

An acoustic system for the individual recognition of insects

Bo Zhang, Jinkui Cheng, Yan Han, and Liqiang Ji^{a)}

Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Beijing 100101, People's Republic of China

Fuming Shi

College of Life Sciences, Hebei University, Baoding 071002, People's Republic of China

(Received 19 August 2011; revised 7 February 2012; accepted 12 February 2012)

Research into acoustic recognition systems for insects has focused on species identification rather than individual identification. In this paper, the feasibility of applying pattern recognition techniques to construct an acoustic system capable of automatic individual recognition for insects is investigated analytically and experimentally across two species of Orthoptera. Mel-frequency cepstral coefficients serve as the acoustic feature, and α -Gaussian mixture models were selected as the classification models. The performance of the proposed acoustic system is promising and displays high accuracy. The results suggest that the acoustic feature and classifier method developed here have potential for individual animal recognition and can be applied to other species of interest.

© 2012 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.3692236>]

PACS number(s): 43.60.Lq, 43.60.Uv, 43.80.Ka [SAF]

Pages: 2859–2865

I. INTRODUCTION

Many animals produce sounds either deliberately for communication purposes (nonincidental sounds), or as a by-product of activities such as eating, moving, or flying (incidental sounds) (Chesmore, 2001). Nonincidental sounds have evolved to be species-specific, and have been used to detect and identify animal species for many years. Many animal calls can also be individual-specific, including those of fish (Crawford *et al.*, 1997), amphibians (Bee *et al.*, 2001), birds (Blumstein and Munos, 2005; Policht *et al.*, 2009), and mammals (Frommolt *et al.*, 1997; Darden *et al.*, 2003; Frommolt *et al.*, 2003). This characteristic enables the study of individual identification using only acoustic modalities. Research into the individual identification of insects using only the acoustic modality is lacking, and some claimed that insects do not exhibit sufficient intraspecific variation in sound production (Chesmore, 2001), which would be necessary to achieve individual acoustic identification. However, developments in computer technology and pattern recognition may have now made this possible for animals such as insects.

Effective acoustic features that show greater variation between, rather than within, species or individuals, and well-developed classification models, are essential for species identification and individual identification. Acoustic features such as time domain signal coding (TDSC) (Chesmore, 2001), linear prediction cepstral coefficients (LPCC) (Akhtar *et al.*, 2001; Lewis and Ramachandran, 2001), Mel frequency cepstral coefficients (MFCC) (Picone, 1993; Campbell, 1997; Mashao and Skosan, 2006), and linear frequency cepstral coefficients (LFCC) (Ganchev and Potamitis, 2007; Ganchev *et al.*, 2007) have been successfully applied to the automatic species and other forms of recognition in insects. For example, using TDSC, Chesmore (2001) and Chesmore

and Ohya (2004) experimented on 25 species of Orthoptera (including grasshoppers) and four British grasshoppers and achieved species recognition accuracies of between 70% and 100%; Chesmore (2004) later implemented the same features on four British Orthoptera species, and recognition accuracies of between 81.8% and 100% were achieved. Tang *et al.* (2005) achieved recognition accuracies of 95% in stored-product insects using LPCC. Zhu *et al.* (2010) applied MFCC to identify several groups of insects, achieving more than 96% recognition accuracy. Using LFCC, Ganchev and Potamitis (2007) and Ganchev *et al.* (2007) produced an identification accuracy exceeding 98% at the levels of suborder and family, and 86% at the level of species (using 313 species). Using the same features, Potamitis *et al.* (2006, 2007) and Potamitis *et al.* (2009) reported high recognition accuracies in crickets and cicadas. Additional features such as peak frequency have also been applied to this field and yielded good identification accuracies (Potamitis *et al.*, 2006, 2007; Pinhas *et al.*, 2008; Potamitis *et al.*, 2009).

Classification models, especially those developed for human speech recognition, have been applied to species recognition in insects. For example, support vector machines (Tang *et al.*, 2005), neural network (including artificial neural network and probabilistic neural network) (Chesmore, 2001; Schwenker *et al.*, 2003; Chesmore, 2004; Chesmore and Ohya, 2004), hidden Markov models (Ganchev *et al.*, 2007), Gaussian mixture models (GMM) (Potamitis *et al.*, 2006; Ganchev and Potamitis, 2007; Ganchev *et al.*, 2007; Potamitis *et al.*, 2007; Pinhas *et al.*, 2008; Potamitis *et al.*, 2009), and vector quantization (Pinhas *et al.*, 2008; Zhu *et al.*, 2010) have been applied to this type of research. Although there has been no research into the individual recognition of insects using only the acoustic modality, studies on the individual recognition of other animals based on vocalizations (Clemins *et al.*, 2005; Trawicki *et al.*, 2005) can benefit our research into individual recognition of insects.

