

APPLICATION OF THE SIGNAL PROCESSING TECHNOLOGY IN THE DETECTION OF RED PALM WEEVIL

Walid Barakat Hussein, Mohamed Ahmed Hussein, and Thomas Becker

Faculty of Life Science Engineering, Technical University of Munich
Weihenstephaner Steig 23, 85350, Freising, Germany
Email: whussein@pa.uni-hohenheim.de

ABSTRACT

The Red Palm Weevil (RPW) is the most destructive pest for the date palm in the world. The economic damage to palm crops due to RPW could be mitigated significantly by bioacoustics recognition in an earlier phase of infestation and applying the appropriate treatment. In this paper, a novel signal processing system is developed to detect the existence of RPW through its bioacoustics features. A large set of features are extracted, including some unconventional features which are temporal roll-off, temporal slope, and temporal spread. Since the recorded sound is divided into time frames, an analysis for the criteria behind the choice of the optimum time frame length, as well as the selection of the suitable window function is provided. The results imply the efficiency of the developed system with the selected representative features, window functions, and frame period to detect the existence of the RPW through its feeding sound.

1. INTRODUCTION

Since 1980s, RPW has rapidly expanded its geographical range westwards. It reached Saudi Arabia and the United Arab Emirates in 1985, spreading throughout the Middle East and into Egypt. In 1994 it was detected in Spain, Israel, Jordan and Palestine in 1999, Italy in 2004, Canary Islands in 2005, Balearic Islands, France, and Greece in 2006, and Turkey in 2007 [1]. All stages (egg, larva, pupa and adult), as described in figure 1, are spent inside the palm itself and the life cycle (about 4 months) cannot be completed elsewhere.

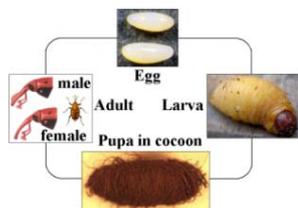


Figure 1- The life cycle stages of the red palm weevil

Larvae and adults destroy the interior of the palm tree, often without the plant showing signs of deterioration unless damage is severe [2]. Hollowing out of the trunk reduces its mechanical resistance, making the plant susceptible to collapse and a danger to the public [3]. In most cases, attack leads to the mortality of trees whatever their sizes. The larvae are large, but the hidden kind of living makes an early detection with traditional methods very difficult. Adults of RPW can be efficiently monitored using pheromone based traps [4]. However, these methods are unsuitable for quarantine

inspections of planting material. Consequently, infested planting material is often transported to a new location before the first detectable symptoms of infestation appear [5].

Bioacoustics technology and x rays enable the detection of the early phase of infestation, however the usage of x rays is expensive, but the acoustic technology has potential for reducing the expense and dangers involved in tree inspection. The acoustic recordings from insects in trees often reveal signals with spectral and temporal features that make them distinctive and easily detectable [6].

Preliminary studies demonstrated that sensitive microphones and dedicated amplifiers enable detection of the movement and feeding sounds of RPW larvae in palm trees [7]. Due to its high reproduction rate, the RPW prefers to live with no other insects in one trunk, which gives a good base for acoustic detection. On the other hand, the greatest difference between Red Palm Weevil and other possible beetle species is that the activities of RPW seem to be very aggressive, as shown in figure 2; the sound stream is a collection of many pulses showing the difficulty to separate them. While the sound of *Copris hispanus* [8] is more distinctive and clear to spot.

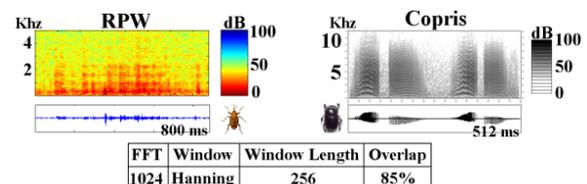


Figure 2- Spectrogram, envelope and detailed waveforms of pulse trains emitted by RPW feeding (left) and *Copris hispanus* sounds (right)

Different sounds of the RPW activities could be recorded, including feeding, moving, and spinning the cocoon sounds. In this paper, the audio detection system for the RPW feeding sound is presented in detail. The following section demonstrates the recording method and the families of features which have been used in the system. Following by a section clarifies the criteria behind the choice of the window function, before applying Fourier Transform to each time frame of the recorded sound. Afterward, an explanation for the selection of the representative features and the optimum frame length is described. Finally, the detection system results are shown and discussed. After the positive determination of the presence of the red palm weevil in a palm, it is necessary to remove the palm tree, because up today no real efficient treatment method exists. But with the removing of an infested palm a strong blow against the Red Palm weevil population will be done.

