

# Classification of Acoustic Emissions Using Modified Matching Pursuit

**Samuel P. Ebenezer**

*Acoustic Technologies, Inc., 1620 Stapley Drive, Mesa, AZ 85204, USA  
Email: samuel.ebenezer@acoustictech.com*

**Antonia Papandreou-Suppappola**

*Department of Electrical Engineering, Arizona State University, P.O. Box 877206, Tempe, AZ 85287-7206, USA  
Email: papandreou@asu.edu*

**Seth B. Suppappola**

*Acoustic Technologies, Inc., 1620 Stapley Drive, Mesa, AZ 85204, USA  
Email: seth.suppappola@acoustictech.com*

*Received 31 May 2003; Revised 22 October 2003; Recommended for Publication by Kenneth Barner*

We propose methodologies to automatically classify time-varying warning signals from an acoustic monitoring system that indicate the potential catastrophic structural failures of reinforced concrete structures. Since missing even a single warning signal may prove costly, it is imperative to develop a classifier with high probability of correctly classifying the warning signals. Due to the time-varying nature of these signals, various time-frequency classifiers are considered. We propose a new time-frequency decomposition-based classifier using the modified matching pursuit algorithm for an actual acoustic monitoring system. We investigate the superior performance of the classifier and compare it with existing classifiers for various sets of acoustic emissions, including warning signals from real-world faulty structures. Furthermore, we study the performance of the new classifier under different test conditions.

**Keywords and phrases:** classification, detection, acoustic emissions, time-frequency representations, matching pursuit decomposition.

## 1. INTRODUCTION

### 1.1. Acoustic monitoring systems

Acoustic monitoring systems are used in many applications to determine the integrity of structures such as buildings and bridges. The purpose of these monitoring systems is to listen for any warning signals that may indicate future catastrophic effects. This premonition of future events can provide invaluable time to prevent catastrophic failures that may follow. Figure 1 shows a typical monitoring system of structures along with a warning event [1, 2, 3, 4]. For example, in real life situations, the warning event could be the breaking sound of metals in steel bridges, the abnormal knocking sound in aircraft and automobile engines, or the breaking sound of metals in prestressed concrete structures with reinforcing metals.

There are numerous reasons why it is desirable to ascertain the condition of a structure to determine if failure is imminent. The failure of a structure may result in the loss of its use which usually implies loss of revenue. In addition to

this, the replacement cost after the failure is generally much more than preventive maintenance costs. As a result of collateral expenses, the cost of structural failure may far exceed the cost of the structure itself. By monitoring for warning signals that indicate future failures, we can prevent the failure of the structure by repairing it before the failure occurs. The exact location of the distress can be determined by employing multiple sensors.

### 1.2. Automatic acoustic classifiers

Even though monitoring systems are intended to record some specific warning sounds, we cannot prevent the sensor from recording all other undesirable acoustic events. For example, in a monitoring system of prestressed concrete structure with reinforcing metals, even though we are looking for the breaking noise of reinforcing metals, the monitoring system will also record other undesirable sounds such as sounds from passing workers, birds, animals, machinery, automobiles, or rain. Hence, it is important to classify the warning signals from all these other signals. This classification can be

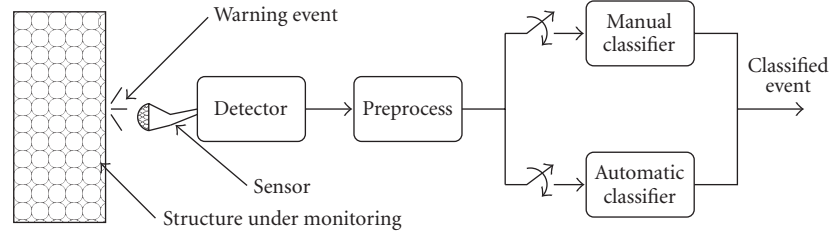


FIGURE 1: Typical acoustic monitoring system of structures.

done manually by listening to all the sounds that are recorded by the monitoring system and using manual signal analysis combined with human judgment to identify the warning signals. However, in practice, it is not uncommon for less than one percent of recorded emissions to be of diagnostic value. Hence, manual classification becomes a tedious process. This necessitates the use of an automatic acoustic classifier.

Automatic classifiers are used in diverse fields ranging from digital communications [5] to surveillance applications [6]. Due to the diverse applications of classification theory, many advancements have been made over the last couple of decades through extensive research. Although many classification theories have been developed, classification of time-varying signals whose frequency content varies with time is still a challenging task. Due to the time-varying nature of the signals recorded by an acoustic monitoring system of a concrete structure, the performance of some well-established one-dimensional (1D) time domain classification methods may not be satisfactory.

Time-frequency representations (TFRs) are two-dimensional (2D) transformations of time and frequency that are powerful tools for analyzing time-varying signals [7, 8, 9, 10]. This is because TFRs can provide the time-localized frequency information of a signal. As a result, TFRs or time-frequency-(TF) based methods have been used for classification of time-varying signals, and they have been shown to outperform the conventional 1D time- or frequency-based methods. Note, however, that this improved performance may not be sufficient.

In our application, an acoustic monitoring system can afford to have a small number of false positive classifications provided that it does not significantly increase the effort of the manual analysis of detected warning signals. However, the system demands that the probability of correct classification of warning signals should be close to one. Since the main objective of an acoustic monitoring system is the premonition of catastrophic failures, missing even a few warning signals may prove to be very costly.

In this paper, we develop an automatic classifier that classifies time-varying acoustic emissions from a monitoring system using TF decomposition-based techniques. In particular, the classifier is based on a modified version of the matching pursuit decomposition (MPD) algorithm [11, 12, 13] that decomposes a signal into basic components throughout the TF plane. As we will demonstrate, the new classifier achieves

higher performance when compared to other TF classifiers proposed in the literature.

This paper is organized as follows. In Section 2, we describe some existing TF-based classifiers. In Section 3, we develop the new classifier using a modified matching pursuit algorithm. The classification results of the various proposed classifiers for real data, together with comparisons to other methods, are discussed in Section 4.

## 2. CURRENT CLASSIFIERS

There exist many classification methods depending on the nature of the signals. In this section, we discuss some TF-based classification methods of acoustic events. The effectiveness of these classification methods is evaluated using acoustic signals from an actual acoustic monitoring system.<sup>1</sup> Figure 2 shows twelve TFRs, each representing the spectrogram of a learning signal from different acoustic event classes. As seen from the figure, only Class 1 and Class 3 acoustic events have high-frequency components, whereas all other classes consist largely of low-frequency signals. For example, in this acoustic monitoring system, low-frequency signals like machinery, human voice, or automobile noises are often recorded. For the example in Figure 2, the empirical data indicates that Class 1 is the warning signal which usually precedes a catastrophic failure, and thus must be classified accurately.

