16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

AUTOMATIC BIOACOUSTIC DETECTION OF RHYNCHOPHORUS FERRUGINEUS

[‡]Ilyas Potamitis, [†]Todor Ganchev, [†]Nikos Fakotakis

[‡]Department of Music Technology and Acoustics, Technological Educational Institute of Crete,
Daskalaki-Perivolia, 74100, Rethymno, Greece, potamitis@stef.teicrete.gr

[†]Department of Electrical and Computer Engineering, University of Patras,
26500 Rion-Patras, Greece, {tganchev, fakotaki}@wcl.ee.upatras.gr

ABSTRACT

The goal of this project is to research and develop a bioacoustic automatic detector of a devastating pest attacking palm trees, which has recently appeared in Mediterranean countries. The method is based on piezoelectric sensors that are inserted in the tree trunk and record locomotion and feeding vibrations of the insect. The obtained signals are amplified, filtered, parameterized and automatically classified by advanced machine learning methods on a portable computer. We report excellent detection results reaching 99.5% on real-field recordings.

1. INTRODUCTION

The world of insects is interesting from both a scientific perspective, for investigating the behaviour and diversity of biological organisms, and a practical perspective, due to financial reasons, as insects act as beneficial organisms in agriculture and forestry (they play a significant role in the food chain of other species and the fertility of plants). However, a number of insect species also have a negative bearing on agricultural economy as they are threat to plants and crops.

Insects are mainly categorized according to their appearance [1-2] and sound production [3-4] that are species-specific. The localization and species recognition of insects are usually carried out manually, using trapping and observation methods. The detection and identification process is in most cases a highly complex procedure because insects are heard more often than seen or trapped (especially those that live in complex environments or demonstrate nocturnal activity).

In brief, acoustic identification of insects is based on their ability to generate sound either deliberately as a means of communication or as a by-product of eating, flying or locomotion. Provided that the bioacoustic signal produced by insects follows a consistent acoustical pattern that is species-specific, it can be employed for detection and identification purposes. Despite some rise of interest in the recent years [1-5], the applicability of machine learning techniques to the insect identification problem is still in its infancy.

In this paper we present ongoing research on detection and early warning of pests dangerous to agriculture. Particularly we focus on a devastating pest for palm plantations, namely *Rhynchophorus ferrugineus* (Red Palm Weevil -- RPW). RPW has been detected in several Mediterranean countries (Spain, Egypt, Greece, Cyprus and Israel). Its presence is confirmed by the official authorities of more than 30 Asian and African countries. Red Palm Weevil (RPW), is the most dangerous and deadly pest to date for coconut, oil, date, sago and other palms. Wherever it has been located its presence caused severe catastrophes in the plantation of palms.

RPW has been detected in the islands of Crete and Rhodes in Greece and in Cyprus. Collected specimens of the weevil have been sent to the Museum of Physical History of London, United Kingdom where their identity has been verified officially (Benakio Phytopathological Institute, November 2005 *internal communication*). The cause of the high rate of spread of this pest is human intervention, by transporting infested young or adult date palm trees and offshoots from contaminated to uninfected areas. Date palm is an important crop in North African and Asian countries and ornamental palms are widely planted as amenity trees in the whole Mediterranean area. The pest attacks palm trees and if left undetected within a few weevil generations results in very severe overall damage in the plantation. The main problem is that the attack by the weevil is visible only when the tree has been fatally wounded and adult insects have escaped and infested other trees (RPW has a strong flying capability). Moreover the treatment with insecticides is not efficient since the trunk protects the pest. If the pest is detected on time, the damage can be minimized. The infested trees once located are destroyed, endangered trees are treated and biological traps are deployed. This treatment saves the rest of the plantation from being infested.

The main thrust of this work is the development of a system for automatic acoustic recognition of the RPW acoustic signal, by employing suitable piezoelectric probes with uncoated waveguides that are inserted in the tree trunk and supervised recognition algorithms that classify recorded sound vibrations. In the present contribution, we minimize human intervention by exploiting machine learning techniques and a signal parameterization method, which was explicitly designed for the needs of insect recognition.

