Computers and Electronics in Agriculture 91 (2013) 154–159

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Automatic detection and identification of brown stink bug, *Euschistus servus*, and southern green stink bug, *Nezara viridula*, (Heteroptera: Pentatomidae) using intraspecific substrate-borne vibrational signals

B.D. Lampson [a,*], Y.J. Han [a], A. Khalilian [b], J. Greene [b], R.W. Mankin [c], E.G. Foreman [c]

[a] School of Agricultural, Forest, and Environmental Sciences, Clemson University, 231 McAdams Hall, Clemson, SC 29634, USA
[b] School of Agricultural, Forest, and Environmental Sciences, Edisto Research and Education Center, 64 Research Drive, Blackville, SC 29817, USA
[c] United States Department of Agriculture, Agricultural Research Service, Center for Medical, Agricultural and Veterinary Entomology, 1600 Southwest 23rd Drive, Gainesville, FL 32608, USA

ARTICLE INFO

Article history:
Received 9 July 2012
Received in revised form 5 December 2012
Accepted 10 December 2012

Keywords:
Bioacoustic recognition
Neural network
Mixture model

ABSTRACT

Stink bugs cost the southeastern US cotton industry millions of dollars each year in crop losses and control costs. These losses are reduced by strategic pesticide applications; however, current methods of monitoring these pests for making management decisions are time-consuming and costly. Therefore, improved methods to identify and monitor these bugs must be investigated in order to optimize pesticide applications. One such method would be to exploit the substrate-borne vibrational signals (SBVSs) of these insects. Recordings of SBVS for two prevalent regional pests, the brown stink bug, *Euschistus servus*, and southern green stink bug, *Nezara viridula*, were segmented into separate pulses of variable duration based on signal energy. For each pulse, the linear frequency cepstral coefficients, dominant frequency, and duration were calculated and used as features. These features were classified using a Gaussian mixture model (GMM) and a probabilistic neural network (PNN) to discriminate these SBVS from incidental sounds and SBVS of different species from each other. Detection of SBVS generated by brown stink bugs was performed with over 92% accuracy for single male–female pairs with both PNN and GMM and with over 86% accuracy for 30 individuals with both PNN and GMM. Detection of SBVS generated by southern green stink bugs was performed with up to 82.5% accuracy with PNN and 68.0% accuracy with GMM for 30 individuals. Also, both PNN and GMM were over 90% accurate in identifying SBVS of brown and southern green stink bugs. Concurrent detection of SBVS from noise and identification of SBVS of brown and southern green stink bugs was 83.3% accurate using PNN and 71.5% accurate using GMM. These results indicated the capability of detecting and identifying stink bug species using their SBVS.

Published by Elsevier B.V.

1. Introduction

Cotton, *Gossypium hirsutum* (L.), containing transgenes from the bacterium *Bacillus thuringiensis* (*Bt*) has significantly reduced the amount of foliar-applied insecticides used to control major pests of the crop. However, because species of stink bugs are no longer being controlled coincidentally by applications of these insecticides, these insects have emerged as major pests and caused economic damage to cotton (Panizzi et al., 2000; Greene et al., 2001). In 2011, losses due to stink bugs were estimated to be over $48 million in the United States, with control costs exceeding $5.3 million (Williams, 2012).

In order to preserve the benefits of *Bt* cotton that minimize the amount of insecticides needed for acceptable control, detection methods must be developed to measure stink bug range, density, and/or damage so pesticides are only applied when a damage threshold is exceeded. Ideally, detection methods also will provide pest species identification so new biological control treatments (Saxena and Kumar, 1980; Čokl and Millar, 2009; Mankin, 2012), such as interference with intraspecific communicatory substrate-borne vibrational signals (SBVSs), will be targeted correctly.

Individuals of many pest species of stink bugs in the southeastern United States communicate intraspecifically using SBVS, including the southern green stink bug, *Nezara viridula* (L.) (Harris et al., 1982; Čokl et al., 2000), the green stink bug, *Chinavia hilaris* (Say) (Čokl et al., 2001), and the brown stink bug, *Euschistus servus* (Say) (Lampson et al., 2010). Stink bugs use SBVS to locate and court mates, with each signal being specific to the species producing it and to the stage in the mating process (Čokl and Virant-Doberlet, 2003).

* Corresponding author. Tel.: +1 (864) 653 4082; fax: +1 (864) 653 0338.
E-mail address: bdettma@clemson.edu (B.D. Lampson).

0168-1699/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.compag.2012.12.010

Previous research on classifying insect sounds can be divided into three categories – detection, monitoring, and identification (Mankin et al., 2011). Detection and monitoring involve classification of sounds of predefined types from background noise over short or long periods. Identification involves classifying predefined sounds into targeted categories, e.g. the species, family, order, or males and females of a particular species that produced them.