### 2.1. Time-domain classifier

For binary linear classification, if  $\mathbf{X}_1$  and  $\mathbf{X}_2$  are the characteristic deterministic feature vectors of two classes embedded in additive white Gaussian noise (AWGN), then the matched filter operation minimizes the probability of misclassification. Any heuristic approach to the choice of feature can be used, provided it leads to a minimum misclassification rate. A common feature that we can consider are the time samples of the signal itself. Hence, the template vectors  $\mathbf{X}_1$  and  $\mathbf{X}_2$  of Class 1 and Class 2, respectively, in AWGN can be defined as follows:

$$\bar{\mathbf{X}}_l = [\bar{x}_l[0] \ \bar{x}_l[1] \ \cdots \ \bar{x}_l[N-1]], \quad l = 1, 2, \quad (1)$$

<sup>1</sup>As the data used throughout this paper is proprietary, the different signal types are generically referred to by class number.

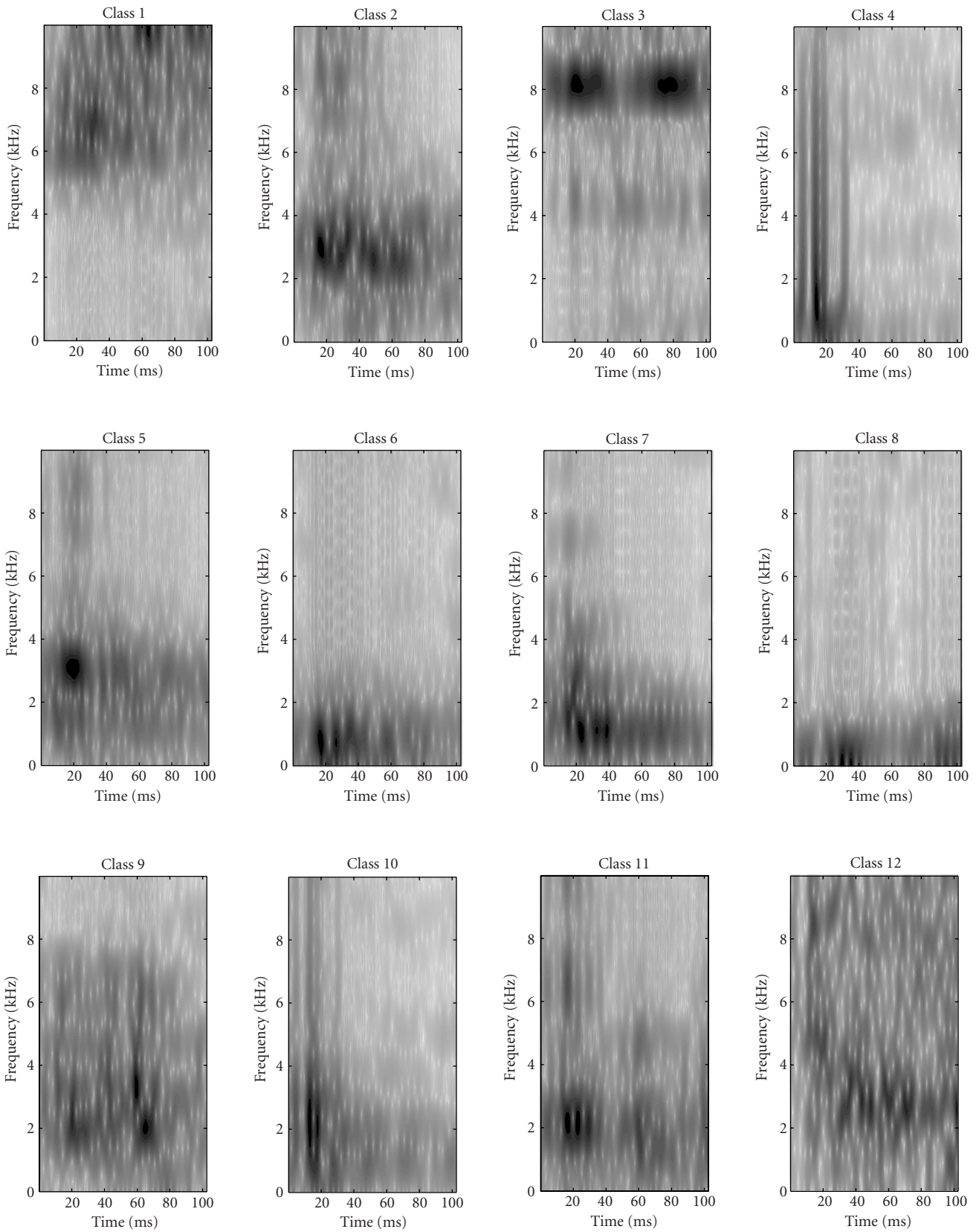


FIGURE 2: Sample spectrograms of acoustic signals from twelve different classes. The dark areas indicate greater energy.

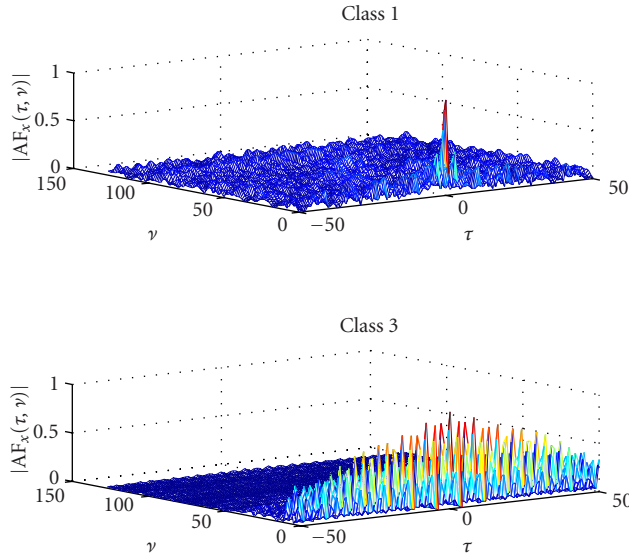


FIGURE 3: The narrowband auto ambiguity function of signals from Class 1 and Class 3.

where  $N$  is the number of samples, and

$$\bar{x}_l[n] = \frac{1}{P_l} \sum_{p=1}^{P_l} x_l^p[n], \quad n = 0, \dots, N-1, \quad (2)$$

denotes the sample mean of the number  $P_l$  of learning signals of Class  $l$ .