^{a)}Author to whom correspondence should be addressed. Electronic mail: ji@ioz.ac.cn

In this paper we aim to construct a recognition system and test the ability of MFCC features combined with α -Gaussian mixture models (α -GMM) (Wu *et al.*, 2009) to individually recognize the two Orthoptera species *Deracantha onos* and *Gampsodeis gratiosa*.

II. DATA AND METHODS

A. Data

The male calls of *Deracantha onos* are trills with dominant frequencies in the range 13–20 kHz. The pulse repetition rate at 23 °C is between 19 and 38 per second (Zhantiev *et al.*, 1995). Male *Gampsodeis gratiosa* song sequences last 4–232 s, dominant frequencies are in the range 2.56–22.03 kHz, and the frequency of the two highest power peaks are 6.87 and 3.25 kHz (Wang *et al.*, 2008).

Four individual's sounds were used for *Deracantha onos* and 15 for *Gampsodeis gratiosa* in this paper. Spectrograms of the songs are illustrated in Figs. 1 and 2.

B. Methods

Effective acoustic features that show greater variation between, rather than within, species or individuals, and excellent classification models, are essential for species and individual identification. Therefore, feature extraction and classification are consequently included in the architecture of our acoustic-driven individual recognition system for insects. Preprocessing of insect sound signals is also included in the architecture (Fig. 3). All insect sounds used in this paper have had silence and human voices deleted.

1. Feature extraction

Among methods that can demonstrate sound information, LPCC and MFCC are the most frequently used. Between LPCC and MFCC, both known as cepstral coefficients, many studies of speech recognition have shown that

MFCC outperforms LPCC in recognition accuracy (Davis and Mermelstein, 1980; Openshaw *et al.*, 1993; Vergin *et al.*, 1996), and is thus widely used for animal sound recognition purposes (Lee *et al.*, 2006; Juang and Chen, 2007; Trifa *et al.*, 2008; Cheng *et al.*, 2010; Zhu *et al.*, 2010). The advantage of MFCC compared to LPCC is that it follows a Mel scale instead of a linear scale (used by LPCC) and is based on no assumptions. These two advantages mean that MFCC is more robust to noise and it is for these reasons that we chose MFCC as the feature vectors for insects.

We obtained feature vectors by combining MFCC and the first-order difference of MFCC. The method we used to get MFCC is shown in Fig. 4. Frame size for the two species is 400 samples (for *Deracantha onos*) and 222 samples (for *Gampsodeis gratiosa*), both corresponding to ~ 10 ms. We gave the dimension of MFCC the value of 15 because previous research has shown almost all the information needed for classification is obtained when the dimension of MFCC is 12 for speech recognition and 16 for speaker recognition (Zhen *et al.*, 2001). We investigated all dimensions of MFCC between 12 and 16 and found that 15 gave the best performance. The number of the Mel scale filter bank was set to 24, which is the general method for speaker and animal recognition. We deleted the first dimension of MFCC because we wanted to eliminate any dependence on the recorders. Last, we obtained the feature vectors by combining MFCC and the first-order difference of MFCC. Our feature vectors are in 28 dimensions.

2. Classification

In the field of speech recognition, there are two classification models: Probabilistic models and nonprobabilistic models. It has been proposed, and research has proved, that probabilistic models, including GMM, are more capable of dealing with noise and channel degradation than are nonprobabilistic models (Fox, 2008). The classification model GMM