Accurate automatic identification of insect communication sounds has been shown using linear frequency cepstral coefficients (LFCCs) as features classified with Gaussian mixture models (GMMs) or probabilistic neural networks (PNNs). Previous research on cicadas, which use a tymbal mechanism to produce sound, classify sounds from family Cicadidae into genera with accuracies of 94.4% using GMM and up to 97.9% using PNN using dominant frequency, segment duration, and 23 LFCCs as features (Ganchev and Potamitis, 2007; Ganchev et al., 2007); however, the previous research only used high quality recordings with high signal-to-noise ratios. In order to develop a model for detection and/or monitoring, the model must be presented with segments of high energy noise and be able to differentiate those noise segments from insect sounds.

Accurate automatic detection of insect sounds has also been performed by distinguishing incidental vibrations (e.g. moving, eating, or tunneling) of different species from noise using cepstral features with GMM classification. Sounds of the red palm weevil were automatically detected from noise with up to 98.8% accuracy using GMM as the classification method and mel-frequency cepstral coefficients as features (Pinhas et al., 2008). Also, sounds of the red palm weevil were automatically detected from noise with over 96.9% accuracy and sounds of the rice weevil were detected with 100% accuracy with GMM classification using dominant frequency and 23 LFCCs as features (Potamitis et al., 2009). However, no research focuses on using these techniques for automatic detection of insect vibrational communication or for concurrent detection and species identification.

The objective of this research was to detect and identify the SBVS of the brown stink bug, *E. servus*, and the southern green stink bug, *N. viridula*, on cotton by utilizing state-of-the-art identification and detection techniques. These techniques included variable-length segmentation, which has been shown to outperform fixed-length segmentation, and the addition of dominant frequency and segment length as features, which have been shown to increase accuracy of species identification (Potamitis et al., 2006, 2009;Ganchev and Potamitis, 2007).

Piezoelectric transducers have replaced microphones for recording SBVS as microphones do not record vibrations through soil and/or woody substrates well. Popular piezoelectric transducers include accelerometers, ultrasonic transducers, and film transducers (Mankin et al., 2011). For this study, a piezoelectric accelerometer was used for recording SBVS.

2. Methods and materials

2.1. Rearing of insects and plants

Over 100 adults of brown and southern green stink bugs were used for this experiment, including some reared from eggs from mated females in the laboratory and some collected from fields of cotton and soybeans at the Edisto Research and Education Center near Blackville, South Carolina, in late spring and early summer of 2008. Stink bugs were reared and/or held in multiple plastic cages and fed a diet of fresh green beans and raw peanuts (Harris and Todd, 1981). The cages were held in a controlled environment room at 30 °C and 70% relative humidity with a photoperiod of 14:10 (L:D) h. Three plants of cotton, *Gossypium hirsutum* (L.), variety DP 164 B2RF (Delta and Pine Land 164 Bollgard 2 Round-up-Ready Flex) were held in plastic black pots 27.9 cm in diameter and 24.1 cm tall. The plants varied in height from 0.6 m to 0.9 m and were blooming and setting bolls during experimentation.

2.2. Recording SBVS

All recordings were made from 23 to 26 June 2008 with sexually mature adult bugs between 0900 and 2100 h in an anechoic chamber located at the USDA Center for Medical, Agricultural, and Veterinary Entomology in Gainesville, Florida. Recordings were made using an accelerometer (Model 4370, Brüel & Kjær, Naerum, Denmark) attached to the cotton plant by alligator clips at 55 cm above the soil. This setup was enclosed in a $0.6 \times 0.6 \times 1.2$ m wire mesh cage to contain the insects.

Ten recordings of 90 s each were made of individual pairs of a male and a female of the brown stink bug (BSB). Then, to consider environments with increased incidental noise, ten recordings of 90 s each were made of 30 individuals of the BSB. Ten recordings of 90 s each were also made of 30 individuals of the southern green stink bug (SGSB). Recordings were amplified 20 dB using a charge amplifier (Model 2635, Brüel & Kjær, Naerum, Denmark) and 30 dB using a secondary amplifier (Model 2610, Brüel & Kjær, Naerum, Denmark) and then bandpass-filtered between 70 and 2000 Hz before being digitized at 44.1 kHz using a speech-analysis system (Model 4300B, Kay Elemetrics Corp., Lincoln Park, NJ).

