

4C-1 ニューラルネットを用いた ピッチ抽出の一手法

加藤 誠巳 鶴飼 敏之 粕谷 治彦
(上智大学理工学部)

1. まえがき

近時、ニューラルネットに関する研究が盛んであり、そのパターン認識への応用が注目され、音声信号認識の分野でも多くの成果が得られている。本稿では、ニューラルネットを用いて楽器音信号のピッチ周期を抽出する手法について検討を行なった結果について述べる。

2. ピッチ周期学習の手法

ここではピッチ周期の抽出に時間波形から計算された短時間自己相関関数を使用した。図1にここで使用したニューラルネットの構造を示す。10kHzでサンプリングした時間信号から256サンプルを切り出してハミング窓を乗じ、それから得られた短時間自己相関関数値 $R(iT)$ (T はサンプリング周期、 $i=0,1,\dots,255$) を $R(0)$ で正規化した $r(iT)$ のうち、最初の128点をニューラルネットの入力ノードに加えた。中間ノード数は25、出力ノード数は1とし、12.8msで正規化したピッチ周期を出力ノードが出力するようにBP法を用いて学習を行わせた。

なお、学習は出力ノードの出力値と教師データとの相対誤差がすべて1%以内になった時点で打ち切った。

学習データとしては、 $C_2(130.81\text{Hz})$ から $B_3(493.88\text{Hz})$ まで半音毎に変化させた計24の異なるピッチ周期のピアノ、フルート、トランペット、オーボエなど20種類の楽器音をシンセサイザにより発生させたものを用いた。図2に学習に用いた同一ピッチ周期の短時間自己相関関数の例を示す。

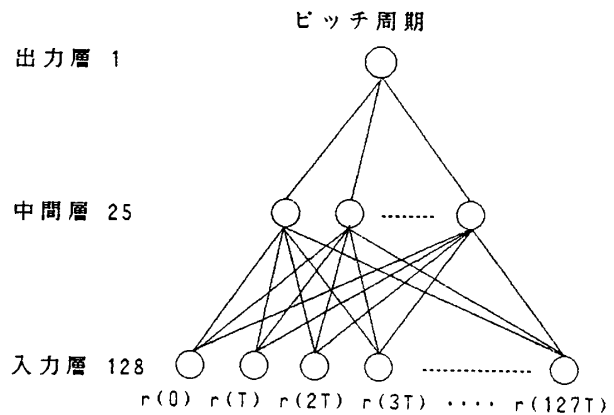


図1 ニューラルネットの構造

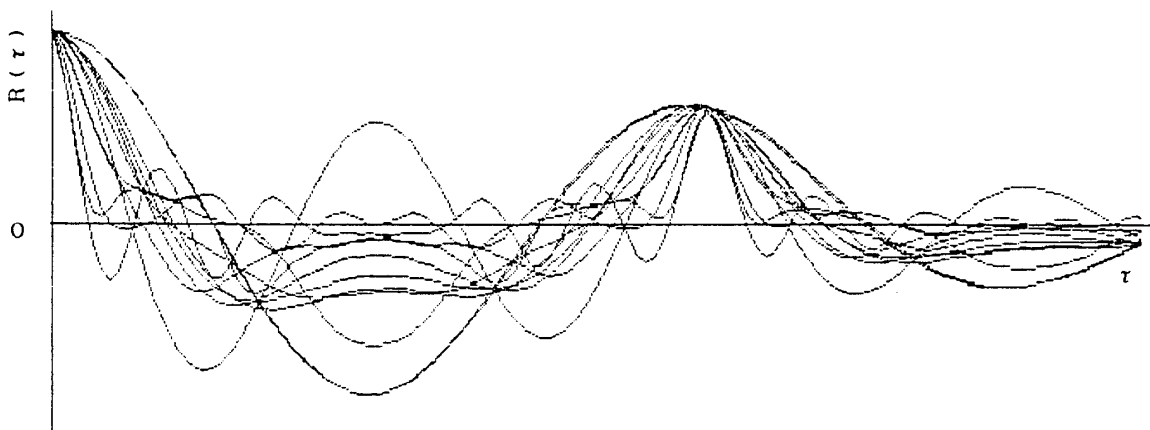


図2 同一ピッチ周期の短時間自己相関関数の例 (音高 C_2)

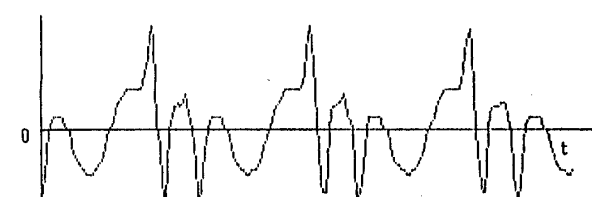


図3 時間波形 ($S/N = \infty$)
(ハ'17'ホカ'ン、C2)

