

6D-8

ニューラルネットワークによる 楽曲のトップダウン処理について

後藤 和之 横田 誠

電気通信大学

1. まえがき

我々は、楽曲が移調によって様々な音の高さで演奏されても、同じ曲であると認識する。逆に、同じ曲であっても、それぞれの声の高さで表現することができる。

楽曲を“正規化⁽¹⁾”すれば、楽曲を移調によらない情報として扱うことができる。また、今回提案する方法によれば、正規化された情報から、具体的な音の高さを持った情報を作り出すことができる(このような操作を“脱正規化”と呼ぶことにする)。

音程に関する正規化(第2節参照)と脱正規化は、人間の脳における音楽などの情報の処理の1つのモデルである。また、正規化されたパターンは、音楽などの情報が脳においてどのように表現されているかのモデルである。耳をボトム、脳において高次の処理を行なう部分をトップと見ると、正規化はボトムアップ処理であり、脱正規化はトップダウン処理と見ることができる。また、正規化された情報は、トップにおいて情報がどのように表現されているかを示している。

今回は、長期記憶器(正規化されたパターンを想起のためのキーとして、次の音に関する音程情報を想起するもの)の存在を前提とし、この長期記憶からのトップダウン処理モデルについて、基礎的考察を行なう。

2. 正規化と音程情報

正規化にはいくつかの方法がある。曲の最初の音(初音と呼ぶ)が特定の周波数 $C^0 (= 2^0 = 256 \text{ Hz})$ となるように、曲全体を移調する正規化を初音正規化と呼んでいる⁽²⁾。このような正規化によれば、楽曲の具体的な音の高さによらない情報を得ることができる。

文献(3)では、最後に聞いた音(逐次終音と呼ぶ)と、それ以前のいくつかの音との音程を処理するニューラルネットワークを構成し、これにより正規化されたパターンを得た。一般に、このような音程情報は、正規化情報と等価の表現となる(このことは、文献(4)においてグラフ論的に示されている)。

次章以降に示すトップダウン処理は、音程抽出により正規化を行なうボトムアップ処理⁽³⁾と対になる関係がある。

3. トップダウン処理モデルの位置付け

今回のモデルは、長期記憶器で想起された正規化された情報と、現在聞いている音の高さとを入力とし、次の音の高さを予測するものである。これは、例えば次のようなことに対応している。他人がいろいろな高さで歌う歌を聞いて、その音の高さに合わせて歌うには、その曲を知っているだけでなく(正規化された情報のみではなく)、その人が歌っている高さの情報を得て、同じ高さに合わせて歌う必要がある。

このトップダウン処理の処理全体における位置付けを示すために、ボトムアップ処理を含めた概念図を図1に示す。ボトムアップ処理では、音程情報などを取り出すことで、正規化を行なう。この正規化された楽曲の情報は、長期記憶器への入力(想起のためのキー)となる。このキーなどにより長期記憶器は記憶していた正規化された情報(次の音を予測するもの)を出力する。トップダウン処理では、この出力と、それまでに聞いた音の高さを利用して、次の音の具体的な音の高さを予測する。想起されるために十分な数の音が入力された後は、想起音高自体をボトムからの入力として想起を持続する。

このような処理は、声帯などの制御の他に、記憶を利用した楽曲認識のための、基本的なモデルであると考えられる。想起された記憶は、ノイズの多い入力や不完全な入力を認識する際に、その入力は何であるか