2.3. Vibrational data analysis and recognition

An algorithm was written using Visual Studio (Microsoft, 2007) to calculate the features of the SBVS recorded. First, each 90-s recording was reduced 10× to a sample frequency of 4.41 kHz and divided into 6 segments of 15 s each to reduce computational demand. These segments were further divided into pulses using the procedure described by Ganchev et al. (2007). The frame energy of each segment was calculated using:

$$E_{\text{use}}(k) = \sum_{i=1}^{K_{\text{frame}}} (x(kL+i))^2 \quad \text{for } k = 0, 1, \ldots, M-1$$

where $x(i)$ is the input signal, k is the frame index, K_{frame} is the frame size, L is the predefined step size that determines the overlap between successive frames, and:

$$M = \text{floor}\left(\frac{\text{N}_{\text{rec}} - K_{\text{frame}} + L}{L}\right)$$

where the floor operator stands for rounding toward the smaller integer value, and N_{rec} is the total number of samples in the recording. A frame size of $K_{\text{frame}} = 44$ and a step size of $L = 5$ was used, which corresponds to a frame duration of 10.0 ms and a step duration of 1.13 ms. The start of a pulse was defined when E_{use} exceeded the upper threshold, T_{high}, and the pulse continued until E_{use} was reduced below the lower threshold, T_{low}. These thresholds were defined as:

$$T_{\text{high}}(k) = 0.96 \cdot T_{\text{high}}(k-1) + 0.04 \cdot E_{\text{use}}(k-1)$$

$$T_{\text{low}}(k) = 0.75 \cdot T_{\text{high}}(k)$$

for $k = 1, 2, \ldots, M$ where $T_{\text{high}}(0) = E_{\text{use}}(0)$. Two pulses separated by 12 ms or less were combined into one pulse.

Then, LFCCs were calculated for each pulse. First, the input signal for each pulse were multiplied by a Hamming window and subjected to a discrete Fourier transform (DFT), and the output energy from the DFT was multiplied by the amplitude gain of each of 32 overlapping, triangular filters with a length of 11.5 Hz, covering

60–250 Hz. The log-energy filterbank outputs were then subjected to a discrete cosine transform (DCT):

$$\mathrm{LFCC}_j = \sum_{i=1}^{B}\left(X_i \cos\left(j\left(i-\frac{1}{2}\right)\frac{\pi}{B}\right)\right) \quad \text{for } i = 0, 1, \ldots, J$$

where X_i are the log-energy filterbank outputs, j is the order of the LFCC, J is the number of LFCCs calculated, i is redefined in this equation as a filter index, and $B = 32$ is the number of filters. The 0th order LFCC was discarded, and the 1st through 24th order LFCCs were selected to be used in the identification process. Various numbers of LFCCs were tested to determine contributions of higher order LFCCs. The dominant frequency and duration were also used as features. Cepstral mean subtraction and dynamic range normalization were applied to all features.

Classification models, GMM and PNN, were made using these features. The 10-fold cross-validation method was used to verify each model, which includes dividing the dataset randomly into 10 subsets and repeating the holdout method ten times using each subset as the testing set exactly once. The classification criterion for classifying feature vector X to population i using PNN was:

$$g_i(X) > g_j(X) \quad \text{for all } j \neq i$$

$$\text{where } g_i(X) = \frac{1}{(2\pi)^{p/2}\sigma^p n_i}\sum_{k=1}^{n_i} e^{\frac{-\|X-X_{ik}\|^2}{2\sigma^2}} \quad \text{for } i = 1, 2, \ldots, N$$

where p is length of each feature vector, σ is the spread, n_i is the number of feature vectors for population i, X_{ik} is the kth feature vector for population i, N is the number of populations, k is redefined in this equation as an index over feature vectors, and i and j are redefined as population indexes. A population was defined as a set of pulses classified as either SBVS from a particular stink bug species or incidental noise. The classification criterion for classifying feature vector X_k to mixture component i for the GMM was:

$$h_i(X_k) > h_j(X_k) \quad \text{for all } j \neq i$$

where $h_i(X_k)$ is the probability the feature vector X_k belongs to mixture component i or posterior probability, i and j are redefined as mixture component indexes, k is a feature vector index, and:

$$h_i(X_k) = \frac{\alpha_i \varphi(X_k|\mu_i, \Sigma_i)}{\sum_{j=1}^{J}\alpha_j \varphi(X_k|\mu_j, \Sigma_j)}$$

where α_i is the weight for the ith mixture component, J is the number of mixture components, and the mixture density function:

$$\varphi(X_k|\mu_i, \Sigma_i) = \frac{1}{\left((2\pi)^p|\Sigma_i|\right)^{1/2}} e^{-\frac{(X_k-\mu_i)'\Sigma_i^{-1}(X_k-\mu_i)}{2}} \quad (1)$$

where p is length of each feature vector, Σ_i is the covariance of the ith mixture component, and μ_i is the mean of the ith mixture component. Parameters α_i, μ_i, Σ_i were calculated iteratively using the Expectation Maximization (EM) algorithm until convergence of the likelihood function:

$$L(X|\alpha_i, \mu_i, \Sigma_i) = \prod_{k=1}^{n}\left(\sum_{j=1}^{J}\alpha_j \varphi\left(X_k|\mu_j, \Sigma_j\right)\right)$$

where n is the number of feature vectors, k is a feature vector index, j and i are mixture component indexes, J is the number of mixture components, α_j is the weight for the jth cluster, and the mixture density function is described in Eq. (1). Two mixture components ($K = 2$) were used for either identification or detection, and three mixture components ($K = 3$) were used for concurrent identification and detection. The training set was used to calculate the initial parameters α_i, Σ_i, and μ_i, but the testing set was used in the EM algorithm.

Both the GMM and PNN were performed using MATLAB (MathWorks, 2010). The PNN was performed using the 'newpnn' and 'sim' functions, using a spread of 0.50. The GMM was performed using the 'gmdistribution.fit' function, with the start parameter given as a structure array using the mean, covariance, and mixing proportion matrices calculated from the training set.

First, classification models were made to detect SBVS of single pairs of BSB, 30 individuals of BSB, and also 30 individuals of SGSB from incidental noise. For detection of BSB and SGSB, all pulses were manually labeled as 'positive' or 'negative'. A 'positive' label indicated a clearly identified stink bug vibrational signal in the pulse. A 'negative' label indicated the pulse did not contain any SBVS. False positives were defined in the classification model as incorrectly classifying a 'negative' pulse as 'positive', and false negatives were defined in the classification model as incorrectly classifying a 'positive' pulse as 'negative'.

Then, classification models were made to identify SBVS of BSB and SGSB, and pulses were manually labeled as 'brown' or 'southern green', respectively. Pulses which did not contain any SBVS were discarded. False match 'A' was defined as the classification model incorrectly identifying a 'brown' pulse as 'southern green', and false match 'B' was defined as the classification model incorrectly identifying a 'southern green' pulse as 'brown'.

Then, classification models were designed to concurrently detect and identify BSB and SGSB from incidental noise. All pulses were manually labeled as 'brown', 'southern green', or 'noise'. A 'brown' or 'southern green' label indicated a clearly identified vibrational signal in the pulse from BSB or SGSB, respectively. A 'noise' label indicated the pulse did not contain any 'brown' or 'southern green' SBVS. False match 'A_B' was defined as incorrectly identifying a 'brown' pulse as 'southern green', and false match 'A_N' was defined as incorrectly identifying a 'brown' pulse as 'noise'. Likewise, false match 'B_A' was defined as incorrectly identifying a 'southern green' pulse as 'brown', and false match 'B_N' was defined as incorrectly identifying a 'southern green' pulse as a 'noise' pulse. In addition, false match 'N_A' was defined as incorrectly identifying 'noise' as a 'brown' pulse, and false match 'N_B' was defined as incorrectly identifying 'noise' as a 'southern green' pulse.

3. Results

3.1. Detection of SBVS of BSB and SGSB from incidental noise

From the recordings of single pairs of BSB, 2009 stink bug pulses and 1352 noise pulses were identified. Pulses were correctly identified with up to 94.1% accuracy with PNN, using dominant frequency, duration, and 1st through 6th order LFCCs as features (Table 1). Pulses were correctly identified with up to 92.5% accuracy with GMM, using dominant frequency, duration, and 1st through 8th order LFCCs as features (Table 2).

Table 1
Accuracy of probabilistic neural network to detect substrate-borne vibrational signals from male–female pairs of *Euschistus servus* using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs).

No. of LFCCs	Correct (%)	False + (%)	False − (%)
4	93.4	3.54	3.09
6	94.1	2.86	3.03
8	93.5	3.03	3.48
12	92.7	3.33	3.93
16	92.3	3.93	3.78
20	92.5	3.66	3.84
24	91.8	4.02	4.17

Table 2
Accuracy of Gaussian mixture model to detect substrate-borne vibrational signals from male–female pairs of *Euschistus servus*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs).

No. of LFCCs	Correct (%)	False + (%)	False − (%)
4	92.5	3.99	3.51
6	92.1	3.27	4.58
8	92.5	3.21	4.28
12	92.3	3.30	4.40
16	91.6	3.63	4.76
20	90.5	4.55	4.91
24	90.7	4.97	4.34

Table 3
Accuracy of probabilistic neural network to detect substrate-borne vibrational signals from 30 *Euschistus servus*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs).

No. of LFCCs	Correct (%)	False + (%)	False − (%)
4	89.1	3.69	7.21
6	90.7	3.49	5.76
8	89.2	4.49	6.35
12	88.9	4.69	6.39
16	88.4	4.87	6.77
20	88.6	4.90	6.52
24	88.2	4.97	6.87

Table 4
Accuracy of Gaussian mixture model to detect substrate-borne vibrational signals from 30 *Euschistus servus*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs).