2. MATERIAL AND METHOD

2.1 Recording Device

An infested date palm sample trunk (70 cm length, 38-43 cm diameters, 117 kg weight) was transferred from Kingdom of Saudi Arabia to Germany, and stored in a quarantine room. The trunk is further divided into small blocks, to focus on the weevil development and investigate the activities of the larvae during recording. This results in a complete picture for the hidden RPW activities in the trunk. Afterwards, the equipment was installed including a mercury steam lamp, heater, humidity device, thermometer, infrared video cameras, as well as the recording device.

Two different recording devices were tested to find the best of them for this task. The first was a Laar Ultrasound Gate[®] hard disk recording System (50 Hz – 250 kHz). The second was Laar WD 60[®] detector with amplifying system and insertion sensors of different types (Contact microphone, airsound ultrasound microphone, contact acceleration sensor and a combined contact airsound probe sensor). The best recordings are obtained via the second device which was attached to a Laar/Avisoft SASLab Pro[®] sound analyzer (50 Hz to 30 kHz). Since there was no contact from the recording person, every recorded sound was caused by the RPW activities in the trunk. For a period of 13 subsequent months (~3 life cycles), covering different environments (i.e., pressure and temperature changes), 251 adult weevils had left the trunk - later were eradicated - and 980 successful feeding sound recordings had been made for the RPW.

2.2 Signal Processing

The recordings were digitized at 11025 Hz sampling rate and then were high-pass filtered with cutoff frequency of 0.2 kHz to eliminate low-frequency background noise. Afterwards, they were divided into time frames with 90% overlap between frames. A program was developed using MATLAB (Math Works[®]); with friendly user graphical interface (GUI), to hold all the signal processing procedures up to the detection step.

2.2.1 Time Domain Features

A list of temporal domain features, shown in [Table 1], was extracted for each time frame in order to have a time distribution of the features along the recorded signal. Three unconventional features were being included: temporal roll-off that describes the time below which 90% of the energy distribution is concentrated, temporal slope of decreasing or increasing of the signal amplitudes which is computed by linear regression, and temporal spread which represents the spread of the signal amplitudes around its mean value.

For extracting the spectral domain features, each time frame was scaled by a suitable window function and then transformed to frequency domain using Fast Fourier Transform (FFT) method, which assumes that a signal is periodic. When the FFT of a non-periodic signal (which is the form of most signals) is computed, the resulting spectrum suffers from spectral leakage and reduction of the frequency resolution.

2.2.2 Windowing

The window function is used for better representation of the frequency spectrum of the data, by reducing the spectral leakage and/or increasing the frequency resolution. Figure 3 shows how a Hanning window function reduced the spectral leakage, when applied to a non-periodic sine wave.

Table 1: The time domain features which are extracted from the recorded signals. $[x(n)]$ is the amplitude at sample n , and N is the total number of samples

