Target Detection and Classification Using Seismic and PIR Sensors

Xin Jin, *Student Member, IEEE*, Soumalya Sarkar, Asok Ray, *Fellow, IEEE*, Shalabh Gupta, *Member, IEEE*, and Thyagaraju Damarla, *Senior Member, IEEE*

Abstract—Unattended ground sensors (UGS) are widely used to monitor human activities, such as pedestrian motion and detection of intruders in a secure region. Efficacy of UGS systems is often limited by high false alarm rates, possibly due to inadequacies of the underlying algorithms and limitations of onboard computation. In this regard, this paper presents a wavelet-based method for target detection and classification. The proposed method has been validated on data sets of seismic and passive infrared sensors for target detection and classification, as well as for payload and movement type identification of the targets. The proposed method has the advantages of fast execution time and low memory requirements and is potentially well-suited for real-time implementation with onboard UGS systems.

Index Terms—Feature extraction, passive infrared sensor, seismic sensor, symbolic dynamic filtering, target detection and classification.

I. INTRODUCTION

UNATTENDED ground sensors (UGS) are widely used in industrial monitoring and military operations. Such UGS systems are usually lightweight devices that automatically monitor the local activities *in-situ*, and transfer target detection and classification reports to the processing center at a higher level of hierarchy. Commercially available UGS systems make use of multiple sensing modalities (e.g., acoustic, seismic, passive infrared, magnetic, electrostatic, and video). Efficacy of UGS systems is often limited by high false alarm rates because the onboard data processing algorithms may not be able to correctly discriminate different types of targets (e.g., humans from animals) [1]. Power consumption is a critical consideration in UGS systems. Therefore, power-efficient

sensing modalities, low-power signal processing algorithms, and efficient methods for exchanging information between the UGS nodes are needed [2].

In the detection and classification problem at hand, the targets usually include human, vehicles and animals. For example, discriminating human footstep signals from other targets and noise sources is a challenging task, because the signal-to-noise ratio (SNR) of footsteps decreases rapidly with the distance between the sensor and the pedestrian. Furthermore, the footstep signals may vary significantly for different people and environments. Often the weak and noise-contaminated signatures of humans and light vehicles may not be clearly distinguishable from each other, in contrast to heavy vehicles that radiate loud signatures [3], [4].

Seismic sensors are widely used for personnel detection, because they are relatively less sensitive to Doppler effects and environment variations, as compared to acoustic sensors [5]. Current personnel detection methods, based on seismic signals, are classified into three categories, namely, time domain [6], frequency domain [3], [4], [7], and time-frequency domain [5], [8]–[10]. Generally, time-domain analysis may not be able to detect targets very accurately because of the interfering noise, complicated signal waveforms, and variations of the terrain [5]. On the other hand, accuracy of frequency domain methods may be degraded due to underlying non-stationarity in the observed signal. Therefore, recent research has relied on time-frequency domain (e.g. wavelet transform-based) methods because of their denoising and localization properties. Passive Infrared (PIR) sensors have been widely used in motion detectors, where the PIR signals are usually quantized into two states, i.e., "on" and "off". PIR signals contain discriminative information in the time-frequency domain and are well-suited for UGS systems due to low power consumption. Although PIR sensors have been used for detection and localization of moving targets [11], similar efforts for target classification have not been apparently reported in open literature.

The work reported in this paper makes use of a wavelet-based feature extraction method, called Symbolic Dynamic Filtering (SDF) [12]–[14]. The SDF-based feature extraction algorithm mitigates the noise by using wavelet analysis, captures the essential signatures of the original signals in the time-frequency domain, and generates robust low-dimensional feature vectors for pattern classification [15]. This paper addresses the problem of target detection and classification using seismic and PIR sensors that monitor the infiltration of humans, light vehicles and domestic animals for border

Fig. 1.   Illustration of the test scenario with three sensor sites.

Fig. 2.   Examples of test scenarios with different targets. (a) Human. (b) Vehicle. (c) Animal led by human.

security. The major contributions of the paper are as follows:

1) Formulation of a hierarchical structure for target detection and classification.
2) Experimental validation of the SDF-based feature extraction method on seismic and PIR sensor data.
3) Performance evaluation of using seismic and PIR sensors in target payload and movement type identification.

The paper is organized into five sections including the present one. Section II describes and formulates the problem of target detection and classification. Section III presents the procedure of feature extraction from sensor time-series. Section IV describes the details of the proposed method and the results of field data analysis. The paper is concluded in Section V along with recommendations for future research.

II. PROBLEM DESCRIPTION AND FORMULATION

The objective is to detect and classify different targets (e.g., humans, vehicles, and animals led by human), where seismic and PIR sensors are used to capture the characteristic signatures. For example, in the movement of a human or an animal across the ground, oscillatory motions of the body appendages provide the respective characteristic signatures.

The seismic and PIR sensor data, used in this analysis, were collected on multiple days from test fields on a wash (i.e., the dry bed of an intermittent creek) and at a choke point (i.e., a place where the targets are forced to go due to terrain difficulties). During multiple field tests, sensor data were collected for several scenarios that consisted of targets walking along an approximately 150 meters long trail, and returning along the same trail to the starting point. Figure 1 illustrates a typical data collection scenario.

The targets consisted of humans (e.g., male and female), animals (e.g., donkeys, mules, and horses), and all-terrain vehicles (ATVs). The humans walked alone and in groups with and without backpacks; the animals were led by their human handlers (simply denoted as "animal" in the sequel) and they made runs with and without payloads; and ATVs moved at different speeds (e.g., 5 mph and 10 mph). Examples of the test scenarios with different targets are shown in Fig. 2. There were three sensor sites, each equipped with seismic and PIR sensors. The seismic sensors (geophones) were buried approximately 15 cm deep underneath the soil surface, and the PIR sensors were collocated with the respective seismic sensors. All targets passed by the sensor sites at a distance of approximately 5 m. Signals from both sensors were acquired at a sampling frequency of 10 kHz.
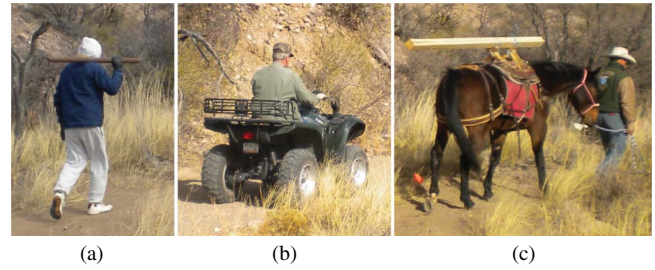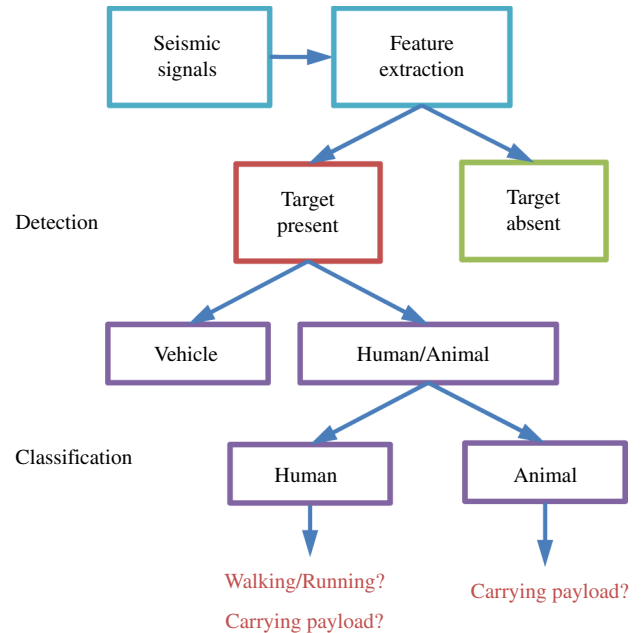

