Generalized perceptual features for animal vocalization classification

Two sets of generalized, perceptual-based features are investigated for use in classifying animal vocalizations. Since many species, especially mammals, share similar physical sound perception mechanisms which vary in size, two features sets commonly used in human speech processing, Mel-Frequency Cepstral Coefficients (MFCCs) and Perceptual Linear Prediction (PLP) analysis, are modified for use in other species. One modification made to the feature extraction process is incorporating the frequency range of hearing and length of the basilar membrane of the animal in order to correctly determine the width and location of the critical band filters used for signal processing. Experimentally determined critical bands (equivalent rectangular bandwidth) and equal loudness curves (audiograms) can also be incorporated directly into the feature extraction process. Experiments are performed on African elephant (Loxodonta africana) vocalizations using a Hidden Markov Model (HMM) based classifier showing increased classification accuracy when using features sets based on the specific animals' perceptual abilities compared to the original human perception-based feature sets.