Based on the matched-filter decision rule [14], a test signal  $x[n]$  with feature vector  $\mathbf{X}$  is classified in Class 1 if

$$\langle \mathbf{X}, \bar{\mathbf{X}}_1 \rangle > \langle \mathbf{X}, \bar{\mathbf{X}}_2 \rangle, \quad (3)$$

where the inner product is given by  $\langle \mathbf{X}, \bar{\mathbf{X}}_l \rangle = \sum_n x[n] \bar{x}_l^*[n]$ . If the signals are time-varying and the initial phase of the signal is random, then the matched filter results in poor performance. In this case, other features that reveal the time-varying information of the signals may be used for classification.

## 2.2. Time-frequency-representation-based classifiers

In [15], TF points were used as a feature to classify time-varying signals. In order to obtain specific TF points to form a feature vector, we use the discrete time and discrete frequency version of a TFR. Consider, for example, the discrete version of the spectrogram TFR [8] given as  $\text{SPEC}_x(nT, kF) = \text{SPEC}_x[n, k]$ , where  $T$  is the sampling period,  $F$  is the frequency separation:

$$\text{SPEC}_x[n, k] = \left| \sum_{m=0}^{N-1} x[m] h^*[m-n] e^{-j2\pi km/N} \right|^2, \quad (4)$$

and  $h[n]$  is the analysis window of the spectrogram. Here, we

assume that each TF point in the TF plane is Gaussian<sup>2</sup> for large  $N$ . Using the TF points of the spectrogram as a feature vector, the authors in [15] extended the matched-filter decision strategy to the TF domain. Specifically, for binary classification, one can deduce that the test signal  $x[n]$  belongs to Class 1 if

$$\langle \text{SPEC}_x, \overline{\text{SPEC}}_1 \rangle > \langle \text{SPEC}_x, \overline{\text{SPEC}}_2 \rangle, \quad (5)$$

where, for a 2D function,  $\langle \text{SPEC}_x, \overline{\text{SPEC}}_l \rangle$  is the 2D inner product  $\sum_n \sum_k \text{SPEC}_x[n, k] \overline{\text{SPEC}}_l[n, k]$ . The average spectrogram of Class  $l$ ,  $\overline{\text{SPEC}}_l$ , is obtained by averaging the spectrogram of all the learning signals in Class  $l$ , that is,

$$\overline{\text{SPEC}}_l[n, k] = \frac{1}{P_l} \sum_{p=1}^{P_l} \text{SPEC}_{x_l^p}[n, k], \quad l = 1, 2, \quad (6)$$

where  $\text{SPEC}_{x_l^p}[n, k]$  is the spectrogram of the  $p$ th learning signal in Class  $l$  and  $P_l$  is the number of learning signals in Class  $l$  [15]. Thus, the inner product between the spectrogram of the test signal and the average spectrogram of the learning signals of Class 1 and Class 2 is used as the test statistic. The test signal will be assigned to the class whose 2D spectrogram correlation with the spectrogram of the test signal is maximum. For an  $L$ -class signal classification, the classification rule is

$$x[n] \in C_j \iff j = \arg \max_l \{ \langle \text{SPEC}_x, \overline{\text{SPEC}}_l \rangle \} \quad (7)$$

for  $l = 1, \dots, L$ , where  $C_j$  is the  $j$ th class.

Instead of using the TF points of a spectrogram as the feature vector, we can also use the TF points of other TFRs. Since most of the acoustic events recorded by the monitoring system are multicomponent, we decided not to use the well-known Wigner distribution (WD) because, for a multicomponent test or learning signal, the WD suffers from oscillatory cross terms [7, 8, 9, 10]. Although cross terms can provide discriminatory information, for our application, the extra information could over resolve individual signal types within a class. The representation would then be specific to the given signal rather than a more generic representation of the class of signals as a whole. There exist other TFRs that do not suffer from cross terms that correspond to smoothed versions of the WD. A highly localized reassigned spectrogram [18, 19] was also chosen for our testing to investigate whether TF localization is important in this TF classification problem. The reassigned spectrogram TFR is a spectrogram whose every TF point is moved to the local centroids of the WD, thus achieving very high localization in the TF plane [18, 19]. The resulting classifier is in the form of (7) except that the spectrogram is replaced by the reassigned spectrogram. As we will show, the higher localization offered by this TFR results in better classification results when compared

<sup>2</sup>Although recent studies [16, 17] showed that the distribution of TF pixels could be, for example,  $\chi^2$ , the Gaussian assumption has worked well for our application.



to the spectrogram that can suffer from spreading due to smoothing. Note that there have been other studies on TF detection and classification techniques, some of which include [20, 21, 22, 23, 24, 25]. Acoustic signal classification has also been considered, for example, in [21, 23, 24].

### 2.3. Ambiguity-function-based classifier

The narrowband ambiguity function (AF) of a continuous time signal  $x(t)$  is defined as follows:

$$AF_x(\tau, \nu) = e^{j\pi\tau\nu} \int x(t)x^*(t - \tau)e^{-j2\pi\nu t} dt, \quad (8)$$

and it corresponds to the 2D Fourier transform of the WD. In the AF domain, the auto terms of a multicomponent signal  $x(t)$  are centered around the origin  $(\tau, \nu) = (0, 0)$ , and all the cross terms are away from the origin [9]. The AFs of the different signals to be classified in our acoustic data are remarkably different. For example, Figure 3 shows the AF of a Class 1 signal and a Class 3 signal. The Class 3 signal is an 8 kHz deterministic signal that is used to check whether the detector is properly detecting the acoustic signals. The AF of each signal corresponds to an auto term, but the origin concentration differs due to the properties of each signal. Since Class 3 is a deterministic signal and its time samples are highly correlated, we can expect an appreciable amount of correlation for nonzero time lags. Moreover, since the Class 3 signal contains most of its energy at 8 kHz, there will be less frequency correlation for nonzero frequency lags. Thus, it is clear from the figure that the AF of the third class signal has peaks all along the  $\tau$ -axis with the maximum peak at the origin, whereas the AF of the first class signal has peaks only around the origin. As the magnitude of the AF, the ambiguity surface, is different for different classes, we define a test statistic as [12]

$$\langle |AF_x|, |\overline{AF}_l| \rangle = \int_{\tau} \int_{\nu} |AF_x(\tau, \nu)| |\overline{AF}_l(\tau, \nu)| d\tau d\nu, \quad l = 1, \dots, L, \quad (9)$$

where

$$|\overline{AF}_l(\tau, \nu)| = \frac{1}{P_l} \sum_{p=1}^{P_l} |AF_{x_l^p}(\tau, \nu)|, \quad (10)$$

$|AF_{x_l^p}(\tau, \nu)|$  is the absolute value of the AF of the  $p$ th learning signal in Class  $l$ , and  $P_l$  is the number of learning (i.e., training) signals in Class  $l$ . The classification rule for an  $L$ -class classifier using the AF surface is given by

$$x[n] \in C_j \iff j = \arg \max_l \langle |AF_x|, |\overline{AF}_l| \rangle, \quad (11)$$

for  $l = 1, \dots, L$ , where  $C_j$  is the  $j$ th class. Based on this heuristic approach, we developed this test statistic using the feature set obtained from the AF of a signal. Note that the AF plane was also considered for classification, for example, in [24] and in radar applications. This classification approach was computationally less expensive than other methods considered as it did not require a large number of correlations.

