

RESEARCH ARTICLE

An information rich subspace separation for non-stationary signal classification

T.A. Ratnayake, D.B.W. Nettasinghe*, G.M.R.I. Godaliyadda, M.P.B. Ekanayake and J.V. Wijayakulasooriya

Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya.

Revised: 28 October 2015; Accepted: 30 December 2015

Abstract: This paper proposes a novel automated approach for the classification of highly non-stationary signals based on a non-principal component analysis (non-PCA) methodology. This method generates an eigen analysis based pseudo-spectrum to emulate the spectral characteristic variations of the non-stationary signals to be classified. Then, the estimated pseudo-spectrum is used to implement a comb like subspace filter structure, which captures the variations of all significant spectral components throughout the whole observation period. It is shown that this filter implementation method yields better results than the existing dimensionality reduction methods, which only utilise the principal k components of the eigen space. Finally, a novel probabilistic approach which creates a signature vector representing each class of signals in the training phase is proposed for the classification process. It is also shown that the proposed method can be effectively used not only for classification but also for the extraction of hidden stationary signature features from a non-stationary signal. Further, it is also proven that the proposed subspace filtering scheme can be used as a dynamic spectral estimation technique, which can eliminate the time frequency resolution tradeoff that exists in techniques such as short-time fourier transform (STFT).

Keywords: Feature extraction, non-stationary signals, pattern recognition, signal classification, speech signals, subspace techniques.

INTRODUCTION

Classification of stationary signals according to their original sources is of high importance and a much studied problem in a vast number of applications such as vehicular classification (Padmavathi *et al.*, 2010; Kandpal *et al.*, 2013), acoustic emission classification (Meng *et al.*, 2011), etc. Although the classification of stationary signals according to their origin is a well

studied problem, the classification of highly non-stationary signals still remains a largely unexplored area due to the difficulty in accurately capturing the variations in both time and frequency domains.

The signals in concern in a vast number of important practical applications are non-stationary in nature. However, currently in many practical applications stationarity is assumed for such non-stationary signals considering small windows or slow fading/variations. Some examples where the analysis and classification of these non-stationary signals is of significant importance include EEG signal classification (Subasi & Erçelebi, 2005), underwater acoustics analysis (Josso *et al.*, 2011), EMG signal classification (Chendeb, 2010), speech signal analysis and classification (Tüske *et al.*, 2012), and bio signal analysis (Avendaño-Valencia *et al.*, 2011), etc.

Two of the most widely used methods for the analysis and classification of non-stationary signals are parametric methods and short-time fourier transform (STFT) based methods. However, parametric power spectral estimation methods such as the auto-regressive method are not appropriate for non-stationary signals (Subasi & Erçelebi, 2005). Even the short-term spectral estimation methods such as STFT are ineffective when the time varying information are useful (Sukittanon *et al.*, 2003) due to the time-frequency tradeoffs and less resolution effects.

In a previous study (Nettasinghe *et al.*, 2013; Pollwaththage *et al.*, 2013; Ratnayake *et al.*, 2013), it has been proven that subspace based signal analysing techniques can be successfully incorporated for signal classification in cases where the signals are highly

* Corresponding author (wijayanthanet@gmail.com)

correlated and noisy. In this paper, we explore how the subspace based signal analysing techniques can be utilised for feature extraction and classification of non-stationary signals.

The major contribution of this paper is the introduction of a novel approach for non-stationary signal classification based on a non-PCA approach. In other words, unlike the widely used PCA based methods, this method selects frequency components from a signal in a manner such that the selected spectral components represent all distinct spectral components that might contain information about the signal (not only the ones that have the highest power). It utilises an eigen-decomposition based methodology to identify a signal specific subspace that contains unique spectral characteristics for the signal in concern. It is shown that this method can be used for hidden signature extraction and classification of non-stationary signals. Further, it can be used to identify and extract stationary components of a non-stationary signal. In other words the non-stationarity of the spectral content can be harvested in the generation of the fingerprint for classification or synthesis of the signals.

SIGNAL DATABASE AND PRE-PROCESSING OF SIGNALS

A collection of signals is utilised to verify the classification accuracy of the proposed methods and to compare their performance with the existing methods. Synthetic signals are mainly used to formulate the method, and naturally occurring signals such as speech signals and elephant rumbles are used to test the accuracy of the proposed methods and to prove their applicability.

The higher degree of non-stationarity in speech signals (Lee & Lee, 2001; Nayak & Bhat, 2003; Tüske *et al.*, 2012) and the fact that they contain a significant amount of important hidden information such as gender specific information (Fagundes *et al.*, 2002), linguistic and acoustic information (Hailemariam & Prahallad, 2007; Lee & Siniscalchi, 2013), are the main reasons for using speech signals.

As shown in Table 1, five speakers were used and fifty signals from each speaker were recorded while they were reciting the same word ('Evaporation'). Figure 1 depicts the time domain waveforms and the spectrograms for one recital by each speaker. As the same word was used, all the word specific features (phonemes and syllables) remain highly correlated between the speakers. Further, the selected speakers have similar English accents, hence increasing the correlated content. Therefore, only the

Table 1: Details of the speech signal database

Speaker	Spoken word	Gender	No. of signals
EVP1	'Evaporation'	Female	50
EVP2	'Evaporation'	Male	50
EVP3	'Evaporation'	Female	50
EVP4	'Evaporation'	Male	50
EVP5	'Evaporation'	Female	50

hidden features related to the vocal cords of a particular speaker will be useful for the identification of that speaker as word specific features and the accent are the same for all five speakers. This database would provide an ideal setup to extract the hidden information about a particular speaker, which will be useful in recognising that speaker.

Figure 1(b) shows that the signals are highly non-stationary since the spectral content of the signals is rapidly changing with time. It is noted that classification by visual observation is not possible due to the high degree of correlation between the signals of different classes. This again highlights the importance and the need for accurate extraction of hidden signature features.

Synthetic signals used in this research are generated by concatenating different sinusoids of different frequencies such as the synthetic signal no.1 shown in Figure 2 (a). These signals are mainly used to formulate the proposed algorithm.

Apart from the synthetic signals and the main database of speech signals, elephant infrasound calls are also used to illustrate how the proposed algorithm can be used for stationary feature extraction from non-stationary signals.

PROPOSED ALGORITHM FOR SPECTRAL ESTIMATION AND FILTER IMPLEMENTATION

In this section, a novel non-PCA based approach for the generation of the pseudo-spectrum of a signal is presented. Then, how the generated pseudo-spectrum can be used to capture the significant spectral components of a signal is described. It is also shown how the proposed method can be used for the extraction of stationary signature features from a non-stationary signal. The significant advantages of the proposed method over some of the popular dynamic spectral estimation methods such as PCA based methods and STFT is also discussed in this section.

