Meas. Sci. Technol. **20** (2009) 123001 (6pp) doi:10.1088/0957-0233/20/12/123001

RAPID COMMUNICATION

Asynchronous data-driven classification of weapon systems

Xin Jin[1], Kushal Mukherjee[1], Shalabh Gupta[1], Asok Ray[1], Shashi Phoha[1] and Thyagaraju Damarla[2]

[1] The Pennsylvania State University, University Park, PA 16802, USA
[2] US Army Research Laboratory, Adelphi, MD 20783, USA

E-mail: xuj103@psu.edu, kum162@psu.edu, szg107@psu.edu, axr2@psu.edu, sxp26@psu.edu and rdamarla@arl.army.mil

Received 22 June 2009, in final form 8 October 2009
Published 29 October 2009
Online at stacks.iop.org/MST/20/123001

Abstract

This communication addresses real-time weapon classification by analysis of asynchronous acoustic data, collected from microphones on a sensor network. The weapon classification algorithm consists of two parts: (i) feature extraction from time-series data using symbolic dynamic filtering (SDF), and (ii) pattern classification based on the extracted features using the language measure (LM) and support vector machine (SVM). The proposed algorithm has been tested on field data, generated by firing of two types of rifles. The results of analysis demonstrate high accuracy and fast execution of the pattern classification algorithm with low memory requirements. Potential applications include simultaneous shooter localization and weapon classification with soldier-wearable networked sensors.

Keywords: weapon systems, pattern classification, symbolic dynamic filtering, language measure, support vector machine

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Sniper attacks have become one of the major sources of casualty in asymmetric warfare, especially in urban environments. A counter-sniper system that assists identification of the shooter's location and weapon class information would significantly reduce the potential peril of both soldiers and civilian population. Counter-sniper systems make use of several different physical phenomena that are related to weapon data (e.g., acoustic, visual or electromagnetic signals). In spite of the wide range of possible measurement devices, acoustic signals (e.g., muzzle blast, shockwave and surface vibration [1]) apparently provide the most convenient and accurate way to identify sniper shots. Hence, a majority of existing counter-sniper systems use acoustic signals as the primary information source [2]. In recent years, several commercial sniper localization systems have been developed [3–5], which have not taken weapon classification into consideration. To this end, Volgyesi *et al* [6] developed a soldier-wearable sensor network system for both shooter localization and weapon classification. It estimates the trajectory, range, caliber and weapon type using data from a single sensor or fusion of multiple time-synchronized sensors, where the weapon classification is dependent on the localization. To alleviate the requirement of time synchronization, Damarla *et al* [7] developed a sniper localization method for a network of sensors, which relies only on the *time difference of arrival* (TDOA) between the muzzle blast and shock wave from multiple single sensor nodes, relaxing the need for precise time synchronization across the network.

The goal of the research work explored in this communication is to formulate a real-time weapon classification algorithm based on asynchronous time series data collected from microphones on a sensor network. The

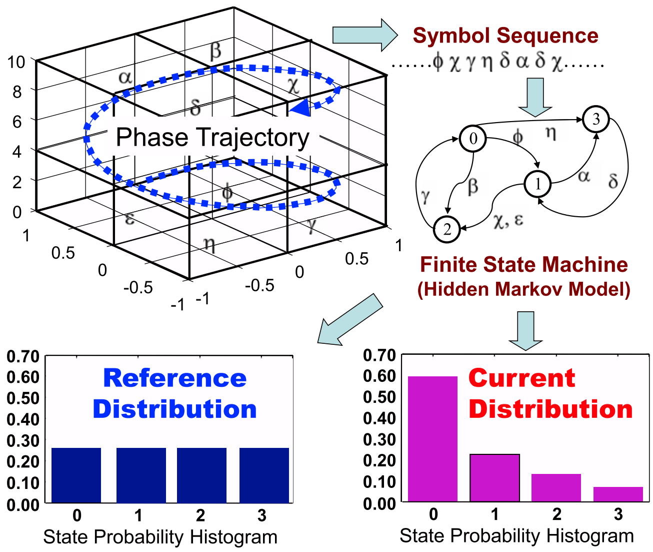


Figure 1. Concept of symbolic dynamic filtering (SDF).

major contribution of this communication is formulation of a weapon classification algorithm that is capable of real-time feature extraction and pattern classification directly from the time series of asynchronous acoustic signals, collected from networked sensors. This algorithm is independent of sniper location and thus is expected to be more reliable under adverse conditions in battlefields than existing weapon classification systems [6].