Fig. 3.   Tree structure formulation of the detection & classification problem.

The tree structure in Fig. 3 shows how the detection and classification problem is formulated. In the detection stage, the pattern classifier detects the presence of a moving target against the null hypothesis of no target present; in the classification stage, the pattern classifiers discriminate among different targets, and subsequently identify the movement type and/or payload of the targets. While the detection system should be robust to reduce the false alarm rates, the classification system must be sufficiently sensitive to discriminate among different types of targets with high fidelity. In this context, feature extraction plays an important role in target detection and classification because the performance of classifiers largely depends on the quality of the extracted features.

In the classification stage, there are multiple classes (i.e., humans, animals, and vehicles); and the signature of the vehicles is distinct from those of the other two classes. Therefore, this problem is formulated into a two-layer classification procedure. A binary classification is performed to detect the presence of a target and then to identify whether the target is a vehicle or a human/animal. Upon recognizing the target as a human/animal, another binary classification is performed to determine its specific class. More information could be derived upon recognition of the target type. For example, if the target
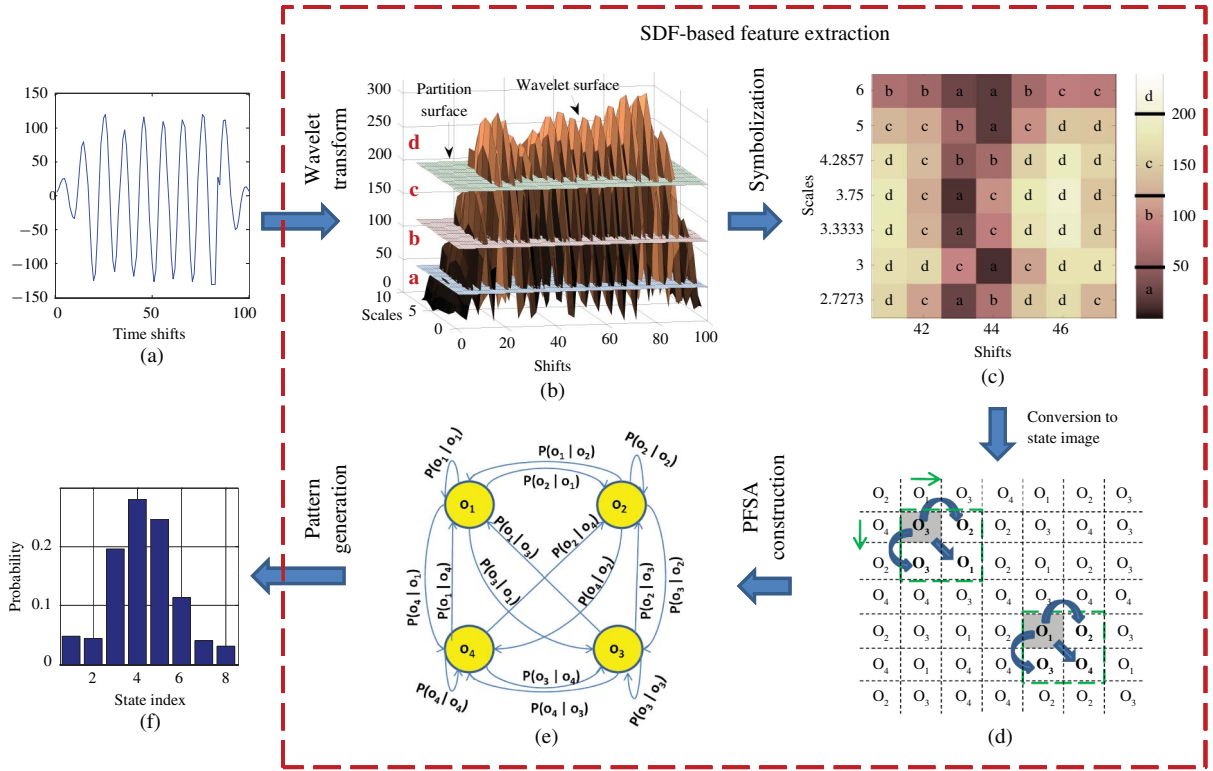
Fig. 4. Overview of the SDF-based feature extraction algorithm. (a) Sensor time series data. (b) Partition of the wavelet coefficients. (c) Symbolized wavelet image (a section). (d) Feature extraction from state image. (e) 4-state PFSA. (f) State probability vector.

is recognized as a human, then further binary classifications are performed to identify if the human is running or walking, and if the human is carrying a payload or not.

III. SYMBOLIC DYNAMICS-BASED FEATURE EXTRACTION

A key step in target detection and classification is feature extraction from sensor signals, which is accomplished by symbolic dynamic filtering (SDF) in this paper. While the details of SDF have been reported in earlier publications [12]–[15], this section briefly reviews the underlying concepts of feature extraction from sensor time series for completeness of this paper.

A. Transformation of Time Series to Wavelet Domain

A crucial step in SDF is partitioning of the transformed data space for symbol sequence generation. In wavelet-based partitioning, the time series is first transformed as a set of wavelet coefficients at different time shifts and scales, where the choice of the wavelet basis function depends on the time-frequency characteristics of the underlying signal, and the (finitely many) wavelet scales are calculated as follows:

$$\alpha^i = \frac{F_c}{f_p^i \, \Delta t} \tag{1}$$

where F_c is the center frequency [16] that has the maximum modulus in the Fourier transform of the signal; and f_p^i's are obtained by choosing the locally dominant frequencies in the Fourier transform.

Figure 4 shows an illustrative example of transformation of the time series (Fig. 4(a)) to a (two-dimensional) wavelet image (Fig. 4(b)). The amplitudes of the wavelet coefficients over the scale-shift domain are plotted as a surface. Subsequently, symbolization of this wavelet surface leads to the formation of a symbolic image as shown in Fig. 4(c).