In this paper, the signal in consideration x ,

$$x = [x(0) \ x(1) \ \dots \ x(i) \ \dots \ x(N)] \quad \dots(1)$$

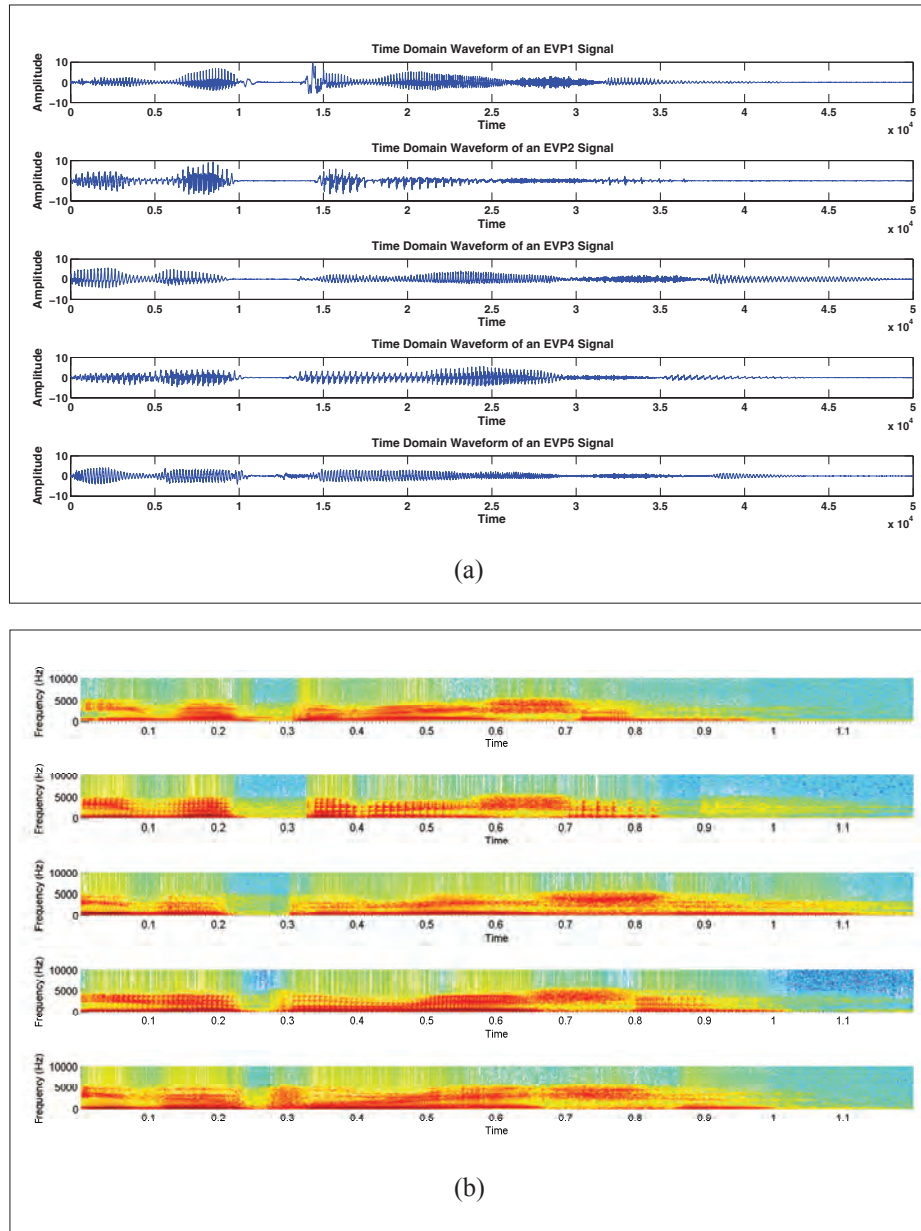


Figure 1: One recorded signal from each speaker when saying the word “evaporation” (a) time domain waveforms; (b) spectrograms

is treated as a set of samples stored in an array where $x(i)$ is the i^{th} sample of the signal obtained by sampling at a frequency of F_s (here F_s is 44100 Hz).

In order to calculate the statistical parameters of the signal x , it is assumed that the signal x is a realisation of an ergodic random process. In other words, it is assumed that every realisation of the random process, which generates the signal x carries the complete statistics of the random process with it.

Next, the autocorrelation matrix (ACM) of the signal x denoted by ϕ_{xx} , is created as,

$$\phi_{xx} = \begin{bmatrix} R_{xx}(0) & \dots & \dots & R_{xx}(N-1) \\ \vdots & \ddots & R_{xx}(\tau) & \vdots \\ \vdots & \dots & \ddots & \vdots \\ R_{xx}(N-1) & \dots & \dots & R_{xx}(0) \end{bmatrix} \dots(2)$$

Here, $R_{xx}(\tau)$ [where $0 \leq \tau \leq (N-1)$] is the autocorrelation function of the signal x , which is defined as,

$$R_{xx}(\tau) = E[x(i)x(i + \tau)] \quad \dots(3)$$

where $x(i)$ is the sample of the signal x . The ergodic assumption allows us to calculate the same result using the time average with the expression,

$$R_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \sum_{i=-T}^T x(i)x(i + \tau) \quad \dots(4)$$

without using multiple realisations (ensemble).

Therefore, all the elements of the ACM can be easily calculated using the expression (4) for any signal.

In the training phase of this proposed approach, the minimum ACM order (maximum number of shifts in the time domain), which is required to model the signals in concern accurately is estimated using a minimisation of variance technique. Here the ACM is constructed initially using a very small model order ($N = 10$). Thereafter, the eigenvalues and eigen-vectors of the ACM are identified. The eigenvalues of the ACM (sorted in the descending order) are stored in the diagonal of the matrix D and their respective eigen-vectors are stored in the matrix V . Therefore the relation between V , D and the ACM (Φ_{xx}) can be expressed as

$$\Phi_{xx} \times V = V \times D \quad \dots(5)$$

using the basic concept of eigen-vectors and eigenvalues.

Then the eigen-vectors are realised as a set of FIR digital filters (or eigenfilters). For example, consider the eigen-vector

$$q_i = [q_{0,i} \quad q_{1,i} \quad \dots \quad q_{N-1,i}]^T \quad \dots(6)$$

which corresponds to the i^{th} largest eigenvalue of Φ_{xx} . Hence, q_i appear in the i^{th} column of V . This will produce the transfer function given by

$$H_i(z) = q_{0,i} + \dots + q_{(N-1),i} z^{-(N-1)} \quad \dots(7)$$

Similarly, the eigenfilters corresponding to all the eigen-vectors of the ACM (all the columns of V) are obtained. The frequency responses of the eigenfilters, $H_i(e^{j\omega})$ are then obtained for all the eigenfilters.

Then, the corresponding eigenvalues are plotted against the main lobe center frequencies of the frequency responses of the obtained eigenfilters. This produces a

pseudo-spectrum (eigen spectrum), which gives an idea about the significant spectral content and the variations of spectral content of the signal in concern. The accuracy of this pseudo-spectrum increases with the filter order or the ACM size (N).

In order to find the optimal order of the ACM (N_{opt}), the order of the ACM is increased from the starting value of 10 by incremental step sizes of 10. For each and every ACM size the number of peaks in the pseudo-spectrum and the frequencies corresponding to those peaks are recorded. When the ACM size is increased this way, the number of peaks in the estimated pseudo-spectrum will also increase and the frequencies corresponding to those peaks will vary. Hence, the ACM order is increased until the number of peaks and the frequencies corresponding to those peaks becomes steady (number of peaks and their frequencies do not vary when the ACM size is further increased).

The condition to select N_{opt} is as follows. If the number of peaks and their corresponding frequencies remain at a steady state for 10 steps starting from a certain value where each step is an increment of 10, the starting model order is selected as N_{opt} .

This method of determining N_{opt} is based on the reasoning that the ACM order, which gives the minimum sufficient resolution for the estimated spectrum to accurately capture all the peaks and their corresponding frequencies is the optimum ACM order. This means that N_{opt} is the model order required for the generated pseudo-spectrum to accurately capture all the dominant spectral components in the actual power spectrum of the signal in concern.

Consider the non-stationary signal for which the spectrogram (variation of frequency with time) is depicted in Figure 2 (a) (labelled hereafter as- synthetic signal no. 1). This signal was generated synthetically, using sinusoidal signals of frequencies 100, 500, 1000, 2000, 3000 and 4000 Hz at different time windows.