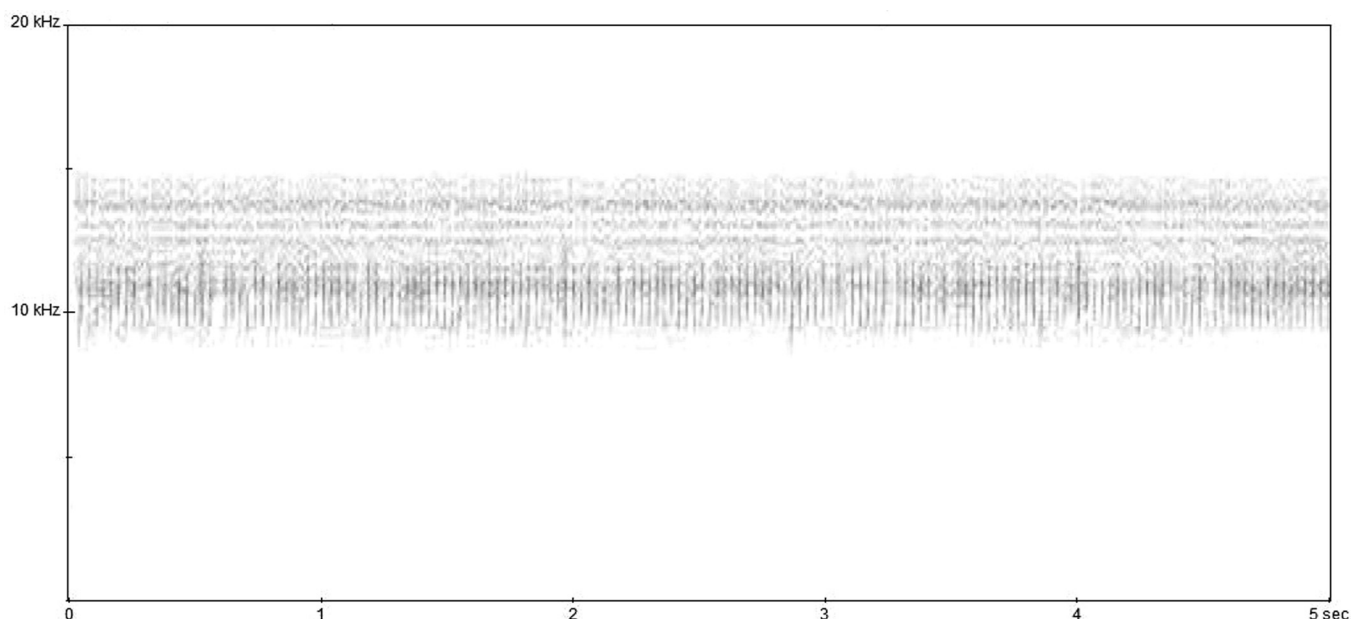


FIG. 1. Spectrogram of the songs of *Deracantha onos*.

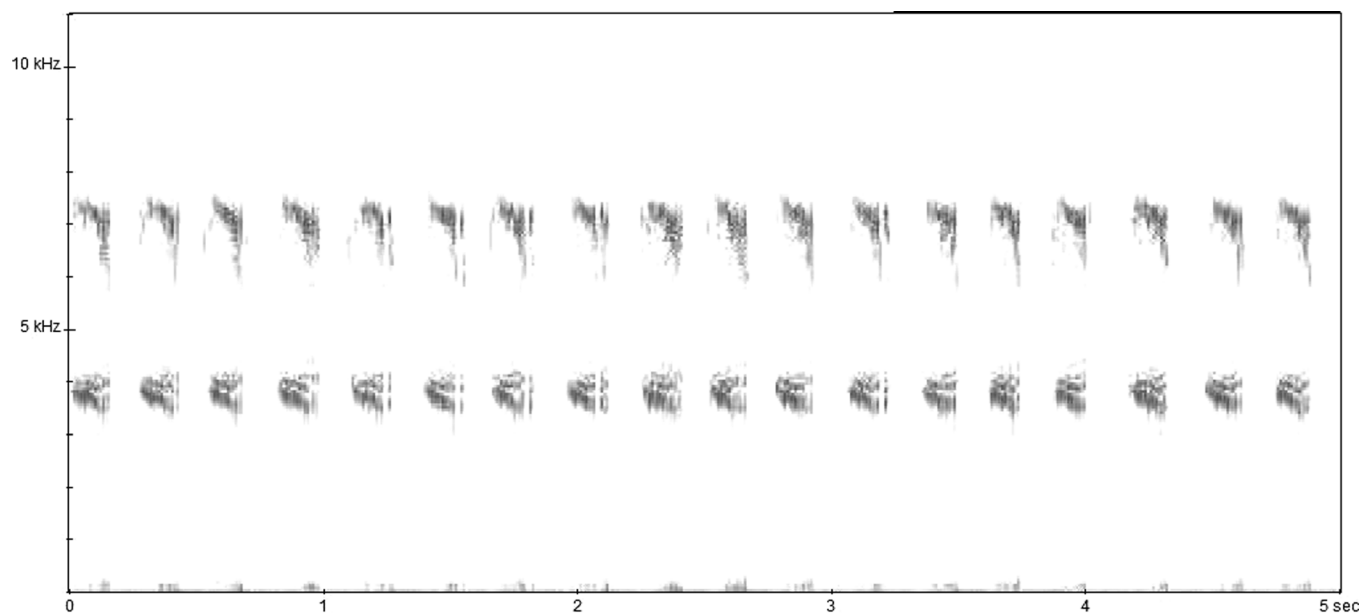


FIG. 2. Spectrogram of the songs of *Gampsodeis gratiosa*.

