

フィードバック機能をもつ自動演奏システムの試み

玉木明和 姚鳳会 加藤清史

九州工業大学 工学部
〒804 北九州市戸畑区仙水町1-1
E-Mail : tamaki@comp.kyutech.ac.jp

人間が音楽を演奏する場合、演奏者も自己の演奏の流れとともに感情が変化する。すなわち、演奏している音楽を聴きそれを演奏に反映させている。自動演奏は単調で機械的であるといわれているが、より良い演奏を行うために音楽を聴くことによって生ずる感情変化をフィードバックすることによって、演奏制御に反映させることを試みた。音声から音楽情報を検出することは難しいので、音楽を聴くことはその演奏のMIDI信号を入力することで代用した。ニューラルネットを用いて、抽出した音楽情報から感情変化を行った。ニューラルネットの学習データは音楽を聴いた被験者から得られたものである。演奏中のMIDI信号から音楽特徴を抽出し、ニューラルネットを用いてその特徴から演奏制御の修正量を算出し、より良い自動演奏ができるように試みる。

On the trial of Automatic Music Performer with Feedback Control

Akikazu Tamaki, Feng-Hui Yao, Kiyoshi Kato

Faculty of Engineering, Kyushu Institute of technology
1-1 Sensui Tobata Kitakyushu, 804, Japan
E-Mail : tamaki@comp.kyutech.ac.jp

The authors try to make an automatic music performer with feed back control. The control information is generated by the pseudo emotion, which is changing by hearing the some musical sound in the simulation model. The MIDI signal is used for the musical input data instead of audio musical sound which is generated by a musical instrument. Therefore, the musical feature is detected from the MIDI signals, by which the performance information is communicated among the electronic musical instrument devices. In this system, the feedback control information is generated by the pseudo emotion which is changed according to the extracted musical feature by using the neural networks.

1. はじめに

情報処理の対象として種々のものが扱われ始め、最近では、感性情報処理といわれる人間の感性に関わる情報の研究が行われている〔1〕。コンピュータミュージックなどに代表される音楽に関する研究は古くから行われてきたが、感性情報処理の一環として種々の方法で音楽とコンピュータに関する研究が盛んに行われている〔2〕。人間の情報処理機構の一部を利用したニューラルネットは多くの分野で応用されているが、音のデータベースにニューラルネットを応用したものもある〔3〕。

著者らはコンピュータを用いた自動演奏システムの開発を行っているが、それは演奏の必要な情報を楽譜情報と演奏者情報に分けて扱っている〔5〕。演奏者情報は演奏時の演奏者の感情変化を反映するための情報であり、自動演奏をより人間的なものに近づけるために用いるものである。

人間は音楽を聴くと感情が音楽の流れとともに変化する〔4〕〔6〕〔7〕。演奏者も自己の演奏の流れとともに感情が変化する。この感情変化は彼が育った音楽環境に左右されるものである。

人間が音楽を演奏する場合、その演奏に乗って演奏することがある。すなわち、演奏している音楽を聴きそれを演奏に反映させている。自動演奏は単調で機械的であるといわれているが、より良い演奏を行うためにフィードバック機能の導入を試みたので報告する。

音楽を聴くことによって生ずる感情変化をフィードバックすることによって、演奏制御に反映させることを試みた。音声から音楽情報を検出することは難しいので、音楽を聴くことはその演奏のMIDI信号を入力することで代用した。感情抽出及び変化機構にはニューラルネットを用い、学習データは音楽を聴いた被験者から得られたものである。ニューラルネットの学習データを音楽環境として考えることもできる。

2. システムの概要

図1に示すように、本システムは演奏制御部、MIDI楽器、感情抽出部、感情変換部から構成される。演奏制御部は楽譜情報、演奏者情報及び感情変換部からの情報をもとに音量、音高を示すMIDI信号を作成し、MIDI楽器に送る。MIDI楽器は送られたMIDI信号に従って音楽を演奏する。また、感情抽出部は同じMIDI信号から感情を抽出し感情変換部に送る。感情変換部は送られた感情から演奏を制御するパラメータ P_F を生成する。