B. Symbolization of Wavelet Surface Profiles

This section presents partitioning of the wavelet surface profile, as shown in Fig. 4(b), which is generated by the coefficients over the two-dimensional scale-shift domain, for construction of the symbolic image in Fig. 4(c). The $x - y$ coordinates of the wavelet surface profiles denote the shifts and the scales respectively, and the z-coordinate denotes the surface height as pixel values of the wavelet coefficients.

The wavelet surface profiles are partitioned such that the ordinates between the maximum and minimum of the coefficients along the z-axis are divided into regions by different planes parallel to the $x - y$ plane. For example, if the alphabet is chosen as $\Sigma = \{a, b, c, d\}$, i.e., $|\Sigma| = 4$, then three partitioning planes divide the ordinate (i.e., z-axis) of the surface profile into four mutually exclusive and exhaustive regions, as shown in Figure 4(b). These disjoint regions form a partition, where each region is labeled with one symbol from the alphabet Σ. If the intensity of a pixel is located in a particular region, then it is coded with the symbol associated with that region. As such, a symbol from the alphabet Σ is assigned to each pixel corresponding to the region where its intensity falls. Thus, the two-dimensional array of symbols,

called *symbol image*, is generated from the wavelet surface profile, as shown in Figure 4(c).

The surface profiles can be partitioned by using different partitioning methods. If the partitioning planes are separated by equal-sized intervals, then it is called the *uniform partitioning* (UP). However, the partitioning would be more reasonable if the information-rich regions of a data set are partitioned finer and those with sparse information are partitioned coarser. To achieve this objective, the *maximum entropy partitioning* (MEP) [13], [14] has been adopted such that the entropy of the generated symbols is maximized. The procedure for selection of the alphabet size $|\Sigma|$, followed by generation of a MEP, has been reported in [13], [14]. In general, the choice of alphabet size depends on specific data set. The partitioning of wavelet surface profiles to generate symbolic representations enables robust feature extraction, and symbolization also significantly reduces the memory requirements.

For the purpose of pattern classification, the reference data set is partitioned with alphabet size $|\Sigma|$ and is subsequently kept constant. In other words, the structure of the partition is fixed at the reference condition and this partition serves as the reference frame for subsequent data analysis [12].

C. Conversion of the Symbol Image to the State Image

This section presents construction of a *probabilistic finite state automaton* (PFSA) for feature extraction based on the symbol image generated from a wavelet surface profile.

For analysis of (one-dimensional) time series, the states of a PFSA represent different combinations of blocks of symbols on the symbol sequence and the edges represent the transition probabilities between these blocks [12]. Therefore, for analysis of (one dimensional) time series, the "states" denote all possible symbol blocks (i.e., words) within a window of certain length. The notion of "states" is now extended for analysis of wavelet surface profiles via construction of a "*state image*" from a "*symbol image*".

In general, the computational requirements increase with the number $\mathcal{Q}$ of states, which must be constrained for real-time applications. As $|\mathcal{Q}|$ increases with the window size $|\mathcal{W}|$ and the alphabet size $|\Sigma|$, a probabilistic state compression method is employed, which chooses m most probable symbols from each state as a representation of that particular state. State compression must preserve sufficient information as needed for pattern classification, albeit possibly lossy coding of the wavelet surface profile.

In this method, each state consisting of $\ell \times \ell$ symbols is compressed to a word of length $m < \ell^2$ symbols by choosing the top m symbols that have the highest probability of occurrence. (Note: If two symbols have the same probability of occurrence, then either symbol may be preferred with equal probability.) This procedure reduces the state set $\mathcal{Q}$ to an effective smaller set $\mathcal{O} \triangleq \{o_1, o_2, \ldots, o_{|\mathcal{O}|}\}$ that enables mapping of two or more different configurations in a window $\mathcal{W}$ to a single state. For example, if $|\Sigma| = 4$, $|\mathcal{W}| = 4$ and $m = 2$, then the state compression reduces the total number of states to $|\mathcal{O}| \leq |\Sigma|^m = 16$ instead of 256. The choice of $|\Sigma|$, ℓ and m depends on application domains, noise level,

and the available computational power, and is made by an appropriate tradeoff between robustness to noise and capability to detect small changes. For example, a large alphabet may be noise-sensitive while a small alphabet may miss the information of signal dynamics. This issue is discussed further in Subsection IV-A.2.

D. Construction of PFSA and Pattern Generation

A probabilistic finite state automaton (PFSA) is constructed such that the states of the PFSA are elements of the compressed state set $\mathcal{O}$ and the edges are transition probabilities between these states. The transition probabilities are defined as:

$$\wp(o_k|o_l) = \frac{N(o_l, o_k)}{\sum_{k'=1,2,\ldots,|\mathcal{O}|} N(o_l, o_{k'})} \quad \forall\, o_l, o_k \in \mathcal{O} \quad (2)$$

where $N(o_l, o_k)$ is the total count of events when o_k occurs adjacent to o_l in the direction of motion. The calculation of these transition probabilities follows the principle of sliding block code [17]. A transition from the state o_l to the state o_k occurs if o_k lies adjacent to o_l in the positive direction of motion. Subsequently, the counter moves to the right and to the bottom (row-wise) to cover the entire state image, and the transition probabilities $\wp(o_k|o_l) \,\forall\, o_l, o_k \in \mathcal{O}$ are computed using Eq. (2). Therefore, for every state on the state image, all state-to-state transitions are counted, as shown in Fig. 4(d). For example, the dotted box in the bottom-right corner contains three adjacent pairs, implying the transitions $o_1 \rightarrow o_2$, $o_1 \rightarrow o_3$, and $o_1 \rightarrow o_4$ and the corresponding counter of occurrences $N(o_1, o_2)$, $N(o_1, o_3)$, and $N(o_1, o_4)$, respectively, are increased by one. This procedure generates the state-transition probability matrix of the PFSA given as:

$$\Pi = \begin{bmatrix} \wp(o_1|o_1) & \ldots & \wp(o_{|\mathcal{O}|}|o_1) \\ \vdots & \ddots & \vdots \\ \wp(o_1|o_{|\mathcal{O}|}) & \ldots & \wp(o_{|\mathcal{O}|}|o_{|\mathcal{O}|}) \end{bmatrix} \quad (3)$$

where $\Pi \equiv [\pi_{jk}]$ with $\pi_{jk} = \wp(o_k|o_j)$. Note: $\pi_{jk} \geq 0 \,\forall j, k \in \{1, 2, \ldots|\mathcal{O}|\}$ and $\sum_k \pi_{jk} = 1 \,\forall j \in \{1, 2, \ldots|\mathcal{O}|\}$.

In order to extract a low-dimensional feature vector, the stationary state probability vector **p** is obtained as the left eigenvector corresponding to the unity eigenvalue of the stochastic transition matrix Π. The state probability vectors **p** serve as the "*feature vectors*" and are generated from different data sets from the corresponding state transition matrices. These feature vectors are denoted as "*patterns*" in this paper.