## 3. TF-DECOMPOSITION-BASED CLASSIFIER

### 3.1. Need for a new classifier

The TF classification techniques that were discussed in Section 2 are better matched to the time-varying nature of the signals and can perform better than the classical 1D methods. However, as will be demonstrated in Section 4, their results need further improvements since the classification of acoustic signals from an acoustic monitoring system requires a very high probability of correct classification of the warning signals with reduced probability of false positives. Since the main objective of an acoustic monitoring system is the premonition of catastrophic failures to be followed, missing some warning signals may prove to be very costly. Hence, we need to design a new classification algorithm that can provide us with a very low misclassification rate (less than 1%). The new technique discussed next is based on an iterative classification algorithm that demonstrates high performance with a trade-off of increased computational time. This can be avoided, however, using some preprocessing classification algorithms.

### 3.2. Modified matching pursuit

Over the last two decades, different types of TFRs were introduced to provide signal analysis with high TF resolution and reduced cross terms [7, 8, 9, 10]. New TFRs were obtained in [11, 26, 27] by decomposing an analysis signal into its elementary components, and then using the weighted sum of the WD of the decomposed components as the resulting TFR. One of these decomposition methods is the MPD [11]. The MPD decomposes any signal into a linear expansion of waveforms that belong to a redundant dictionary. The dictionary of the MPD consists of a collection of TF atoms which are the scaled, time-shifted, and frequency-shifted versions of a single basic atom  $d(t)$ . Specifically,  $I$  different elements are obtained as  $d(t; i) = d(\beta_i t - \tau_i) e^{j2\pi\nu_i t}$ ,  $i = 1, \dots, I$ , by varying the time shift  $\tau_i$ , the frequency shift  $\nu_i$ , and the scaling factor  $\beta_i$ . Although the MPD is an iterative nonlinear algorithm, it can preserve signal energy due to its use of orthogonal expansions and can guarantee convergence [11]. The MPD has been used in many applications including analysis and classification [28, 29, 30].

When the analysis signal has multiple components with different TF structures, the MPD uses many dictionary elements to decompose it, and thus becomes very computationally intensive. In [26], rotated Gaussian atoms were used as dictionary elements to decompose linear chirps more efficiently and with fewer waveforms. In order to process signals with different types of TF characteristics, a modified MPD (MMPD) algorithm was used in [31, 32], where the analysis signal was expanded into dictionary waveforms that matched it in structure. These dictionary elements include signals that may have linear phase functions or nonlinear (such as logarithmic or power) phase functions. The advantage of using a dictionary that is matched to the analysis data is that only a small number of elements is used to decompose a signal leading to a parsimonious representation, and hence fewer number of iterations is required.

We propose to design the TF-decomposition-based classifier using the MMPD for classifying time-varying signals. For our MMPD approach, we need to decompose and classify acoustic signals. Thus, we must form a dictionary<sup>3</sup> from the acoustic signals for a successful decomposition. To ensure completeness, we use TF-shifted versions of the learning signals from each class to form our dictionary. Our main objective in using the MMPD is to obtain some distinctive feature set for each class during the iterative procedure of the signal decomposition algorithm and to use these parameters to classify time-varying signals.

### 3.3. Design of the MMPD classifier

In order to design the MMPD classifier, we need to select appropriate learning signals for each class and then form our dictionary by TF shifting these learning signals. By observing the representative spectrograms of the different classes in Figure 2, the range of frequency shifts is limited. Thus, we only choose a small number of frequency shifts to decrease the misclassification rate. In order to reduce the complexity of our MMPD classifier, we incorporate the time shift of the dictionary elements in the cross-correlation computation described in the algorithm steps below. That is, rather than storing a lot of time-shifted atoms, we instead use a cross-correlation between the atom and the residue. This virtually positions the atoms at every lag in the cross-correlation, thereby effecting time shifts without having to store time-shifted copies of the atom in the dictionary. The following steps demonstrate how to classify any test signal<sup>4</sup> from the given acoustic system with the MMPD technique.

*Step 1.* The test signal  $x[n] = x^0[n]$  is cross-correlated with all the dictionary elements.

*Step 2.* The dictionary element whose cross-correlation with the test signal is maximum is chosen. This element, denoted by  $d_l^1[n]$ , best matches the test signal and belongs to Class  $l$ . The corresponding cross-correlation  $a_l^1 = \langle x^0, d_l^1 \rangle$  quantifies the similarity between  $x^0[n]$  and  $d_l^1[n]$ .

*Step 3.* The chosen dictionary element is subtracted from the test signal, and the class number of the chosen dictionary element is noted. The residual signal after the first iteration is given by

$$x^1[n] = x^0[n] - a_l^1 d_l^1[n]. \quad (12)$$

*Step 4.* As the MMPD is iterative, we continue to decompose the residual signal from the previous iteration. At the  $k$ th iteration,

$$x^k[n] = x^{k-1}[n] - a_l^k d_l^k[n], \quad k = 1, \dots, K, \quad (13)$$

where  $x^k[n]$  is the residual signal after the  $k$ th iteration, and  $d_l^k[n]$  is the best matched dictionary element at the  $k$ th iteration that happens to belong to Class  $l$ . The correlation coefficient  $a_l^k$  at the  $k$ th iteration is given by  $a_l^k = \langle x^{k-1}, d_l^k \rangle$  for  $k = 1, \dots, K$ . Step 4 is repeated for  $K$  iterations. The algorithm stopping criterion is a minimum of the correlation coefficient or a maximum number of iterations. Specifically, when the correlation coefficient is much smaller than the threshold, then further iteration only results in adding an error to the decomposition. A similar concept is true for a large number of iterations.

*Step 5.* When the algorithm stops, the net contribution of the correlation coefficients from each class is used as the test statistic. Specifically, the test statistic for Class  $l$  is given by

$$\gamma_l = \sum_{m=1}^M |a_l^{k_m}|, \quad (14)$$

where  $k_m$  is the iteration number in which Class  $l$  was chosen during the decomposition, and  $M$  is the number of times in which Class  $l$  was chosen. Finally, the unknown test signal is classified in Class  $j$  based on

$$x[n] \in C_j \iff j = \arg \max_l \gamma_l, \quad l = 1, \dots, L. \quad (15)$$

As the residual signal  $x^k[n]$  is uncorrelated with the best matched dictionary element  $d_l^k[n]$ , the MMPD is an orthogonal decomposition which guarantees energy conservation.