Figure 2 (b) illustrates how the number of peaks and the frequencies corresponding to those peaks vary with the ACM order. It is observed that when N is greater than 230, the number of peaks does not vary even though the ACM order increases. However, the frequencies corresponding to the peaks show a slight fluctuation until 480. After 480, for 10 consecutive increments of the ACM order, the frequencies of the peaks remain exactly the same. Therefore, a value slightly greater than 480 is chosen to be the N_{opt} . Here, the chosen value for N_{opt} is 490 (480 + 10). The selected value does not necessarily have to be 490 and any value, which is greater than the experimentally

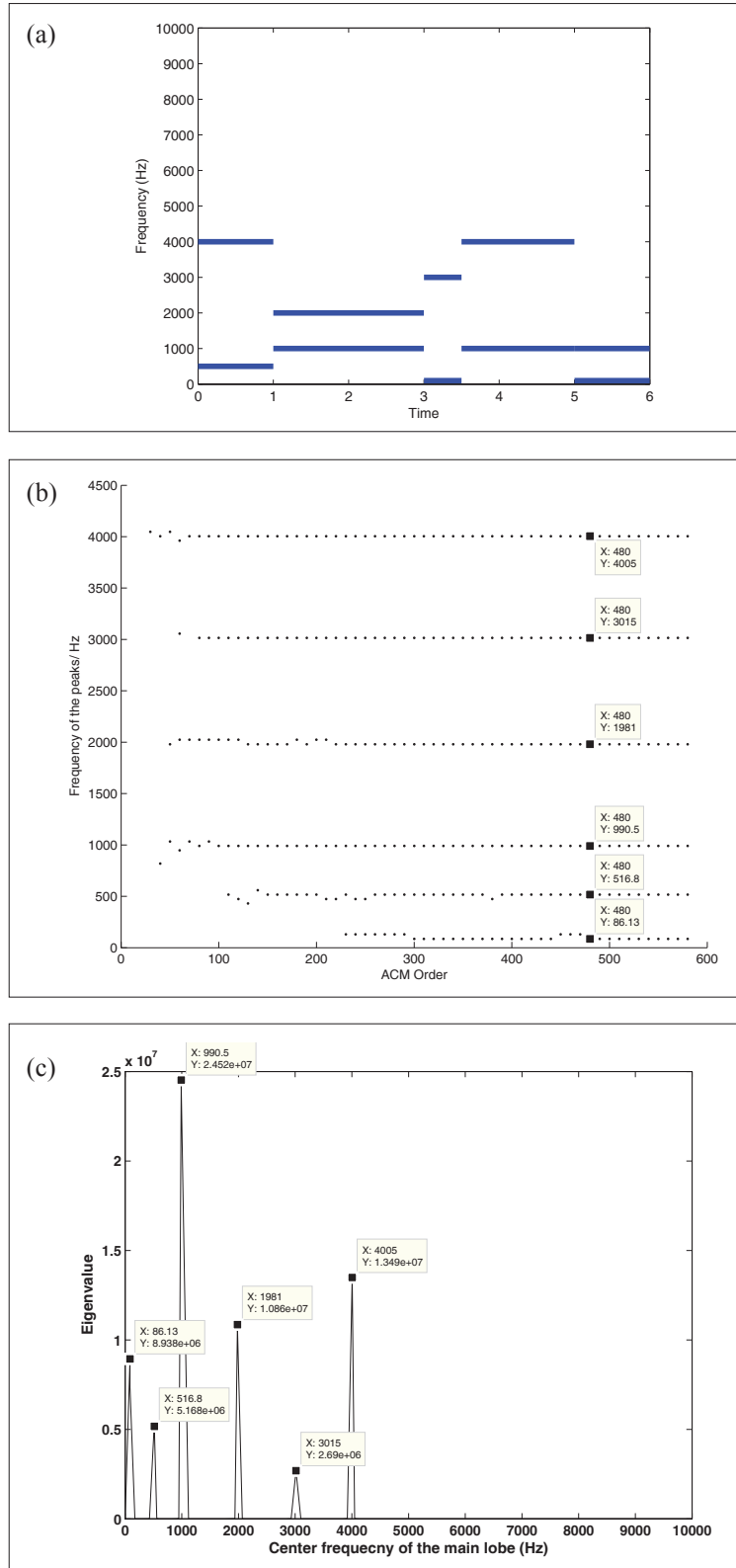


Figure 2: (a) Spectrogram; (b) variation of the frequencies of the peaks of the estimated power spectrum with the ACM order and (c) estimated pseudo-spectrum of the synthetic signal no. 1

selected lower bound of 480 would be suitable. However, to avoid wasting computational power unnecessarily, it is better to select a value which is slightly higher.

Figure 2(c) depicts the estimated pseudo-spectrum with the chosen N_{opt} value. It can be seen that the peak locations of this pseudo-spectrum are approximately equal to the frequencies of the constituent sinusoids of the synthetic signal no. 1 shown in Figure 2 (a).

The most significant advantage of this proposed pseudo-spectrum based method is its capability to extract all the distinct spectral components of a signal. This is achieved by selecting the subspaces that correspond to the peaks of the pseudo-spectrum. Further, it should also be noted that not all these subspaces which correspond to the peaks of the pseudo-spectrum are principal components of the eigen space. However, the proposed method is capable of extracting these principal components as well as non-principal subspaces corresponding to all distinct spectral components unlike the widely used PCA based methods. The overall subspace formed by the selected eigenfilters/vectors is essentially a comb that captures any subspace, which carries unique information irrespective of whether it is a principal or non-principal component.

In the traditional PCA based dimensionality reduction methods, the selection of principal K subspaces (the K subspaces that has the highest eigenvalues) out of the total of N subspaces is performed. But this does not guarantee the capturing of all the distinct spectral components. This is due to the fact that two (or more) adjacent subspaces (i.e. eigenfilters with adjacent eigenvalues in the eigenvalue spread) may contain center frequencies, which are in close proximity to each other, hence containing no new spectral information about the signal. In other words, a batch of adjacent eigenfilters may all be part of one spectral lobe of the signal. The proposed method is capable of latching only on to the required subspaces located at distinct points of the spectrum given that the resolution is adjusted to N_{opt} in the training phase.

For example, Figure 3 shows the plot of the center frequencies of the top 6 eigenfilters and their corresponding eigenvalues. It can be seen that although six principal eigenfilters are selected, they would capture only three distinct frequencies. In the proposed non-PCA based approach, the selected 6 filters correspond to the six distinct frequency components, which essentially describe the spectral content of the signal more completely, thus eliminating the weakness of the PCA based traditional approach.

Another important aspect of this proposed method is its capability to extract constituent stationary components from a non-stationary signal. This is highly useful in many applications such as elephant infrasound call detection (which will be explained later in this paper), where detection and classification is based on the hidden stationary signature features.

In order to highlight the said capability of the proposed method to extract stationary components, the synthetic signal number 1 shown in Figure 2 (a) is considered again. Here, the whole signal is filtered using the eigenfilters corresponding to the peaks of the estimated pseudo-spectrum.

The frequency responses of the selected filters are shown in Figure 4 (a) and the outputs of those filters are shown in Figure 4 (b). It can be clearly seen that the outputs of the selected filters provide an indication of how the total power of the signal is distributed among different frequencies at any given time instant. The popularly used method for this purpose is the short-time fourier transform (STFT). One weakness of the STFT based method is the high time-frequency resolution tradeoff (Quatieri, 2002; Marks II, 2008; Nam *et al.*, 2010). This tradeoff makes the STFT based methods incapable of achieving high resolution simultaneously in both time and frequency domains. The reason for this weakness is the fact that a series of segments over a sliding window are used to obtain the STFT of a signal. The second weakness of the STFT based method is that it only indicates the variation of the spectral content with time and it is not capable of extracting the different spectral components.

This proposed method, as can be seen from Figure 4(b), is capable of significantly reducing both of these weaknesses that exist in the STFT method by identifying the time windows where a particular frequency is present. This is due to the fact that the proposed method incorporates a stationary filtering scheme as opposed to the sliding window method used in STFT. The filters used in this scheme are chosen to be the filters corresponding to the peaks of the pseudo-spectrum. Hence, these selected filters cover all the significant spectral components of the signal. The outputs of these filters give the power of all the selected spectral components at any given time instant. Since the filtering process is stationary, the time frequency resolution effect does not exist in this proposed method unlike the STFT based method. Further, this method is also capable of extracting the distinct spectral components as opposed to merely representing their

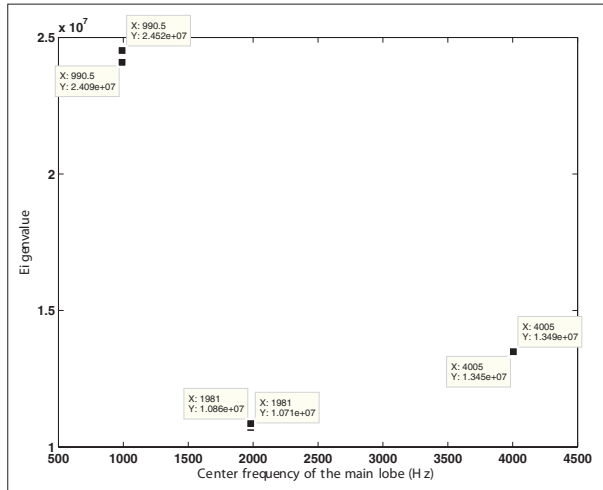


Figure 3: Spectral locations of the six most dominant eigenfilters with respect to the main lobe center frequency

power as in the STFT based spectrogram method. Thus, the proposed method is capable of significantly reducing two major drawbacks of the commonly used STFT method.