A model of musical top-down process by neural network

Kazuyuki GOTO and Makoto YOKOTA

University of Electro-communications

を示す仮説となり、この仮説に対応する部分を強調、あるいは抑制する処理⁽⁵⁾などが考えられる。

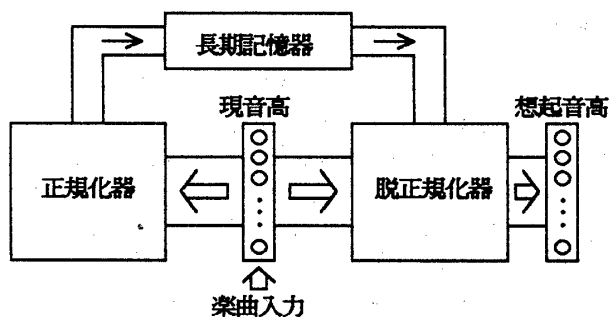


図1 ボトムアップ処理を含めた概念図

4. トップダウン処理の回路モデル

ニューラルネットワークによるトップダウン処理モデルを図2に示す。

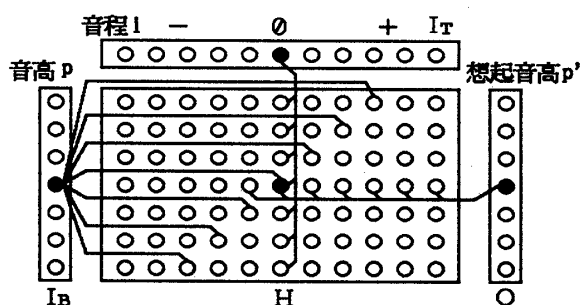


図2 トップダウン処理回路

(I_B : 入力層(音高), I_T : 入力層(音程)
H: 中間層, O: 出力層)

入力層は2つあり、一方は、音の高さ p を入力するための層 I_B で、ボトム側からの情報を受け取る。他方は、現在の音と次の音との間の音程 i を入力するための層 I_T であり、長期記憶器からの情報を受け取る。いずれの入力層においても、1次元的にニューロンが並んでいる。

中間層の各ノードは、それに対応する音の高さ p と音程 i があり、これに対応する2つの入力があるときに、そのノードが発火する。このために、中間層の1個のノードには、各入力層の1個づつのニューロンからの枝がある。

出力層では、次の音の高さ $p' (= p + i)$ に対応するノードが発火する。出力層の1個のノードは、その音の高さ p' に対応する、すべての中間層のノードから入力を受ける。

このモデルの中間層の各ニューロンには、それに1

対1に対応する、文献(3)の中間第2層のニューロンがある。

図3に、この回路による計算例を示した。第 N 音(休符は数えない)まで入力された時の出力が順に並べて示してある。

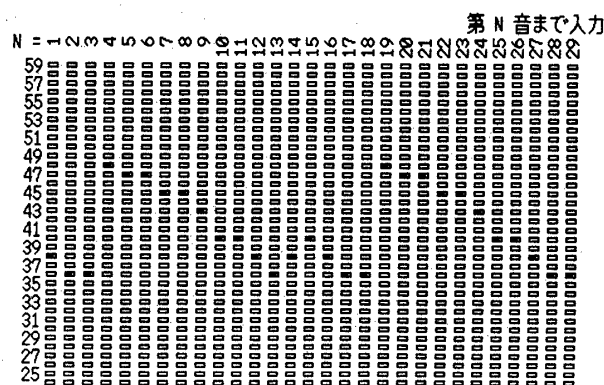


図3 トップダウン処理計算例("Annie Laurie")

5. まとめ

ボトムアップ処理(正規化)と対をなす、トップダウン処理(脱正規化)について検討し、ニューラルネットワークによるトップダウン処理の例を示した。このようなトップダウン処理の考え方は、和音のような並列の情報や、時間情報(リズム)などに対しても適用できると考えられる。

今回は、楽曲を順に想起する長期記憶器を仮定したが、短期記憶や、音楽的な終止感⁽⁶⁾や好みを出力するような回路に対してもこのような処理が考えられる。

参考文献

- (1) 横田 誠: "楽曲構造の音楽的正規化について", 音楽音響研究会資料, 7, 7(1989).
- (2) 横田 誠: "MP (ミュージカル・プレーン) 上の平均律12音系と実行的初音による正規化楽曲構造の分類について", 音響講論集, 2-6-8(1990-3).
- (3) 後藤和之, 横田 誠: "楽曲構造分析のための, ニューラルネットワークを用いた, 逐次終音による正規化について", 信学技報, PRU90-81(1990).
- (4) 後藤和之, 横田 誠: "音程による正規化楽曲構造の音楽的グラフ論的表現について", 音響講論集, (1991-3).
- (5) 福島邦彦: "神経回路と情報処理", 朝倉書店, pp. 163-165(1989).
- (6) 阿部純一: "旋律はいかに処理されるか", 音楽と認知(波多野誼余夫 編), 東京大学出版会, pp. 41-68(1987).