



## Automated bioacoustic identification of species

DAVID CHESMORE

Intelligent Systems Research Group, Department of Electronics, University of York  
Heslington, York, YO10 5DD, England

*Manuscript received on January 15, 2004; accepted for publication on February 5, 2004.*

### ABSTRACT

Research into the automated identification of animals by bioacoustics is becoming more widespread mainly due to difficulties in carrying out manual surveys. This paper describes automated recognition of insects (Orthoptera) using time domain signal coding and artificial neural networks. Results of field recordings made in the UK in 2002 are presented which show that it is possible to accurately recognize 4 British Orthoptera species in natural conditions under high levels of interference. Work is under way to increase the number of species recognized.

**Key words:** automated identification, Orthoptera, bioacoustics, time domain signal coding, biodiversity informatics.

### INTRODUCTION

Recognition of insect, animal and bird species from their calls has been employed for many years for identifying individuals and locating animals. However, such "manual" surveys are slow, time consuming and rely heavily on the surveyor's expert knowledge of the group under investigation. Surveys also generally take place at infrequent intervals primarily due to the time required, leading to difficulties in interpreting long-term trends. Rapid advances in computing and electronics are leading to the development of automated recognition systems capable of providing long-term continuous unattended monitoring in inhospitable regions. These systems can be designed for hand-held use and applications range from rapid biodiversity assessment especially in acoustically rich habitats (Riede 1993), electronic identification guides, acoustic autecology and the detection and recognition of pest species. Research into automated bioacoustic species iden-

tification is more mature in some fields than others. Table I gives some examples of bioacoustic research.

This paper describes the development of a novel bioacoustic signal recognition system (IBIS – Intelligent Bioacoustic signal Identification System) and its application to the recognition of British Orthoptera. The technique employed is a purely time domain method known as Time Domain Signal Coding (TDSC) which, when coupled with an artificial neural network (ANN) classifier, provides a powerful vehicle for bioacoustic signal analysis and recognition. It has been successfully tested on 25 species of British Orthoptera with 99% recognition accuracy (Chesmore et al. 1997, Chesmore 2000, 2001, Chesmore and Nellenbach 2001) and 10 species of Japanese bird with 100% accuracy (Chesmore 1999, 2001). However, these results were for high signal to noise ratio (SNR) signals. This paper describes results of field trials where the SNR is more variable and sounds are corrupted by interference from other natural and man-made noise sources.

**TABLE I**  
**Examples of automated bioacoustic species identification.**

Application	Reference(s)
Orthoptera	Chesmore et al. 1997, Chesmore 2001, Chesmore and Nellenbach 2001, Schwenker et al. 2003
Cicadas	Ohya and Chesmore 2003
Mosquitoes	Campbell et al. 1996
Birds (general)	McIlraith and Card 1995, Anderson et al. 1996
Individual bird recognition	Terry and McGregor 2002
Nocturnal migrating birds	Mills 1995
Frogs and other amphibia	Taylor et al. 1996
Cetaceans	Murray et al. 1998
Deer	Reby et al. 1997
Elephants (individuals and vocalizations)	Clemins and Johnson 2002
Bats	Vaughan et al. 1997, Parsons and Jones 2000, Parsons 2001

## MATERIALS AND METHODS

### SOUND RECORDINGS AND TEST DATA

Recordings were made between June and September 2002 at a variety of sites and habitats in North and East Yorkshire, England. Sounds were recorded on a Sony MZ-R90 portable minidisc recorder with a Sony ECM-MS907 condenser microphone and transferred to a PC (Dell Inspiron 8100) via a standard sound card. The sounds were sampled at 44.1 kHz and stored as 16-bit signed mono .wav format files using Avisoft-SASLab Pro software package. Table II lists the 6 species encountered during the recording sessions; only 4 are used in the acoustic study.

Each recording was “manually” examined for echemes (first order assemblage of syllables) and songs of varying quality; these were extracted to separate files for training and testing purposes.