An important advantage of the proposed method is its capability to perform well even with signals, which have spectral components located very near each other. In this scenario, N_{opt} will be auto-adjusted in the training phase to fit the resolution requirement to identify all the closely located distinct spectral components. Hence, this method can be easily adapted to many types of applications which have signals with different types of spectral properties (biological, acoustic, underwater, etc.).

Another important aspect of the proposed algorithm is its behaviour with the signals, which has spectral changes that occur gradually instead of sudden changes. For example, consider the synthetic signal no. 2, which is illustrated in Figure 5 (a).

Figure 5(a) shows the synthetic signal no. 2 which consists of changes in spectral components that occurs linearly and logarithmically. It can be seen from Figure 5(b) that the ACM order, which captures all the spectral components and their corresponding frequencies is 500. After 500, the frequencies of the peaks do not vary for 10 consecutive increments. Hence 500 is the optimum model order according to the proposed method. It further shows that some peaks occur and then disappear when the ACM order is in the range of 150 – 370. They occur due to the linear and logarithmic frequency transients. The proposed algorithm is capable of finding a suitably higher model order that latches on only to the stationary regions of the signal.

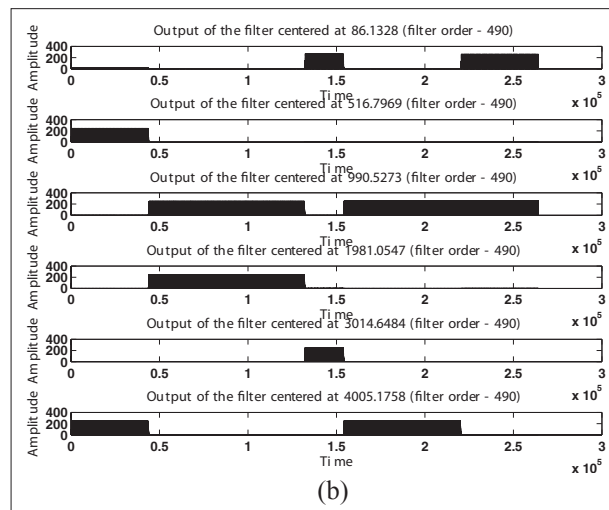
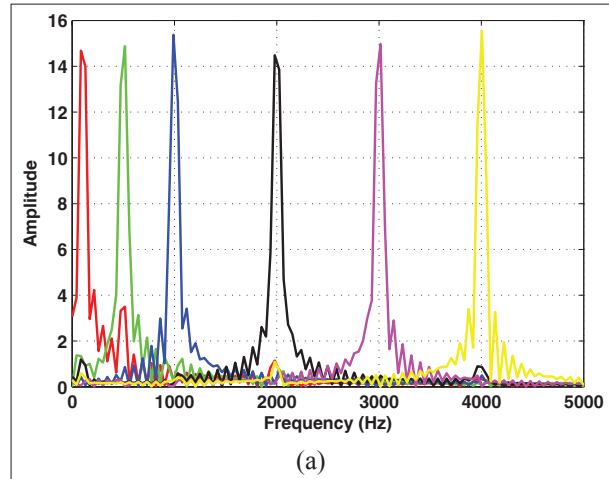
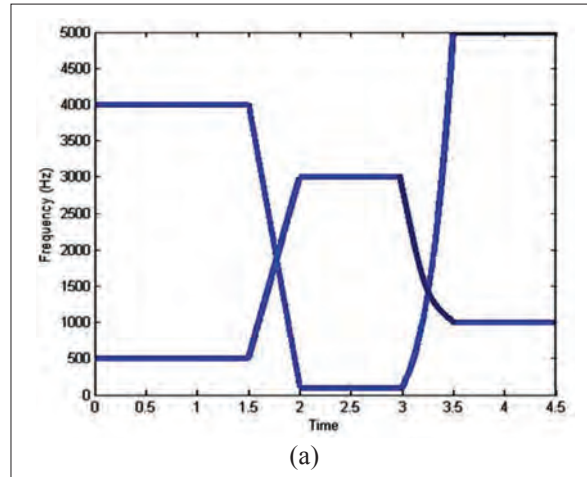


Figure 4: (a) Filter responses of the selected filters for synthetic signal no. 1 and (b) outputs of the selected filters for synthetic signal no. 1

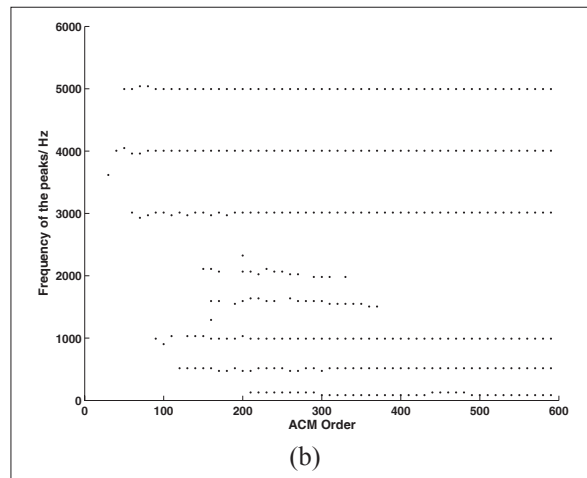
Figure 5(c) depicts the estimated spectrum of the synthetic signal no. 2 with the estimated optimum model order. It can be seen from Figure 5 (c) that all the stationary spectral components are captured. Essentially, the transient components raise the floor of the spectrum around the peaks, enabling the identification of such a transient non-stationarity through the floor level.

AUTOMATED CLASSIFICATION OF NON-STATIONARY SIGNALS USING THE PROPOSED APPROACH

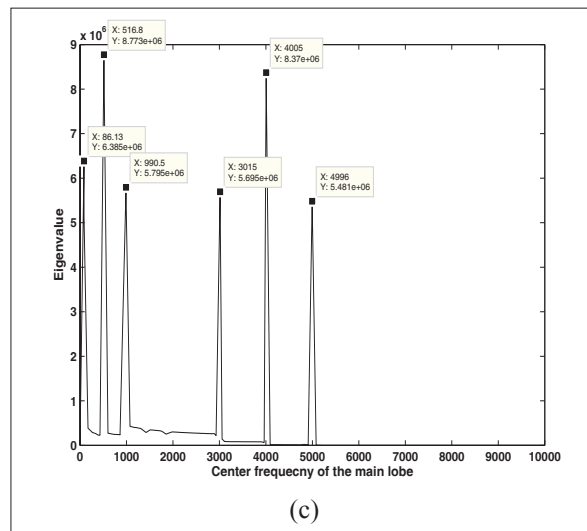
In the previous section, various aspects of the proposed algorithm and how it reduces certain weaknesses of the currently used STFT based methods was discussed. This section describes how the proposed method can be



(a)



(b)



(c)

Figure 5: (a) Spectrogram; (b) variation of the frequencies of the peaks of the estimated power spectrum with the ACM order and (c) estimated pseudo-spectrum of the synthetic signal no. 2

adapted for automated classification of non-stationary signals.

The database of signals described in Table 1 and illustrated in Figure 1 have been used to show how automated classification can be performed using the proposed methods.

Initially, one signal from each class (each speaker) is used to identify the optimum ACM order that can be used to model the signals in concern. The obtained results are shown in Table 2.

Table 2: Minimum ACM orders for each class of signals

Speaker	Minimum ACM order
EVP1	1170
EVP2	1930
EVP3	490
EVP4	1130
EVP5	470

The reason for the difference of this model could be that the vocal system highly differs from person to person. Since the highest required model order is 1930, the optimum model order was selected to be 2000, which is slightly greater than the highest required model order. This selection would ensure that the selected model order accurately models all the five types of signals.

