Automatic Classification of Bird Species From Their Sounds Using Two-Dimensional Cepstral Coefficients

李建興, 韓欽銓, 莊清乾
Computer Science & Information Engineering
Computer Science and Informatics
chlee@chu.edu.tw

Abstract

This paper presents a method for automatic classification of birds into different species based on the audio recordings of their sounds. Each individual syllable segmented from continuous recordings is regarded as the basic recognition unit. To represent the temporal variations as well as sharp transitions within a syllable, a feature set derived from static and dynamic two-dimensional Mel-frequency cepstral coefficients are calculated for the classification of each syllable. Since a bird might generate several types of sounds with variant characteristics, a number of representative prototype vectors are used to model different syllables of identical bird species. For each bird species, a model selection method is developed to determine the optimal mode between Gaussian mixture models (GMM) and vector quantization (VQ) when the amount of training data is different for each species. In addition, a component number selection algorithm is employed to find the most appropriate number of components of GMM or the cluster number of VQ for each species. The mean vectors of GMM or the cluster centroids of VQ will form the prototype vectors of a certain bird species. In the experiments, the best classification accuracy is 84.06% for the classification of 28 bird species.

Keyword：Birdsong classification, Gaussian mixture models, two-dimensional Mel-frequency cepstral coefficients