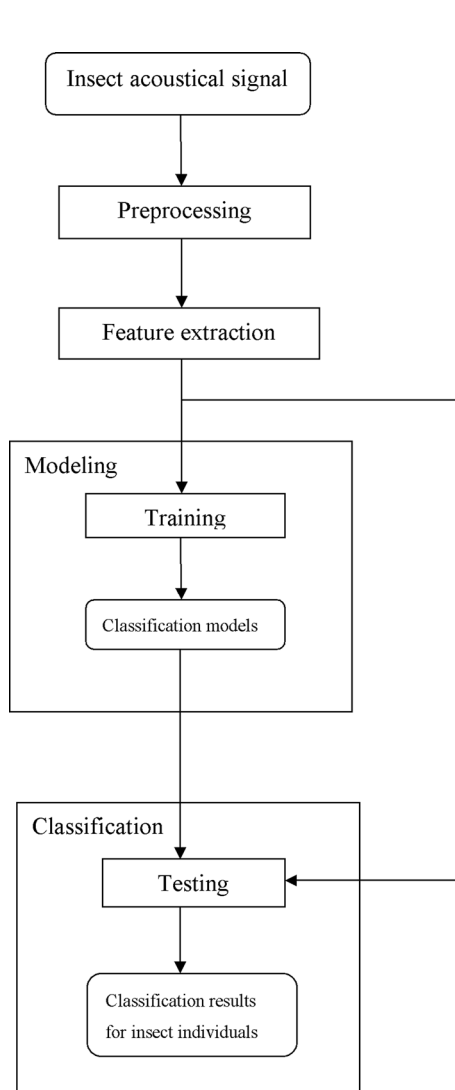


FIG. 3. Architecture of our acoustic system for the individual recognition of insects.

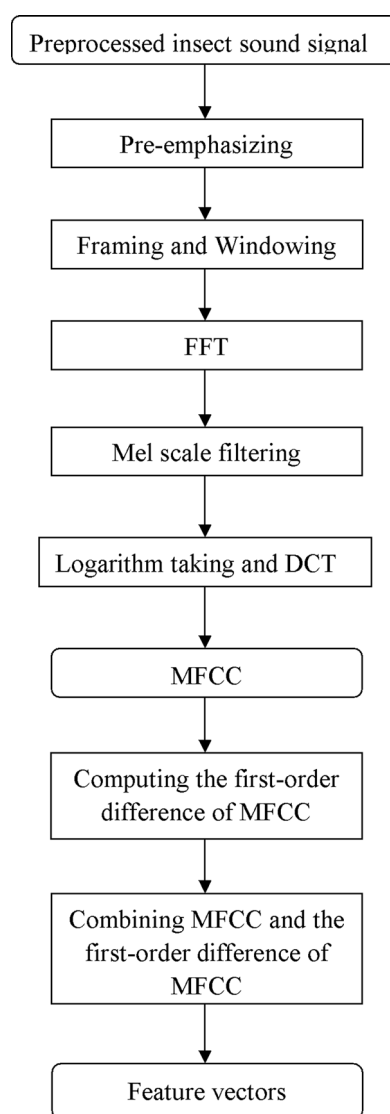


FIG. 4. Feature vectors extraction from insect sounds used in our study.

has the ability to model arbitrary densities (Redner and Walker, 1984; Reynolds and Rose, 1995) and has been widely used in the fields of both speech recognition and animal sound recognition (Cheng et al., 2010; Zhu et al., 2010). However, α -GMM, a newly proposed probabilistic model, has shown better performance than GMM in speech recognition (Wu et al., 2009). Given that α -GMM is a more sophisticated model for integrating stochastic modeling components in a nonlinear way (conventional GMM can be regarded as combining its components in a linear way), α -GMM has the ability to render a recognition system more robust to noise (Wu et al., 2009). Therefore, we chose α -GMM as the classification model for our system.

a. Modeling of α -GMM. The main difference between α -GMM and conventional GMM lies in the fact that α -GMM is obtained through α -integration based on conventional GMM. α -GMM can be defined as follows (Wu et al., 2009):

$$p_{\alpha}(X|S) = \begin{cases} c \left(\sum_{i=1}^K w_i N_i(X|S)^{(1-\alpha)/2} \right)^{2/(1-\alpha)}, & \alpha \neq 1 \\ c \exp \left(\sum_{i=1}^K w_i \log N_i(X|S) \right), & \alpha = 1 \end{cases} \quad (1)$$

where c is a normalization constant, K is the number of Gaussians, and $N_i(X|S)$ is the i th Gaussian distribution for a given insect sound. Equation (1) is thus a conventional GMM when $\alpha = -1$.