Initially, each signal of each class is modelled using the determined model order (2000) and their pseudo-spectra are estimated using the previously proposed method. The estimated spectra for one signal from each class are shown in Figure 6 (in blue colour).

It can be seen from Figure 6 (in blue colour) that the estimated frequency spectra contain peaks at various frequencies, which reveal all the frequencies that a particular signal contains. Therefore these frequencies that correspond to the peaks can be used for the identification of the class that a particular signal belongs.

Frequency responses of all the signals of each class are estimated in the training phase for the identification of the class that a particular signal belongs. Then, all the frequencies that contain powers (eigenvalues), which are smaller than $1/10^{\text{th}}$ of the maximum power value of that spectrum are neglected. After neglecting those peaks, the pseudo-spectrum look as indicated by the red coloured plots in Figure 6 and the word pseudo-spectrum will refer to these plots henceforth. The purpose of doing this is to neglect the eigenfilters that has negligibly smaller eigenvalues, which may correspond to noisy elements.

Then the frequencies corresponding to the peaks of the spectrum are recorded for each signal. The recorded values for signals of each class are used to find out how many signals of a particular class have peaks at a particular frequency. In other words, a plot with the y-axis indicating the normalised (divided by total number of signals -50) number of times a particular spectral content corresponding to an eigenfilter is present in each class is created. Figure 7 depicts the generated plots for each class. For example, the green circled marker located at 473.7, 0.5 indicates that 25 (50×0.5) signals of the EVP2 class has peaks at 473.7 Hz of the pseudo-spectrum.

These plotted values for each class can be used as primary keys, which distinguish each class of signals (different speakers in this case). Those primary keys can be expressed using a generalised expression as in,

$$P_{EVP:k} = [n_{k,f_1} \quad n_{k,f_2} \quad \dots \quad n_{k,f_z}] \quad \dots(8)$$

where n_{k,f_i} is the spectral presence of k^{th} class at the frequency f_i .

These primary keys can be used as means of distinguishing the class that a particular signal belongs. This is due to the fact that the elements of any primary key $P_{EVP:k}$ can be considered as a set of sensors placed at various points (f_i for all i) of the power spectrum. Further, n_{k,f_i} can be considered as a measure of the reliability of the sensor located at f_i since it indicates the probability of a signal of class k having a peak at the frequency f_i .

After the construction of primary keys for the classes, the next phase is the use of constructed keys for the classification of signals when the class is unknown.

When such unknown signal is considered, initially the spectrum of that signal is estimated using the approach proposed previously in this paper. Here, the optimum model order (2000) selected in the training phase is used for the estimation. Then a vector U ,

$$U = [b_{f_1} \quad b_{f_2} \quad \dots \quad \dots \quad b_{f_N}] \quad \dots(9)$$

containing binary values (1s and 0s) to indicate the presence and non-presence of peaks at f_i (for all i) of the spectrum is created.

As indicated in equation (9), the vector U comprises binary elements b_{f_i} , which indicates the presence or non-presence of peaks at the frequencies f_i of its estimated spectrum.

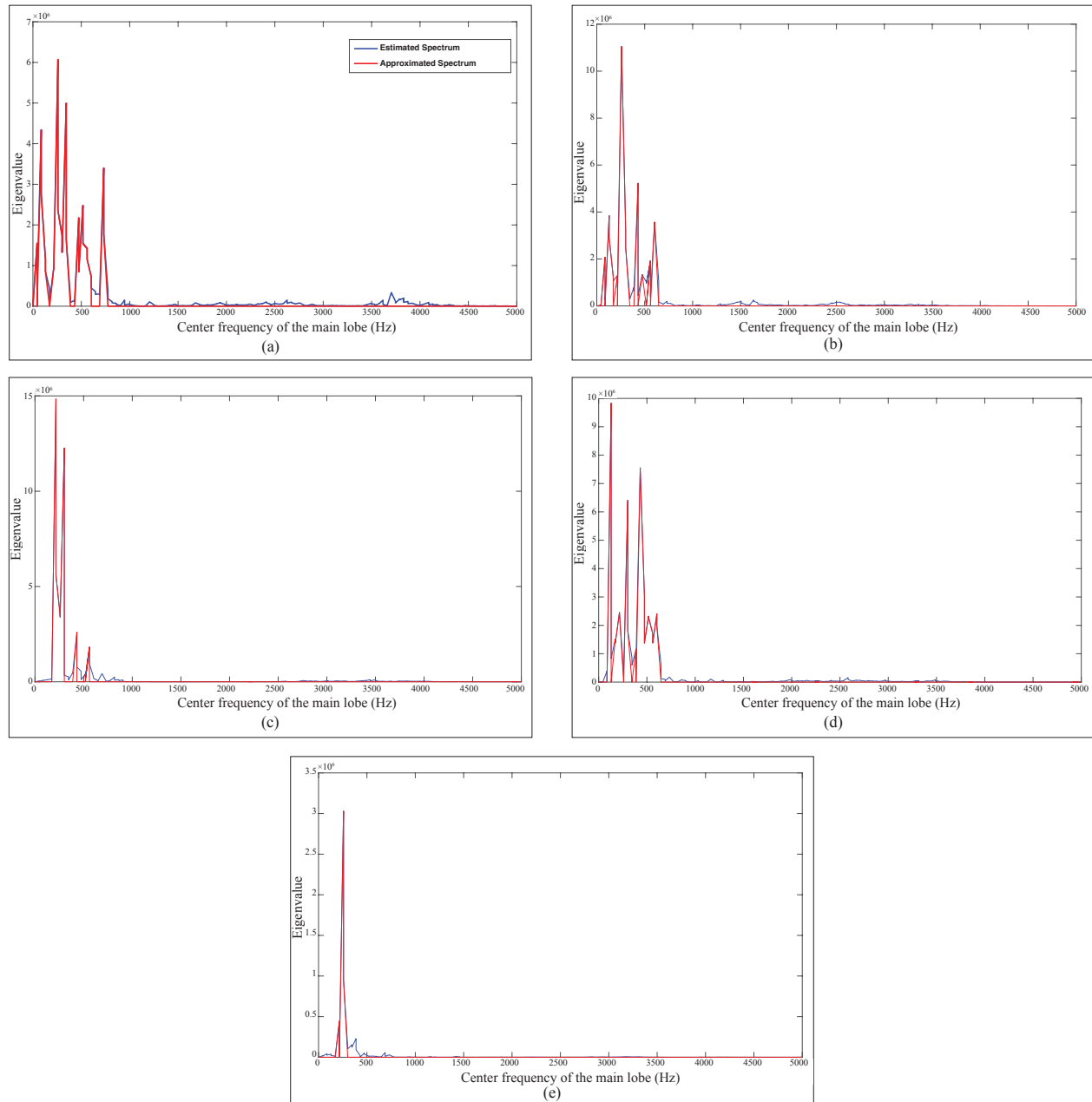


Figure 6: Estimated power spectra for a signal of (a) EVP1 class; (b) EVP2 class; (c) EVP3 class; (d) EVP4 class and (e) EVP5 class

Then, the elements of vector U can be thought as the binary outputs of the sensors that are placed at various points of the spectrum in the form of five primary keys $P_{EVP.k}$. Since the primary key $P_{EVP.k}$ consists of the best set of sensors to identify the signals of class K , the correlation between U and the five primary keys can be used for classification of the unknown signal. Therefore, the dot product can be used to find out the primary key, which is most correlated with the vector U . The problem of utilising the dot product is that although it would take

the presence of a peak at a particular frequency into consideration, the non-presence, which is indicated by a '0' element in the vector U would not be considered (because the dot product will not be affected by zero elements). Therefore, to rectify this problem 0.5 is subtracted from all the elements of primary keys as well as the elements of vector U . This would make both the presence and non-presence of a peak at a particular frequency of the spectrum of the unknown signal useful for the determination of its class.

Therefore the five scalars are generated using the expression,

$$S_k = \frac{(U-0.5) \cdot (P_{EVP:k} - 0.5)}{|(U-0.5)| \times |(P_{EVP:k} - 0.5)|} \dots(10)$$

and the signal is classified to the class corresponding to the largest such scalar. Here, the scalars will be in the range of 0 to 1 since the normalised vectors are used to obtain the scalars.