2. Review of underlying mathematical concepts

The weapon classification algorithm is built upon the following major concepts: (i) symbolic dynamic filtering (SDF) [8, 9] for feature extraction, and (ii) language measure (LM) [10, 11] and support vector machine (SVM) [12, 13] for pattern classification. While the theories of SDF, LM and SVM are reported in detail in the existing literature, this section briefly presents the underlying concepts that have significant relevance to this communication.

2.1. Symbolic dynamic filtering (SDF) for feature extraction

Figure 1 pictorially elucidates the concepts of partitioning a finite region of the phase space of the dynamical system under consideration and a mapping from the partitioned space into the symbol alphabet, where the symbols are indicated by Greek letters (e.g., $\alpha, \beta, \gamma, \delta, \ldots$). It also shows conversion of the symbol sequence into a probabilistic finite-state automaton (PFSA) and generation of the state probability vectors at the current and the reference conditions. The states of the PFSA and the associated histograms in figure 1 are indicated by numerics (i.e., 0, 1, 2 and 3). Feature extraction from training data consists of the following steps.

2.1.1. Collection of time series data. Sensor time series data, generated from a physical system or its dynamical model, are collected as training or test sets over the range of operation. A compact (i.e., closed and bounded) region $\Omega \in \mathbb{R}^n$ (i.e., the n-dimensional real space), where $n \in \mathbb{N}$ (i.e., the set of positive integers), within which the quasi-stationary time series is circumscribed, is identified. Let the space of time series data sets be represented as Y, where the number of data points in each time series is sufficiently large for convergence of statistical properties within a specified threshold [9]. Then, $\mathbf{y}^i = \{y_1^i, y_2^i, \ldots\} \in Y$ denotes a time series for the data set $i \in \{0, 1, \ldots, l-1\}$, where l is the number of time series data sets under consideration.

2.1.2. Space partitioning. Time series data are encoded by introducing a partition $\mathbb{B} \equiv \{B_0, \ldots, B_{(m-1)}\}$ that consists of m mutually exclusive (i.e., $B_j \cap B_k = \emptyset \; \forall j \neq k$), and exhaustive (i.e., $\cup_{j=0}^{m-1} B_j = \Omega$) cells. Let each cell be labeled by symbols $\sigma_j \in \Sigma$ where $\Sigma = \{\sigma_0, \ldots, \sigma_{m-1}\}$ is called the *alphabet*. This process of coarse graining can be executed by appropriate partitioning (e.g., uniform or maximum entropy [14]) of the data set. Then, the data points of the reference time series $\mathbf{y}^0$ which visit the cell B_j are assigned the corresponding symbol as $\sigma_j \; \forall j = 0, 1, \ldots, m-1$. This step enables transformation of the reference time series $\mathbf{y}^0$ to a symbol sequence $\mathbf{s}^0 = \{s_1^0, s_2^0, \ldots\}$, where each $s_j^0 \in \Sigma$. To alleviate the difficulties associated with noisy time series, symbolization is carried out by Hilbert transform-based analytical signal space partitioning (ASSP) [14], which is an essential ingredient of SDF analysis in the proposed weapon classification algorithm. Symbol sequences $\mathbf{s}^1, \ldots, \mathbf{s}^{l-1}$ are generated from the respective time series, $\mathbf{y}^1, \ldots, \mathbf{y}^{l-1}$, using the same partitioning for generation of the symbol sequence $\mathbf{s}^0$ from the reference time series $\mathbf{y}^0$.