Step 5 is demonstrated next with an example, with  $L = 8$  different classes. For a given test signal, if the MMPD yields the following correlation coefficients after 8 iterations:  $a_1^1 = 0.9$ ,  $a_2^1 = 0.8$ ,  $a_3^1 = 0.7$ ,  $a_4^1 = 0.6$ ,  $a_5^1 = 0.5$ ,  $a_6^1 = 0.4$ ,  $a_7^1 = 0.3$ , and  $a_8^1 = 0.2$ , then the test statistics for each class are  $\gamma_1 = 0.9 + 0.8 + 0.7 = 2.4$ ,  $\gamma_2 = 0.6$ ,  $\gamma_3 = 0.5$ ,  $\gamma_4 = 0.4$ ,  $\gamma_5 = 0.3$ ,  $\gamma_6 = 0.2$ ,  $\gamma_7 = 0$ , and  $\gamma_8 = 0$ . Recall that  $a_3^1$  is the correlation coefficient from the third iteration, and the best matched dictionary element at the third iteration belongs to Class 1;  $\gamma_1$  is the test statistic for Class 1. Based on the classification rule in (15), this test signal is classified in Class 1. Figure 4 shows the basic block diagram of the algorithm of this classifier.

## 4. CLASSIFICATION RESULTS

For our real-data classification, we compare the performance of the four classifiers presented in Sections 2 and 3. Specifically, we use the spectrogram-based classifier in (7), the reassignment spectrogram-based classifier (with the spectrogram replaced with the reassigned spectrogram in (7)), the AF-based classifier in (11), and the new MMPD classifier in (15). For the MMPD classifier stopping criteria, we set the correlation coefficient threshold value at 0.15, and we use a maximum number of  $K = 10$  iterations. All the test and learning signals used were sampled at 20 kHz, and the duration of each data vector is 0.1024 seconds (2,048 data

<sup>3</sup>Wave-based dictionaries were also used in [30] to process scattering data.

<sup>4</sup>For accurate classification, both the test signals and the learning signals must originate from monitoring hardware of the same design.

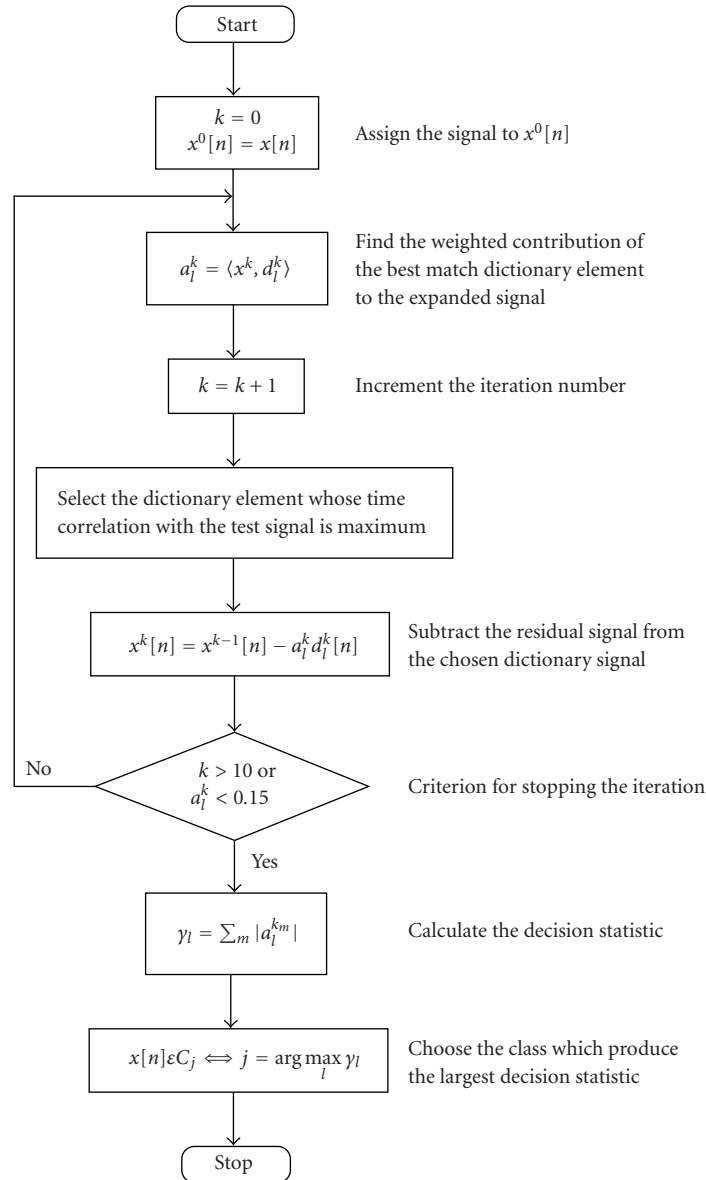


FIGURE 4: Implementation of the MMPD classification algorithm.

samples). Based on this, the size of both the spectrogram and reassignment spectrogram was  $1024 \times 128$ , whereas the AF was only  $50 \times 256$ . Twelve classes of signals were used to compare the different techniques. Figure 2 shows the spectrogram of a characteristic sample signal from each of the twelve different types of signals. Table 1 shows the number of learning and test signals used for each class. We have configured Class 1 to be the major class of interest (warning signal) and the other classes to be nuisance events. As we are interested in the classification of the warning signals, we have used more learning signals from Class 1. By having more learning signals, the probability of correctly classifying a Class 1 signal of varying TF structure increases. Moreover, we have used a large number of test signals of Class 1 to make a fair conclusion about the performance of each classifier. Note that we

are not using large data sets because the probability of correct classification is more pertinent to our problem than the probability of false positive. Specifically, we used all the true positive data (warning signals) we had, some for training and some for testing, and we used a random assortment of false positive data (nuisance signals). In our classification, even probabilities of false positive as high as 0.5 were accepted as it would reduce half the amount of human analysis required; it was important that the probability of correct classification be kept high so as not to miss any warning signals.

#### 4.1. Twelve-Class classifier

In general, of utmost importance is the differentiation of warning signals from nonwarning signals in order to detect the imminent failure of the structure. However, further

TABLE 1: Number of learning and test signals in each class. Learning signals are used to train the classifier and test signals are used to evaluate the classifier.