Name	Equation	Set
Short time energy	$\int_{frame} (x(n))^2 dn$	1
Energy root mean square	$\sqrt{\frac{\sum_{n=1}^N (x(n))^2}{N}}$	1
Temporal centroid	$\frac{\sum_{n=1}^N n x(n) ^2}{\sum_{n=1}^N x(n) ^2}$	1
Temporal entropy	$-\sum_{n=1}^N \left(\frac{ x(n) }{\sum_{n=1}^N x(n) } \right)^2 \ln \left(\frac{ x(n) }{\sum_{n=1}^N x(n) } \right)^2$	
Temporal roll-off	$\sum_{n=1}^N x(n) ^2 = 0.9 \sum_{n=1}^N x(n) ^2$	1
Temporal slope	$\frac{N \sum_{n=1}^N (n \cdot x(n)) - \sum_{n=1}^N n \sum_{n=1}^N x(n)}{\sum_{n=1}^N x(n) (\sum_{n=1}^N n^2 - (\sum_{n=1}^N n)^2)}$	1
Temporal flux	$\sum_{n=1}^N \left(\frac{ x(n) }{\max(x(n))} \right) - \frac{ x(n) }{\max(x(n))} \Big _{t-1}^t$	
Temporal flatness	$\frac{(\prod_{n=1}^N x(n))^{1/n}}{\frac{1}{N} \sum_{n=1}^N x(n) }$	
Temporal crest-factor	$\frac{\max(x(n))}{\frac{1}{N} \sum_{n=1}^N x(n) }$	1
Temporal decrease	$\left(\frac{1}{\sum_{n=2}^N x(n) } \sum_{n=2}^N \frac{ x(n) }{n-1} \right)$	
Temporal tonality	$\min \left(\frac{10 \log(\text{temporal flatness})}{-60}, 1 \right)$	
Temporal spread	$\int_{t(1)}^{t(N)} (t(n) - t(\text{tempcentroid}))^2 tpmf(n). dt$, $tpmf(n) = \frac{ x(n) }{\sum_{n=1}^N x(n) }$	1
Temporal skewness	$\frac{\text{temporal spread}^{3/2}}{m_3}$, $m_3 = \int_{t(1)}^{t(N)} (t(n) - t(\text{tempcentroid}))^3 tpmf(n). dt$	
Temporal smoothness	$\sum_{n=2}^{N-1} \frac{20 \log x(n) - \frac{\log x(n-1) + \log x(n) + \log x(n+1) }{3} }{3}$	
Zero crossing rate	$\frac{1}{2N} \sum_{n=2}^N \text{sign}(x(n)) - \text{sign}(x(n-1)) $, $\begin{cases} \text{sign} = 1 & ; x(n) > 0 \\ \text{sign} = 0 & ; x(n) = 0 \\ \text{sign} = -1 & ; x(n) < 0 \end{cases}$	1

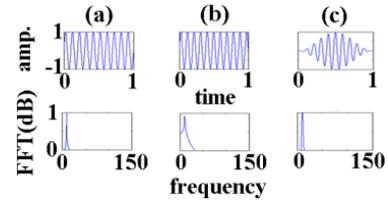


Figure 3- The FFT computation assumes that a signal is periodic, there is integer number of cycles in the data block (as in (a)). The FFT for a non-periodic signal leads to a spectrum that suffers from leakage (as in (b)). When a Hanning window is applied, the leakage is reduced and the resultant spectrum is a sharp narrow peak (as in (c)).

The most common window functions are rectangular, Hanning, flat top and Blackman windows. Each window function has slightly different wave form from the others, as shown in figure 4, which results in a different performance in the frequency domain. In figure 4, the tighter the main lobe width, the better the frequency resolution of the window functions, as for the rectangular window function. On the other hand, the higher the reduction rate between the main lobe and the first side lobe, the less the spectral leakage around the central frequency, as for the Blackmann window.

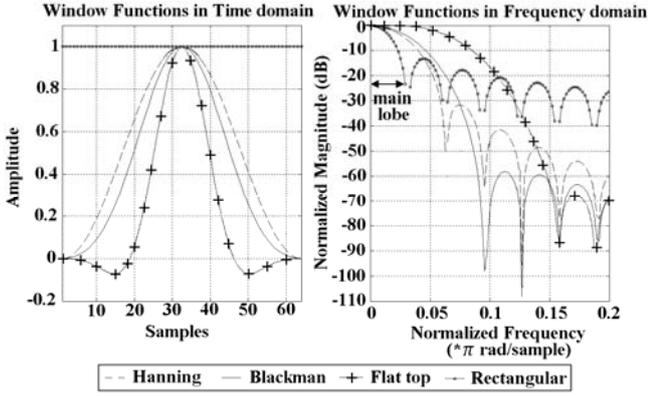


Figure 4- The distribution of the most common window functions in time domain (left) and frequency domain (right)

The efficiency of each window function according to their frequency resolution and spectral leakage reduction criteria are displayed in table 2. In this study, the rectangular window was used to calculate the spectral features that require high frequency resolution (e.g. spectral roll-off), while the Blackmann window is applied to calculate the spectral features that require less spectral leakage (e.g. spectral centroid).