2.1.3. Construction of probabilistic finite state automata (PFSA). Probabilistic finite state automata (PFSA) are constructed [8] with a chosen depth D, and the corresponding $(r \times r)$ state transition matrices $\mathbf{\Pi}^i = [\pi_{jk}^i]$ are generated by running the symbol sequences through the PFSA structure; the pair of subscripts $j, k \in \{1, 2, \ldots, r\}$ denotes a state transition from j to k and the superscript i denotes the ith training data set, $i \in \{0, 1, \ldots, l-1\}$. Since $\pi_{jk}^i \geqslant 0$ is the transition probability from state j to state k, $\mathbf{\Pi}^i$ is a stochastic matrix, i.e., $\sum_k \pi_{jk}^i = 1 \;\; \forall j \in \{1, 2, \ldots, r\}$.

2.2. Pattern classification

A threshold-based binary classifier is constructed in terms of a signed scalar measure [10, 11] of the language generated by the PFSA $\mathbf{\Pi}$. The classification logic is as follows:

$$\text{class} = \begin{cases} C_1: & \text{if } \nu < 0 \\ C_2: & \text{if } \nu > 0, \end{cases} \tag{1}$$

where ν is the language measure obtained by assigning a weight to each state of the PFSA $\mathbf{\Pi}$.

Definition 2.1. *The characteristic vector* $\boldsymbol{\chi}$ *assigns a signed real weight to each of the r states of the PFSA, where larger weights are assigned to relatively more desirable states. The* $(1 \times r)$ *characteristic vector is defined as*

$$\boldsymbol{\chi} = [\chi_1 \; \chi_2 \cdots \chi_r]. \tag{2}$$

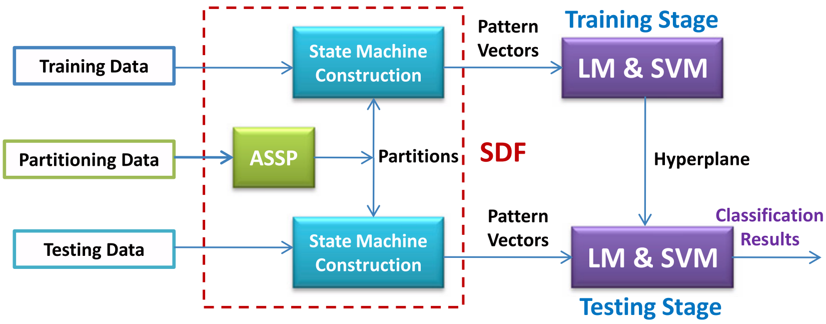


Figure 2. Flow chart of the proposed methodology.

Definition 2.2. *Measure of the language generated by a PFSA in terms of its characteristic vector* χ *is defined as*

$$\bar{\nu}(\theta) = \theta[I - (1-\theta)\mathbf{\Pi}]^{-1}\chi^T \quad \text{where} \quad \theta \in (0, 1). \tag{3}$$

Proposition 2.1. *The measure* $\bar{\nu}(0) \triangleq \lim_{\theta\to 0^+} \bar{\nu}(\theta)$ *exists and is bounded as* $\|\bar{\nu}(0)\|_\infty \leqslant \|\chi\|_\infty$.

Proof. Given in [11]. □

Proposition 2.2. *Given a primitive (i.e., irreducible and acyclic) state transition matrix* $\mathbf{\Pi}$*, the measure in equation (3) reduces to* $\bar{\nu}(0) = \nu\mathbf{1}$ *in the limit, where* $\mathbf{1} \triangleq [1\,1\,\ldots\,1]^T$*. Then, the scalar measure* ν *is denoted as [11]*

$$\nu = \boldsymbol{p}\chi^T, \tag{4}$$

where $\boldsymbol{p} = [p_1, p_2, \ldots, p_r]$ *is the* $(1\times r)$ *state probability vector that is the (sum-normalized) left eigenvector of* $\mathbf{\Pi}$ *corresponding to its unique unity eigenvalue [15].*