No. of LFCCs	Correct (%)	False + (%)	False − (%)
4	78.3	15.1	6.63
6	85.5	10.5	4.00
8	86.0	9.15	4.90
12	85.5	8.53	5.94
16	74.2	21.0	4.76
20	70.8	24.9	4.25
24	71.1	24.5	4.42

Table 5
Accuracy of probabilistic neural network to detect substrate-borne vibrational signals from 30 *Nezara viridula*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs).

No. of LFCCs	Correct (%)	False + (%)	False − (%)
4	82.5	8.61	8.86
6	81.4	7.31	11.3
8	79.2	8.24	12.5
12	77.1	9.09	13.8
16	75.8	9.48	14.7
20	76.1	9.37	14.5
24	76.3	9.45	14.2

From the recordings of 30 BSB, 1373 'positive' pulses and 1524 'negative' pulses were identified. The average dominant frequency for 'negative' pulses was 104 Hz and the average dominant frequency for 'positive' pulses was 105 Hz. Pulses were correctly identified using PNN with up 90.7% accuracy, using dominant frequency, duration, and 1st through 6th LFCCs as features (Table 3). Pulses were correctly identified with up 86.0% accuracy with GMM, using dominant frequency, duration, and 1st through 8th order LFCCs as features (Table 4). The average dominant frequency for false positives was 115 Hz, and the average dominant frequency for false negatives was 112 Hz for the PNN model, using dominant frequency, duration, and 1st through 6th LFCCs as features. The average dominant frequency for correct 'positive' pulses was 104 Hz, and the average dominant frequency for correct 'negative' pulses was 93.7 Hz for the PNN model, using dominant frequency, duration, and 1st through 6th LFCCs as features.

From the recordings of 30 SGSB, 2166 stink bug pulses and 1389 noise pulses were identified. The average dominant frequency for 'negative' pulses was 89.5 Hz, and the average dominant frequency for 'positive' pulses was 114 Hz. Pulses were correctly identified with up to 82.5% accuracy with PNN, using 1st through 4th order LFCCs as features (Table 5) and up to 68.0% accuracy with GMM, using dominant frequency, duration, and 1st through 12th LFCCs as features (Table 6). The average dominant frequency for false positives was 111 Hz, and the average dominant frequency for false negatives was 93.4 Hz for the PNN model, using dominant frequency, duration, and 1st through 4th LFCCs as features. The average dominant frequency for correct 'positive' pulses was 118 Hz, and the average dominant frequency for correct 'negative' pulses was 83.5 Hz for the PNN model, using dominant frequency, duration, and 1st through 4th LFCCs as features.

Table 6
Accuracy of Gaussian mixture model to detect substrate-borne vibrational signals from 30 *Nezara viridula*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs).

No. of LFCCs	Correct (%)	False + (%)	False − (%)
4	61.9	7.45	30.7
6	61.9	7.88	30.2
8	62.1	8.19	29.8
12	68.0	10.4	21.6
16	66.0	21.5	12.5
20	62.4	22.1	15.5
24	48.9	13.8	37.3

Table 7
Accuracy of probabilistic neural network to identify substrate-borne vibrational signals of *Nezara viridula* and *Euschistus servus*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs). False match 'A' was defined as incorrectly identifying an *E. servus* signal as a *N. viridula* signal, and false match 'B' was defined as incorrectly identifying an *N. viridula* signal as a *E. servus* signal.

No. of LFCCs	Correct (%)	False A (%)	False B (%)
4	91.6	5.43	3.00
6	93.3	4.80	1.89
8	93.8	4.44	1.78
12	93.4	4.41	2.15
16	93.1	4.58	2.35
20	92.8	4.55	2.63
24	92.9	4.78	2.29

3.2. Identification of SBVS of BSB and SGSB

Overall, 3482 pulses were classified- 2166 belonging to SGSB and 1373 belonging to BSB. Pulses were correctly identified with up to 93.8% accuracy with PNN (Table 7) and up to 90.6% accuracy with GMM (Table 8), using dominant frequency, duration, and 1st through 8th order LFCCs as features.

3.3. Concurrent detection and identification of SBVS of BSB and SGSB from incidental noise

Overall, 6452 pulses were classified- 2913 noise pulses, 2166 pulses belonging to SGSB and 1373 pulses belonging to BSB. Pulses were correctly identified with up to 83.3% accuracy with PNN, using dominant frequency, duration, and 1st through 6th order LFCCs as features (Table 9), and up to 71.5% accuracy with GMM, using dominant frequency, duration, and 1st through 8th LFCCs as features (Table 10).