Table 2: Comparison between the window functions of figure 5 in their frequency resolution and spectral leakage analyses

Window	Frequency Resolution	Spectral Leakage Reduction
Hanning	Good	Good
Blackman	Poor	Best
Flat top	Poor	Good
Rectangular	Best	Poor

2.2.3 Frequency Domain Features

After applying the FFT for the windowing time frame, a list of spectral domain features, as shown in table 3, were extracted. Adding these spectral features to the previously calculated temporal features, a group of 31 features were extracted for each time frame.

2.2.4 Features' Selection

From the above extracted features, there are some features which are better than others in detection of the feeding sound of the RPW. On the other hand, some features may be useless to perform this particular task. The purpose of applying a features' selection algorithm is to select the representative features from a large set of features. Feature selection algorithms typically fall into two categories; feature ranking [9] and subset selection [10]. Feature ranking ranks the features by a metric and eliminates all features that do not achieve an adequate score. Subset selection searches the set of possible features for the optimal subset by removing most irrelevant and redundant features from the data. In this study the feature ranking method was applied. To handle this, the temporal and spectral features were extracted for 765 out of the 980 RPW feeding sound recordings. Subsequently, the average feature value is calculated for its 765 values (based on the 765 recordings). If more than 80% of the recordings have feature value that lie within 5% deviation from the corresponding average feature value, this feature will be selected, and stored in set(1). The algorithm is explained in

figure 5, and the selected features of set(1) are indicated beside each feature in table 1 and table 3.

Table 3: The frequency domain features which are extracted from the recorded signals. $|X(n)|$ is the spectrum at bin number n , and N is the total number of frequency bins].

Name	Equation	Set
Spectral roll-off	$\sum_{n=1}^{n_r} X(n) = 0.9 \sum_{n=1}^N X(n) $	1
Spectral flux	$\sum_{n=1}^N \left(\frac{ X(n) }{\max(X(n))_t} - \frac{ X(n) }{\max(X(n))_{t-1}} \right)^2$	
Spectral centroid	$\frac{\sum_{n=1}^N n \cdot X(n) }{\sum_{n=1}^N X(n) }$	1
Spectral band width	$\sqrt{\frac{\sum_{n=1}^M (n - C_t)^2 \cdot (X(n))^2}{\sum_{n=1}^M (X(n))^2}}$, $M = \text{find} \left(X \geq \frac{\max(X)}{2}, \text{last}' \right)$	1
Spectral root mean square	$\sqrt{\frac{\sum_{n=1}^N (X(n))^2}{N}}$	1
Spectral slope	$\frac{1024 \sum_{n=1}^{1024} n X(n) - \sum_{n=1}^{1024} n \sum_{n=1}^{1024} X(n) }{\sum_{n=1}^{1024} X(n) \left(\sum_{n=1}^{1024} n^2 - \sum_{n=1}^{1024} n \right)}$	1
Spectral flatness	$\frac{(\prod_{n=1}^{1024} X(n))^{1/n}}{\frac{1}{1024} \sum_{n=1}^{1024} X(n) }$	
Spectral crest factor	$\frac{\max(X(n))}{\frac{1}{1024} \sum_{n=1}^{1024} X(n) }$	
Spectral energy	$\int_{n=1}^{1024} X(n) ^2 \cdot dn$	
Spectral tonality	$\min \left(\frac{10 \log(\text{spectral flatness})}{-60}, 1 \right)$	
Spectral smoothness	$\sum_{n=2}^{1023} 20 \left \log X(n) - \frac{\log X(n-1) + \log X(n+1) }{3} \right $	
Spectral decrease	$\left(\frac{1}{\sum_{n=2}^{1024} X(n) } \right) \sum_{n=2}^{1024} \frac{ X(n) - X(1) }{n-1}$	1
Spectral spread	$\int_{n=1}^{1024} (n - C_t)^2 \cdot pmf(n) \cdot dn$, $pmf(n) = \frac{ X(n) }{\sum_{n=1}^{1024} X(n) }$	1
Spectral skewness	$\frac{\int_{n=1}^{1024} (n - C_t)^3 \cdot pmf(n) \cdot dn}{\text{spectral spread}^{3/2}}$	1
Spectral kurtosis	$\frac{\int_{n=1}^{1024} (n - C_t)^4 \cdot pmf(n) \cdot dn}{\text{spectral spread}^2}$	1
Spectral entropy	$-\sum_{n=1}^{1024} pmf(n) \cdot \ln(pmf(n))$	1