Proof. Given in [11]. □

The scalar measure ν is of the form

$$\begin{aligned}\nu = \boldsymbol{p}\chi^T &= \sum_{i=1}^{r} p_i\chi_i = \sum_{i=1}^{r-1} p_i\chi_i + p_r\chi_r \\ &= \sum_{i=1}^{r-1} p_i\chi_i + \left(1 - \sum_{i=1}^{r-1} p_i\right)\chi_r \\ &= \sum_{i=1}^{r-1} p_i(\chi_i - \chi_r) + \chi_r = \sum_{i=1}^{r-1} p_i a_i + b,\end{aligned} \tag{5}$$

where $a_i \triangleq (\chi_i - \chi_r)$ and $b \triangleq \chi_r$. Therefore, the scalar measure ν is an affine transformation of the $(r-1)$ independent state probabilities, where the rth state probability may be expressed in the form $1 - \sum_{i=1}^{r-1} p_i$. Further, the equi-measure surfaces in space of probability vectors are hyperplanes as described by equation (5). The classifier construction involves computation of the values of $a_i, i = 1, \ldots, r-1$ and b, that in turn may be used to evaluate the characteristic vector $\chi = [\chi_1 \cdots \chi_r]$. The classifier follows the form of equation (1).

The support vector machine (SVM) [12, 13] estimates χ and simultaneously maximizes the separation margin between the two classes. In the linearly separable case, the binary SVM classifier selects a hyperplane to separate the two classes by maximizing the margin that is defined as the sum of the distances of the hyperplane to the closest points of the two classes. If the two classes are non-separable, positive slack variables are introduced to allow some of the training samples to fall on the wrong side of the separating hyperplane. The SVM then finds the hyperplane that maximizes the margin and simultaneously minimizes a quantity directly related to the number of classification errors. This procedure, called the 'soft margin' method, is an extension of the linear SVM [16].

Figure 2 depicts the flow chart of the proposed classification algorithm that is constructed based on the theories of SDF, LM and SVM. Upon collection of acoustic time series data, symbol sequences are generated by analytic signal space partitioning (ASSP) [14] that is invariant for both training and testing stages. In the training stage, l_{train} sets of rifle-shot time series data with known class labels are selected. A PFSA structure is constructed using SDF; subsequently, a feature vector $\boldsymbol{p}^i, i \in \{0, \ldots, l_{\mathrm{train}} - 1\}$ is generated for each set of the rifle-shot data. The time series data belong to exactly one of the two known classes C_1 and C_2.

The feature vectors are inputs to the LM and SVM module that generates a hyperplane that maximizes the margin and minimizes classification errors between feature vectors of the training data. A linear kernel has been used in this communication. In the testing stage, the feature vectors $\boldsymbol{p}^i, i \in \{0, \ldots, l_{\mathrm{test}} - 1\}$ are generated from SDF with unknown class labels and are then separated by the hyperplane obtained in the training stage. The SM and SVM algorithm yields a binary output (i.e., C_1 or C_2) as the class labels of the testing data.

3. Results of field data analysis

Figure 3 shows a data collection scenario for classification of two weapon types, namely, Rifle 1 and Rifle 2 that were fired from two different locations that are ~60 m apart. Eight microphone sensors are distributed over a region of ~30 m × 30 m around each of the three aim-points that are ~250 m down range.