Table 8
Accuracy of Gaussian mixture model to identify substrate-borne vibrational signals of *Nezara viridula* and *Euschistus servus*, using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs). False match 'A' was defined as incorrectly identifying an *E. servus* signal as a *N. viridula* signal, and false match 'B' was defined as incorrectly identifying an *N. viridula* signal as a *E. servus* signal.

No. of LFCCs	Correct (%)	False A (%)	False B (%)
4	86.3	7.12	6.58
6	89.8	6.70	3.48
8	90.6	7.04	2.35
12	90.5	8.19	1.30
16	85.6	6.08	8.28
20	63.1	3.67	33.2
24	63.3	3.81	32.9

Table 9
Accuracy of probabilistic neural network to concurrently detect and identify substrate-borne vibrational signals of *Nezara viridula* and *Euschistus servus*, from incidental noise using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs). False match 'A_B' was defined as incorrectly identifying an *E. servus* signal as a *N. viridula* signal, false match 'A_N' was defined as incorrectly identifying an *E. servus* signal as noise, false match 'B_A' was defined as incorrectly identifying a *N. viridula* signal as a *E. servus* signal, false match 'B_N' was defined as incorrectly identifying a *N. viridula* signal as noise, false match 'N_A' was defined as incorrectly identifying noise as an *E. servus* signal, and false match 'N_B' was defined as incorrectly identifying noise as a *N. viridula* signal.

No. of LFCCs	Correct (%)	A_B (%)	A_N (%)	B_A (%)	B_N (%)	N_A (%)	N_B (%)
4	82.1	1.38	3.70	1.01	6.63	1.32	3.91
6	83.3	1.22	2.70	0.589	7.24	1.29	3.69
8	82.2	1.02	2.90	0.682	7.41	1.52	4.29
12	80.0	0.961	2.91	0.899	8.11	1.86	5.21
16	80.0	0.914	2.87	0.682	8.77	1.88	4.88
20	79.5	0.992	2.96	0.760	8.62	1.89	5.24
24	79.4	0.961	2.93	0.775	8.52	1.89	5.52

Table 10
Accuracy of Gaussian mixture model to concurrently detect and identify substrate-borne vibrational signals of *Nezara viridula* and *Euschistus servus*, from incidental noise using dominant frequency, duration, and 4–24 linear frequency cepstral coefficients (LFCCs). False match 'A_B' was defined as incorrectly identifying an *E. servus* signal as a *N. viridula* signal, false match 'A_N' was defined as incorrectly identifying an *E. servus* signal as noise, false match 'B_A' was defined as incorrectly identifying a *N. viridula* signal as a *E. servus* signal, false match 'B_N' was defined as incorrectly identifying a *N. viridula* signal as noise, false match 'N_A' was defined as incorrectly identifying noise as an *E. servus* signal, and false match 'N_B' was defined as incorrectly identifying noise as a *N. viridula* signal.

No. of LFCCs	Correct (%)	A_B (%)	A_N (%)	B_A (%)	B_N (%)	N_A (%)	N_B (%)
4	61.2	2.62	2.48	2.87	13.2	8.21	9.41
6	67.4	2.40	2.03	1.39	13.1	6.71	7.01
8	71.5	2.96	1.41	1.41	9.55	5.69	7.49
12	69.9	3.55	1.47	0.729	8.73	5.16	10.5
16	59.5	3.63	1.02	1.08	7.44	8.91	18.4
20	41.9	1.39	0.667	18.5	3.75	13.0	20.8
24	43.4	1.43	0.636	18.1	4.48	19.9	12.0

4. Discussion

Both classification models, GMM and PNN, successfully classified BSB signals from noise. For both the single pairs of BSB and the 30 individuals of BSB, both the GMM and PNN models were over 86.0% accurate in classifying SBVS of BSB from noise. However, the PNN outperformed the GMM by 2.0% for pulses from single pairs of BSB and by 5.2% for pulses from 30 individuals of BSB, using dominant frequency, duration, and 1st through 6th order LFCCs as features. The average dominant frequency of false positives and negatives was higher than the average dominant frequency of the correctly identified pulses.

It also should be noted that the model accuracy decreased from up to 94.1% for single pairs of BSB to up to 90.7% for 30 individuals of BSB. This implies that increased instances of incidental noise reduce model accuracy. Some reductions in model accuracy could have been due to a high rate of occurrence of noise pulses occurring at the same time as signal pulses, resulting in incorrectly classified pulses.

