AUTOMATIC BIRD VOCALIZATION IDENTIFICATION BASED ON FUSION OF SPECTRAL PATTERN AND TEXTURE FEATURES

Sai-hua Zhang, Zhao Zhao, Zhi-yong Xu, Kristen Bellisario, Bryan C. Pijanowski

School of Electronic and Optical Engineering, Nanjing University of Sci. & Tech., 210094, China

Department of Forestry and Natural Resources, Purdue University, West Lafayette IN47907, USA

Abstract

Automatic bird species identification from audio field recordings is studied in this paper. We first used a Gaussian mixture model (GMM) based energy detector to select representative acoustic events. Two different feature sets consisting of spectral pattern and texture features were extracted for each event. Then, a ReliefF-based feature selection algorithm was employed to select distinguishing features. Finally, classification was performed using support vector machine (SVM). The main focus of the proposed method lies in the fusion of a spectral pattern feature with several texture descriptors, which extends our previous work. Experiments used an audio dataset comprised of field recordings of 11 bird species, containing 2762 bird acoustic events and 339 detected "unknown" events (corresponding to noise or unknown species vocalizations). Experimental results demonstrate superior classification performance compared with that of the state-of-the-art method, which renders the proposed method more suitable for real-field recording analysis.

Proposed method

1. Automated segmentation

2. Feature extraction

2.1 Spectral pattern feature

- sp: spectral pattern (SP) feature. We employed a Mel band-pass filter bank on the spectrograms of each event. The output in each channel, i.e. a time-series containing variable-band limited energy, was parameterized by an autoregressive (AR) model, which resulted in a parameterized feature consisting of all model coefficients.

2.2 Texture descriptors

- lbp: uniform local binary pattern (ULBP)
- lpq: local phase quantization (LPQ)
- hac: heterogeneous auto-similarities of characteristics (HASC)
- gf: Gabor filter (GF)

2.3 Feature selection

\[ cf_{ij} = \frac{1}{K} \sum_{k=1}^{K} lbp_{ij} \cdot hac_{ij} \cdot GF_{ij} \] (1)

where \( K \) is the total number of events.

It is worth noting that this simple concatenation leads to a high-dimensional feature vector. In this context, the ReliefF feature selection algorithm, aiming to choose a discernible subset of features to reduce the fusion feature length with the lowest information loss, can be employed to remove irrelevant and/or redundant features. With the help of ReliefF, an attribute weight was calculated for each feature, ranging from -1 to 1 with a high positive weight assigned to an important attribute. Then we sorted out \( N \) most important features as the effective subset. According to our preliminary study, we selected \( N=40 \).

3. SVM-based classification

In this work, we employed multi-class SVM based on the "one-versus-one" strategy for species classification. Besides, we used the radial basis function (RBF) as the kernel function. The two parameters of the RBF kernel, gamma \( \gamma \) and cost \( C \), were optimized by grid search to obtain the best models. For each feature, we employed the ReliefF feature selection algorithm to remove irrelevant and/or redundant features. This step reduces the dimensionality of the feature space. The averaged performance metrics for each species between the proposed method with the baseline method are presented in Table 1. At this stage, we selected 11 bird species in the recordings that can be divided into five types based on sound unit shapes, including constant frequency (CF), frequency modulated whistles (FM), broadband pulses (BP), chirping trill (CT), and song (S).

Conclusions

Aiming to improve the audio parameterization process in bird species identification tasks, we proposed an automatic acoustic classification method based on feature fusion in this work. Two different feature sets, the spectral pattern feature and texture descriptors, were extracted. As the combination of the two sets, ReliefF-based feature selection algorithm was employed to select a distinguishing feature subset. Experimental results using real-world recordings showed that the proposed method outperformed the state-of-the-art robust approach.

Acknowledgments

This work was supported by the National Natural Science Foundations of P.R. China under Grants No. 61402103 and No. 61771167, and the State Scholarship Fund of China [grant number 201660040023].

Table 1. Details of species used in this work.

- **Bird species**: Common Yellowthroat, Chipping Sparrow, Cedar Waxwing, American Crow, American Robin, Blue Jay, Song Sparrow, American Yellowthroat, Great Blue Heron, and Rusty Blackbird.
- **Data source**: Xeno-canto archive.
- **Classification accuracy (%):** 93.7, 96.8, 89.6, 94.7, 91.4, 95.1, 97.9, 95.6, 93.8, 96.8, 96.3.
- **F**: 0.973, 0.951, 0.973, 0.950, 0.953, 0.953, 0.973, 0.953, 0.953, 0.953, 0.953.
- **Baseline method**: SVM with all 40 features.

Table 2. Comparison of various features in terms of classification accuracy (CA) equipment with the same SVM classifier.

- **Features**: Baseline method, GMM-based feature + LBP, GMM-based feature + LBP + LPQ, GMM-based feature + LBP + LPQ + HASC, GMM-based feature + LBP + LPQ + HASC + GF.
- **CA (%)**: Baseline method 75.0, GMM-based feature + LBP 79.5, GMM-based feature + LBP + LPQ 85.7, GMM-based feature + LBP + LPQ + HASC 93.0, GMM-based feature + LBP + LPQ + HASC + GF 97.9.

The experimental results demonstrate superior classification performance compared with that of the state-of-the-art method, which renders the proposed method more suitable for real-field recording analysis.

Fig. 1. Overview block diagram of the proposed method.

Table 3. The averaged performance metrics for each species between the proposed method with the baseline method.

- **Species**: American Robin, Blue Jay, Song Sparrow, American Yellowthroat, Great Blue Heron, Rusty Blackbird, American Crow, Cedar Waxwing, American Robin, and Rusty Blackbird.
- **CA (%)**: American Robin 96.3, Blue Jay 99.0, Song Sparrow 93.8, American Yellowthroat 91.4, Great Blue Heron 95.6, Rusty Blackbird 94.7, American Crow 97.9, Cedar Waxwing 94.7, American Robin 99.0, and Rusty Blackbird 96.8.
- **F**: 0.953, 0.973, 0.953, 0.953, 0.953, 0.953, 0.953, 0.953, 0.953, 0.953.