

特定ピアノ演奏者の特徴抽出による 自動演奏システムの構築

4G-7

○木野弘樹 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

一流のピアノ演奏者は、今まで培った高度な技術、経験、譜面の解釈、そしてその表現方法を駆使して演奏を行うことにより芸術性豊かな演奏を行い、聞く人に感動を与える。我々はその演奏者の特徴や弾き癖を抽出することで、演奏者の特徴を持った演奏を実現できる自動演奏システムを構築してきた^{(1),(2),(3)}。このシステムの核となるものは、演奏データ生成システム PDGS (Performance Data Generating System) であり、我々は 2 基の ES-PDGS (エキスパートシステムを使用した PDGS) 及び、3 基の NN-PDGS (ニューラルネットワークを使用した PDGS) を多段階に用いて構成される拡張型自動演奏システム E-APS(Enhanced Auto Performance System)の構築を行っている。これは、譜面中のメロディラインだけでなく、全部の音符を対象とする画期的なシステムである。本稿では、このシステムに前処理を導入し、和音部分をニューラルネットワークに入力する際の入力データを決定する。また、特徴抽出の対象となる MIDI パラメータである Velocity, Steptime, Gatetime それぞれに個別のニューラルネットワークを使用することにより、このシステムをより洗練させる。

2. E-APS の概要

図1に E-APS の構成を示す。演奏者の特徴を持った演奏データは演奏データ生成システム PDGS により生成される。ここで、この E-APS の処理過程について説明する。まず、ES-PDGS1 及び ES-PDGS2

により、右手と左手の基本的な演奏ルールをエキスパートシステムを用いて導出する。次にそれらの演奏ルールから得られた演奏データを NN-PDGS1, NN-PDGS2 に入力して、右手と左手の特徴をニューラルネットワークによって学習する。最後にそれぞれの NN-PDGS から得られた演奏データを NN-PDGS3 に入力して、全部の音符をニューラルネットワークにより学習する。ひとたび学習が完了すれば、その演奏者にとって未知の曲であっても、楽譜データを与えるだけであたかも演奏者が演奏したような演奏データを得ることができる。

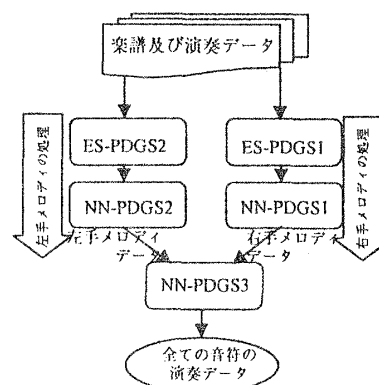


図1 自動演奏システム E-APS の概要

3. 前処理

ここでは、前処理の方法を詳しく述べる。E-APS の問題の一つとして、NN-PDGS3 に入力する和音データが不足するということが挙げられる。すなわち、NN-PDGS1,2 それぞれから出力されるデータはメロディラインのみのデータであることから、NN-PDGS3 に入力する際にメロディラインのデータは存在しても、その他の和音構成音のデータが不足しているのである。従来は、これを補うために、和音構成音のデータはその和音に属するメロディの音符と同一のデータを使用していた。しかしながら、譜面上では同時刻に発音される和音も、実時間では多

Construction of the Automatic Performance System by the Extraction of Piano Performer's Characteristics.

Hiroki Kino,

Yoshinobu Kajikawa, Yasuo Nomura

Department of Electronics, Faculty of Engineering,

Kansai University

3-3-35, Yamate-cho, Suita-shi, Osaka, 564-8680, Japan

少のずれがあるのである。よって、この演奏者の曲を前処理用の曲として多数用意し、和音に属するメロディとその他の和音構成音の打鍵位置の差がどれほどあるか統計を取り、NN-PDGS3に入力する和音部分のデータを、その統計データの確率分布により決定する。

4. シミュレーション

4.1 学習曲認識による過去未来音符数の決定

ニューラルネットワークの入力として、現在の音符はもちろん、過去の音符、未来の音符も入力しているのであるが、調査の結果 Velocity, Steptime, Gatetime それぞれのパラメータに最適な過去、未来の数というのが違うことがわかった。よって、今回は過去未来の数を変化させることによって7種類の過去、未来の数を用意し(表1参照)、それぞれについて学習曲を学習後のニューラルネットワークに認識させ、相関係数を調べることで、認識に使用する過去、未来の数を決定した。これはすなわち、それぞれのNN-PDGS内部に3つのニューラルネットワークを存在させるということである。シミュレーションには、「ブルグミュラー25の練習曲」の中から特徴抽出用の学習曲、及び未知演奏データ生成用の未知曲を選択した。

学習曲を認識させた結果を表1に示す。なお、実際はNN-PDGS1~3において、それぞれ相関係数を調べ、最適な過去、未来の音符数を決定するのであるが、今回は紙面の都合上、NN-PDGS3の結果のみを載せる。なお、網掛けの部分が、それぞれ選択された過去未来数であることを示す。

表1 学習曲認識による、システムと演奏者の相関係数

両手(NN-PDGS3)			
0-6	0.8063	0.8918	0.9049
1-5	0.8238	0.9040	0.9231
2-4	0.8065	0.8904	0.9266
3-3	0.8175	0.8736	0.9060
4-2	0.8246	0.8460	0.9094
5-1	0.8188	0.8425	0.8758
6-0	0.8290	0.8602	0.8785

左端の数値は、(過去の音符数) - (未来の音符数) を示す。

4.2 未知曲認識結果

未知曲を認識させ、従来のシステムとの比較を行う。なお、本来はE-APSの出力(=NN-PDGS3の出力)は両手の音符が混在して出力されるのであるが、右手なら右手、左手なら左手といった音符列のつながりが重要であると考え、評価するにあたり、右手、左手それぞれに分けることにする。そしてその結果を表2に示す。

表2 従来と今回のE-APSの数値比較

		相関係数	二乗誤差
Velocity	(従来)	0.2957	3.3537E-02
右手	(今回)	0.4979	3.1232E-02
Velocity	(従来)	0.4560	5.3951E-02
左手	(今回)	0.5690	4.7585E-02
Steptime	(従来)	0.3116	1.8929E-02
右手	(今回)	0.3644	1.7325E-02
Steptime	(従来)	0.6672	3.3416E-02
左手	(今回)	0.6839	3.2906E-02
Gatetime	(従来)	0.8269	1.3347E-02
右手	(今回)	0.8801	1.6270E-02
Gatetime	(従来)	0.9187	8.5637E-02
左手	(今回)	0.9199	1.3722E-01

表2より、二乗誤差で若干今回のシステムの方が数値の悪い所があるものの、やはり今回のシステムが優れているということがわかる。Velocity用、Steptime用、Gatetime用とそれぞれにニューラルネットワークを用意したので、当然といえば当然の結果であるが、かなり飛躍的に向上した部分もあり、今回のシステム改善は有効であるといえる。

5. まとめと今後の方針

本稿では、E-APSに前処理を加えることによりNN-PDGS3入力時における問題を解決し、それぞれのNN-PDGSにおいて3つのパラメータそれぞれに、個別のニューラルネットワークを使用することでシステムの改善を行った。その結果、従来に比べて良い結果を出すことができた。しかしながら、Steptimeの右手の認識率が悪いなどといった問題もあるので、それを改善することが大きな課題となる。

【参考文献】

(1)村上他：第54回情処全大,Vol2,pp.269-270,

March.1997

(2)波多他：第55回情処全大,Vol2,pp.4-5,Sep.1997

(3)岡西他：第58回情処全大,Vol2,pp.95-96, March.1999