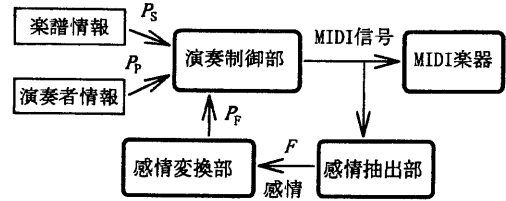


図1 システムの概略

3. 演奏制御部

楽譜がもつ楽譜情報は音高、音量 (v)、テンポ (t)、奏法 (s) などを指示し、楽譜に忠実に演奏すると機械的な演奏といわれる。楽譜で指示された音高は変えることはできないが、より芸術的に演奏するために演奏者は他のものがある程度自由に選んで演奏する。これらのパラメータを $P = (v \ t \ s)^T$ で表す。このパラメータの値は演奏の進行に伴って変化するものである。本システムではパラメータ P を適当に変化することにより、より良い演奏を行うことを目指している。人間の演奏者によって楽譜を解釈することにより得られるパラメータのシーケンスを演奏者情報と呼ぶ。楽譜情報及び演奏者情報は、演奏時における時系列データとして演奏制御部に入力される。楽譜情報のパラメータを P_s 、演奏者情報のものを P_p 、感情変換部から得られるものを P_f とすると、実際に演奏制御に用いられるパラメータ P はそれらの総和で、

$$P = P_s + P_p + P_f \quad (1)$$

である。 P のもつ情報によって機械的な演奏をより人間的なものにすることができる。

4. 感情抽出部

音としての音楽から音楽感情を抽出すべきであるが、実際の音声から音楽情報を抽出するのが困難である。したがって、感情抽出部は演奏されるMIDI信号から感情の変化を抽出する。感情抽出部は図2に示すように、信号変換機構、特徴抽出機構、感情抽出機構、感情変化機構などから構成される。ここでは音楽を聴いたときに感じる感情として、「明るい-暗い」、「軽快-退屈」、「和やか-緊張」、「安心-淋しさ」の4対の感情の組を用いた。4つの感情対を、それぞれ、「明暗」、「軽退」、「和緊」、「安淋」とも書く。

・信号変換機構

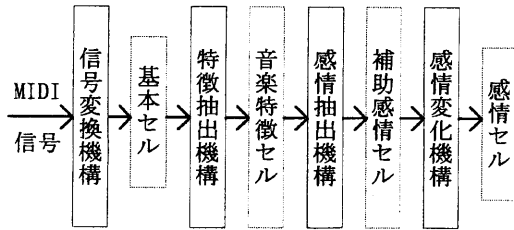


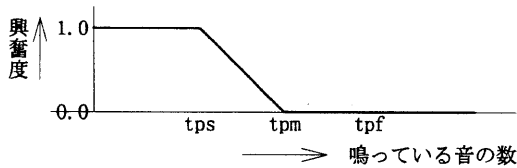
図2 感情抽出部のブロック図

基本セルは音階上のそれぞれの音高に対応する。対応する音高の音が鳴っていれば興奮し、音の大きさに応じて0～1の値をとる。信号変換機構はMIDI信号から余韻を考慮して基本セルを興奮させる。

