

Hopfield 型ニューラルネットワークを用いた音源同定

Sound Source Identification Using Hopfield-Type Neural Networks

久保田 優

Yuu KUBOTA

早稲田大学理工学部 電子・情報通信学科

甲藤 二郎

Jiro KATTO

School of Science and Engineering, Waseda University

1. まえがき

複数の音響信号が重畳された信号を入力とし、それに含まれる楽器音を推定する音源同定に関する検討が盛んである [1][2]。本稿では、音源同定の問題を組み合わせ最適化問題と位置づけ、Hopfield 型のニューラルネットワークを用いて最適解を求める方式を提案する。MIDI 音源を用いた単音・2 和音の認識実験を行い、良好な結果を得た。

2. 従来手法と問題点

これまでに提案されている音源同定アルゴリズムでは、音源分離を行った後で、各単音についてあらかじめ用意した波形やテンプレートと比較するものが多い。これらの方法は、実際の認識段階で入力信号がテンプレートと同じ単音の形になっているため認識が比較的容易な反面、正確に分離を行うためには複雑な処理が必要となり、かつ分離に失敗するとその後の処理が信頼できないものとなる。

そこで本稿では、入力信号を分離することなく、逆に音源同定をテンプレートの組み合わせ最適化問題（エネルギー最小化問題）と考え、Hopfield 型のニューラルネットワークを用いて最適解（同定結果）の探索を試みる。

3. 提案手法

図 1 に、本提案手法の処理の流れを示す。提案手法では、まず準備段階として、音源同定を行う対象の信号に含まれている可能性のある全ての楽器・音高の単音の演奏データを録音し、ウェーブレット変換を施した後に、周波数スペクトルとしてテンプレートを得る。合わせて、テンプレート間の相互相関 W_{ij} を計算しておく。一方、認識段階では、まず入力された演奏データのウェーブレット変換結果に対して、各テンプレートとの相互相関 I_i を計算する。そして、この相互相関値 I_i と、事前に計算したテンプレート間の相互相関値 W_{ij} とを Hopfield 型のニューラルネットワークへの入力とし、次式に表されるエネルギー E を最小にするテンプレートの組み合わせとして、同定結果を出力する。

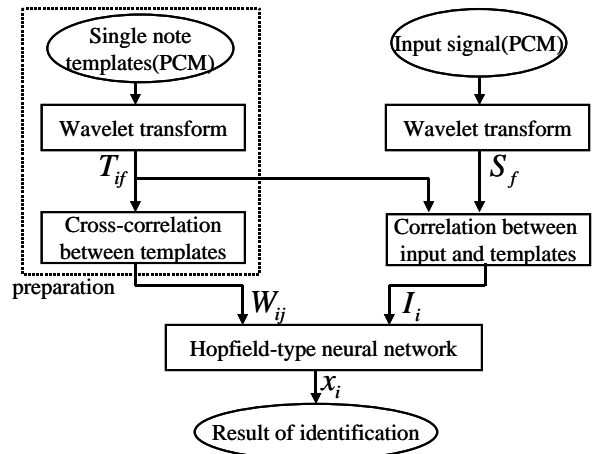
$$E = K_1 \cdot \frac{1}{2} \sum_{f=1}^F \left(\sum_{i=1}^N x_i T_{if} - S_f \right)^2 + K_2 \cdot \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij}^2 x_i (1 - x_j) \quad (1)$$

上式において、 F はスペクトルの点数、 N は総テンプレート数、 x_i はニューロン i の出力、 T_{if} は i 番目のテンプレートのスペクトルにおける f 番目の成分、 S_f は入力信号のスペクトルにおける f 番目の成分、そして K_1 と

K_2 は各項の寄与率である。また後項はネットワークの収束性を上げるための補正項である。

上記手法の特徴は、演奏されている楽器数を限定せずに、複数の音を同時に認識できる点にある。しかし、その一方で、出力音数に制限がないために、例えば 2 和音の入力に対して 3 和音出力されるといった「出力音数誤り」が新たな問題となる。ただし、時系列として見れば、このような事象が連続して発生することは少ない。

そこで、本稿ではさらに、突然出力音数が変わったフレームから後続するフレームの出力音数をチェックし、その出力音数の変化が突発的なものであれば、「出力音数誤り」とみなして訂正するようにした。さらにはまた、ニューラルネットワークの局所解の問題を軽減するために、楽器音間の遷移確率を考慮した補正を行った。



4. 実験

図 1 処理の流れ

MIDI 音源を用いた単音・2 和音音源同定実験の認識結果を表 1 に示す。入力信号は単音・2 和音の曲であり、テンプレートを組み合わせて作ったものである。

表 1 認識結果

単音認識率 (%)		100.0
2 和音認識率 (%)	2 音認識	81.3
	1 音認識	18.7
	誤認識	0.0

5. おわりに

本稿では、Hopfield 型のニューラルネットワークを用いて音源同定を行う方法を提案し、その有効性を確認した。

6. 参考文献

- [1] 柏野邦夫他, “音楽情景分析の処理モデル OPTIMA おける単音の認識,” 信学論, Vol. J79-D- , No.11, pp.1751-1761, 1996.
- [2] 村瀬樹太郎他, “ニューラルネットワークを用いた複数楽器の音源同定処理,” 情処学音楽情報科学研報, MUS-39-10, 2001