IV. RESULTS OF FIELD DATA ANALYSIS

Field data were collected in the scenario illustrated in Fig. 1. Multiple experiments were made to collect data sets of all three classes, i.e., human, vehicle and animal. The data were collected over three days at different sites. A brief summary is given in Table I showing the number of runs of each class.

Each data set, acquired at a sampling frequency of 10 kHz, has 1×10^5 data points that correspond to 10 seconds of the experimentation time. In order to test the capability of the proposed algorithm for target detection, another data set

	Day 1	Day 2	Day 3	Total
No target	50	36	32	118
Vehicle	0	8	0	8
Human	30	22	14	66
Animal	20	6	18	44

was collected with no target present. The problem of target detection is then formulated as a binary pattern classification, where no target present corresponds to one class, and target present (i.e., human, vehicle or animal) corresponds to the other class. The data sets, collected by the channel of seismic sensors that are orthogonal to the ground surface and the PIR sensors that are collocated with the seismic sensors, are used for target detection and classification. For computational efficiency, the data were downsampled by a factor of 10 with no apparent loss of information.

Figure 5 depicts the flow chart of the proposed detection and classification algorithm that is constructed based on the theories of symbolic dynamic filtering (SDF) and support vector machines (SVM) [18]. The proposed algorithm consists of four main steps, namely, signal preprocessing, feature extraction, detection, and classification, as shown in Fig. 5.

In the signal conditioning step, the DC component (i.e., the constant offset) of a seismic signal is eliminated by subtracting the average value and the resulting (zero-mean) signal is normalized to unit-variance with division by its standard deviation. The rationale for normalization to unit variance is to make pattern classification independent of the signal amplitude and any discrimination should be solely texture-dependent. For example, the amplitude of the seismic signal of an animal with a heavy payload walking far away could be comparable to that of a pedestrian passing by at a closer distance, although these two signals are of different texture. However, for PIR signals, only the DC component is removed and the normalization is not carried out because the range of the PIR signals is not changed during the field test experiments.

In the feature extraction step, SDF captures the signatures of the preprocessed sensor time-series for representation as low-dimensional feature vectors. Based on the spectral analysis of the ensemble of seismic data at hand, a series of pseudo-frequencies from the 1-20 Hz bands have been chosen to generate the scales for wavelet transform, because these bands contain a very large part of the footstep energy [8]. Similarly, a series of pseudo-frequencies from the 0.2-2.0 Hz bands have been chosen for PIR signals to generate the scales. Upon generation of the scales, continuous wavelet transforms (CWT) are performed with an appropriate wavelet basis function on the seismic and PIR signals. The wavelet basis *db*7 is used for seismic signals since it matches the impulsive shape of seismic signals very well, and *db*1 is used for the PIR case since PIR signals are close to square waves. A maximum-entropy wavelet surface partitioning is then performed. Selection of the alphabet size $|\Sigma|$ depends on the characteristics of the signal; while a small alphabet is robust against noise and environmental variations, a large alphabet has more discriminant
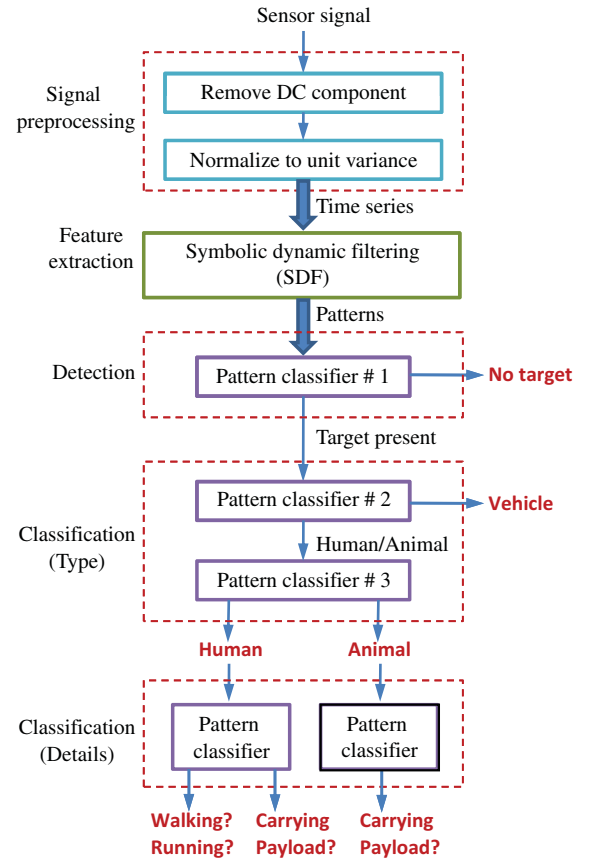

Fig. 5. Flow chart of the problem of target detection and classification.

power for identifying different objects. The same alphabet is used for both target detection and classification. The issues of optimization of the alphabet size and data set partitioning are not addressed in this paper. Subsequently, the extracted low-dimensional patterns are used for target detection and classification. One pattern is generated from each experiment, and the training patterns are used to generate the separating hyperplane in SVM.

A. Performance Assessment Using Seismic Data

This section presents the classification results using the patterns extracted from seismic signals using SDF. The leave-one-out cross-validation method [18] has been used in the performance assessment of seismic data. Since the seismic sensors are not site-independent, they require partial information of the test site, which is obtained from the training set in the cross-validation. Results of target detection and classification, movement type and target payload identification are reported in this section.