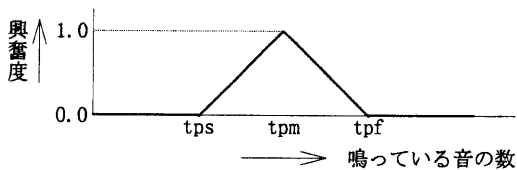
・特徴抽出機構

特徴抽出機構は基本セルの興奮状態によって、テンポ、音高、和音などの音楽特徴44個を抽出しニューラルネットワークに入力する。

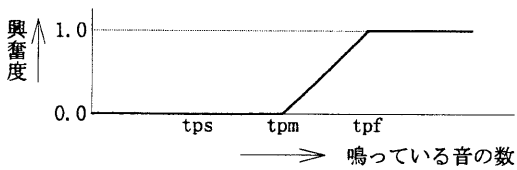
テンポについては、一定に時間（ここでは、3秒間）に鳴っている音の数が多いほど速いものとした。低速、中速、高速を表すセルを3つ用い、その興奮度を図3に示す。図3における速度域の区切りは、tps=5, tpm=10, tpf=20とした。



(a) 低速域セル

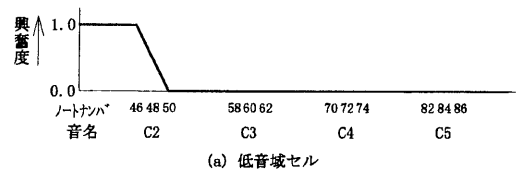


(b) 中速域セル

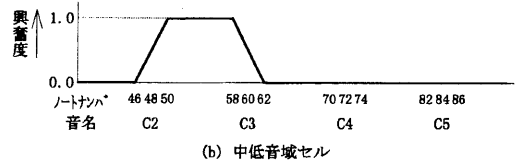


(c) 高速域セル

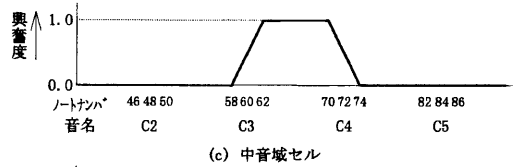
図3 特徴抽出機構におけるテンポ



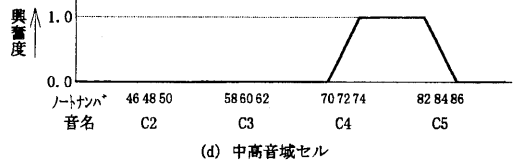
(a) 低音域セル



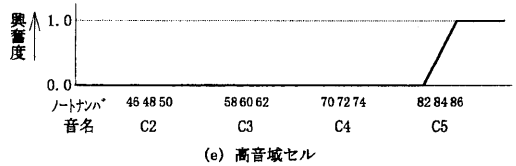
(b) 中低音域セル



(c) 中音域セル



(d) 中高音域セル



(e) 高音域セル

図4 特徴抽出機構における音高域

音高域については、ノートナンバによって5つの音域に分けた。各音域を表すセルの興奮度を図4に示す。

音高については、1オクターブ内の12音に圧縮し、12個のセルを用いた。また、和音については、その12音についてメジャー及びマイナーを考え、24個のセルを用いた。

・感情抽出機構

図5に感情抽出機構を示す。感情抽出には3層ニューラルネットを用い、抽出された音楽特徴から4つ感情を抽出する。出力層の4つのセルの興奮度によって4つの感情対「明暗」、「軽退」、「和緊」、「安淋」の強さを表す。ニューラルネットの学習はBP法を用いた。学習データの作成はつぎのように行った。

ある被験者に2,3小節のフレーズを演奏したものを聴かせ、感じた感情を回答するアンケートを行った。すなわち、4対の感情「明るい-暗い」、

「軽快-退屈」, 「和やか-緊張」, 「安心-淋しさ」に対して, 「明るい」, 「軽快」, 「和やか」, 「安心」, または「暗い」, 「退屈」, 「緊張」, 「淋しさ」, または「どちらでもない」を回答する. 「明るい」, 「軽快」, 「和やか」, 「安心」を回答した場合にはそのセルの興奮度を1.0に, 「暗い」, 「退屈」, 「緊張」, 「淋しさ」と回答した場合には, 興奮度を0.0に, 「どちらでもない」と回答したばあいには, 興奮度を0.5とする.

感情抽出機構の学習データを作成するために, 21曲のフレーズを被験者聴かせアンケートを行った. そのアンケート結果から, 学習データを作成し, 感情抽出機構の学習を行った.