### SIGNAL ANALYSIS AND RECOGNITION

The basic principle of TDSC is to characterize the “shape” of the waveform between successive zero-crossings of the signal (termed an “epoch”). Full

details of the algorithm can be found in Chesmore (2001). The output of the coding process is a stream of codewords (1 per epoch) describing changes in the shape of the waveform over time. Further processing is carried out in 2 ways: accumulation of the frequency of occurrence of each codeword – the S-matrix, and the frequency of occurrence of pairs of codewords – the A-matrix. The A-matrix is employed in this application. Recognition of sounds via A-matrices is carried out using an artificial neural network (ANN) which takes the A-matrix as input and has an output for each species (or sound) to be recognized. The system operates in 2 phases, training phase and operational phase. In the training phase, high quality examples of sounds that are to be identified (known as exemplars) are used to train the ANN so that the correct ANN output is activated. Training occurs by repeated presentation of the sounds and modification of the weights within the network in such a way as to reduce the overall error between the current outputs and desired outputs. Training continues until the overall error is below a given threshold. Once trained, the system is ready to use and unknown sounds can be classi-

TABLE II  
Orthoptera species recorded in Yorkshire in 2002.

Vernacular name	Scientific name
Common Green Grasshopper	<i>Omocestus viridulus</i> (Linnaeus, 1758)
Meadow Grasshopper	<i>Chorthippus parallelus</i> (Zetterstedt, 1821)
Lesser Marsh Grasshopper	<i>Chorthippus albomarginatus</i> (De Geer, 1773)
Mottled Grasshopper	<i>Myrmeleotettix maculatus</i> (Thunberg, 1815)
Short-winged Cone-head	<i>Conocephalus dorsalis</i> (Latreille, 1804)*
Common Groundhopper	<i>Tetrix undulata</i> (Sowerby 1806)**

\* Species not used in this acoustic study. \*\* Does not produce any acoustic signal.

fied. Each of the outputs will give a value between 0.0 (zero match) and 1.0 (perfect match); the unknown sound being recognized as the output with the highest value. The type of ANN used in this application is a standard multilayer perceptron (MLP) with backpropagation training. Upon listening to the recordings, it was discovered that there were many other sounds present, mainly man-made and it was decided to include these sounds for recognition. Representative sounds for each category (insect, animal, man-made) were selected, stored as separate .wav files and used to train the ANN. The following 13 sound sources were used in training:

- 4 grasshopper species;
- 1 blow fly sound (wing beats of unknown species);
- 4 bird sounds (3 different alarm calls of undetermined origin and Chiffchaff *Phylloscopus collybita*);
- 2 vehicle (car) sounds (metaled road and dirt road);
- 1 single engine light aircraft sound;
- 1 background sound (sound when no other sources present – includes wind noise).

## RESULTS

Testing of the recognition system was carried out in 3 ways: recognition of single echemes, recognition of whole songs and recognition of sounds in 2s

intervals. The latter approach does not rely on *a priori* knowledge of the signals (e.g. start of echeme or song) but simply allocates a sound to 2s intervals; this leads to the possibility of generating continuous sound maps.

### RECOGNITION OF SINGLE ECHES

Echeme duration for the 4 species under consideration is approximately 2s. Echemes were manually extracted from the recordings and stored as separate .wav files. Table III gives results for the 4 species which were recognized from 13 sounds. The threshold is used to remove any recognition results below the threshold to reduce low accuracy results. It is evident that recognition accuracy for a threshold of 0.9 is between 81.8% (*C. parallelus*) and 100% (*O. viridulus*) whereas with no threshold they drop to 64.3% and 97% respectively, and with *M. maculatus* dropping is from 90% to 41.2%.

### RECOGNITION OF WHOLE SONGS

Figure 1 shows results of whole song recognition with varying threshold. Recognition for a threshold of 0.9 is between 80% and 100% for the 4 species. *O. viridulus* has a song that can last for more than 30s so a 10s segment was selected.