1) Target Detection and Classification: Figure 6 shows the normalized seismic sensor signals (top row) and the corresponding feature vectors (bottom row) extracted by SDF of the three classes of targets and the no target case. In the top row of Fig. 6, the unit of the ordinate axes is dimensionless due to normalization of the seismic signals, where the original data were recorded in the unit of *volt* by microphones for storage in a digitized format. Each feature vector in the bottom row of Fig. 6 consists of 8 elements since the alphabet size $|\Sigma| = 8$
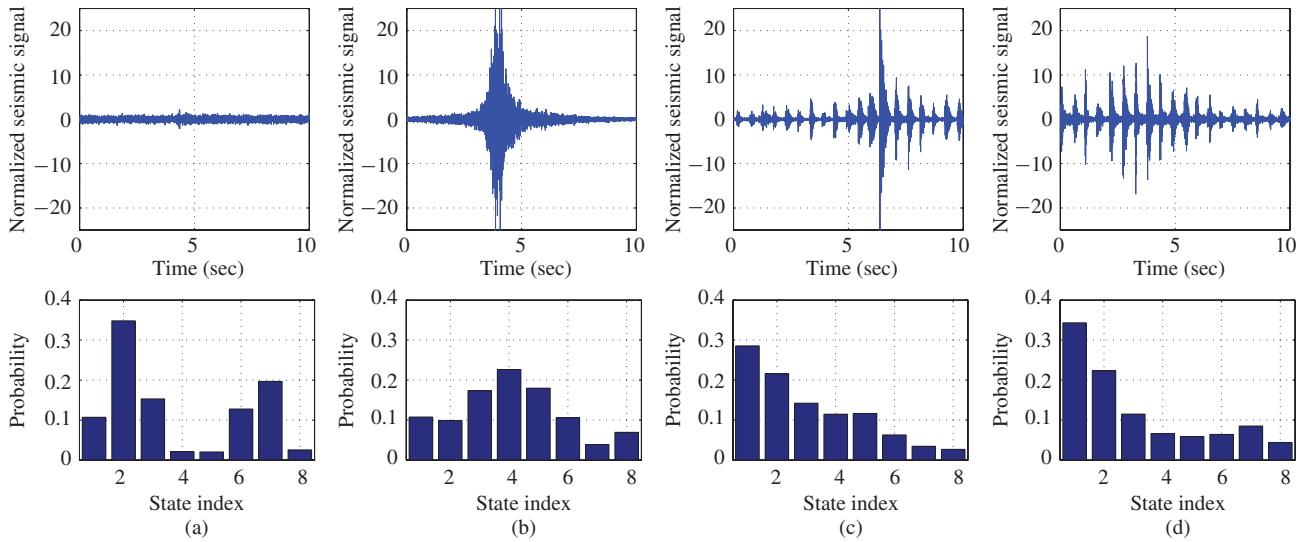
Fig. 6. (a) No target. (b) Vehicle. (c) Human. (d) Animal. Examples of seismic sensor measurements (top) and the corresponding feature vectors extracted by SDF of the four classes (bottom).

TABLE II

CONFUSION MATRICES OF THE LEAVE-ONE-OUT CROSS-VALIDATION RESULTS USING SDF AND KURTOSIS ANALYSIS

SDF	No target	Vehicle	Human	Animal
No target	114	1	1	2
Vehicle	0	7	1	0
Human	3	0	61	2
Animal	0	0	1	43

Kurtosis	No target	Vehicle	Human	Animal
No target	102	0	5	11
Vehicle	0	0	7	1
Human	1	0	47	18
Animal	1	0	30	13

TABLE III

COMPARISON OF THE DETECTION AND CLASSIFICATION ACCURACY BY USING SDF AND KURTOSIS ANALYSIS

	Detection	Classification	
		Vehicle versus Others	Human versus Animal
SDF	97.0%	99.1%	97.2%
Kurtosis	92.4%	93.1%	55.6%

TABLE IV

CONFUSION MATRICES OF THE LEAVE-ONE-OUT CROSS-VALIDATION RESULTS FOR MOVEMENT TYPE IDENTIFICATION

	Human Walking	Human Running
Human Walking	47	1
Human Running	5	13

TABLE V

CONFUSION MATRICES OF THE LEAVE-ONE-OUT CROSS-VALIDATION RESULTS FOR TARGET PAYLOAD IDENTIFICATION

		Human		Animal	
		no payload	payload	no payload	payload
Human	no payload	45	4	1	0
	payload	8	7	1	0
Animal	no payload	0	1	6	7
	payload	0	0	2	28

and the sum of all the elements in each feature vector is 1. It is observed that the feature vectors are quite different among no target, vehicle and human/animal case. The feature vectors of human and animal are similar and yet still distinguishable. In the feature vector plots in Fig. 6, the states with small index number corresponds to the wavelet coefficients with large values, and vice versa.

For the purpose of comparative evaluation, kurtosis analysis [6], a benchmarking technique of footstep detection, is also used for target detection and classification. Kurtosis analysis is useful for footstep detection because the kurtosis value is much higher in the presence of impulsive events (i.e., target present) than the case of no target [6]. The results of SDF and kurtosis analysis are shown in Table II using confusion matrices, where the rows are the actual classes and the columns are the predicted classes. Similar notations are followed in the sequel in Tables IV, V, and VI. The shaded area in Table II represents the confusion matrices of target classification. The detection and classification accuracy is summarized in Table III. It is observed kurtosis analysis has slightly worse but comparable performance with SDF

in target detection and vehicle classification, whereas SDF outperforms kurtosis analysis in distinguishing human from animal.

The execution of the MATLAB code takes 2.27 seconds and 43.73 MB of memory for SDF and SVM on a desktop computer to process a data set of 1×10^4 points and perform pattern classification with the following parameters: alphabet size $|\Sigma| = 8$, number of scales $|\alpha| = 4$, window size $\ell \times \ell = 2 \times 2$, number of most probable symbol $m = 1$, and *quadratic*

TABLE VI

CONFUSION MATRIX OF THE THREE-WAY CROSS-VALIDATION

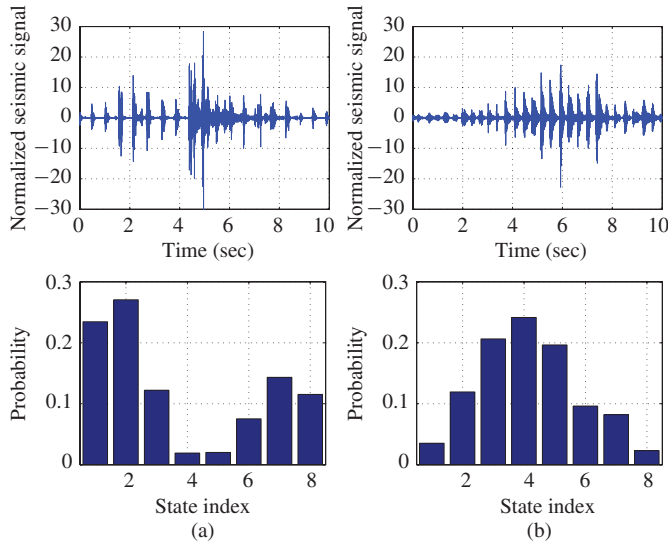|	No target	Human		Animal
		Walking	Running	
No target	110	0	0	0
Human — Walking	1	33	7	7
Human — Running	0	5	13	0
Animal	0	2	0	42

Fig. 7. Examples of seismic sensor measurements (top) and the corresponding feature vectors extracted by SDF (bottom) for human walking and running. (a) Walking. (b) Running.

kernel for SVM. Pattern classification consumes about 80% of the total execution time because by using leave-one-out cross-validation, the pattern classifier need to be trained with all the remaining patterns (e.g., 235 in detection stage). The choice of *quadratic* kernel in SVM improves the performance of the classifier; however, it also increases the computation in training the classifier. It is expected the execution time and memory will be reduced significantly if fewer training patterns are used.