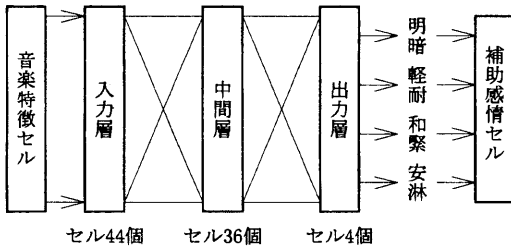


図5 感情抽出機構

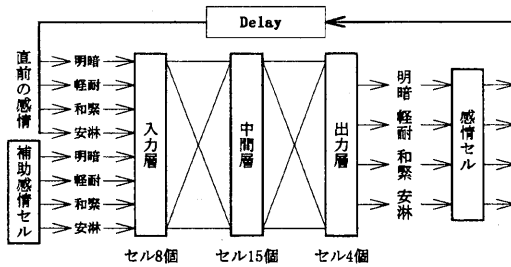


図6 感情変化機構

・感情変化機構

図6に感情変化機構を示す. 感情変化機構も3層ニューラルネットを用い, 学習もBP法を用いる. 出力層の4つのセルは音楽を聴いて感ずる4つの感情対の興奮度を表す.

感情変化機構は保持している直前の感情と感情抽出機構によって抽出された感情から次に保持すべき感情に変化させる.

4対の感情組に対して感情変化機構の学習データは同じものを用い, 各感情対に対して独立に学習させた. 表1にその学習データを示す. 例えば, 「明るい-暗い」において, セルの直前の出力が0.0, 抽

出された出力が1.0の場合, 「明るい-暗い」を表す感情セルの値を0.5となるように学習させる.

表1 感情変化機構の学習データ

直前感情	抽出感情		
	0.0	0.5	1.0
0.0	0.0	0.5	0.5
0.5	0.0	0.5	1.0
1.0	0.5	0.5	1.0

5. 感情変換部

抽出された4つの感情の強さから演奏制御に使用されるパラメータ P_F を作成し, 演奏制御部に送る. パラメータ P_F は変換行列 A を用いて次のように作成される.

$$\begin{aligned}
 P_F &= AF \\
 &= (v_F \quad t_F \quad s_F)^T \\
 &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} (f_1 \quad f_2 \quad f_3 \quad f_4)^T \\
 &\quad (2)
 \end{aligned}$$

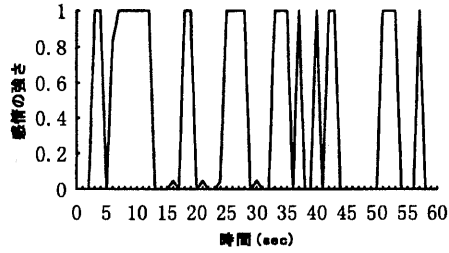
ここで, $(f_1 \quad f_2 \quad f_3 \quad f_4)$ は感情抽出部より送られる4つの感情セル「明るい-暗い」, 「軽快-退屈」, 「和やか-緊張」, 「安心-淋しさ」の強さを表すベクトルである.

変換行列は, 被験者のアンケートをもとにつぎの変換行列を作成した.

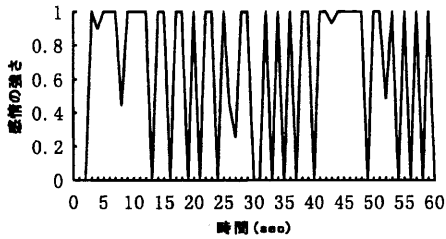
$$A = \begin{pmatrix} +3 & +0.5 & -2 \\ +3 & +1 & -2 \\ -3 & -0.5 & +2 \\ -3 & -1 & +2 \end{pmatrix} \quad (3)$$

なお, フィードバック信号 P_F の値による演奏制御部に与える音量, 速度, 奏法の変化は, 以下のようになる. ただし, 数値による変化の度合は経験的な値を用いて変換行列を設定した.

