BIOSOUND – Birdsong recognition on mobile devices

Timos Papadopoulos[1,2], Stephen Roberts[2], and Kathy Willis[1]

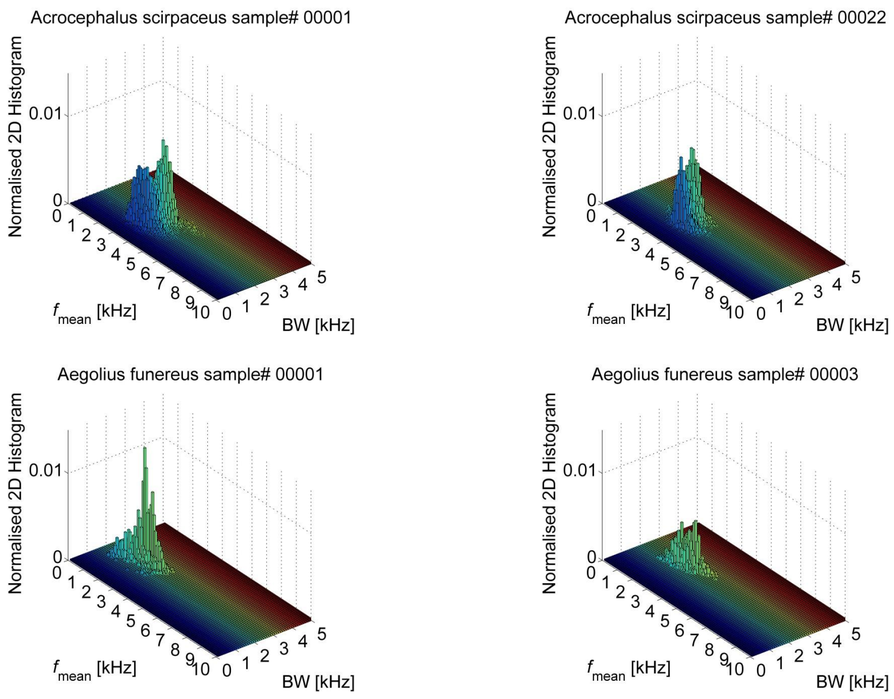
(1) Biodiversity Institute, Department of Zoology, University of Oxford

(2) Machine Learning Research Group, Department of Engineering Science, University of Oxford

Overview

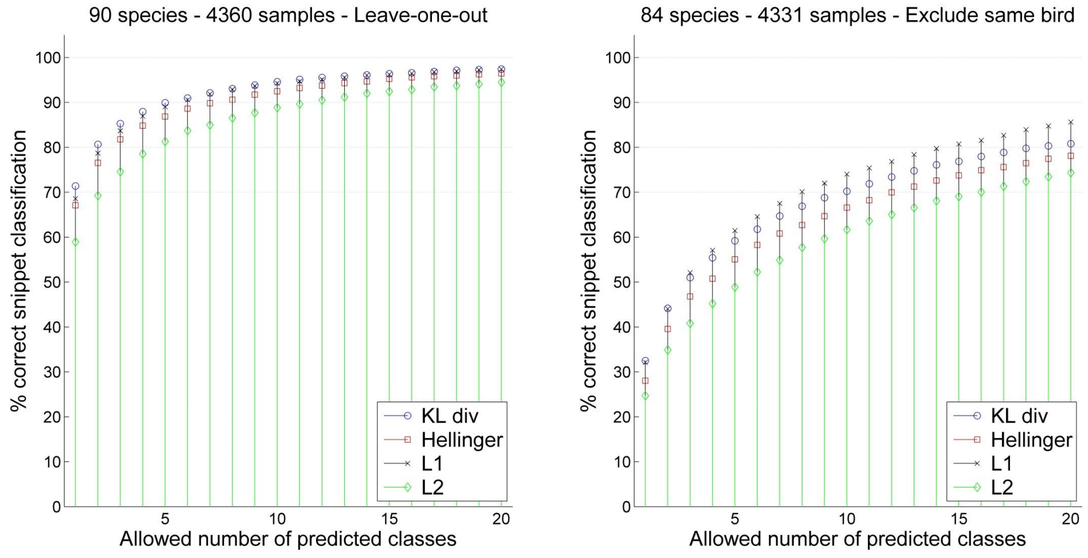
- **Main objectives and specification**
  - Limit computational load for feature extraction and classification to meet mobile device platform limitations
  - Avoid expert syllable segmentation but exploit crowd-sourcing for whole song/call segmentation
  - Surpass existing methods and increase the number of considered labels (species) from O(10) to O(100)
  - Offer multiple label suggestions and fold-in human classification results
- **First (baseline) classification algorithm**
  - Simple two-dimensional vector of spectral features (mean and bandwidth of STFT time slices)
  - Feature space spanned by 2D histograms rather than points corresponding to time-averaged features [1]
  - Nearest Neighbour classifier implemented on a PC - KL divergence, Hellinger, L1 and L2 distance metrics examined