In particular, by adopting statistical learning techniques, such as Gaussian mixture models (GMM), we aim at early detection of RPW acoustic activity. To the knowledge of the authors, the proposed combination of spectral analysis and advanced machine learning techniques constitute a totally novel approach to this kind of problem. Moreover this work demonstrates that the paradigm of GMM classifiers and the feature extraction process that is currently very popular in speech/speaker recognition can be used very efficiently for general bioacoustic recognition tasks.

2. RPW AND ITS SOUND EMISSIONS

There are a specific number of behavioural modes that have been observed in connection with sound production of RPW that are closely connected to eating, flying and locomotion. Adult females lay eggs in the crown of palm trees, larvae then penetrate the crown and later to most parts of the upper trunk, making tunnels of up to 1 meter long. Many a time the insect completes several generations inside the crown or trunk feeding on the inner tissues until the trunk or crown becomes hollow and tree gets killed. In the Greek climate three generations are completed per year with an average of 100-2000 adult insects coming out of each tree (depending on the size of the tree). After the death of the palm adult weevils come out and seek fresh trees to attack. Three months after the infestation we have evidence of the palm decadence. Severely attacked palm trees show a total loss of the palms and rotting of the trunk which lead to the sure death of the tree.

RPW has already infested a large number of palms in the area of Khersonisos in the island of Crete. As all palms are suitable for the development of the RPW there is a grave danger for Phoenix Theophrastii that is found in the palm tree forest of Vai in Crete. Vai is the last palm tree forest on Earth that is composed of this specific plant and is characterized as a monument of physical beauty.

2.1 The role of piezoelectric sensors

The function of a piezoelectric sensor that is in contact with the inside mass of the tree, is to convert the sound waves of the acoustic emission of the pest into small variations of voltage. Subsequently, a preamplifier coupled to the sensor amplifies the recorded signal. The amplified signal is a continuous, low amplitude background noise with occasional or regular acoustic emission hits. The detection, conversion and signal amplification process results in a measurable acoustic emission signal that can be classified automatically by machine learning techniques or heard by a human operator to provide important information about the activity (or lack of activity) inside the trunk. In this work the signal output is fed to a portable computer and the signal is automatically classified without human intervention. The benefit of the proposed approach is the rapid survey that becomes obvious when one has to check out a large number of trees (sometimes hundreds) or when there is strong background noise that interferes with the human operator trying to make out through the microphone if there is activity inside the tree. Sound emission created by insects or larvae feeding on palm mass sounds like tearing and snapping at intermittent intervals and at different intensities. If the probe is centimeters away from the infection squealing sounds can be detected as well as clicks. A single acoustic event caused by an RPW larva is illustrated in Fig. 1. The density of these events depends on the population of larvae and how far the probe is from the source of the activity. Activity rates can vary from a few hits every 1-2 seconds, to over 100 per second in very active areas. Due to the fibre structure and high humidity of the internal of the trunk, the distance over which detection may occur is about 1 meter form the source of activity (usually the crown).