The nature of the idea of α -GMM is its extension of the classical GMM in modeling capacity. More powerful integrated models can be obtained by introducing a new factor of α . By using negative α -values, the integrated α -GMM de-emphasizes small values but emphasizes large values; this increases the tolerance for mismatches between training and testing and makes a recognition system built on α -GMM more robust to noise.

b. Parameter estimation of α -GMM. Parameter estimation of α -GMM was conducted through training using insect sounds. As in estimating the parameters of conventional GMM, an expectation maximization (Dempster et al., 1977) algorithm can be used to estimate the parameters of α -GMM for the case of $\alpha \neq 1$, but with a minor change as the definition of the models has been altered (Wu, 2009).

c. Decision. With every set of parameters corresponding to each α -GMM belonging to a certain individual insect having been fixed, the decision was made by computing the posteriori probability after putting test vectors into our system in the classification process. The decision criterion for insect individual identification is as follows:

$$i = \arg \max \sum_{t=1}^T \log(p_{\alpha}(x_t|\lambda_j)), \quad j = 1, 2, \dots, M \quad (2)$$

where λ_j denotes the j th α -GMM corresponding to a certain individual insect, $\{x_t\}$ is a frame sequence of one insect sound to be tested, and M is the number of insect individuals.

III. EXPERIMENTS

A. Test on individuals under different environmental conditions: *Deracantha onos*

We first conducted the experiment on *Deracantha onos*. Sounds from different individuals were not recorded under the same environmental conditions (such as temperature and humidity). Four individual's sounds were recorded using a Sony MZ-R55 Digital CD recorder in Fengning and Yi counties, Hebei, China. The sampling rate was 40 000 Hz.

We split each audio file in two equal nonoverlapping parts: 50% for training and 50% for testing. In the process of testing, a 1 s testing signal was used (audio files for testing were therefore divided into a series of 1 s samples).

B. Test on individuals under the same environmental conditions: *Gampsodeis gratiosa*

To make the results more convincing, we then repeated the experiment on *Gampsodeis gratiosa*, where all individuals' sounds were recorded at the same time of day. This ensured that the environmental conditions were almost the same for different individuals during recording. Fifteen individuals' sounds were recorded at Olympic Forest Park, Beijing, China. These recordings were made using a WM-D6c professional recorder (Sony Corporation, Tokyo, Japan) with a directional microphone (Sennheiser, Wedemark, Germany) placed 10–15 cm from a singing insect. Recordings were converted to a digital medium at 22 050 Hz sampling frequency and saved in 8 bit wave format using Batsound v3.10 (Pettersson Elektronik AB, Uppsala, Sweden). Each individual was recorded three times (morning, noon, and evening) and this yielded three groups of data: Morning *Gampsodeis gratiosa*, noon *Gampsodeis gratiosa*, and evening *Gampsodeis gratiosa*.

As we did on *Deracantha onos*, we also split each audio file in two equal nonoverlapping parts: 50% for training and 50% for testing. In the process of testing, a 1 s testing signal was used (audio files for testing were therefore divided into a series of 1 s samples). Data from all three groups were tested using the proposed system.

IV. RESULTS

In order to provide an efficient evaluation of our automatic individual identification system with respect to the value of α , we modeled each insect by letting the value of α equal 0, -1 , -2 , \dots , -12 , similar to methods for speaker recognition (Wu et al., 2009). We did this because α -GMM emphasizes the large values of its probability density function components and de-emphasizes the small values when $\alpha < 1$, which means a small α is more effective in increasing the accuracy of our identification system. In order to investigate an α -GMM model with component densities that exhibit the best performance, each insect species was modeled using 4, 8, 16, 32, 48, and 64 component densities.

Identification performances for different α values and model orders in both of the two species are shown in Tables I and II. From our results it is clear that the performance of α -GMMs fails to remain constant for all α values as the

TABLE I. α -GMM performance across varied α values and model orders in *Deracantha onos*.

α values	Model order					
	4	8	16	32	48	64
0	71.7%	84.1%	86.7%	75.2%	77.0%	71.7%
-1	88.5%	90.3%	89.4%	88.5%	90.3%	89.4%
-2	88.5%	89.4%	90.3%	89.4%	91.2%	89.4%
-3	88.5%	90.3%	91.2%	90.3%	91.2%	89.4%
-4	88.5%	90.3%	91.2%	90.3%	91.2%	89.4%
-5	88.5%	91.2%	91.2%	90.3%	89.4%	89.4%
-6	88.5%	91.2%	92.0%	91.2%	91.2%	89.4%
-7	88.5%	91.2%	90.3%	90.3%	89.4%	89.4%
-8	88.5%	91.2%	91.2%	90.3%	90.3%	89.4%
-9	88.5%	90.3%	93.8%	90.3%	90.3%	89.4%
-10	88.5%	90.3%	92.0%	90.3%	90.3%	89.4%
-11	88.5%	90.3%	92.9%	90.3%	89.4%	89.4%
-12	88.5%	90.3%	93.8%	89.4%	89.4%	89.4%

number of mixture components changes from 4 to 64. The best model order is 16 for *Deracantha onos*, and 16 and 32 for *Gampsodeis gratiosa*. Similarly, the α -GMMs are also sensitive to α values. The best values of α are -9 and -12 for *Deracantha onos*, and -4 and -11 for *Gampsodeis gratiosa*. Performance data reflect that higher accuracies can be achieved by selecting a proper value of α , which is the essence of α -GMM.