Previous research detected locomotion and feeding sounds of the red palm weevil from noise with up to 98.8% accuracy using GMM as the classification method and mel-frequency cepstral coefficients as features in a sound-isolated box (Pinhas et al., 2008) and with up to 99.5% accuracy using GMM as the classification method and 23 LFCCs and dominant frequency as features in a field setup (Potamitis et al., 2009). Reduced accuracy in the present study may be due to a lower signal-to-noise ratio. Red palm weevil larvae are much larger and produce louder signals than BSB and SGSB adults; consequently, their vibrations have higher signal-to-noise ratios.

The PNN showed acceptable accuracy in detecting SBVS of SGSB from noise with >82% accuracy. However, GMM only correctly identified SBVS of SGSB from noise pulses with 68.0% accuracy. It should be noted that the average dominant frequency of false positives was closer to the average dominant frequency of correctly identified 'positive' pulses than that of 'negative' pulses. Also, the average dominant frequency of false negatives was closer to the average dominant frequency of correctly identified 'negative' pulses than that of 'positive' pulses. Both of these factors may have contributed to the misclassifications.

Both classification models, GMM and PNN, showed acceptable accuracy in identifying species of stink bugs. Both models were over 90% accurate in classifying SBVS of BSB and SGSB. However, the PNN outperformed the GMM by 3.2% using dominant frequency, duration, and 1st through 8th order LFCCs as features. Previous research classifying sounds of Hemipteran insects of a family into genera have shown accuracies of 94.4% using GMM and up to 97.9% using PNN for family Cicadidae, using dominant frequency, segment duration, and 23 LFCCs as features; however, while the methodology and results are comparable to the present study, it should be noted the research of Ganchev and Potamitis (2007) and Ganchev et al. (2007) classified sounds from family Cicadidae into 4 genera and only used high quality recordings with high signal-to-noise ratios.

The PNN showed acceptable accuracy in concurrent identification and detection of SBVS of BSB and SGSB from noise with up to 83.3% accuracy, using dominant frequency, duration, and 1st through 6th order LFCCs as features. However, the GMM only correctly classified SBVS of BSB and SGSB from noise pulses with up to 71.5% accuracy, using dominant frequency, duration, and 1st through 8th order LFCCs as features. Most of the incorrect classifications were made when distinguishing SBVS of SGSB from noise. These accounted for 10.9% of all classifications and for 65.5% of all classification errors made with the PNN classification model, using dominant frequency, duration, and 1st through 6th order LFCCs as features. With the GMM classification model, these classifications accounted for 17.0% of all classifications and 59.8% of all classification errors made using dominant frequency, duration, and 1st through 8th order LFCCs as features.

The PNN outperformed the GMM in both detection and identification. Because the PNN classifies unknown feature vectors according to the weighted distance from all training feature vectors instead of the distance from the average training feature vectors for each mixture component, it is better at resolving clusters of outliers than the GMM. Because stink bug species have been shown to have a repertoire of songs, this may have given the PNN advantage over GMM. Although previous research has shown over 96.9% accuracy in distinguishing feeding and movement sounds of larvae in palm trees from noise in field conditions using GMM as the classification method with dominant frequency and 23 LFCCs as features (Potamitis et al., 2009), this repertoire of songs may have contributed to the reduced accuracy of the GMM classification method for stink bug species as compared to previous studies. However, because the features of these songs change with geographic region (Čokl et al., 2000), the GMM may still be useful for adapting this technology to various locations because training feature vectors are unnecessary if the mean, variance, and weight can be appropriately estimated.

For both methods, the most accurate models included 4–12 LFCCs. However, most models showed the highest accuracy using 6–8 LFCCs. Negligible contributions were made using higher order LFCCs.

Part of the reason that the inclusion of increasingly higher-order, frequency-sensitive LFCCs in the models failed to increase the accuracy of identification may be due to variations in the rates of transmission of different frequencies over distance as the signals moved along the plant from the insect to the sensor (Cocroft and Rodríguez, 2005; Čokl et al., 2005; Mankin et al., 2011). Because the insects were able to move freely during and between each recording, the resultant LFCCs were averages over the different spectral distributions resulting from the various configurations of the insects during the experiments.

A limitation of this study is that it only includes incidental noises of the insects themselves and does not include other low-frequency noise such as wind, bird sounds, or voices. Future studies may include in-field testing where pest density is correlated with the rate of positive detections of these insects.

5. Conclusion

The accuracy of the LFCC–PNN model indicated an acceptable method to detect pulses of BSB and SGSB from incidental noise and to identify pulses of BSB and SGSB. The accuracy of LFCC–GMM model indicates an acceptable method to detect pulses of BSB from incidental noise to identify pulses of BSB and SGSB. This research showed the capability of an algorithm to identify and detect stink bug species using their SBVS and the feasibility of using insect sound identification and detection techniques to detect and/or identify stink bug species in cotton.

Acknowledgements

Technical Contribution No. 6043 of the Clemson University Experiment Station. This material is based upon work supported by NIFA/USDA, under Project number SC-1700289.