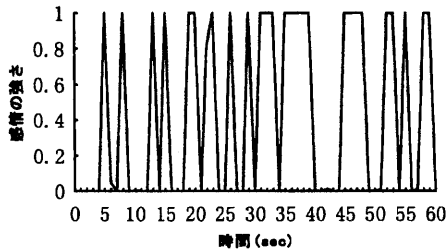
- ・音量(v_F) 正の値: 音量を大きくする.
負の値: 音量を小さくする.
- ・速度(t_F) 正の値: 速度を遅くする.
負の値: 速度を速くする.
- ・奏法(s_F) 正の値: 1つの音符で鳴っている時間が長くなる.



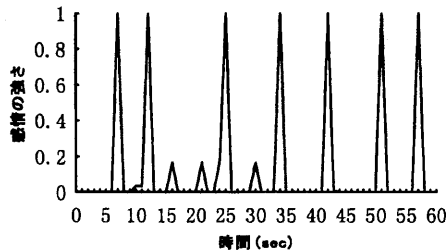
(1) 「明るさ-暗さ」の感情の強さの変化



(2) 「軽快-退屈」の感情の強さの変化



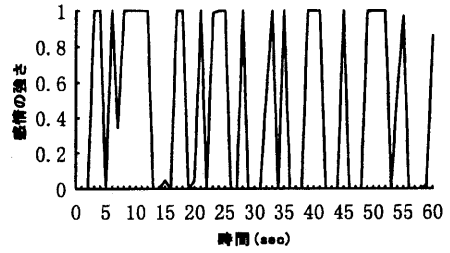
(3) 「和やか-緊張」の感情の強さの変化



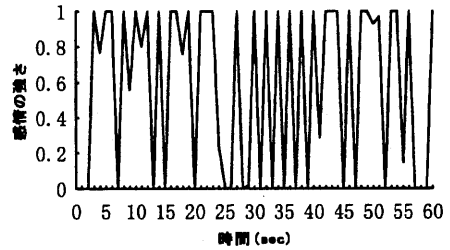
(4) 「安心-淋しさ」の感情の強さの変化

図7 フィードバックのない場合の出力結果

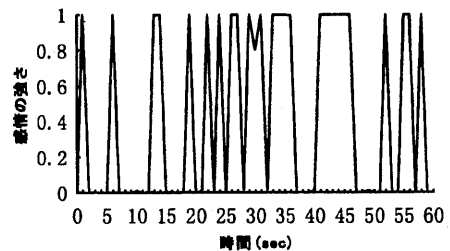
負の値：1つの音符で鳴っている時間が短くなる。



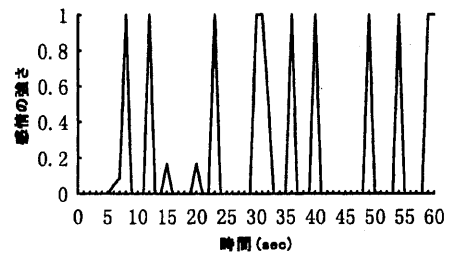
(1) 「明るさ-暗さ」の感情の強さの変化



(2) 「軽快-退屈」の感情の強さの変化



(3) 「和やか-緊張」の感情の強さの変化



(4) 「安心-淋しさ」の感情の強さの変化

図8 フィードバックのある場合の出力結果

表2 楽曲「時の扉」の演奏に対する被験者の回答

感情の組	明るい-暗い	軽快-退屈	和やか-緊張	安心-淋しさ
被験者の回答	どちらでもない	軽快	緊張	淋しさ

6. 実験結果

曲名「時の扉」という音楽に対して演奏制御の実験を行った。

まず、被験者に対して、この曲を聴いて感ずる感情のアンケートを行った。すなわち、4つの感情対「明るい-暗い」、「軽快-退屈」、「和やか-緊張

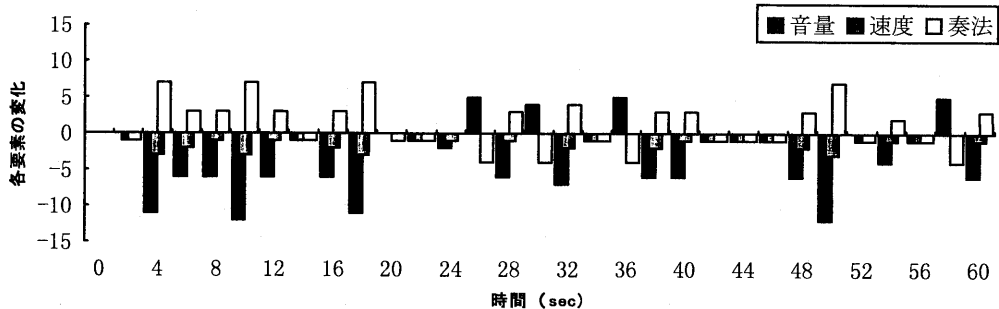


図9 フィードバック信号の値の変化

張」,「安心-淋しさ」に対して,どの感情を感じたかを回答させた。被験者の回答を表2に示す。

つぎに,フィードバックを掛けずに,曲名「時の扉」という音楽の演奏制御を行った。そのときの感情抽出部で抽出された結果を図7に示す。図7によれば,抽出された4つの感情対に対して小刻みな変化がみられる。