Figure 8 depicts the values of the five scalars for the 50 signals of each of the five classes in separate plots.

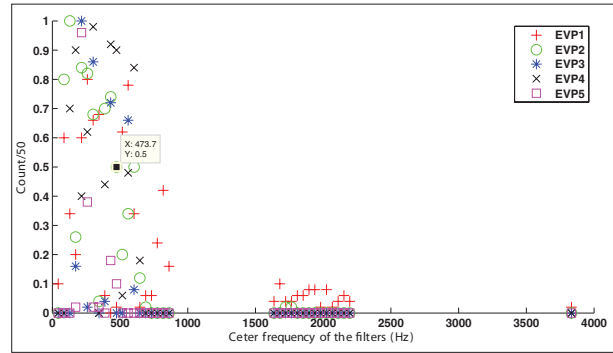


Figure 7: Spectral presence of each class

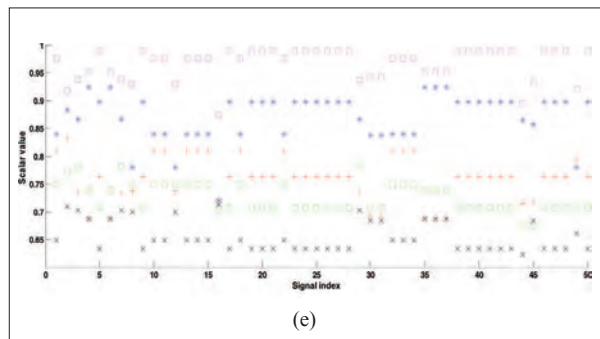
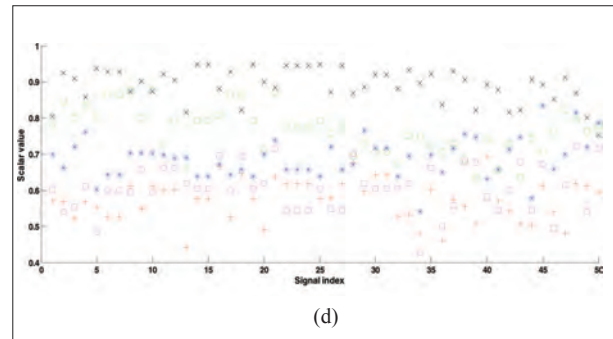
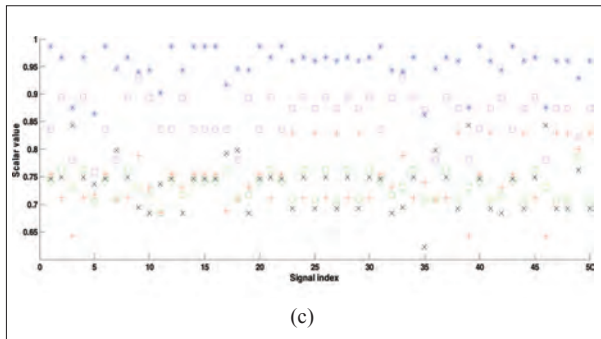
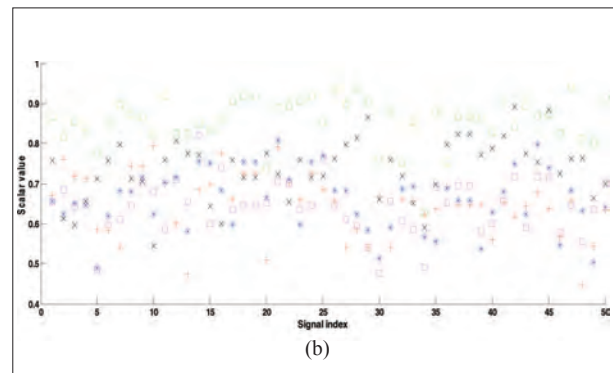
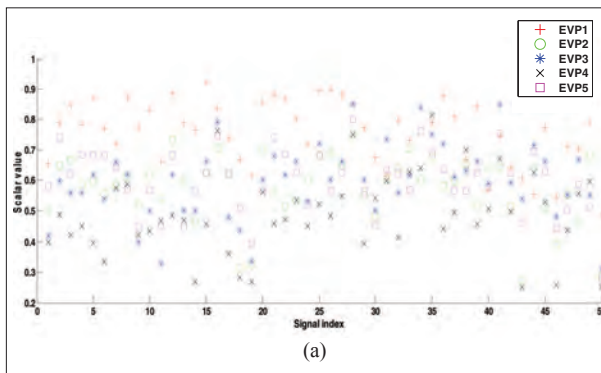


Figure 8: Values of the five scalars S_k for the 50 signals of (a) EVP1 class; (b) EVP2 class; (c) EVP3 class; (d) EVP4 class and (e) EVP5 class

It can be seen that S_k is the largest scalar for most of the signals of the k^{th} class. This proves the accuracy and feasibility of the proposed approach for classification. For example, for the signals of the EVP1 class, the S_1 is the largest scalar as depicted by Figure 8 (a).

Table 3 shows the classification accuracies of each class for the proposed approach. It can be seen that the proposed approach is capable of performing a robust classification of highly non-stationary signals with high classification accuracy.

Table 3: Classification accuracy of the proposed method for each class

Class	Total signals	True positives	Percentage accuracy
EVP1	50	43	86 %
EVP2	50	46	92 %
EVP3	50	49	98 %
EVP4	50	46	92 %
EVP5	50	50	100 %

PERFORMANCE COMPARISON BETWEEN THE PROPOSED METHOD AND SOME EXISTING METHODS

This section highlights the importance and the superior performance of the proposed method compared to the existing techniques.

Parametric techniques, which are widely used for signal classification require the selection of one or more distinct parameters that are capable of distinguishing signals according to their respective classes. The database of signals given in Table 1 is used to identify such distinct parameters for classification of speakers.

The box plots shown in Figure 9 illustrates the distributions of eight dominant parameters (first 4 formant frequencies and their corresponding bandwidths) extracted from the signals of the five speakers. It can be seen from these box plots that all the parameters are overlapped with different classes of signals and hence these parameters cannot be utilised for classification of these signals accurately. Therefore, it can be concluded that the parametric methods tend to fail in classification of highly non-stationary signals.

In our previous study (Nettasinghe *et al.*, 2013; Ratnayake *et al.*, 2013), a subspace technique has been proposed for signal classification in cases where the

signals are highly correlated and noisy. This technique is based on the minimisation of cross-correlation between signals of different classes by selecting the eigenfilters, which correspond to the uncorrelated information. This technique was proven to give highly accurate results for stationary signals. The results obtained by this technique for the classification of non-stationary signals (speech signals) are illustrated in Figure 10.

Figure 10 shows that the majority of the scalars obtained by vector projection through subspace filtering are overlapped and possess similar values among classes. The results illustrated in these figures are numerically presented in Table 4. According to the table it can be clearly seen that the accuracy of the classifier for non-stationary signals is very low compared to the results obtained from the method proposed in this paper, which is more appropriate for non-stationary signals.

Table 4: Classification accuracy using the correlation minimisation method

Class	Total signals	True positives	Percentage accuracy
EVP1	50	6	12 %
EVP2	50	31	62 %
EVP3	50	50	100 %
EVP4	50	15	30 %
EVP5	50	32	64 %

Elephant infrasound call detection using the proposed method

In this section, the applicability of the method proposed in this paper to extract hidden signature features from noisy, non-stationary signals is described. The elephant infrasound call detection is used as the primary application for this purpose.

Previous research (Wijayakulasooriya, 2011; Dissanayake *et al.*, 2012) has proven that elephants use low frequency elephant calls for long distance communication. These infrasound calls can be utilised for elephant detection due to a comb-like pattern observed in the time frequency distribution of these signals. This comb-like pattern, which is illustrated in Figure 11 is a signature of elephant calls.

Figure 11 shows that the signature feature of the elephant call is masked by noise. Thus, it is crucial to extract the elephant call from surrounding noise such as other elephant sounds, wind and vehicular noise. The comb-like pattern observed at specific frequencies of these elephant calls makes it ideal for applying the proposed method.