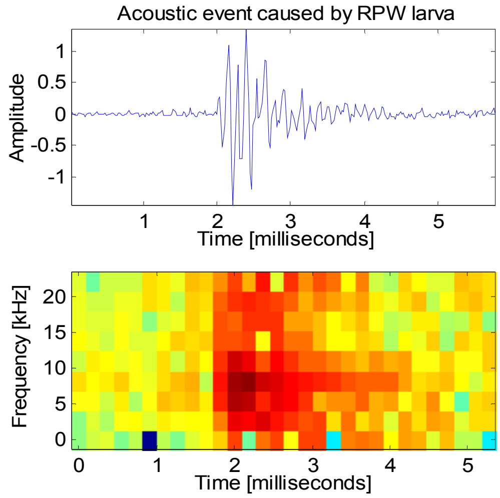

Figure 1 - Acoustic emission of RPW

2.2 Handling noises

Since the detection framework is based on contact probes which are inserted inside the tree the overall level of background noise is relatively low. However, field-recordings have revealed an unforeseen source of noise. The slightest interference from wind or frictional noises at the contact or noises introduced from the ambient environment (traffic, bird songs and voices) introduce fibre vibrations that interfere with the objective of monitoring. Most of these noises are of impulsive or transient nature and their spectral signature is very similar as in the case of RPW larvae moving/feeding when the probe is not close to the infestation location. Moreover, the human operator experiences much fatigue in his effort to carefully examine dozens of trees in the presence of a variety of background noises. All the aforementioned types of noise have been isolated from the field recordings and have formed the training corpora of the background noise model of the automatic detection system.

3. ACOUSTIC EMISSION DETECTION

In Fig. 2 we present a diagram illustrating the operation of the RPW detector. As the figure illustrates, the RPW detection process consists of two main steps: signal parameterization and pattern recognition. While the parameterization aims at computing descriptors which capture the acoustic

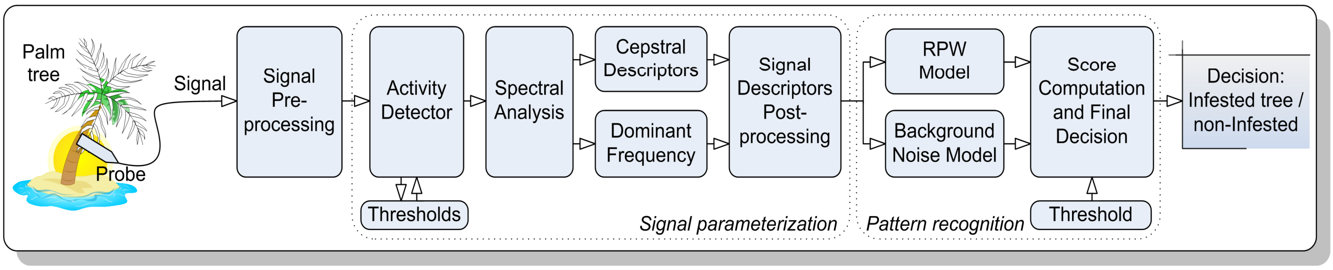

Figure 2 - Architecture of the detector of RPW acoustic activity

signature of the RPW, the pattern recognition step estimates the degree of match between an unknown input and the predefined models of the RPW and the background acoustic environment. Finally, a decision threshold is applied on the probabilistic match and a final decision is made: either the input signal originates from an infested tree or not. We briefly outline the four steps of signal acquisition and parameterization:

Step 1: *Signal acquisition and pre-processing:* The output of the piezoelectric probe is a continuous signal of acoustic emission hits, which is amplified 40 dB. The pre-processing of the input consists of mean value removal, which is performed on the time domain signal, for eliminating the dc-offset that might have occurred during signal acquisition, and amplification. Furthermore, in order to reduce the influence of the environment, a high-pass filtering is applied on the amplified signal. It was experimentally found that a Butterworth filter of order 50 with cut-off frequency 2 kHz effectively reduces the low-frequency interferences originating from interference vibrations.

Step 2: *Activity detection and variable-length segmentation:* The energy-based detector of acoustic activity [5] is adapted here for localizing RPW acoustic emissions and detecting their boundaries. These boundaries serve for obtaining a variable-length segmentation, which preserves for further processing only these portions of the signal that correspond to some kind of acoustic activity. In brief, the energy-based detector of acoustic activity functions as follows: In its relaxed state, the activity detector outputs zero. When there is some acoustic activity, which causes the short-term energy to reach or to exceed the threshold θ_{high}, the activity detector output triggers to value one. It remains in that state until the value of short-term energy drops below the threshold θ_{low}, when the activity detector relaxes to its inactive state. In the present work, both the thresholds are dynamically adjusted with respect to the mean energy of the signal observed for a time interval ~5 seconds, and the short-term energy is computed for groups of samples corresponding to ~1 millisecond with a skip step ~90 microseconds. These values provide a good trade-off between temporal resolution and computational demands.