### CONTINUOUS SOUND RECOGNITION

It is possible to simply recognize sound on a short time scale without any *a priori* knowledge of the signals thus reducing computational overheads in

TABLE III

Identification accuracy of grasshopper vocalizations for different threshold levels, single echeme samples. Assumes all outputs less than threshold are rejected.

Threshold		<i>O. viridulus</i>	<i>M. maculatus</i>	<i>C. parallelus</i>	<i>C. albomarginatus</i>
sample size		34	17	14	16
<b>0.5</b>	<b>rejected</b>	2	4	1	1
	<b>accuracy</b>	100%(32/32)	76.9%(10/13)	69.2%(9/13)	86.7% (13/15)
<b>0.6</b>	<b>rejected</b>	2	5	2	1
	<b>accuracy</b>	100%(32/32)	75% (9/12)	75% (9/12)	86.7% (13/15)
<b>0.7</b>	<b>rejected</b>	3	5	2	1
	<b>accuracy</b>	100%(31/31)	75% (9/12)	75% (9/12)	86.7% (13/15)
<b>0.8</b>	<b>rejected</b>	4	5	3	2
	<b>accuracy</b>	100%(30/30)	75% (9/12)	81.8%(9/11)	85.7% (12/14)
<b>0.9</b>	<b>rejected</b>	4	7	3	2
	<b>accuracy</b>	100%(30/30)	90% (9/10)	81.8%(9/11)	85.7% (12/14)
<b>None</b>	<b>accuracy</b>	97% (33/34)	41.2% (7/17)	64.3 (9/14)	87.5% (14/16)

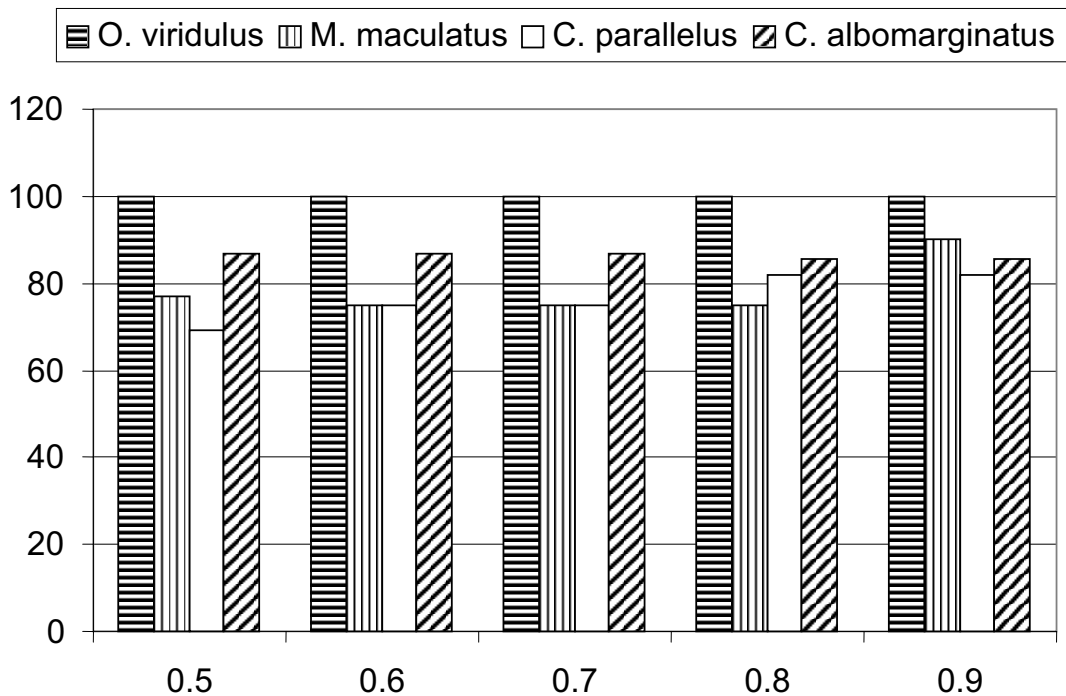


Fig. 1 – Graph of whole song recognition with varying threshold. Y-axis is accuracy in percentage and x-axis is threshold level (0.5-0.9).

