig.3: Generation of phrase data.

3.2 データの設定

まず用いるニューラルネットワークは4層構造の階層型であり、学習アルゴリズムは誤差逆伝播法である。次に、ニューラルネットワークに入力するパラメータを Table 1 に示す。これらを入力するわけだが演奏者が演奏する場合、これまでの演奏やこれからの演奏も影響を受けると考えられる。よって現在のフレーズの他に過去2フレーズ、未来4フレーズ分も入力する。また、これまでの演奏である演奏済みデータも入力する。認識時には、演奏済みデータの代わりにニューラルネットワークの出力を用いる。

Table 1: Input parameter.

音高レベルとスロープ	<i>mf</i> 等の演奏記号
演奏済みレベルとスロープ	<i>cresc.</i> 等の演奏記号
フレーズの位置	コード
繰り返し回数	曲の構造

4 学習と未知曲に対する結果

以上の手法を用いて実際にニューラルネットワークに学習させた結果を述べる。今回のシミュレーションに使用した曲はブルグミュラー作曲「25の練習曲」の中から学習曲を4曲、未知曲を2曲として行った。それぞれの曲名を Table 2 に示す。また、演奏者には苦瓜瑞恵講師（常盤女子短期大学）の演奏データを用いた。

未知曲を認識させた場合の相関係数と二乗誤差を Table 3 に示す。Table 3 を見ると、相関係数が0.7以上であり、二乗誤差もある程度小さいものとなっている。実際にどのようにになっているかを Fig.4 (Tarentelle) に示す。これを見ると、演

Table 2: The tunes used for this simulation.

	Tune name
Learning tunes	Innocence
	La gracieuse
	Tendure fleur
	Douce plainte
Unknown tune	Adiue
	Tarentelle

Table 3: The values of correlation coefficient and square error between a model player and the system for each result.

Tune name	Correlation coefficient	Square error
Adiue	7.355E-01	5.851E+01
Tarentelle	7.168E-01	7.465E+01

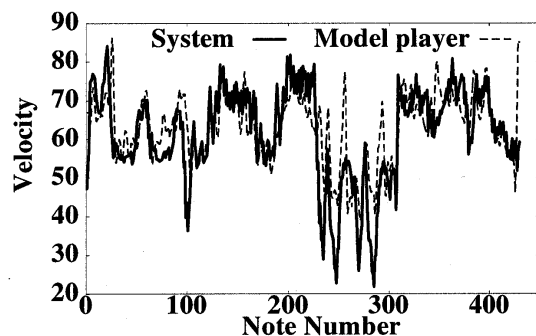


Fig.4: The result of simulation for the unknown tune, Tarentelle.

奏者と異なる点があるものの、演奏者の特徴を反映した結果が得られているといえる。

5 まとめと今後の方針

本稿では、演奏者が持つ大局的揺らぎをフレーズという概念を用いてニューラルネットワークに認識をさせることに成功した。しかしながら、実際の演奏を聞く上で何が人間に影響を与えるかということ、ヴェロシティというよりむしろテンポであると言えよう。今後は、ヴェロシティを考慮しつつテンポに関する研究を進める所存である。

【参考文献】

- (1) 宮田他：平4 春季音講論集,2-7-4,Mar. 1992
- (2) 小田他：情処学音楽情報科学研報,pp.7-12, May,1995
- (3) 小田他：平6 秋季音講論集 2-9-12(平 6.10)