Identification performances shown in Table II are for the morning *Gampsodeis gratiosa* group only because the results based on the noon *Gampsodeis gratiosa* and evening *Gampsodeis gratiosa* groups were similar to the morning *Gampsodeis gratiosa* group.

The accuracy for each species obtained through our system reached 93.8% for *Deracantha onos*, and 97.89% for *Gampsodeis gratiosa*. The confusion matrix based on the highest identification accuracy of *Deracantha onos*, through which we can clearly see how many samples are correctly classified and how many are not, is shown in Table III. The same confusion matrix for *Gampsodeis gratiosa* is presented in Table IV.

TABLE II. α -GMM performance across varied α values and model orders in *Gampsodeis gratiosa*.

α values	Model order					
	4	8	16	32	48	64
0	96.95%	97.42%	96.71%	94.60%	89.44%	89.44%
-1	96.95%	96.95%	97.65%	90.14%	92.02%	87.79%
-2	96.95%	97.65%	97.42%	93.90%	92.72%	84.74%
-3	96.95%	96.95%	96.48%	94.37%	89.67%	89.67%
-4	96.95%	97.42%	97.89%	96.71%	89.44%	85.68%
-5	97.18%	96.95%	97.18%	91.08%	91.08%	88.26%
-6	97.18%	96.95%	97.42%	95.77%	93.66%	83.10%
-7	97.18%	97.65%	97.65%	95.54%	90.61%	88.26%
-8	96.95%	96.95%	97.42%	95.77%	92.25%	84.74%
-9	97.18%	97.18%	96.71%	96.71%	90.61%	84.51%
-10	96.95%	97.65%	97.18%	94.13%	94.13%	86.62%
-11	97.18%	97.42%	97.18%	97.89%	90.85%	88.50%
-12	97.18%	97.42%	97.65%	95.77%	90.14%	88.73%

TABLE III. Confusion matrix of the recognition results in *Deracantha onos*.

Known class	Predicted class by α -GMM			
	1	2	3	4
1	23	2	1	0
2	0	30	0	0
3	0	0	28	2
4	0	0	2	25

V. DISCUSSION

We succeeded in illustrating that the individual recognition of insects through the sole use of acoustic information, is achievable. MFCC and α -GMM provide potential robustness to applications under various noise conditions, and their combination in our system results in even greater robustness to noise and high identification accuracies.

The number of individual recordings for the two species, although not ideal, was sufficient and acceptable for our system. It was difficult to obtain more recordings because research into individuals within insect species is extremely rare.

The system we present here for the automatic individual recognition using Mel-frequency cepstral coefficients and α -Gaussian mixture models has shown promising results. Accuracies of 93.8%–97.89% were achieved and the acoustic feature and classifier methods developed here have strong potential for use in individual animal recognition and can be applied to other species. This is the first study into automatic individual recognition for insects using only the acoustic modality, and is also the first application of these techniques to the individual recognition of insects. The results challenge the prevailing conception that it is impossible to achieve individual identification of insects using only the acoustic modality and, prove that the individual recognition of insects is possible.

TABLE IV. Confusion matrix of the recognition results in *Gampsodeis gratiosa*.

Known class	Predicted class by α -GMM														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	29	0	0	0	0	0	0	0	0	0	0
6	0	0	5	0	0	11	0	0	0	2	0	2	0	0	0
7	0	0	0	0	0	0	29	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	39	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	27	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25

Although the use of α -GMM did not yield a significant improvement over the regular GMM approach, it did improve system performance. As the main advantage of α -GMM compared to the regular GMM is its robustness to noise, we may not have detected a remarkable improvement because our recordings contained little noise. However, recordings of insect sounds are more likely to include some noise and it is necessary to use the α -GMM approach. We did not investigate the degree to which the α -GMM approach outperformed the regular GMM approach when recordings of insect sounds contain noise because this is not the focus of our paper. The purpose of our study was an assessment of the feasibility of applying pattern recognition techniques to construct an automatic acoustic system that can reach the goal of automatic individual recognition for insects.

Table I may help explain that individual recognition of insects through the sole use of acoustic information is achievable, but it cannot exclude the contribution of environmental conditions. It is possible that it is the different environmental conditions that tell these individuals apart and so we conducted experiments on *Gampsodeis gratiosa*, whose recordings from different individuals were recorded under the same environmental conditions. A good identification performance was achieved under all three environmental conditions (Table II), and proved that the individual recognition of insects is possible and viable.

