ACOUSTICS2008/1155
Automatic classification of vocalizations with Gaussian Mixture Models and Hidden Markov Models

Judith Brown[a] and Paris Smaragdis[b]
[a]MIT Media Lab, 37 Indian Ridge Rd, Natick, MA 01760, USA
[b]Adobe Systems Inc., 275 Grove St., Newton, MA 02466, USA

The automatic classification of marine mammal sounds is very attractive as a means of assessing massive quantities of recorded data, freeing humans and offering rigorous and consistent output. Calculations on a set of vocalizations of Northern Resident killer whales using Dynamic Time Warping have been reported recently. [Brown, J. C., and P.J.O. Miller, "Automatic classification of killer whale vocalizations using dynamic time warping," J. Acoust. Soc. Am. 122, 1201-1207, (2007).] Since this method requires the time-consuming pre-processing measurement of the frequency contours, we have explored the use of Gaussian Mixture Models (GMM) and Hidden Markov Models (HMM). These methods can be applied directly to time-frequency decompositions of the recorded signals. Calculations have been made on a set of 75 calls previously classified perceptually into 7 call types. Preliminary results give an agreement of roughly 85% agreement with the perceptual classification for the GMM and over 90% for an HMM.