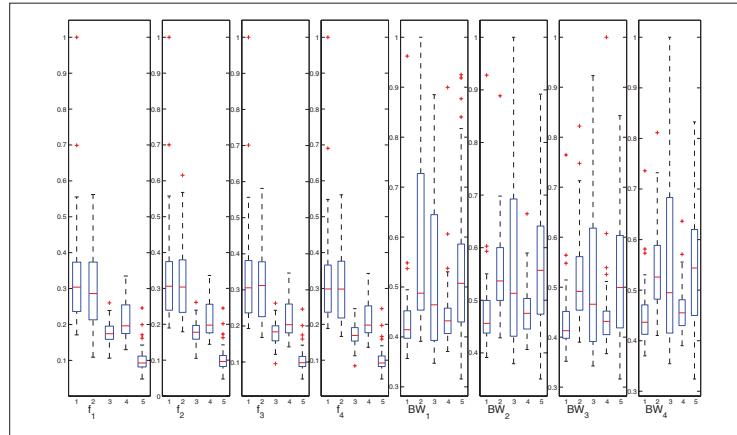


Figure 9: The box plots of eight dominant parameters extracted from 50 signals of the 5 speakers

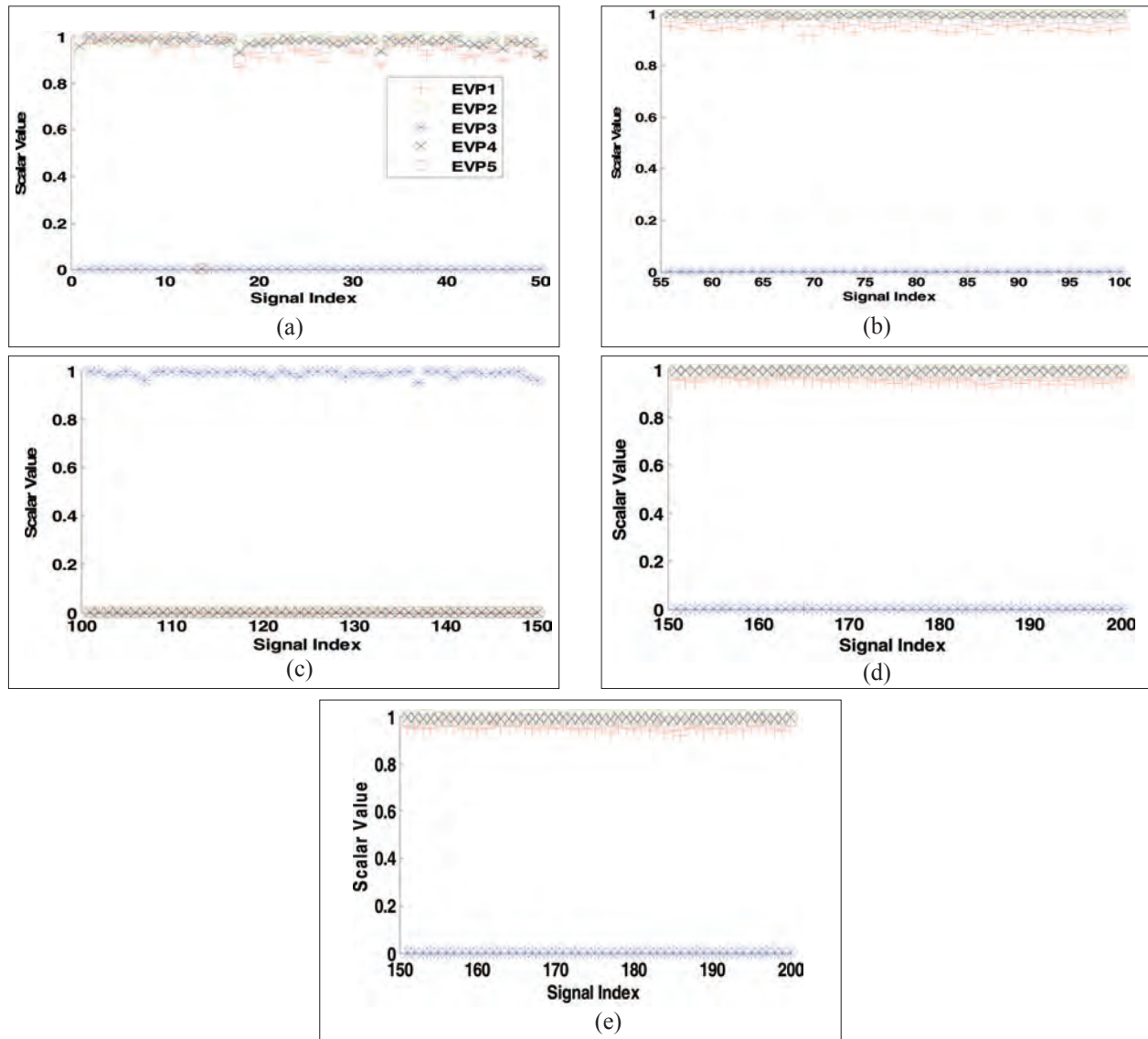


Figure 10: Scalars for the 50 signals of (a) EVP1 class; (b) EVP2 class; (c) EVP3 class; (d) EVP4 class and (e) EVP5 class

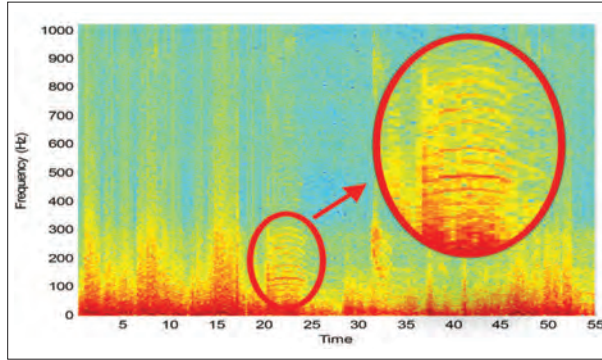


Figure 11: Spectrogram of an elephant infrasound call

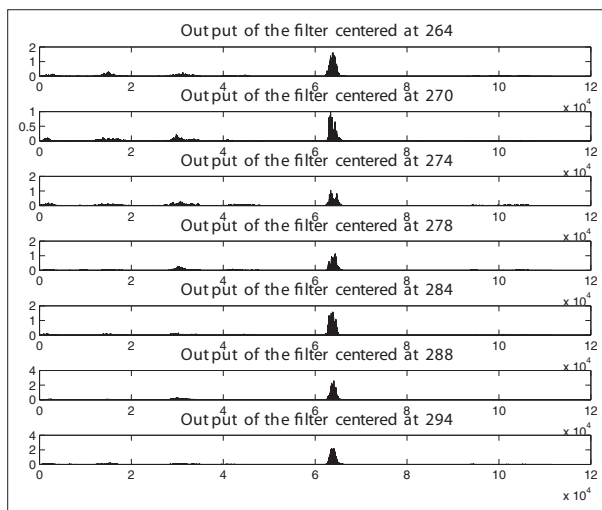


Figure 12: The filter outputs of the 7 chosen eigenfilters

By applying the proposed method for an acoustic signal that contains an elephant call, the eigenfilters which can extract the signature frequencies are identified. For the signal illustrated in Figure 11, seven such eigenfilters whose center frequencies are matched with the signature frequencies of the elephant call are found. By filtering the whole signal using these seven eigenfilters separately, the existence of the corresponding center frequency in the signal can be checked. Figure 12 illustrates the seven outputs of the eigenfilters and from this we can clearly see that in all seven outputs a high output power is filtered out in the region where an elephant call exists (highlighted in red).

From Figure 12 it is apparent that the signature frequencies of the elephant call is extracted from these seven selected eigenfilters. A binary array is constructed utilising the outputs of the selected eigenfilters. This is achieved by setting '1' when the output signal power from each eigenfilter is greater than the threshold value,

and '0' when it does not. This binary array is considered as the outputs of seven sensors and thus fused together in order to make a decision on the presence of an elephant.