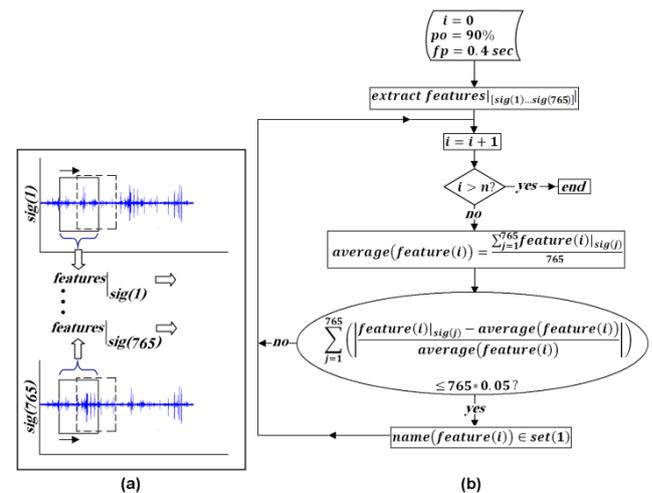


Figure 5- (a) Schematic diagram for the features selection algorithm. (b) Flow chart for the features selection algorithm. (po) is the percentage of overlapping, (fp) is the frame period, and (n) is the total number for the extracted features out of the signal.

2.2.5 Optimum Frame Length

As mentioned above, the recorded signal is divided into a number of time frames. The time frame length depends on the total duration of the desired sound (i.e., the feeding sound), the percentage of overlapping, and the required number of time frames. On the other hand there is an optimum selection for the frame length which is affected by the intended representative features. In depth, if the mean zero crossing rate is considered as identification feature for the feeding sound of the RPW, the frame period should ensure that the value of the mean zero crossing rate is approximately the same for all recorded feeding sounds of the RPW. In figure 6, the predetermined set of features (i.e., set(1)) was extracted for two audio signals (signal (1), and signal (2)), as well as for a reference signal of a typical feeding sound of the RPW, for different choices of frame lengths. Signal (1) contains a feeding sound for a RPW, while signal (2) does not. The mean square error between the selected features of each signal and the reference signal features was calculated and scanned with the corresponding frame length. As a result of figure 6, an optimum frame length of 0.42 sec. is selected.

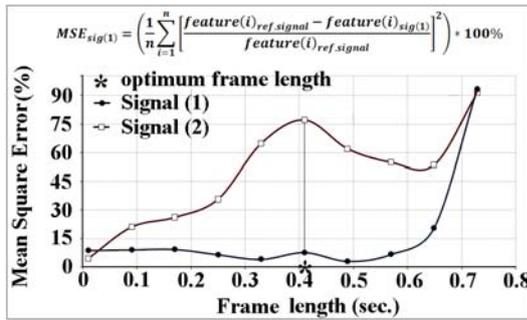


Figure 6- Mean square error between recorded sounds and a reference feeding sound of RPW versus the frame length, based on features set(1). Signal (1) contains a feeding sound while signal (2) does not.

3. RESULTS AND DISCUSSION

3.1 Developing of The System

The selected temporal as well as the selected spectral features were extracted for 765 out of the 980 recordings, using forms of table 1 and table 3; respectively. The distributions of these features with time are shown in figure 7 for a typical RPW feeding sound. In figure 7, some features distributions suffer from approximately equal standard deviation; e.g. temporal roll-off and temporal centroid, spectral slope and spectral centroid, and spectral kurtosis and spectral skewness. Meanwhile the values of the mean for those distributions are different. As a conclusion, the mean value of most features is more representative than the standard deviation. This remark necessitates careful attention to keep out redundant values in the choice of the related statistic for each feature. Equation (1) and equation (2) were used to calculate the mean and the standard deviation; respectively, while table 4 shows which statistic assigned to each feature.

$$f_{mean} = \frac{\sum_{i=1}^K f_i}{K} \quad (1), \quad f_{std} = \sqrt{\frac{\sum_{i=1}^K (f_i - f_{mean})^2}{K - 1}} \quad (2)$$

Where f_i is the value of the feature on time frame i , and K is the total number of frames on the recorded signal.