2) Movement Type Identification: Upon recognition of human, more information can be derived by performing another binary classification to identify whether the human is running or walking. The physical explanations are: i) the cadence (i.e., interval between events) of human walking is usually larger than the cadence of human running; ii) the impact of running on the ground is much stronger than that of walking, and it takes longer for the oscillation to decay. Figure 7 shows the seismic signal and corresponding feature vectors of human walking and running. The feature vectors of human walking and running are very different from each other, which is a clear indication that the SDF-based feature extraction method is able to capture these features (cadence and impact). It is noted that the feature vectors shown in Fig. 7 are different from those in Fig. 6 because different partitions are used in the target classification and movement type identification stages.

Ideally, the identification of movement type should be performed based on the results of human classification. However, in order to assess the performance of SDF in this particular application, a binary classification between human walking and human running is directly performed. The results are listed in Table IV, where the proposed feature extraction algorithm and SVM are able to identify the human movement type with an accuracy of about 91%.

As stated in Subsection III-C as well as in earlier publications [13], [14], the alphabet Σ in the SDF algorithm plays an important role for target detection and classification. An example illustrating the effects of the alphabet size $|\Sigma|$ on human movement type identification is presented in Fig. 9, where the human movement type identification was performed with $|\Sigma|$ varying from 2 to 20. It is seen that that the classification accuracy is consistent within the range of $|\Sigma|$ from 2 to 20. The rationale is that the information loss increases with a smaller $|\Sigma|$ and robustness to noise decreases with a larger $|\Sigma|$.

3) Target Payload Identification: Similar with the movement type identification shown above, the target payload information can also be derived by performing another binary classification for both animal and human targets. Figure 8 shows the seismic signals and feature vectors of human/animal with and without payload examples. It is observed that the feature vectors extracted by SDF has large inter-class separation while small intra-class variance, and yet the intra-class differences between the with payload and without payload cases are still distinguishable.

Table V shows the results of the human/anmal payload identification. The shaded area in Table V represents the payload identification. It is seen that the proposed method is able to distinguish human from animal with high accuracy (97.3%). The payload identification result is also reasonable (human: 81.3%, animal: 79.1%); however, more than half samples in the human with payload and animal without payload cases are incorrectly classified. Three factors may contribute to low classification rate for these two classes: i) the payloads are not the same throughout all the experiments; ii) the weight of the payload is only a small fraction of the weight of human/animal target, so difference between the two classes (with payload/without payload) are not obvious; iii) Unbalanced number of samples in each class. The first two issues are related with data collection; the last issue may be resolved by increasing the weight of the class with fewer samples when generating the separating hyperplane in SVM.

B. Performance Assessment Using PIR Data

PIR sensors are widely used for motion detection. In most applications, the signals from PIR sensors are used as discrete variables (i.e., on or off). This may work for target detection, but will not work well for target classification because the time-frequency information is lost in the discretization. In this paper, the PIR signals are considered to be continuous signals, and continuous wavelet transform (CWT) is used to reveal the distinction among different types of targets in the time-frequency domain. Since a PIR sensor does not emit an
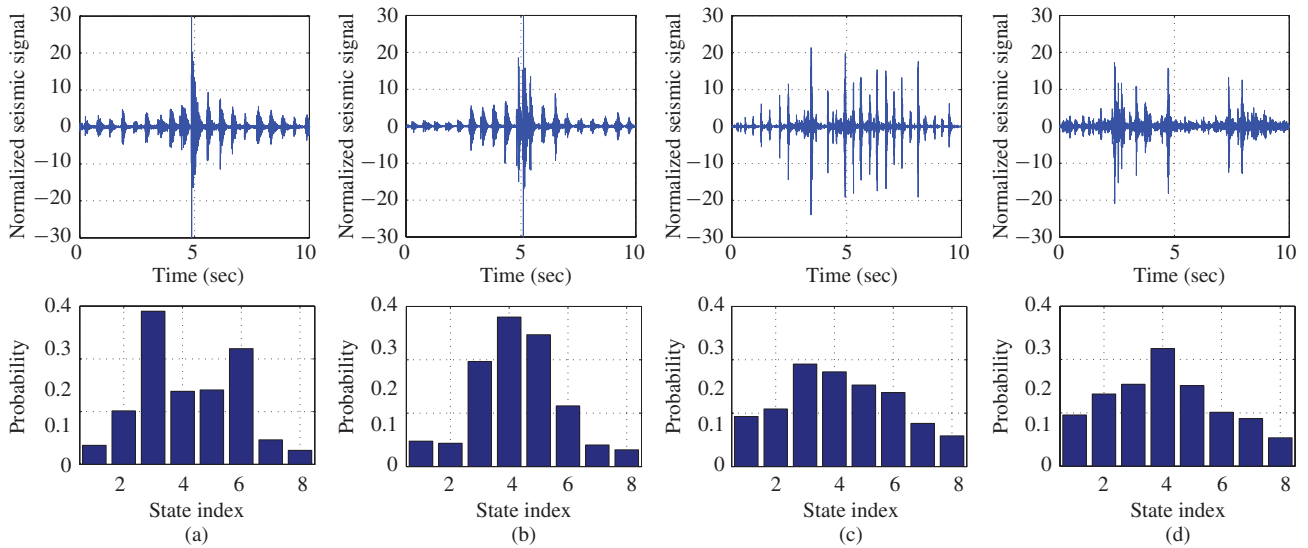
Fig. 8. (a) Human without payload. (b) Human with payload. (c) Animal without payload. (d) Animal with payload. Examples of seismic sensor measurements (top) and the corresponding feature vectors extracted by SDF (bottom) for payload identification.
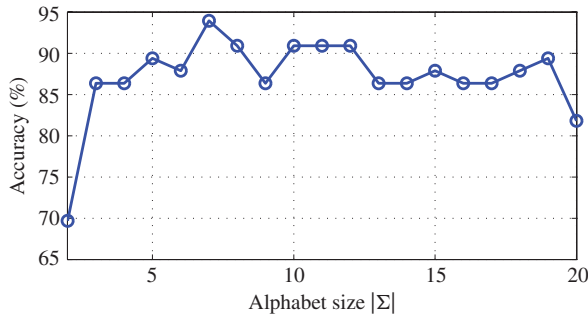

Fig. 9. Effect of alphabet size on human movement type identification.

infrared beam but merely passively accepts incoming infrared radiation, it is less sensitive to environmental variations (i.e., variation in test sites) than the seismic sensor that is ground-based. A three-way cross-validation [18] is used for the performance assessment of PIR data. The data are divided into three sets by date (i.e., Day 1, Day 2, and Day 3) and three different sets of experiments are performed:

1) Training: Day 1 + Day 2; Testing: Day 3
2) Training: Day 1 + Day 3; Testing: Day 2
3) Training: Day 2 + Day 3; Testing: Day 1.