Class no.	No. of learning signals	No. of test signals
Class 1	59	381
Class 2	10	42
Class 3	4	45
Class 4	5	44
Class 5	8	19
Class 6	9	42
Class 7	8	54
Class 8	8	42
Class 9	12	31
Class 10	4	4
Class 11	9	41
Class 12	6	7

subclassification of the nonwarning signals can provide information useful to field technicians. For example, if an excessive number of biological noises (such as bird chirps) is detected, then it may be necessary for a field technician to provide supplemental acoustic insulation at the monitoring site. This will reduce the total number of unclassified detections and permit longer monitoring due to less acquired data. Moreover, due to the cost involved, it would be prohibitive to circumvent the problem by arbitrarily acoustically insulating every monitoring station regardless of whether or not there is a problem.

Table 2 provides the number of misclassified test signals for each of the twelve possible classes using the four different classification methods. Specifically, the second row of Table 2 compares the performance of the different classifiers in classifying warning signals (Class 1). This table also shows the number of false classification events that correspond to the number of test signals from Class 2 to Class 12 (nuisance signals) that were classified in Class 1 (warning signals). Table 3 provides the probability of correct classification ( $P_{CC}$ ) of Class 1 test signals using the four different methods. Note that by combining the information from Tables 1, 2, and 3, it can be observed that the MMPD-based classifier misclassifies only 3 out of 381 Class 1 test signals which yields  $P_{CC} = 0.9921$ . However, the TF-based classifiers using the spectrogram, reassigned spectrogram, and AF misclassify 19, 21, and 39 Class 1 test signals, respectively. Moreover, only 7 test signals are misclassified in Class 1 yielding a probability of false classification  $P_{FC} = 0.0188$ . The remaining misclassified events of nuisance test signals are misclassified in ten other possible classes. By comparing the misclassification rate for other class test signals in Table 2, the MMPD-based classifier has the best performance. Thus, the MMPD-based classifier appears well suited for applications like this, where we require a high probability of correct classification of warning signals.

TABLE 2: Number of misclassified test signals for four different classifiers: spectrogram (SPEC), reassigned spectrogram (RSPEC), ambiguity function (AF), and modified matching pursuit decomposition (MMPD).

Class no.	SPEC	RSPEC	AF	MMPD
Class 1	19	21	32	3
Class 2	2	2	2	2
Class 3	0	0	1	4
Class 4	0	0	1	1
Class 5	5	3	4	0
Class 6	6	6	12	6
Class 7	3	7	5	2
Class 8	3	3	5	4
Class 9	9	15	13	0
Class 10	1	1	1	1
Class 11	12	8	23	11
Class 12	2	3	1	5
No. of FC	0	0	39	7

TABLE 3: Probability of correct classification ( $P_{CC}$ ) and false classification ( $P_{FC}$ ) for Class 1 test signals using all four methods (for 12 classes).

Classifier	$P_{CC}$	$P_{FC}$
MMPD	0.9921	0.0188
SPEC	0.9501	0
RSPEC	0.9449	0
AF	0.9160	0.1049

The poor performance of the 2D classifiers using the spectrogram, reassigned spectrogram, and AF can be attributed to the fact that the template for each class is obtained by averaging the TFR or AF of the learning signals. Because of this averaging, the templates may not be able to represent any subtle differences that exist between learning signals of the same class. However, in the case of the MMPD, there is no such averaging involved in the classification process. We can expect better results using these 2D signal representations by individually computing the 2D inner product of the TFR or AF of a test signal with the TFR or AF of all the learning signals. Then a test signal will be classified in Class  $l$  when its 2D inner product with a Class  $l$  learning signal is maximum. However, this individual computation of 2D inner products may not be practically feasible due to memory and processing time constraints. It was possible, however, to reduce processing time (without degrading classifier performance) by applying some simple preclassification rules for classes which are easily separated from other classes. For example, Class 3 signals consist of bandpass signals highly concentrated around 8 kHz. Detecting high levels of energy in this band was sufficient for classifying these signals without resorting to the MMPD.



## 4.2. Binary classifier

In Section 4.1, we have seen the classification results of TF- and MMPD-based 12-class classifiers, where it is desirable to classify all the test signals in 1 of 12 possible classes. For example, it is required to distinguish automobile noise from bird chirping noise or dog barking noise. However, this kind of classification may not be necessary for an acoustic monitoring system. Hence, consider a scenario where it is sufficient to determine whether a test signal is a Class A warning signal or not. In this scenario, there are only 2 classes, namely, Class A (warning signals as in Class 1 in Table 1) and Class B (nuisance signals). Two kinds of binary classifiers are discussed in this section based on whether we use learning signals from Class B or not.

### 4.2.1. MMPD with only Class A learning signals

The dictionary of the MMPD in Section 4 is composed of TF-shifted learning signals from 12 classes. However, in this binary classifier, the dictionary is composed of only TF-shifted versions of learning signals from Class A. Since only the learning signals of Class A are used in this classifier, we cannot use the decision strategy in (15). Instead, we use either the residual energy or the correlation coefficient  $a_l^k$  of the MMPD after some specific number of iterations. Since there are no representative signals for Class B, we can anticipate that the residual energy of a test signal that belongs to Class B after some specific number of iterations will be higher when compared to a test signal from Class A. Alternatively, we expect to have a low correlation value coefficient when the test signal actually belongs to Class B.

The performance of the classifier using either the correlation or the residual energy as a stopping criterion gives poor results. For example, we found that for a  $P_{CC} = 0.99$ , the corresponding  $P_{FC}$  is close to 0.8 when the residual energy criterion is used. Thus, if we want to achieve a Class A misclassification rate of 1% (4 wrong out of 381 test signals), then the Class B misclassification rate will be 0.8 (297 out of 372 test signals). We can infer that neither of the two MMPD parameters are sufficient to clearly differentiate Class A signals from Class B signals. This poor classification performance is obtained since there are no representative signals from Class B. However, the MMPD will still choose a best matched dictionary element for any test signal including one from Class B.

### 4.2.2. MMPD with learning signals from both classes

To improve classification performance, we added 31 Class B learning signals to the dictionary. The decision strategy for this binary classifier is the same as in (15) but with  $l = 1, 2$  corresponding to A, B, respectively. Table 4 shows the classification results of this method using (a) 5 and (b) 10 iterations as the stopping criterion. The table shows that the MMPD with Class B learning signals performs much better than the MMPD with no learning signals from Class B. This method misclassifies only 3 out of 381 ( $P_{CC} = 0.9921$ ) Class A test signals, and the number of false classifications (signals from Class B that were classified in Class A) is 4 out of 372

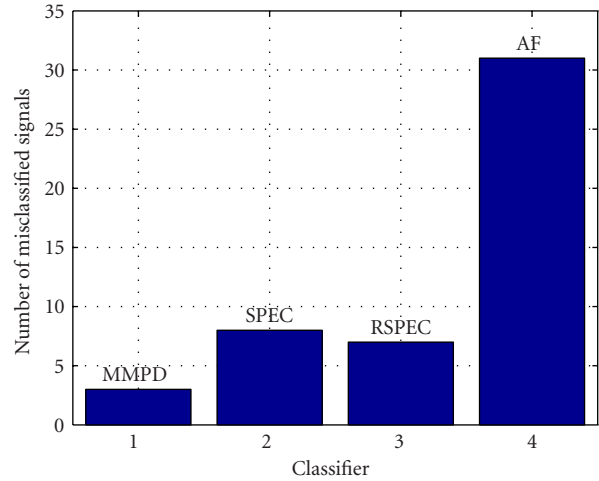


FIGURE 5: Comparison of the four different classifiers (for 2 classes) based on the number of misclassified warning signals which equals 381.