Step 3: *Computation of the signal descriptors:* Each segment detected so far is zero-padded to 2048 samples and then becomes subject to the parameterization procedure. Specifically, we compute two alternative sets of spectral descriptors: The first one is based on the short-time discrete Fourier transform and leads to a set of linear frequency cepstral coefficients (LFCC), and the second one is based on the discrete wavelet packet transform (DWPT) and leads to a set of discrete wavelet packet features (DWPF). In both cases the frequency range of interest is [2, 16] kHz. For computing the LFCC, a filter-bank consisting of 140 equal-bandwidth and equal-height filters is applied on the logarithmically compressed power spectrum. The centres of the linearly spaced filters are displaced 100 Hz one from another, and serve as boundary points for the corresponding neighbouring filters. For computing the DWPF we relied on Daubechies wavelet function and a uniform wavelet packet tree obtained at level 8 of the DWPT. The leaves of the tree that correspond to the frequency range [0, 2] kHz were discarded. To compute the filter-bank output, we applied a logarithm on the squared sum of coefficients for each sub-band. This way, we obtained 112 frequency bands, each with resolution 125 Hz. Next, the log-energy filter-bank outputs are decorrelated via the discrete cosine transform to compute the final set of spectral descriptors.

A series of feature selection tests have shown that the first 24 cepstral coefficients are sufficient to carry out the recognition task and the contribution of the higher order coefficients has a minor impact on the recognition performance. Since we want to reduce the dependence on the field recordings setup, the 0-th cepstral coefficient was excluded from the feature vector. Thus, for each active segment the final feature vector is composed of the logarithm of the dominant frequency (*Fd*) of the spectrum that is estimated via search of the maximum magnitude in the power spectrum, and 23 LFCCs or DWPFs.

Step 4: *Post-processing of the signal descriptors:* Cepstral mean subtraction and dynamic range normalization is applied to the LFCC and DWPF feature vectors.

Finally, the normalized feature vectors obtained so far are fed to a GMM classifier [6]. The class-specific diagonal covariance GMMs are trained by employing a standard version of the expectation maximization algorithm [7].

4. EXPERIMENTS AND RESULTS

The training data for the RPW model consisted of 120 seconds of recordings originating from infested palms. These palms where definitely known to be infested as they were tagged by experienced biologists. Moreover the infestation was evident by the number of pests escape from wounds of the trunk. The background noise model (BNM) was trained by using 3 minutes of recordings, collected from 10 non-infested trees. These recordings were taken from areas of Crete where no transportation of plants has taken place. The test data consisted of fifty 5-second segments collected from 10 infested palm trees, and thirty-two 5-second segments form a number of non-infested palms.

During the evaluation of the RPW detector, each test segment (i.e. the unknown input) was parameterized and the feature vectors obtained to this end were fed to the RPW and BNM models. The log-likelihood ratio between the scores computed for the two models was subsequently compared to a predefined threshold. In preliminary experimentations with validation dataset we realized that a decision threshold in the range [-2, 0] leads to the optimal detection performance. In all following experimentations we report the RPW detection accuracy obtained for a decision threshold set to -1.

Results for experiments with noisy real-field recordings are presented Tables 1 and 2. The percentage values in each cell in the tables correspond to the RPW detection accuracy for the specific condition. Specifically, Table 1 presents results for the DWPF computed for 3 different wavelet functions, namely Daubechies of order 8, 32, and 44. As the results show, the Daubechies 32 wavelet function is the most appropriate among all tested wavelets. Furthermore, the feature vector that includes the dominant frequency *Fd* demonstrated superior performance when compared to the one for the feature vector that contains only the DWPF. Finally, the overall performance observed for the LFCC parameters (Table 2) was higher than the one of the DWPFs.

