

音声情報によるニューラルネットワークを用いた夜行性野鳥の識別

東谷 幸治†
Koji Higashitani

三田 長久‡
Nagahisa Mita

牧野 洋平‡
Yohei Makino

1. 研究背景

近年、急激な技術革新に伴い、環境問題が深刻になっている。それにつれて、環境評価の研究が盛んに行われている。本研究では、環境の変化に敏感な野鳥について、鳴き声を使って種類識別を行っている。本研究では、対象を夜行性野鳥とし、鳴き声の識別、及び、長時間録音データから鳴き声の部分の抜き出しを行う。

2. ニューラルネットワーク

ニューラルネットワークとは、ある入力値が入力層に入力されたとき、所望の出力が出力されるように、入力層から中間層、中間層から出力層への重みを学習させ、それを利用して、識別を行うものである。今回は、3層のニューラルネットワークを利用し、誤差逆伝播法で学習を行った。入力ユニットは32、出力ユニットは野鳥の種類である12種類で行っている。

3. 周波数帯域パワーの計算方法

周波数帯域に変換された音声データを $X(p)$ と表し、周波数 $f_n \sim f_{n+1}$ の領域に含まれるパワーを P_n とすると次式で周波数帯域パワーを求める。

$$P_n = \sum_{p=f_n}^{f_{n+1}} |X(p)|^2 \quad (1)$$

ここで重要になるのが周波数の分割の方法である。夜行性の野鳥は低い周波数領域で鳴く野鳥が多いため、低い周波数の分割数を大きく、高い周波数の分割数を少なくした。決定した分割を示す。

0.25~1[kHz]	0.25[kHz]間隔
1~3[kHz]	0.4[kHz]間隔
3~5[kHz]	0.5[kHz]間隔
5~8[kHz]は、	0.75[kHz]間隔

また、野鳥の声の高さの個体差に対応するため、図2による窓関数をかけて(1)を計算した。

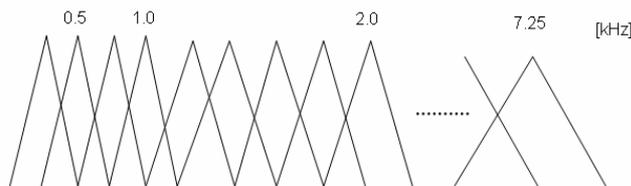


図2 周波数帯域パワーを求める時に使用した窓関数の図

4. シミュレーション

シミュレーション条件について述べる。今回は、12種類の野鳥の識別を行った。ニューラルネットワークのパラメータを以下に示す

学習データ	60個
中間ユニット	32
学習係数	0.05
学習回数	20000

5. 結果

前章のシミュレーション条件による結果を以下に示す。

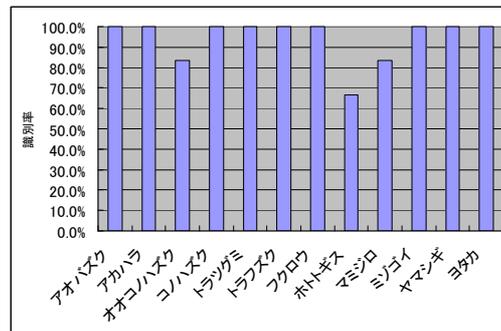


図3 識別対象種類毎の識別率

学習時間は5分40秒前後だった。(Pentium4 HT 3.0GHz) なお、ニューラルネットワークのユニットの値が最大になったユニットを識別結果とするが、閾値を0.5とし、全てのユニットがそれ以下になった場合は結果なしとした。

全体の識別率は94.29%で、高い識別率を得ることができた。誤識別の検討だが、誤識別している野鳥は鳴き方のバリエーションがあり、学習のサンプル数が十分でないと推測される。そのため、別の周波数帯域にピークが現れることもあるので、例えばテストデータにおいて、周波数を少し上下に振らすといった対策が考えられる。

6. まとめ

今回提案した周波数帯域パワーを用いたニューラルネットワークによる識別は、夜行性鳥類において高い識別率を得ることができ、有用性が確認できた。今回は地域性を考慮していなかったが、同じ品種でも地域により鳴き方が異なったりするため、今後サンプルを増やし、地域ごとに識別器を構成する必要がある。本研究は一部を環境技術開発等推進費の補助を受けて実施した。

† 熊本大学大学院 自然科学研究科電気システム専攻

‡ 熊本大学大学院 自然科学研究科 情報電気電子工学専攻