The main characteristics differentiating the two species are their dominant frequencies. As their frequencies span a wide range, many other insects whose dominant frequencies are within the range are expected to have good identification performances.

A remaining question is how to optimize the parameter of α automatically, a problem that also remains unresolved in the field of speaker identification. Our model still has room to improve and improvements may be achieved by applying other methods used in speaker identification to the individual recognition of insects. Further, modifications to these methods before applying them to the individual recognition of insects may also contribute to future model improvements.

ACKNOWLEDGMENTS

We thank the anonymous referees for their valuable comments that improved the representation and quality of this paper. We acknowledge Shan Wu for providing recordings of *Deracantha onos*. We thank Jiangning Wang and other members of the Biodiversity Informatics Group, Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, for support and help.

- Akhtar, S., Elshafei-Ahmed, M., and Ahmed, M. S. (2001). "Detection of helicopters using neural nets," *IEEE Trans. Instrum. Meas.* **50**, 749–756.
- Bee, M. A., Kozich, C. E., Blackwell, K. J., and Gerhardt, H. C. (2001). "Individual variation in advertisement calls of territorial male green frogs, *Rana clamitans*: Implications for individual discrimination," *Ethology* **107**, 65–84.
- Blumstein, D. T., and Munos, O. (2005). "Individual, age and sex-specific information is contained in yellow-bellied marmot alarm calls," *Anim. Behav.* **69**, 353–361.
- Campbell, J. P. (1997). "Speaker recognition: A tutorial," *Proc. IEEE* **85**, 1437–1462.

- Cheng, J. K., Sun, Y. H., and Ji, L. Q. (2010). "A call-independent and automatic acoustic system for the individual recognition of animals: A novel model using four passerines," *Pattern Recogn.* **43**, 3846–3852.
- Chesmore, D. (2004). "Automated bioacoustic identification of species," *Ann. Acad. Bras. Cienc.* **76**, 435–440.
- Chesmore, E. D. (2001). "Application of time domain signal coding and artificial neural networks to passive acoustical identification of animals," *Appl. Acoust.* **62**, 1359–1374.
- Chesmore, E. D., and Ohya, E. (2004). "Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition," *Bull. Entomol. Res.* **94**, 319–330.
- Clemins, P. J., Johnson, M. T., Leong, K. M., and Savage, A. (2005). "Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations," *J. Acoust. Soc. Am.* **117**, 956–963.
- Crawford, J. D., Cook, A. P., and Heberlein, A. S. (1997). "Bioacoustic behavior of African fishes (Mormyridae): Potential cues for species and individual recognition in *Pollimyrus*," *J. Acoust. Soc. Am.* **102**, 1200–1212.
- Darden, S. K., Dabelsteen, T., and Pedersen, S. B. (2003). "A potential tool for swift fox (*Vulpes velox*) conservation: Individuality of long-range barking sequences," *J. Mammal.* **84**, 1417–1427.
- Davis, S. B., and Mermelstein, P. (1980). "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans. Acoust. Speech* **28**, 357–366.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via em algorithm," *J. R. Stat. Soc. Ser. B (Methodol.)* **39**, 1–38.
- Fox, E. J. S. (2008). "A new perspective on acoustic individual recognition in animals with limited call sharing or changing repertoires," *Anim. Behav.* **75**, 1187–1194.
- Frommolt, K. H., Goltzman, M. E., and MacDonald, D. W. (2003). "Barking foxes, *Alopex lagopus*: Field experiments in individual recognition in a territorial mammal," *Anim. Behav.* **65**, 509–518.
- Frommolt, K. H., Kruchenkova, E. P., and Russig, H. (1997). "Individuality of territorial barking in arctic foxes, *Alopex lagopus* (L., 1758)," *Z. Säugetierkd.* **62**, 66–70.
- Ganchev, T., and Potamitis, I. (2007). "Automatic acoustic identification of singing insects," *Bioacoustics* **16**, 281–328.
- Ganchev, T., Potamitis, I., and Fakotakis, N. (2007). "Acoustic monitoring of singing insects," *2007 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. IV, pp. 721–724.
- Juang, C. F., and Chen, T. M. (2007). "Birdsong recognition using prediction-based recurrent neural fuzzy networks," *Neurocomputing* **71**, 121–130.
- Lee, C. H., Chou, C. H., Han, C. C., and Huang, R. Z. (2006). "Automatic recognition of animal vocalizations using averaged MFCC and linear discriminant analysis," *Pattern Recogn. Lett.* **27**, 93–101.
- Lewis, M. A., and Ramachandran, R. P. (2001). "Cochannel speaker count labelling based on the use of cepstral and pitch prediction derived features," *Pattern Recogn.* **34**, 499–507.
- Mashao, D. J., and Skosan, M. (2006). "Combining classifier decisions for robust speaker identification," *Pattern Recogn.* **39**, 147–155.
- Openshaw, J. P., Sun, Z. P., and Mason, J. S. (1993). "A comparison of composite features under degraded speech in speaker recognition," *ICASSP-93: 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vols. 1–5, pp. B371–B374.
- Picone, J. W. (1993). "Signal modeling techniques in speech recognition," *Proc. IEEE* **81**, 1215–1247.
- Pinhas, J., Soroker, V., Hetzroni, A., Mizrach, A., Teicher, M., and Goldberger, J. (2008). "Automatic acoustic detection of the red palm weevil," *Comput. Electron. Agric.* **63**, 131–139.
- Policht, R., Petru, M., Lastimozo, L., and Suarez, L. (2009). "Potential for the use of vocal individuality as a conservation research tool in two threatened Philippine hornbill species, the Visayan Hornbill and the Rufous-headed Hornbill," *Bird Conserv. Int.* **19**, 83–97.
- Potamitis, I., Ganchev, T., and Fakotakis, N. (2006). "Automatic acoustic identification of insects inspired by the speaker recognition paradigm," *Interspeech 2006 and Ninth International Conference on Spoken Language Processing*, Vols. 1–5, pp. 2126–2129, 2633.
- Potamitis, I., Ganchev, T., and Fakotakis, N. (2007). "Automatic acoustic identification of crickets and cicadas," *2007 Ninth International Symposium on Signal Processing and Its Applications*, Vols. 1–3, pp. 752–755, 1471.