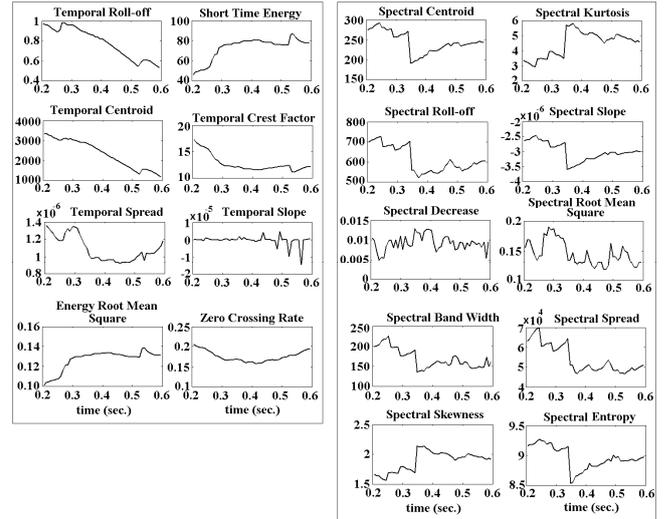


Figure 7- Distribution of the selected temporal features (left) and spectral features (right) with time for a typical RPW feeding sound of total length 0.8 sec., divided into 10 frames each of 0.42 sec. by means of 90% overlapping.

Therefore, one representing value was considered for each feature over one recording, and this value was averaging for the 765 recordings. The resultant system contained 18 representative features for the feeding sound of the RPW, and these features values were saved as the system *features*.

Table 4: The related statistic and the window function assigned to each feature

Feature	Related statistic	Window Function
Spectral Roll-off	mean	Rectangular
Spectral Centroid	mean	Blackman
Spectral Band Width	mean	Blackman
Spectral Root Mean Square	mean	Blackman
Zero Crossing Rate	mean	---
Short Time Energy	std	---
Energy Root Mean Square	mean	---
Temporal Centroid	std	---
Temporal Roll-Off	mean	---
Spectral Slope	mean	Blackman
Spectral Decrease	mean	Blackman
Spectral Spread	mean	Blackman
Spectral Skewness	mean	Blackman
Spectral Kurtosis	mean	Blackman
Spectral Entropy	mean	Blackman
Temporal Slope	mean	---
Temporal Crest-Factor	mean	---
Temporal Spread	mean	---

3.2 Validation of The System

The remaining 215 out of the 980 feeding sound recordings were used for the validation of the developed system. At which the pre-selected 18 features are extracted for each recording, and then were compared to the *system features*. 203 records were successfully detected since their extracted features were within 5% from those of the *system features*. On the other hand, 12 records were not detected since more than 10 of their extracted features had more than 12% percent deviation from the corresponding *system features*. This means that the system efficiency is higher than 94% in the detection of the RPW feeding sounds, when the feeding sound was defined to be the one with features that have a maximum of 5% deviation from the corresponding *system features*.

3.3 Evaluation of The System

In order to test the power of the developed system to detect the existence of the RPW in a field test, an audio stream of length 5 minutes was recorded on one infected palm tree, and the sound features were extracted for each 0.8 second batch, with 90 % overlaps. The extracted features were scanned for each batch as shown in figure 8, as well as the *system features* (the horizontal solid lines). The grey regions in figure 8 represent the detected feeding sound by an expert.