For each of the three aim-points, the data set of Rifle 1, which is generated by firing Rifle 1 from location B, has been used for partitioning; these data sets are not used in the training stage or testing stage. Referring to section 2, a maximum-entropy analytic signal space partitioning (ASSP) [14] is generated in the radial direction with $|\Sigma_R| = 4$ segments and, in the angular direction, with $|\Sigma_A| = 1$ segment [14]; thus, the

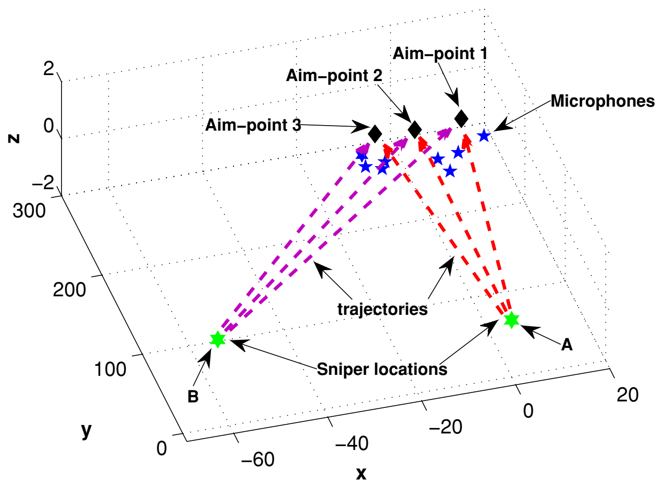
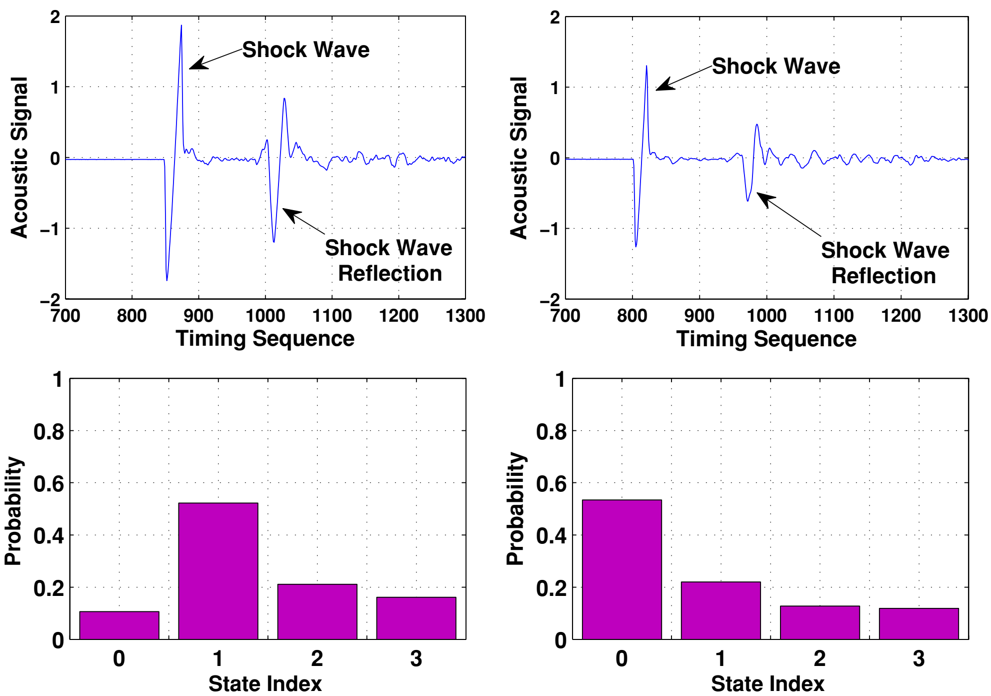


Figure 3. A data collection scenario with eight microphones and two rifle locations.