DISCUSSION

When the number of sensors which produce '1' as the output increases, the probability of the decision being a true-positive increases, i.e., the accuracy of the decision increases. Hence, the number of sensors which produce the same output as the decision made ('0' or '1', or in other words presence or the non-presence of an elephant), can be used as a measure of confidence of the decision made.

CONCLUSION

This paper proposes a novel non-PCA based approach for automated classification of highly non-stationary signals. This approach generates an information rich comb-like subspace structure utilising an eigen analysis based pseudo-spectral estimation technique. The optimum model order for the spectral estimation is determined to be the lowest order, which captures all the distinct and dominant spectral information.

It is shown that the proposed method is capable of latching only on to the required subspaces located at distinct points of the spectrum, while the commonly used PCA based methods does not guarantee the capturing of all the distinct spectral information. It is further shown that the proposed approach has the capability of reducing two major drawbacks that exist in the STFT based spectral estimation methods, namely, the time-frequency resolution tradeoff and the inability to extract stationary signal components separately.

A novel classification approach was also presented, which is based on the proposed spectral estimation method. The proposed method was tested using a highly non-stationary speech signal database and the results were compared with the existing methods such as parametric methods and correlation minimisation methods to highlight the better performance.

REFERENCES

1. Avendaño-Valencia L.D., Acosta-Medina C.D. & Castellanos-Domínguez G. (2011). Time-frequency based feature extraction for non-stationary signal classification. *Applied Biomedical Engineering* (eds. G. Gargiulo & A. McEwan), InTech, Rijeka, Croatia. DOI: <http://dx.doi.org/10.5772/19101>
2. Chendeb M. (2010). Classification of non stationary signals using multiscale decomposition. *Journal of Biomedical*

- Science and Engineering* 3: 193 – 199.
DOI: <http://dx.doi.org/10.4236/jbise.2010.32025>
3. Dissanayake C.M., Kotagiri R., Halgamuge M.N., Moran B. & Farrell P. (2012). Propagation constraints in elephant localization using an acoustic sensor network. *Proceedings of the IEEE 6th International Conference on Information and Automation for Sustainability (ICIAfS)*, 27 – 29 September, Beijing, China, pp. 101 – 105.
DOI: <http://dx.doi.org/10.1109/iciafs.2012.6419889>
 4. Fagundes R.D.R., Martins A.A.C., Comparsi de Castro F. & Felippetto de Castro M.C. (2002). Automatic gender identification by speech signal using eigenfiltering based on Hebbian learning. *Proceedings of the VII Brazilian Symposium on Neural Networks*, 11 – 14 November, Pernambuco, Brazil, pp. 212 – 216.
DOI: <http://dx.doi.org/10.1109/SBRN.2002.1181476>
 5. Hailemariam S. & Prahallad K. (2007). Extraction of linguistic information with the AID of acoustic data to build speech systems. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 15 – 20 April, Honolulu, Hawaii, pp. IV-717 – IV-720.
DOI: <http://dx.doi.org/10.1109/ICASSP.2007.367013>
 6. Josso N.F., Zhang J.J., Papandreou-Suppappola A., Ioana C. & Duman T.M. (2011). Nonstationary system analysis methods for underwater acoustic communications. *EURASIP Journal on Advanced Signal Processing* 9: 1 – 14.
DOI: <http://dx.doi.org/10.1155/2011/807472>
 7. Kandpal M., Kakar V.K. & Verma G. (2013). Classification of ground vehicles using acoustic signal processing and neural network classifier. *Proceedings of the International Conference on Signal Processing and Communication (ICSC)*, 12 – 14 December, Noida, India, pp. 512 – 518.
DOI: <http://dx.doi.org/10.1109/ICSPCom.2013.6719846>
 8. Lee C.H. & Siniscalchi S.M. (2013). An information-extraction approach to speech processing: analysis, detection, verification, and recognition. *Proceedings of the IEEE* 101: 1089 – 1115.
DOI: <http://dx.doi.org/10.1109/JPROC.2013.2238591>
 9. Lee K.Y. & Lee J. (2001). Recognition of noisy speech by a nonstationary AR HMM with gain adaptation under unknown noise. *IEEE Transactions on Speech Audio Processing* 9: 741 – 746.
DOI: <http://dx.doi.org/10.1109/89.952492>
 10. Marks II R.J. (2008). *Handbook of Fourier Analysis and its Applications*. Oxford University Press, Oxford, UK.
 11. Meng X., Liu W. & Ding E. (2011). The research of acoustic emission signal classification. *Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, 14 – 16 October, Dalian, China, pp. 41 – 44.
DOI: <http://dx.doi.org/10.1109/iihmsp.2011.101>
 12. Nam J., Mysore G.J., Ganseman J., Lee K. & Abel J.S. (2010). A super-resolution spectrogram using coupled PLCA. *11th Annual Conference of the International Speech Communication Association 2010 (INTERSPEECH 2010)*, 26 – 30 September, Makuhari, Chiba, Japan, pp. 1696 – 1699.
 13. Nayak J. & Bhat P.S. (2003). Identification of voice disorders using speech samples. *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, volume 3, 15 – 17 October, pp. 951 – 953.
DOI: <http://dx.doi.org/10.1109/TENCON.2003.1273387>
 14. Nettasinghe D.B.W., Ratnayake T.A., Pollwaththage N.N., Godaliyadda G.M.R.I., Wijayakulasooriya J.V. & Ekanayake M.P.B. (2013). A robust subspace classification method for highly correlated acoustic signals. *Fifth International Conference on Computational Intelligence, Modelling and Simulation (CIMSIm)*. 24 – 25 September, Seoul, South Korea, pp. 225 – 230.
DOI: <http://dx.doi.org/10.1109/CIMSIm.2013.43>
 15. Padmavathi G., Shanmugapriya D. & Kalaivani M. (2010). Acoustic signal based feature extraction for vehicular classification. *3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, 20 – 22 August, Chengdu, China, pp. V2-11 – V2-14.
DOI: <http://dx.doi.org/10.1109/ICACTE.2010.5579804>
 16. Pollwaththage N.N., Nettasinghe D.B.W., Ratnayake T.A., Godaliyadda G.M.R.I., Ekanayake M.P.B. & Wijayakulasooriya J.V. (2013). Extraction and classification of non-stationary acoustic signals via dynamic subspace filtering. *IEEE 8th International Conference on Industrial and Information Systems (ICIIS)*. 17 – 20 December, Peradeniya, Sri Lanka, pp. 415 – 420.
DOI: <http://dx.doi.org/10.1109/ICIInfS.2013.6732020>
 17. Quatieri T.F. (2002). *Discrete-Time Speech Signal Processing: Principles and Practice*. Pearson Education, India.
 18. Ratnayake T.A., Pollwaththage N.N., Nettasinghe D.B.W., Godaliyadda G.M.R.I., Wijayakulasooriya J.V. & Ekanayake M.P.B. (2013). Material based acoustic signal classification - a subspace-based approach. *TENCON 2013 – 2013 IEEE Region 10 Conference (31194)*, 22 – 25 October, Xi'an, China, pp. 1 – 4.
DOI: <http://dx.doi.org/10.1109/TENCON.2013.6718974>
 19. Subasi A. & Erçelebi E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine* 78: 87 – 99.
DOI: <http://dx.doi.org/10.1016/j.cmpb.2004.10.009>
 20. Sukittanon S., Atlas L.E., Pitton J.W. & McLaughlin J. (2003). Non-stationary signal classification using joint frequency analysis. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003 (ICASSP '03)*, volume VI, 6 – 10 April, pp. 453 – 456.
DOI: <http://dx.doi.org/10.1109/ICASSP.2003.1201716>
 21. Tüske Z., Drepper F.R. & Schlüter R. (2012). Non-stationary signal processing and its application in speech recognition. *Workshop on Statistical and Perceptual Audition (SAPA)*, Portland, USA.
 22. Wijayakulasooriya J.V. (2011). Automatic recognition of elephant infrasound calls using formant analysis and hidden Markov model. *6th International Conference on Industrial and Information Systems (ICIIS)*, 16 – 19 August, Kandy, Sri Lanka, pp. 244 – 248.
DOI: <http://dx.doi.org/10.1109/ICIINFS.2011.6038074>