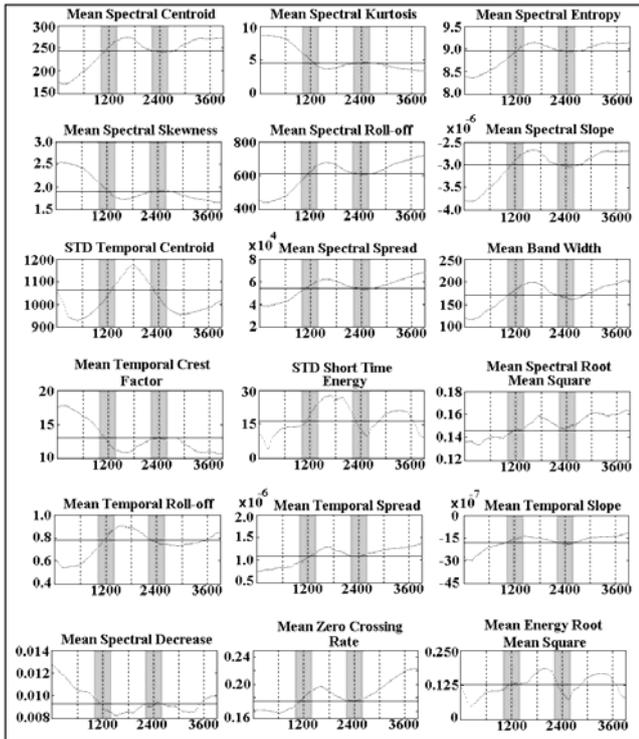


Figure 8- The extracted features from an audio stream of 5 minutes length are scanned with the number of batch. The stream is divided into 3750 batches of length 0.8 sec. with 90% overlapping. The solid horizontal line represents the system features for the RPW feeding sound, while the grey regions are detected by an expert to contain RPW feeding sound.

From figure 8, the features distributions at some batches have the same values of those for the reference feeding sound. These batches lie on the same times that were detected by an expert as a RPW feeding sound. Mean temporal roll-off, mean spectral decrease, STD short time energy, and STD temporal root mean square introduce one more feeding sound near batch number 3600 [ca 4.75 minutes]. But it is not approved by the other features. Although all the above features were used in the developed system, some of them only are sufficient to detect the RPW feeding sound in the 5 minutes stream, as mean spectral entropy, mean spectral roll-off, and mean zero crossing rate. The results imply the efficiency of the developed system with the selected representative features, window functions, and frame period to detect the existence of the RPW through its feeding sound.

4. CONCLUSION

The economic damage to palm crops due to RPW could be mitigated significantly by bioacoustics recognition in an earlier phase of infestation and applying the appropriate treatment. In this study, before dividing the sound wave into time frames, the optimum frame length - for the case of the RPW feeding sound - was investigated, and found to be 0.42 sec. with 90% overlapping. For each frame; the selected 8 time domain features were extracted. In addition; the criteria behind the selection of the window function were studied, followed by the extraction of the selected 10 frequency domain features, using their appropriate window functions. The developed system was validated to have efficiency higher than 94%, therefore; it was used to powerfully detect the existence of the RPW feeding sound in a five minutes sound stream.

ACKNOWLEDGMENT

This work is supported by the Deutsche Bundesstiftung Umwelt (DBU); grant number AZ 25124-34. Finally, we are grateful to Benedikt von Laar who has helped by providing recordings of the red palm weevil sounds, and has manually determined the feeding sounds in these recordings.

REFERENCES

- [1] C. Malumphy and H. Moran, "Plant Pest Notice: Red palm weevil *Rhynchophorus ferrugineus*", *Central Science Laboratory*, Sand Hutton, York, U.K., May 2007
- [2] D. Blumberg, A. Navon, and etal, "Date palm pests in Israel early second millennium", *Alon Hanotea*, 55, pp. 42-48., 2001.
- [3] G. Davis and R. Moore, "Insects on Palms", *Borers of palms*, Howard, CABI Publishing, Wallingford, UK, pp. 267-305, 2001.
- [4] V. Soroker, D. Blumberg, and etal, "The current status of red palm weevil infestation in date palm plantations", *Israel Phytoparasitica*, vol. 33, pp. 97-106, 2005.
- [5] R. Mankin, A. Mizrach, and etal, "Sounds of Wood-Boring Beetle Larvae", *Florida Entomologist*, vol. 91(2), pp. 241-248, 2008.
- [6] R. Mankin, W. L. Osbrink, and etal, "Acoustic detection of termite infestations in urban trees". *Journal of Econ. Entomol*, pp. 981-988, 1995.
- [7] A. Hetzroni, A. Mizrach, and etal, "Developing spectral model to monitor activity of red palm weevil". *Alon Hanotea*, 58, pp. 466-469, 2004.
- [8] C. Palestini, G. Pavan, and etal., "Acoustic signals in *Copris incertus* Say (Coleoptera Scarabaeidae Coprinae)", *Ethology, Ecology & Evolution*, Special Issue 1, pp. 143-146, 1991.
- [9] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27(8), pp.1226-1238, 2005.
- [10] S. Singhi and H. Liu, "Feature subset selection bias for classification learning", *Proceedings of the 23rd international conference on Machine learning*, Pittsburgh, Pennsylvania, June 25-29. 2006, pp. 849-856.