References

Cocroft, R.B., Rodríguez, R.L., 2005. The behavioral ecology of insect vibrational communication. Bioscience 55, 323–334.

Čokl, A., Millar, J.G., 2009. Manipulation of insect signaling for monitoring and control of pest insects. In: Ishaaya, I., Horwitz, A.R. (Eds.), Biorational Control of Arthropod Pests. Springer Science + Business Media BV, New York, pp. 279–316.

Čokl, A., Virant-Doberlet, M., 2003. Communication with substrate-borne signals in small plant-dwelling insects. Annu. Rev. Entomol. 48, 29–50.

Čokl, A., McBrien, H.L., Millar, J.G., 2001. Comparison of substrate-borne vibrational signals of two stink bug species, *Acrosternum hilare* and *Nezara viridula* (Heteroptera: Pentatomidae). Ann. Entomol. Soc. Am. 94, 471–479.

Čokl, A., Virant-Doberlet, M., Stritih, N., 2000. The structure and function of songs emitted by southern green stink bugs from Brazil, Florida, Italy, and Slovenia. Physiol. Entomol. 25, 196–205.

Čokl, A., Zorović, M., Žunič, A., Virant-Doberlet, M., 2005. Tuning of host plants with vibratory songs of *Nezara viridula* L (Heteroptera: Pentatomidae). J. Exp. Biol. 208, 1481–1488.

Ganchev, T., Potamitis, I., 2007. Automatic acoustic identification of singing insects. Bioacoustics 16, 281–328.

Ganchev, T., Potamitis, I., Fakotakis, N., 2007. Acoustic monitoring of singing insects. Int. Conf. Acoust., Speech. Signal Proc. 4, 721–724.

Greene, J.K., Turnipseed, S.G., Sullivan, M.J., May, O.L., 2001. Treatment thresholds for stink bugs (Hemiptera: Pentatomidae) in cotton. J. Econ. Entomol. 94, 403–409.

Harris, V.E., Todd, J.W., 1981. Rearing the southern green stink bug, *Nezara viridula*, with relevant aspects of its biology. J. Georgia Entomol. Soc. 16, 203–210.

Harris, V.E., Todd, J.W., Webb, J.C., Benner, J.C., 1982. Acoustical and behavioral analysis of the songs of the southern green stink bug, *Nezara viridula*. Ann. Entomol. Soc. Am. 75, 234–249.

Lampson, B., Han, Y., Khalilian, A., Greene, J., Mankin, R.W., Foreman, E., 2010. Characterization of substrate-borne vibrational signals of *Euschistus servus* (Heteroptera: Pentatomidae). Am. J. Agric. Biol. Sci. 5, 32–36.

Mankin, R.W., 2012. Applications of acoustics in insect pest management. CAB Reviews 7, 001.

Mankin, R.W., Hagstrum, D.W., Smith, M.T., Roda, A.L., Kario, M.T.K., 2011. Perspective and promise: a century of insect acoustic detection and monitoring. Am. Entomol. 57, 30–44.

MathWorks, 2010. MATLAB. Ver. 7.11.0.584. The MathWorks, Inc., Natick, MA

Microsoft, 2007. Microsoft Visual Studio 2008. Ver. 9.0.21022.8. Microsoft Corporation, Redmond, WA.

Panizzi, A.R., McPherson, J.E., James, D.G., Javahery, M., McPherson, R.M., 2000. Stink bugs (Pentatomidae). In: Schaefer, C.W., Panizzi, A.R. (Eds.), Heteroptera of Economic Importance. CRC Press, Boca Raton, pp. 421–474.

Pinhas, J., Soroker, V., Hetzroni, A., Mizrach, A., Teicher, M., Goldberger, J., 2008. Automatic acoustic detection of the red palm weevil. Comput. Electron. Agr. 63, 131–139.

Potamitis, I., Ganchev, T., Fakotakis, N., 2006. Automatic acoustic identification of insects inspired by the speaker recognition paradigm. In: Interspeech, Pittsburgh, PA, USA, pp. 2126–2129.

Potamitis, I., Ganchev, T., Kontodimas, D., 2009. On automatic bioacoustic detection of pests: the cases of *Rhynchophorus ferrugineus* and *Sitophilus oryzae*. J. Econ. Entomol. 102, 1681–1690.

Saxena, K.N., Kumar, H., 1980. Interruption of acoustic communication and mating in a leafhopper and a planthopper by aerial sound vibrations picked up by plants. Experientia 36, 933–936.

Williams, M.R., 2012. Cotton insect losses-2011. In: Proceedings of the Beltwide Cotton Conference, National Cotton Council of America, Memphis, pp. 1013–1057.