locating specific signals. In these tests, each recording was analyzed in approximately 2s blocks; block length will depend on the typical characteristics of

the acoustic environment. Figure 2 shows the results for an 18s segment, each 2s block has the recognition results shown below the graph. It is evident

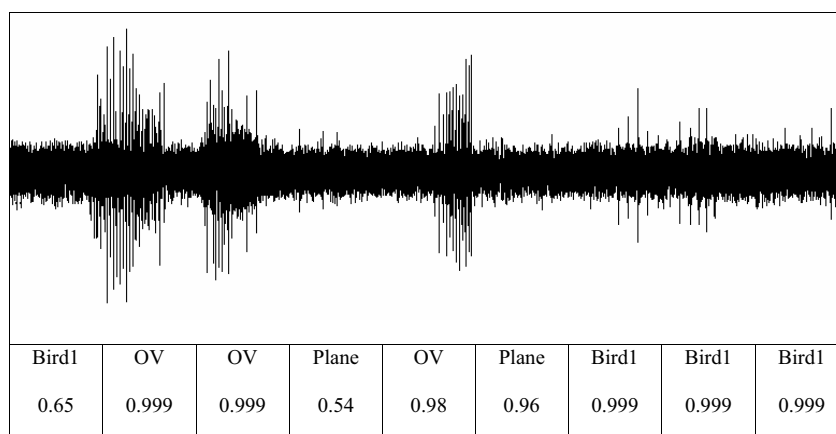


Fig. 2 – Classified sounds from an 18s sequence on a 2s interval recorded at Allerthorpe Common, East Yorkshire on 15 July 2002. The system correctly recognizes 3 short songs by *O. viridulus* (OV), a light aircraft (Plane) and a bird alarm call (Bird1).

that the grasshopper (*O. viridulus*) has been recognized correctly, as have the light aircraft and a bird alarm call. This approach has considerable potential for general sound mapping applications where both sound pressure level and sound type could be monitored.

### CONCLUSIONS

This paper has shown that it is possible to accurately and reliably recognize sounds in a noisy field environment. One important aspect of this research is that the techniques employed are suitable for implementation on hand-held or stand-alone field deployable devices leading to the potential for long-term continuous monitoring. Much work has still to be carried out, in particular better wave shape descriptors and investigation into separation of multiple simultaneous calls. TDSC is not limited to insect sounds and a real-time hand-held recognition system is being developed for British bats.

### ACKNOWLEDGMENTS

The author would like to acknowledge English Nature, the Forestry Commission, East Riding County Council of Yorkshire and the Yorkshire Wildlife Trust for granting access to some of the sites.

### RESUMO

Pesquisas sobre a identificação automatizada de animais através da bioacústica estão se ampliando, principalmente em vista das dificuldades para realizar levantamentos diretos. Este artigo descreve o reconhecimento automático de insetos Orthoptera utilizando a codificação de sinal no domínio temporal e redes neurais artificiais. Resultados de registros sonoros feitos no campo no Reino Unido em 2002 são apresentados, mostrando ser possível reconhecer corretamente 4 espécies britânicas de Orthoptera em condições naturais com altos níveis de interferências. Estão em andamento trabalhos para aumentar o número de espécies identificadas.

**Palavras-chave:** identificação automatizada, Orthoptera, bioacústica, codificação de sinais temporais, informática da biodiversidade.