Feature Space



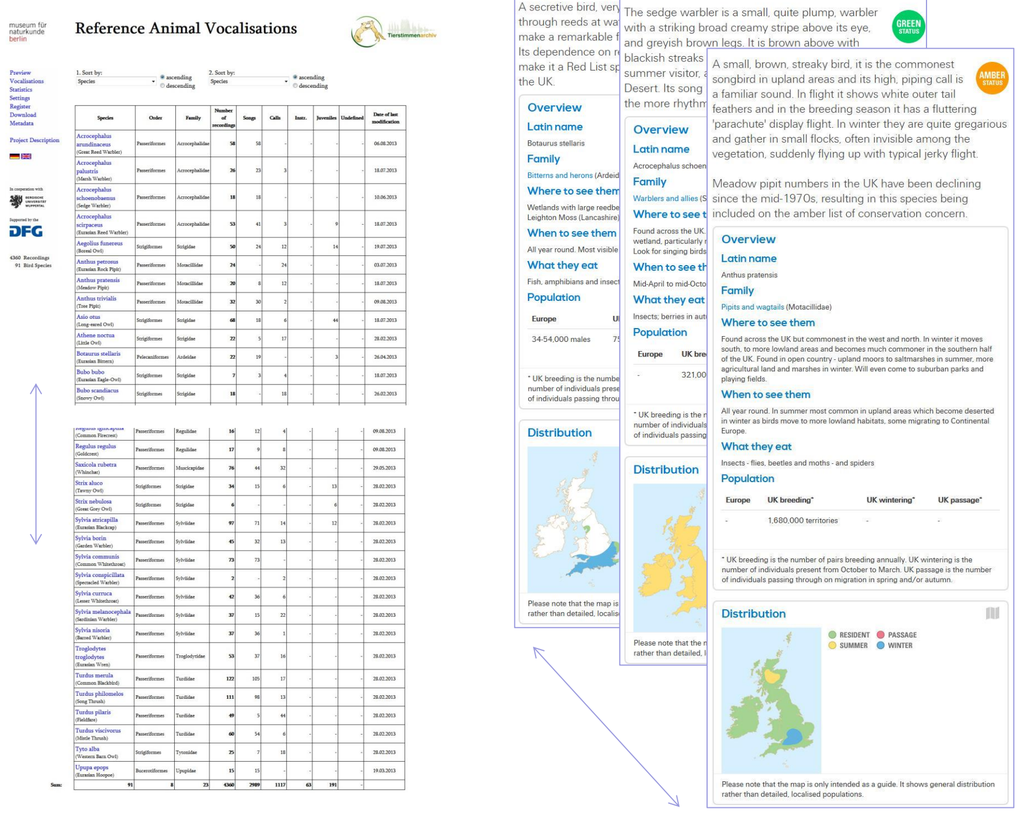
Animal Sound Archive Dataset [2]

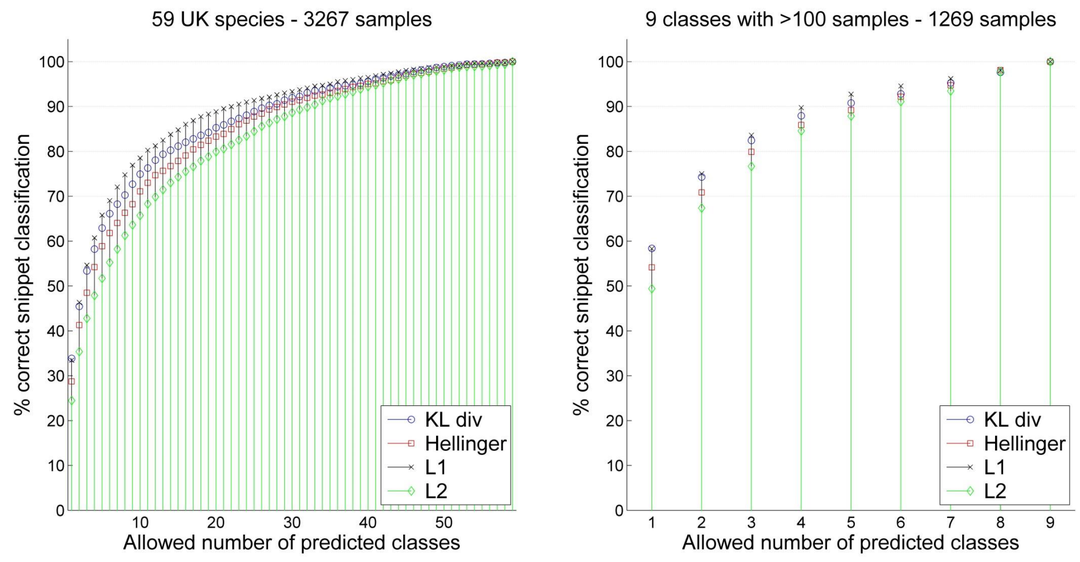
- Data collected over nearly 50 years in numerous locations mainly in Germany
- Varying recording conditions, equipment and parameters
- Approximately 4400 segmented call/song intervals from 91 species
- Ground truth metadata on type of call, individual bird and recording conditions



UK species

The Animal Sound Archive Dataset covers around 60 of the 259 UK species listed by the RSPB [3]





Current Work and Future Plans

- Utilise geographic position, season and time of day information to introduce prior obtained from biodiversity distribution websites to the classifier.
- Evaluation of different options on baseline classification algorithm (different feature sets and feature vector dimensions, distance metrics, introduction of uncertainty estimates, nonparametric estimation of distribution in feature space, increased number of classes based on call/song subcategories)
- Investigation of generalisation properties (external datasets, real world conditioned recordings, online competition datasets)
- Prototype application implemented on laptop PC and tested in real conditions before been ported to mobile device platform
- Investigation of relation between distances in feature space and species phylogenetic relations

References - Acknowledgements

1. Briggs et al. “Audio Classification of Bird Species: A Statistical Manifold Approach” IEEE 9th ICDM (2009)
2. The Animal Sound Archive of the Museum für Naturkunde Berlin http://www.animalsoundarchive.org/RefSys/ProjectDescription.php
3. The Royal Society for the Protection of Birds (RSPB) http://www.rspb.org.uk/wildlife/

We gratefully acknowledge Dr Karl-Heinz Frommolt and the Museum für Naturkunde Berlin for providing the Animal Sound Archive Dataset