

Continuous Birdsong Recognition Using Gaussian Mixture Modeling of Image Shape Features

李建興, 徐勝斌, 石昭玲, 周智勳

Computer Science & Information Engineering

Computer Science and Informatics

chlee@chu.edu.tw

Abstract

Traditional birdsong recognition approaches used acoustic features based on the acoustic model of speech production or the perceptual model of the human auditory system to identify the associated bird species. In this paper, a new feature descriptor that uses image shape features is proposed to identify bird species based on the recognition of fixed-duration birdsong segments where their corresponding spectrograms are viewed as gray-level images. The MPEG-7 angular radial transform (ART) descriptor, which can compactly and efficiently describe the gray-level variations within an image region in both angular and radial directions, will be employed to extract the shape features from the spectrogram image. To effectively capture both frequency and temporal variations within a birdsong segment using ART, a sector expansion algorithm is proposed to transform its spectrogram image into a corresponding sector image such that the frequency and temporal axes of the spectrogram image will align with the radial and angular directions of the ART basis functions, respectively. For the classification of 28 bird species using Gaussian mixture models (GMM), the best classification accuracy is 86.30% and 94.62% for 3-second and 5-second birdsong segments using the proposed ART descriptor, which is better than traditional descriptors such as LPCC, MFCC, and TDMFCC.

Keyword : Angular radial transform (ART), Birdsong recognition, Gaussian mixture models (GMM)