感情対「明るい-暗い」について,被験者は「どちらでもない」と回答しているが,図7(1)では,「明るい」と「暗い」を同程度抽出している。感情対「軽快-退屈」について,被験者は「軽快」と回答しているが,図7(2)では,「軽快」を抽出している方が多いように見える。感情対「和やか-緊張」について,被験者は「緊張」と回答しているが,図7(3)では,「和やか」と「緊張」を同程度抽出している。感情対「安心-淋しさ」について,被験者は「淋しさ」と回答しているが,図7(4)では,「淋しさ」を抽出している方が多いように見える。

そして,同じ曲についてフィードバックを掛けた演奏制御を行った。そのときの感情抽出結果を図8に,フィードバック量の変化を図9に示す。図7と図8を比較すると演奏の違いにより,抽出された感情に違いが見られる。このフィードバックによる演奏制御で演奏に変化が生ずることは確認できるが,それが良い演奏であるとはいえない。

7. おわりに

抽出された音楽特徴からの感情変化をもとに演奏を制御するフィードバック機能を導入した。この機能の導入によって演奏に変化がみられたが,より良い演奏になったかの判断はできなかった。変換行列の値を大きくすれば,フィードバックを用いない場合との差違が大きくなった。今回,抽出した感情が音楽演奏に適したものか疑問が残り,音楽演奏に適した感情についての検討が必要である。また,変換行列の選定など改良を行い,より良い自動演奏システムにしていきたい。

参考文献

- [1] 井口, 他, “感性情報処理”, オーム社, 1994
- [2] 大照, 橋本, “仮想音楽空間”, オーム社, 1994
- [3] Feiten, B., Gunzel, S., “Automatic Indexing of a Sound Database Using Self-organizing Neural Nets”, Computer Music Journal, Vol. 18, Num. 3, pp. 53-65
- [4] 重松, 水本, “ニューラルネットワークを用いた音楽情報からの感情抽出”, 平成5年度九州工業大学工学部電気工学科卒業論文, 1994
- [5] 金子, “自動演奏の高度化に関する研究”, 平成6年度九州工業大学工学研究科設計生産工学専攻修士論文
- [6] 尾野, “ニューラルネットワークを用いた音楽情報からの感情抽出”, 平成6年度九州工業大学工学部電気工学科卒業論文
- [7] 玉木ら, “MIDI信号を用いた演奏音楽による感情変化シミュレーションの試み”, 1995年電子情報通信学会総合大会p295