Training and testing on feature vectors from different days is very meaningful in practice. In each run of the cross-validation, no prior information is assumed for the testing site or the testing data. The classifiers' capability to generalize to an independent data set is thoroughly tested in the three-way cross-validation. In this section, four types of targets are considered, namely, no target, human walking, human running, and animal led by human. Following Fig. 5, the following cases are tested:

1) Detection of target presence against target absence;
2) Classification of target type, i.e., Human vs. Animal;
3) Classification of target movement type (i.e., walking vs. running) upon recognition of the target as human.

Figure 10 shows the PIR sensor measurements (top) and the corresponding feature vectors extracted by SDF (bottom) of the four classes. For the no target case, the PIR signal fluctuates around zero and no information is embedded in the wavelet coefficients, thus the states in the middle (i.e., states 3-10) are occupied; whereas for the target present cases, the PIR sensors are excited by the presence of the targets, so states 1-2 and 11-12 that correspond to the crests and troughs in the PIR signals are more populated than other states.

The following parameters are used in SDF and SVM for processing the PIR signals: alphabet size $|\Sigma| = 12$, number of scales $|\alpha| = 3$, window size $\ell \times \ell = 2 \times 2$, number of most probable symbol $m = 1$, and *quadratic* kernel for SVM. The execution of SDF and SVM takes 1.13 seconds and 39.83 MB of memory on a desktop computer to process a data set of 1×10^4 points, which is a clear indication of the real-time implementation capability for onboard UGS systems.

Table VI shows the confusion matrix of the three-way cross-validation results using PIR sensors. The shaded area represents the target classification stage. It is seen in Table VI that the proposed feature extraction algorithm works very well with the PIR sensor; the target detection accuracy is 99.5%, the human/animal classification accuracy is 91.7%, and the human movement type classification accuracy is 79.3%. Leave-one-out cross-validation usually underestimates the error rate in generalization because more training samples are available; it is expected that the classification accuracy will further improve for the PIR signals if leave-one-out cross-validation is used.

C. Field Deployment of Seismic and PIR Sensors

Seismic and PIR sensors have their own advantages and disadvantages for target detection and classification. The seismic sensor is omnidirectional and has a long range of detection (up to 70 m) [10], whereas a PIR sensor has a typical range of less than 6 m and has a limited field of view (less than 180°), which restricts the sensor from detecting target moving behind it. The seismic sensor is not site-independent and is
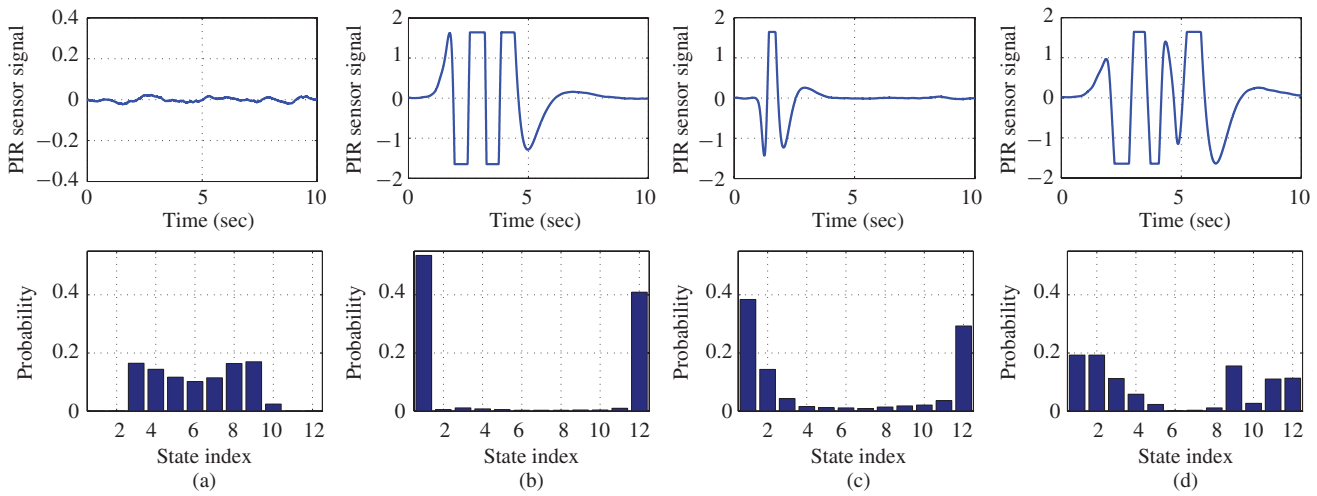
Fig. 10.   Examples of PIR sensor measurements (top) and the corresponding feature vectors extracted by SDF (bottom) of the four classes. (a) No target. (b) Human walking. (c) Human running. (d) Animal led by human.

vulnerable to variations in sensor sites, whereas a PIR sensor merely passively accepts the incoming infrared radiation and is independent of the sensor site. In order to improve the detection and classification accuracy while reducing the false alarm rate, it is recommended that the seismic and PIR sensor should be used together to provide complementary information to each other. Information fusion techniques are needed to combine the outputs of the two sensing modalities, and this is a topic of future research.

Field deployment of sensors largely depends on the tasks and terrains. To enhance the perimeter security [7] in an open field, the sensors are usually deployed linearly or in circles. Since the intruder may approach the secure region from any direction, the worst case scenario is when the intruder approaches the secure region exactly half-way between two adjacent sensors along a straight path perpendicular to the sensor picket line. To ensure intruder detection, the maximum sensor spacing should be less than the effective range of the sensor. Therefore, sensor deployment could be very expensive, because the detection range of PIR sensors is less than 6 *m*. Another critical application is sensor deployment at choke points; since the targets are forced to pass the choke point due to the terrain difficulties, a single node of UGS system can be sufficient to cover the entire region.

V. Conclusion

This paper presents a symbolic feature extraction method for target detection and classification, where the features are extracted as statistical patterns by symbolic dynamic modeling of the wavelet coefficients generated from time series of seismic and PIR sensors. By appropriate selection of wavelet basis and scale range, the wavelet-transformed signal is denoised relative to the original time-domain signal. In this way, the symbolic images generated from wavelet coefficients capture the signal characteristics with larger fidelity than those obtained directly from the time domain signal. The symbolic images are then modeled using probabilistic finite state automata (PFSA) that, in turn, generate low-dimensional

statistical patterns, also called feature vectors. A distinct advantage of the proposed feature extraction method is that the low-dimensional feature vectors can be computed in-situ and communicated in real time over a limited-bandwidth wireless sensor network with limited-memory nodes.

The proposed method has been validated on a set of field data collected from different locations on multiple days. A comparative evaluation is performed on the seismic signals between SDF and kurtosis analysis using leave-one-out cross-validation. Results show that SDF has superior performance over kurtosis analysis, especially in the human/animal classification. In addition, the capabilities for identifying movement type and target payload are examined for the seismic sensor. A three-way cross-validation has been used to assess the performance of PIR sensors for target detection and classification. Results show that PIR sensors are very good for target detection, and has comparable performance with seismic sensors for target classification and movement type identification.