Comparative results of the variable length signal segmentation and the traditional fixed-frame-size approach are presented in Table 2 for the case of LFCC descriptors. As the table presents, the variable length segmentation outperformed the fixed-frame-size approach. Variable length segmentation shows the best recognition results and induces a multi-fold decrease of the computational demands required at the signal parameterization and pattern recognition stages, when compared to the traditional fixed-frame-size approach. Again the dominant frequency *Fd* proved useful, and contributed for enhancing the recognition accuracy.

Finally, it was observed that GMMs with four mixture components in combination with the {*LFCC*, *Fd*} feature vector offer the best RPW detection accuracy.

5. CONCLUSION AND FUTURE WORK

We presented an automatic detector for the RPW that is based on acoustic emissions which the larvae produce during eating and feeding inside the trunk of the palm tree. Our approach demonstrated to be highly accurate in detecting RPW when the probe is inserted within 1-meter form the point of infestation and cost-effective compared to human operators that are usually employed for bioacoustic identification. The prototype system is portable (a notebook computer and the probes) and was deployed and tested on-site in a real-field setup. It proved to provide a quick check-up of a large number of palm trees with minor human intervention.

Table 1. Comparison among various wavelet functions

# GMM comp.	*Daubechies 8*		*Daubechies 32*		*Daubechies 44*	
	DWPF	*DWPF+Fd*	*DWPF*	*DWPF+Fd*	*DWPF*	*DWPF+Fd*
2	93.5%	94.9%	91.5%	95.0%	86.5%	90.7%
3	94.9%	95.6%	**96.7%**	96.7%	83.9%	92.9%
4	**94.9%**	**96.5%**	93.6%	96.6%	87.5%	**96.6%**
5	94.9%	95.6%	95.0%	**98.2%**	**94.2%**	96.6%
6	94.1%	95.6%	94.1%	96.5%	88.2%	96.6%

Table 2. Variable length segmentation vs. fixed-frame approach

# GMM comp.	1 ms *fixed frame size*		*variable length segment*	
	LFCC	*LFCC+Fd*	*LFCC*	*LFCC+Fd*
2	**92.4%**	**95.8%**	**97.8%**	96.9%
3	89.3%	85.3%	94.7%	99.1%
4	89.3%	88.7%	94.7%	**99.5%**
5	85.8%	88.9%	93.3%	98.7%

We postulate that the results of this study will benefit potential pest monitoring applications as our approach is directly expandable to other pests.

6. ACKNOWLEDGMENT

We acknowledge the contribution of Dr. D. Kododimas from the Benakio Phytopathological Institute of Athens Greece for his guidance during this research. The acoustic emission detection hardware was based on the AED 2000L® system.

REFERENCES

[1] K. Gaston and M.A. O'Neill, “Automated species identification -- *why not*?” *Philosophical Transactions of the Royal Society B*, vol. 359, no. 1444, pp. 655–667, 2004.

[2] A.T. Watson, M.A. O'Neill and I.J. Kitching, “A qualitative study investigating automated identification of living macrolepidoptera using the Digital Automated Identification SYstem (DAISY)”, *Systematics and Biodiversity*, vol.1, no. 1, pp. 287–300, 2003.

[3] E. Chesmore, “Application of time domain signal coding and artificial neural networks to passive acoustical identification of animals”, *Applied Acoustics*, vol. 62, pp. 1359–1374, 2001.

[4] C. Dietrich, “Temporal Sensor fusion for the Classification of Bioacoustic Time Series”, PhD thesis, University of Ulm, Department of Neural Information Processing, 2004.

[5] T. Ganchev and I. Potamitis, “Automatic Acoustic Identification of Singing Insects”, *Bioacoustics*, ISSN 0952-4622, vol. 16, pp. 281–328, 2007.

[6] D. Reynolds and R. Rose, “Robust text-independent speaker identification using Gaussian mixture speaker models”, *IEEE Trans. on Speech and Audio Processing*, vol. 3, no. 1, pp. 72–83, 1995.

[7] G.J. McLachlan and T. Krishnan, *The EM algorithm and extensions*, Wiley Series in Probability & Statistics, New York: Wiley, 1997.