### REFERENCES

- ANDERSON SE, DAVE AS AND MARGOLIASH D. 1996. Template-based automatic recognition of birdsong syllables from continuous recordings. *J Acoust Soc Amer* 100: 1209-1217.
- CAMPBELL RH, MARTIN SK, SCHNEIDER I AND MICHELSON WR. 1996. Analysis of mosquito wing beat sound. 132<sup>nd</sup> Meeting of the Acoustical Society of America, Honolulu.
- CHESMORE ED. 1999. Technology Transfer: Applications of Electronic Technology in Ecology and En-

- tomology for Species Identification. *Nat Hist Res* 5: 111-126.
- CHESMORE ED. 2000. Methodologies for automating the identification of species. Proceedings of 1<sup>st</sup> BioNet-International Working Group on Automated Taxonomy, July 1997: 3-12.
- CHESMORE ED. 2001. Application of time domain signal coding and artificial neural networks to passive acoustical identification of animals. *Applied Acoustics* 62: 1359-1374.
- CHESMORE ED AND NELLENBACH C. 2001. Acoustic Methods for the Automated Detection and Identification of Insects. *Acta Horticulturæ* 562: 223-231.
- CHESMORE ED, SWARBRICK MD AND FEMMINELLA OP. 1997. Automated analysis of insect sounds using TESPAN and expert systems – a new method for species identification. In: BRIDGE P. (Ed), *Information Technology, Plant Pathology and Biodiversity*. Wallingford, UK: CAB International, p. 273-287.
- CLEMINS PJ AND JOHNSON MT. 2002. Automatic type classification and speaker verification of African Elephant vocalizations. *Vrije Universiteit, Netherlands: International Conference on Animal Behavior, Proceedings*.
- MCILRAITH AL AND CARD HC. 1995. Birdsong recognition with DSP and neural networks. Winnipeg, Canada: IEEE WESCANEX'95 Proceedings, p. 409-414.
- MILLS H. 1995. Automatic detection and classification of nocturnal migrant bird calls. *J Acoust Soc Amer* 97: 3370-3371.
- MURRAY SO, MERCADO E AND ROITBLAT HL. 1998. The neural network classification of False Killer Whale *Pseudorca crassidens* vocalizations. *J Acoust Soc Amer* 104: 3626-3633.
- OHYA E AND CHESMORE ED. 2003. Automated identification of grasshoppers by their songs. Iwate University, Morioka, Japan: Annual Meeting of the Japanese Society of Applied Entomology and Zoology.
- PARSONS S. 2001. Identification of New Zealand bats in flight from analysis of echolocation calls by artificial neural networks. *J Zool London* 253: 447-456.
- PARSONS S AND JONES G. 2000. Acoustic identification of 12 species of echolocating bats by discriminant function analysis and artificial neural networks. *J Exp Biol* 203: 2641-2656.
- REBY D, LEK S, DIMOPOULOS I, JOACHIM J, LAUGA J AND AULAGNIER S. 1997. Artificial neural networks as a classification method in the behavioural sciences. *Behav Proc* 40: 35-43.
- RIEDE K. 1993. Monitoring biodiversity: analysis of Amazonian rainforest sounds. *Ambio* 22: 546-548.
- SCHWENKER F, DIETRICH C, KESTLER HA, RIEDE K AND PALM G. 2003. Radial basis function neural networks and temporal fusion for the classification of bioacoustic time series. *Neurocomputing* 51: 265-275.
- TAYLOR A, GRIGG G, WATSON G AND MCCALLUM H. 1996. Monitoring frog communities, an application of machine learning. Eighth Innovative Applications of Artificial Intelligence Conference. Portland, Oregon, USA: AAAI Press.
- TERRY AMR AND MCGREGOR PK. 2002. Census and monitoring based on individually identifiable vocalizations: the role of neural networks. *Anim Conserv* 5: 103-111.
- VAUGHAN N, JONES G AND HARRIS S. 1997. Identification of British bat species by multivariate analysis of echolocation call parameters. *Bioacoustics* 7: 189-207.