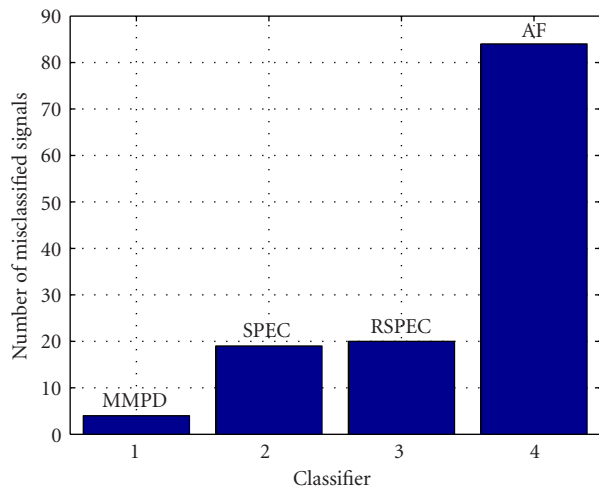


FIGURE 6: Comparison of the four different classifiers (for 2 classes) based on the number of misclassified nuisance signals which equals 372.

( $P_{FC} = 0.0108$ ). The average processing time using 5 iterations was 19 seconds per test signal file and for 10 iterations, it was 33 seconds per test signal file.<sup>5</sup> However, both iteration numbers gave the same results in this particular real-data application. Figures 5 and 6 compare the classification results of the binary MMPD classifier with the other three classification methods discussed in Section 2 based on the number of misclassified signals in Class A and Class B. As can be seen in the figures, the binary MMPD classifier performs significantly better than the other three methods.

<sup>5</sup>Processing was done with Matlab running under Linux on a 800 MHz processor.

TABLE 4: Binary classifier misclassification rate with Classes A and B learning signals.

Class name	10 iterations	5 iterations
A (warning signals)	3/381 = 0.0079	3/381 = 0.0079
B (nuisance signals)	4/372 = 0.0108	4/372 = 0.0108

### 4.3. Other classification methods

A method not discussed in this paper uses TFR log deviation [15] as a test statistic to classify time-varying signals. Another test statistic is used in [33, 34, 35] which depends upon taking the reciprocal of the TF points of a TFR. Note that it is difficult to use these two methods in signal classification when the signals to be classified (like the acoustic events) have no common nonzero values within a band of frequencies. As a result, we did not compare our results with these methods. Moreover, the log deviation method limits the type of TFR, because only positive TFRs like the spectrogram or the reassigned spectrogram can be used.

## 5. CONCLUSION

We have used existing as well as new techniques to classify real acoustic signals from an acoustic monitoring system. For signals similar to those used in this report, our new MMPD-based method provides promising classification results. The only drawback of the MMPD is the processing time. Since signal processing can be done offline in an acoustic monitoring system of concrete structures, the processing time is not of great concern when considered with the highly accurate classification performance. For situations where the processing time is critical, some preclassifier methods can be applied before using the MMPD. However, the performance of such preclassifiers can increase only when the number of classes that can be preclassified with more confidence increases.

We have illustrated the performance of the MMPD-based classifier for multiclass and binary classification. The flexibility of the MMPD-based classifier allows one to implement both cases by simply changing the learning signal sets in the dictionary. Thus, the implementation of the MMPD-based classifier can be easily modified to classify signals from other acoustic monitoring systems by changing the learning signal set. However, care should be taken in selecting the range of frequency shifts of the learning signals used to form the dictionary, since a higher range of frequency shift may increase the misclassification rate. This is due to the fact that when a learning signal of a particular class is frequency shifted by a large amount, then it is possible that the TF structure of this frequency-shifted learning signal may resemble a learning signal of another class.

The automatic classifiers discussed in this paper act as tools that can be used in acoustic monitoring to separate events which are important for further analysis, and to reject a rather large number of signals that are of no diagnostic value. With this capability, it is possible to significantly reduce the amount of manual classification often required in acoustic monitoring.

## REFERENCES

- [1] W. C. Michie, G. Thursby, D. Walsh, B. Culshaw, and M. Konstantaki, "Distributed sensing of physical and chemical parameters for structural monitoring," in *IEE Colloquium on Optical Techniques for Smart Structures and Structural Monitoring* (Digest no. 1997/033), pp. 3/1–3/9, London, UK, February 1997.
- [2] E. A. Johnson, P. G. Voulgaris, and L. A. Bergman, "Methods of system identification for monitoring slowly time-varying structural systems," in *Proc. Intelligent Information Systems*, pp. 569–573, Grand Bahama Island, Bahamas, December 1997.
- [3] R. M. Measures, "Fiber optic structural monitoring of bridges," in *Proc. IEEE Instrumentation and Measurement Technology Conference*, pp. 600–602, Ontario, Canada, May 1997.
- [4] W. Worthington, "Prestressed concrete pipe inspection and monitoring methods," in *Proc. Nondestructive Evaluation of Civil Structures and Materials Conference*, Boulder, Colo, USA, May 1992.
- [5] J. G. Proakis, *Digital Communications*, McGraw-Hill, NY, USA, 2001.
- [6] J. C. Hassab, *Underwater Signal and Data Processing*, CRC Press, Fla, USA, 1989.
- [7] A. Papandreou-Suppappola, Ed., *Applications in Time-Frequency Signal Processing*, CRC Press, Fla, USA, 2002.
- [8] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, NJ, USA, 1995.
- [9] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21–67, 1992.
- [10] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, Calif, USA, 1999.
- [11] S. G. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [12] S. P. Varma, A. Papandreou-Suppappola, and S. B. Suppappola, "Detecting faults in structures using time-frequency techniques," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 6, pp. 3593–3596, Utah, May 2001.
- [13] S. P. Ebenezer, A. Papandreou-Suppappola, and S. B. Suppappola, "Matching pursuit classification for time-varying acoustic emissions," in *Proc. 35th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 715–719, Calif, USA, November 2001.
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*, Prentice-Hall, NJ, USA, 1998.
- [15] I. Vincent, C. Doncarli, and E. Le Carpentier, "Non stationary signal classification using time-frequency distributions," in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 233–236, Philadelphia, Pa, USA, October 1994.
- [16] P. Duvaux and D. Declercq, "Statistical properties of the pseudo-Wigner-Ville representation of normal random processes," *Signal Processing*, vol. 75, no. 1, pp. 93–98, 1999.
- [17] C. Hory, N. Martin, and A. Chehikian, "Spectrogram segmentation by means of statistical features for non-stationary signal interpretation," *IEEE Trans. Signal Processing*, vol. 50, no. 12, pp. 2915–2925, 2002.
- [18] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing*, vol. 43, no. 5, pp. 1068–1089, 1995.
- [19] F. Auger and P. Flandrin, "The why and how of time-frequency reassignment," in *Proc. IEEE-SP International Sym-*