While there are many research issues that need to resolved before exploring commercial applications of the proposed method, the following topics are under active research:

1) Enhancement of target detection and classification performance by fusion of seismic and PIR sensor signals
2) Real-time field implementation of the proposed method on low-cost low-power microprocessors for different types of deployment (e.g., UGS fencing to secure a region).

References

[1] G. L. Graham, "Detection and classification for unattended ground sensors," in *Proc. Inform. Dec. Control*, 1999, pp. 419–424.
[2] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 17–29, Mar. 2002.
[3] A. G. Richard, M. B. Bennett, and D. T. O'Brien, "Vehicle and personnel detection using seismic sensors," in *Proc. Sens. C3I Inform. Training Technol. Law Enforcement*, Boston, MA, 1999, pp. 74–85.
[4] J. Altmann, "Acoustic and seismic signals of heavy military vehicles for co-operative verification," *J. Sound Vibrat.*, vol. 273, nos. 4–5, pp. 713–740, 2004.

[5] Y. Tian and H. Qi, "Target detection and classification using seismic signal processing in unattended ground sensor systems," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, Orlando, FL, 2002, p. 4172.

[6] G. P. Succi, D. Clapp, R. Gampert, and G. Prado, "Footstep detection and tracking," *Proc. SPIE*, vol. 4393, pp. 22–29, Apr. 2001.

[7] J. Lacombe, L. Peck, T. Anderson, and D. Fisk, "Seismic detection algorithm and sensor deployment recommendations for perimeter security," *Proc. SPIE*, vol. 6231, p. 623109, Apr. 2006.

[8] M. Kenneth and P. D. McGaffigan, "Spectrum analysis techniques for personnel detection using seismic sensors," *Proc. SPIE*, vol. 5090, pp. 162–173, Sep. 2003.

[9] Y. Shoji, T. Takasuka, and H. Yasukawa, "Personal identification using footstep detection," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, 2004, pp. 43–47.

[10] L. Peck and J. Lacombe, "Seismic-based personnel detection," in *Proc. IEEE 41st Ann. Int. Carnahan Conf. Security Technol.*, Ottawa, ON, Canada, 2007, pp. 169–175.

[11] Z. Zhang, X. Gao, J. Biswas, and K. K. Wu, "Moving targets detection and localization in passive infrared sensor networks," in *Proc. 10th Int. Conf. Inform. Fusion*, Jul. 2007, pp. 1–6.

[12] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Signal Process.*, vol. 84, no. 7, pp. 1115–1130, 2004.

[13] V. Rajagopalan and A. Ray, "Symbolic time series analysis via wavelet-based partitioning," *Signal Process.*, vol. 86, no. 11, pp. 3309–3320, Nov. 2006.

[14] S. Gupta and A. Ray, "Symbolic dynamic filtering for data-driven pattern recognition," in *Pattern Recognition: Theory Application*. Hauppage, NY: Nova Science Publishers, 2007.

[15] X. Jin, S. Gupta, K. Mukherjee, and A. Ray, "Wavelet-based feature extraction using probabilistic finite state automata for pattern classification," *Pattern Recognit.*, vol. 44, no. 7, pp. 1343–1356, 2011.

[16] P. Abry, *Ondelettes et Turbulence, Multirésolutions, Algorithmes de Décomposition, Invariance Déchelles*. Paris, France: Diderot Editeur, 1997.

[17] D. Lind and M. Marcus, *An Introduction to Symbolic Dynamics and Coding*. Cambridge, U.K.: Cambridge Univ. Press, 1995.

[18] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.

Xin Jin (S'09) received the B.S. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in July 2007, and the M.S. degree in electrical engineering and mechanical engineering from Pennsylvania State University, University Park, where he is currently pursuing the Ph.D. degree in mechanical engineering.

His current research interests include machine learning, signal processing, control systems, energy management, and robotics.

Mr. Jin is a Student Member of the American Society of Mechanical Engineers, the International Society of Automation, and the American Nuclear Society.

Soumalya Sarkar received the B.S. degree in mechanical engineering from Jadavpur University, Kolkata, India, in 2010. He is currently pursuing the M.S. degree in mechanical engineering from Pennsylvania State University, University Park.

His current research interests include machine learning, sensor networks, and fault diagnostics.

Asok Ray (SM'83–F'02) received the Ph.D. degree in mechanical engineering from Northeastern University, Boston, MA, and Graduate degrees in the disciplines of electrical engineering, mathematics, and computer science.

He joined Pennsylvania State University (Penn State), University Park, in July 1985, and is currently a Distinguished Professor of Mechanical Engineering and a Graduate Faculty Member of electrical engineering. Prior to joining Penn State, he held research and academic positions with the Massachusetts Institute of Technology, Cambridge, and Carnegie-Mellon University, Pittsburgh, PA, as well as research and management positions with GTE Strategic Systems Division, Washington D.C., Charles Stark Draper Laboratory, Cambridge, and MITRE Corporation, Bedford, MA. He has authored or co-authored over 500 research publications including 250 scholarly articles in refereed journals and research monographs.

Dr. Ray is a Fellow of the American Society of Mechanical Engineers and the World Innovative Foundation. He had been a Senior Research Fellow with the NASA Glenn Research Center under a National Academy of Sciences Award.

Shalabh Gupta (S'03–M'06) received the M.S. degree in mechanical and electrical engineering from Pennsylvania State University (Penn State), University Park, and the Ph.D. degree in mechanical engineering from Penn State in August 2006.

He joined the Department of Electrical and Computer Engineering, University of Connecticut (UConn), Storrs, as an Assistant Professor in August 2011. Before joining UConn, he was a Post-Doctoral Research Scholar and as a Research Associate with the Mechanical Engineering Department, Penn State. His research efforts have been instrumental in opening new fields of data understanding and pattern discovery, via interfacing multidisciplinary concepts derived from statistical mechanics, symbolic dynamics, and automata theory. His current research interests include science of autonomy, swarm robotics, intelligent systems, machine learning, network science, and fault detection and isolation in complex systems.

Dr. Gupta is a member of the American Society of Mechanical Engineers.

Thyagaraju Damarla (M'84–SM'95) received the B.S. and M.S. degrees from the Indian Institute of Technology, Kharagpur, India, and the Ph.D. degree from Boston University, Boston, MA.

He is currently an Electronics Engineer with the U.S. Army Research Laboratory, Adelphi, MD. He was with the Indian Space Research Organization, Sriharikota, India, from 1973 to 1979, and the Indian Institute of Technology, Kanpur, India, from 1979 to 1982. He was an Assistant Professor with the University of Kentucky, Lexington. He published more than 100 technical papers in various journal and conferences and has three U.S. patents. His current research interests include signal processing and sensor fusion for situational awareness.