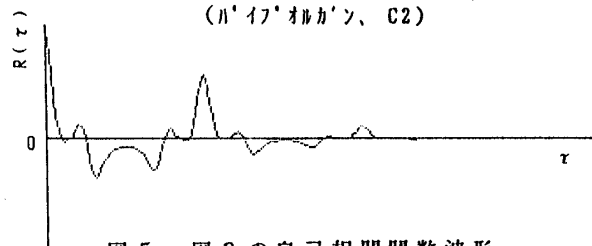


図5 図3の自己相関関数波形

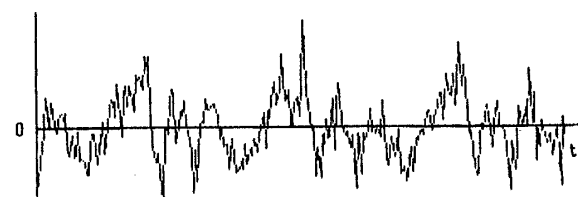


図4 時間波形 ($S/N = 5 \text{ dB}$)

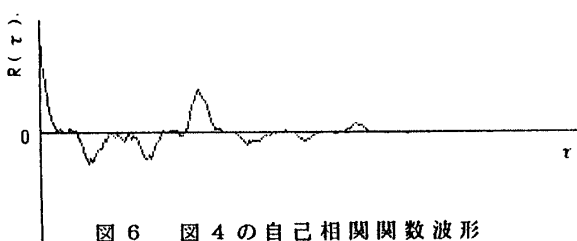


図6 図4の自己相関関数波形

3. S/N に対する特性の評価

上述の条件で学習を行わせたニューラルネットにおいて、白色ガウス性雑音に対する特性評価の実験を行った。

図3に $S/N = \infty$ の時間波形例、図4に $S/N = 5 \text{ dB}$ の時間波形例、図5および図6に夫々の短時間自己相関関数波形を示す。図7は、図3に示す時間波形の S/N を変化させたときのニューラルネットの出力ノードから得られるピッチ周期の相対誤差を示す。図8は、 S/N をパラメータにとって20種類の楽器音の24の音高すべてに対して所望のピッチ周期とニューラルネットの出力ノードの出力値との関係を示したものである。

これらから本ニューラルネットの S/N 特性はかなりすぐれていることがわかった。さらに学習データを増やせば、誤差の小さいネットワークを作ることが期待される。

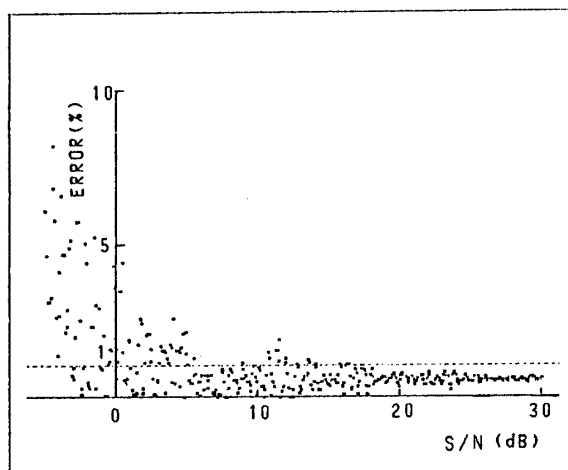
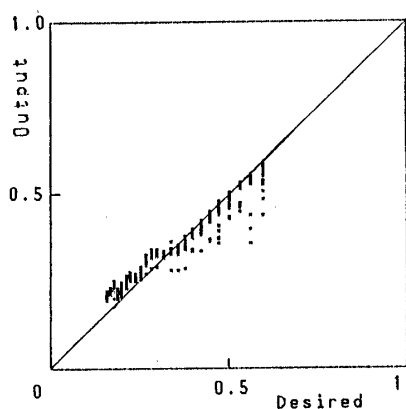
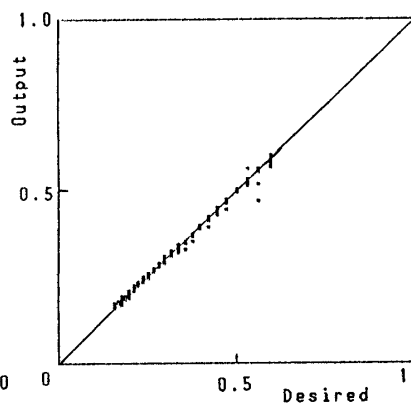


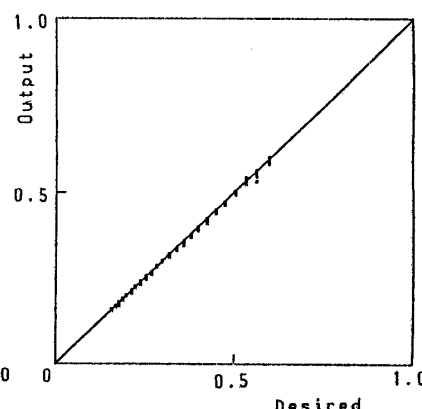
図7 S/N に対するピッチ周期の相対誤差



(a) 0dB



(b) 10dB



(c) 20dB

図8 所望の出力に対するニューラルネットの出力値

4. むすび

入力として短時間自己相関関数を、出力としてピッチ周期を与えるニューラルネットの S/N 特性について検討を行った。終わりに御討論戴いた本学マルチメディアラボの諸氏に謝意を表す。