- posium on Time-Frequency and Time-Scale Analysis*, pp. 197–200, Philadelphia, Pa, USA, October 1994.
- [20] P. Flandrin, “A time-frequency formulation of optimal detection,” *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, no. 9, pp. 1377–1384, 1988.
- [21] C. Heitz and J. Timmer, “Using optimized time-frequency representations for acoustic quality control of motors,” *Acta Acoustica*, vol. 83, no. 6, pp. 1053–1064, 1997.
- [22] G. Matz and F. Hlawatsch, “Time-frequency formulation and design of optimal detectors,” in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 213–216, Paris, France, June 1996.
- [23] M. Davy, C. Doncarli, and G. F. Boudreaux-Bartels, “Improved optimization of time-frequency based signal classifiers,” *IEEE Signal Processing Letters*, vol. 8, no. 2, pp. 52–57, 2001.
- [24] B. W. Gillespie and L. E. Atlas, “Optimizing time-frequency kernels for classification,” *IEEE Trans. Signal Processing*, vol. 49, no. 3, pp. 485–496, 2001.
- [25] C. Richard, “Time-frequency-based detection using discrete-time discrete-frequency Wigner distributions,” *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2170–2176, 2002.
- [26] A. Bultan, “A four-parameter atomic decomposition of chirplets,” *IEEE Trans. Signal Processing*, vol. 47, pp. 731–745, March 1999.
- [27] S. Qian and J. M. Morris, “Wigner distribution decomposition and cross-terms deleted representation,” *Signal Processing*, vol. 27, no. 2, pp. 125–144, 1992.
- [28] R. Gribonval, “Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 3, pp. 3057–3060, Orlando, Fla, USA, May 2002.
- [29] R. Gribonval and E. Bacry, “Harmonic decomposition of audio signals with matching pursuit,” *IEEE Trans. Signal Processing*, vol. 51, no. 1, pp. 101–111, 2003.
- [30] M. R. McClure and L. Carin, “Matching pursuits with a wave-based dictionary,” *IEEE Trans. Signal Processing*, vol. 45, no. 12, pp. 2912–2927, 1997.
- [31] A. Papandreou-Suppappola and S. B. Suppappola, “Adaptive time-frequency representations for multiple structures,” in *Proc. 10th IEEE Workshop on Statistical Signal and Array Processing*, vol. 3, pp. 579–583, Pocono Manor, Pa, USA, August 2000.
- [32] A. Papandreou-Suppappola and S. B. Suppappola, “Analysis and classification of time-varying signals with multiple time-frequency structures,” *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 92–95, 2002.
- [33] G. Roberts, A. M. Zoubir, and B. Boashash, “Classification of non-stationary random signals using multiple hypotheses testing,” in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, vol. 3, pp. 245–248, Paris, France, June 1996.
- [34] G. Roberts, A. M. Zoubir, and B. Boashash, “Time-frequency discriminant analysis for non-stationary Gaussian signals,” in *Proc. International Symposium on Signal Processing and Its Applications*, vol. 3, pp. 33–36, Brisbane, Australia, August 1996.
- [35] G. Roberts, A. M. Zoubir, and B. Boashash, “Time-frequency classification using a multiple hypothesis test: an application to the classification of humpback whale signals,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 1, pp. 563–566, Munich, Germany, April 1997.

**Samuel P. Ebenezer** received his B.E. degree in 1997 from the Manonmaniam Sundaranar University, India, in electronics and communication engineering, and his M.S. degree in 2001 from the Arizona State University, Phoenix, USA, in electrical engineering. He worked as a Telecommunication Engineer and Senior Telecommunication Engineer at Tenet Laboratories, Indian Institute of Technology, Madras, from 1997 to 2000. During this period, he was involved in the design and development of digital enhanced cordless telecommunication (DECT) based wireless in local loop system. During 2000–2001, he worked as a Research Assistant at the Telecommunication Research Center of Arizona State University. Currently, he is a Senior Engineer at Acoustic Technologies Inc. in Mesa, Ariz, USA, where he is involved in acoustical and speech signal processing for wireless applications. His current research interest includes time-varying and stochastic signal processing-based applications, adaptive filters, and detection and estimation theory.



**Antonia Papandreou-Suppappola** received her Ph.D. degree in electrical engineering in May 1995 from the University of Rhode Island where she then held a research faculty position with Navy funding. In August 1999, she joined Arizona State University as an Assistant Professor. Her current research interests are in the areas of time-varying signal processing, signal processing for wireless communications, integrated sensing and processing, and estimation and detection theory. She has a large publication record consisting of more than sixty refereed journal articles, book chapters, and conference papers. Through CRC Press, she published a book entitled *Applications in Time-Frequency Signal Processing* in October 2002. She has received research funding from numerous agencies including the National Science Foundation (NSF), the Defense Advanced Research Projects Agency (DARPA), the Office of Naval Research (ONR), and the Naval Undersea Warfare Center (NUWC), and she is the recipient of an NSF CAREER Award. She is a member of the IEEE Signal Processing Society, Communications Society, and Women in Engineering Society.



**Seth B. Suppappola** was born in Warwick, RI, in 1967. He received his B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Rhode Island in 1989, 1991, and 1994, respectively. From 1988 to 1999, he worked at the Naval Undersea Warfare Center in Newport, RI, where he developed systems and algorithms for both active and passive sonar signal processing. From 2000 to 2001, he was the Director of Research and Development at Pipeline Technologies in Scottsdale, Arizona. There he developed acoustical systems and algorithms for the nondestructive testing of prestressed concrete cylinder pipe. In 2001, he joined Acoustic Technologies Inc. in Mesa, Arizona, where he develops voice processing algorithms for cellular communications. His current research interests are acoustics and signal processing.

