

鳴き声を用いた鳥の自動分類における最適パラメータの検討

T140501 前川 志帆

指導教員 三好 力 教授

1. はじめに

野生動物の生息調査では、捕縛以外のより良い調査の方法が求められている。捕縛以外での調査の難しい野鳥においては、目視と鳴き声による調査が主となっている。特に野鳥の多いイギリスでは、近年この作業の自動化に注目が向けられている。同国では、ダン氏らの Warblr チームが大学等の支援を受け、イギリスの野鳥の鳴き声を学習し分類するアプリ Warblr[1]を開発している。しかし、このような野生動物の計算機を用いての調査の研究はまだ少なく、Warblr もまだ開発途中とされている。

そこで、本研究では、鳥の鳴き声による自動分類を考え、鳥の鳴き声の特徴をより表すことができる新しいパラメータを提案することを目的とする。

2. 新しいパラメータの試案

2.1 Bird's MFCC

鳥は人間に比べて高い音を出す種が多い為、その高音域の違いをより細かく表した特徴量を考える。人間の音声特徴量として用いられる MFCC[2]では、低音域の音の特徴を細かく表す為に、メルスケールが用いられる。これに着目し、処理を行う音域を上げることで、高音域の特徴を細かく表す特徴量 Bird's MFCC を試案する。

2.2 Bird's MFCC の妥当性

3 種類の鳥 (Long-tailed cuckoo・Kea・Morepork) の鳴き声 1 回分 (3~5 秒) が収録された wav ファイルに対して、MFCC と Bird's MFCC を算出する。それらを Weka3.8.1[3]を用いて k -means 法で 3 クラスにクラスタリングを行い、結果を比較検討する。

3. 結果と考察

クラスタリングの結果、MFCC の正答率が

73.3%、Bird's MFCC の正答率が 95.8% と正答率は Bird's MFCC が MFCC を大きく上回る結果となった。

表 1: 全クラスタリング結果のまとめ

	Long-tailed cuckoo	Kea	More pork	総合
最小値 0Hz(MFCC)	50%	70%	100%	73%
最小値 500Hz	56%	67%	100%	74%
最低値 1000Hz	88%	100%	100%	96%
最低値 1500Hz	50%	30%	100%	60%
最低値 2000Hz	50%	88%	100%	79%

また、MFCC では不正解データが大きく高さの異なるクラスに分類されているのに対して、Bird's MFCC では高さの近いクラスに分類されている。このことから、Bird's MFCC は MFCC では表せなかった鳥の鳴き声の特徴を表すことができると考えられる。

4. おわりに

本研究では、MFCC に注目し Bird's MFCC を試案、検討した。既存手法より音の高さの違いをより表すことができた。今後は、対応できる鳥の種類を増やす為に、低音域を切り捨てずに高音域を強調できる新しい Bird's MFCC が必要となる。また、Bird's MFCC は高さの違いを表わすパラメータである為、更に他の要素を表わすパラメータの検討が必要となると考える。

参考文献

[1] warblr 社, "Warblr : the birdsong recognition app"

<https://warblr.net/>, (参照 2017-8-16)

[2] Beth Logan, "Mel Frequency Cepstral Coefficients for Music Modeling", http://ismir2000.ismir.net/papers/logan_paper.pdf, (参照 2017-8-20)

[3] waikato 大学, "Weka 3 : Data Mining Software in Java", <http://www.cs.waikato.ac.nz/ml/weka/>, (参照 2017-11-13) .