- Potamitis, I., Ganchev, T., and Kontodimas, D. (2009). "On automatic bioacoustic detection of pests: The cases of *Rhynchophorus ferrugineus* and *Sitophilus oryzae*," J. Econ. Entomol. **102**, 1681–1690.
- Redner, R. A., and Walker, H. F. (1984). "Mixture densities, maximum-likelihood and the em algorithm," SIAM Rev. **26**, 195–237.
- Reynolds, D. A., and Rose, R. C. (1995). "Robust text-independent speaker identification using Gaussian mixture speaker models," IEEE Trans. Speech Audio Process. **3**, 72–83.
- Schwenker, F., Dietrich, C., Kestler, H. A., 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.
- Tang, F. M., Chen, M. Y., and Wang, Z. D. (2005). "Stored-product insect recognition based on support vector machines," J. Huazhong Univ. Sci. Technol. **33**, 34–36 (in Chinese).
- Trawicki, M. B., Johnson, M. T., and Osiejuk, T. S. (2005). "Automatic song-type classification and speaker identification of Norwegian Ortolan Bunting (*Emberiza hortulana*) vocalizations," 2005 IEEE Workshop on Machine Learning for Signal Processing (MLSP), pp. 277–282.
- Trifa, V. M., Kirschel, A. N. G., Taylor, C. E., and Vallejo, E. E. (2008). "Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models," J. Acoust. Soc. Am. **123**, 2424–2431.
- Vergin, R., OShaughnessy, D., and Gupta, V. (1996). "Compensated Mel frequency cepstrum coefficients," Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vols. 1–6, pp. 323–326.
- Wang, S. G., Mao, Z. M., Lu, R. S., Zheng, Z. M., and Shi, F. M. (2008). "A comparative study on the songs structures of *Gampsodeis ussubiensis* and *Gampsocleis gratiosa* Brunner (Orthoptera, Tettigoniidae)," Acta Zootaxon. Sin. **33**, 553–557 (in Chinese).
- Wu, D. L. (2009). "Parameter estimation for α -GMM based on maximum likelihood criterion," Neural Comput. **21**, 1776–1795.
- Wu, D. L., Li, J., and Wu, H. Q. (2009). " α -Gaussian mixture modelling for speaker recognition," Pattern. Recogn. Lett. **30**, 589–594.
- Zhantiev, R. D., Korsunovskaya, O. S., and Bysov, S. D. (1995). "Acoustic communication in desert bush-cricket (Orthoptera, Bradyporidae)," Zool. Zh. **74**, 58–72.
- Zhen, B., Wu, X. H., Liu, Z. M., and Chi, H. S. (2001). "On the importance of components of the MFCC in speech and speaker recognition," Acta Sci. Nat. Univ. Peking **37**, 371–378 (in Chinese).
- Zhu, L. Q., Wang, H. B., and Zhang, Z. (2010). "Automatic acoustical identification of insects based on MFCC and VQ," Acta Entomol. Sin. **53**, 901–907 (in Chinese).