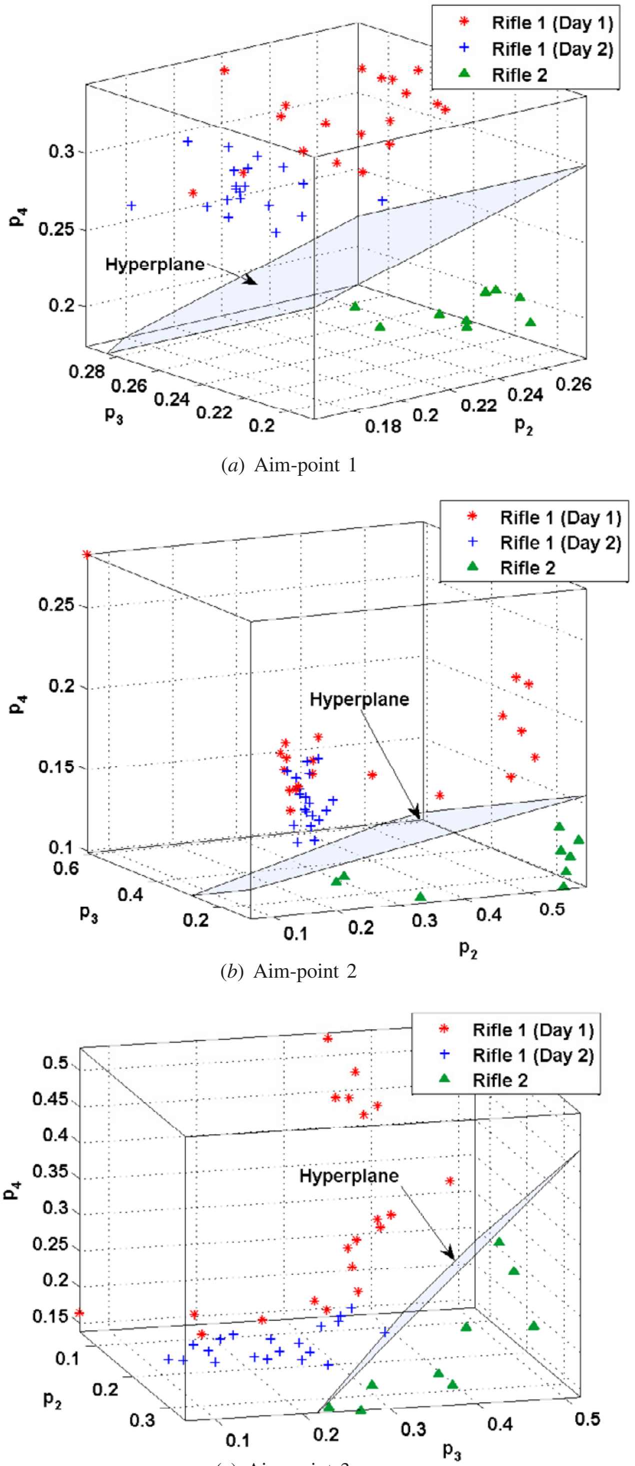
alphabet size is $|\Sigma| \triangleq |\Sigma_A| \times |\Sigma_R| = 4$. Thus, for each aim-point, there is a unique partition that is kept invariant in both training and testing stages. For construction of PFSA from symbol sequences, a depth of $D = 1$ is found to be adequate to capture sufficient information for pattern classification in this application [8]. Each feature vector obtained as the (stationary) state probability vector of the PFSA represents the respective signature of the rifle acoustic signal. The top row of figure 4 exhibits a typical time series of acoustic data from Rifle 1 and Rifle 2, which are collected from the same microphone; examples of feature vectors for the two types of rifles are shown as histograms in the bottom row of figure 4, which display how the structure of the underlying probability distribution varies with the rifle type.

Three plots, arranged as a vertical column in figure 5, present the results of classification for two types of rifles firing from location A at aim-points 1, 2 and 3, respectively, based on the time series data of Microphone 4. No occurrence of pattern classification errors is observed in these experiments for this specific choice of feature extraction and pattern classification parameters. For each aim-point, the respective plot in figure 5 exhibits the results of classification for a single set of Rifle 2 data and a combination of two different sets of Rifle 1 data that have been collected on two different days. For each aim-point, about a half of the data in each set are used for training the SVM classifier and the remaining data for testing the classification algorithm. As stated earlier, the partition of individual data sets for each aim-point is generated based on the respective data set of Rifle 1 fired from the building as indicated by firing location B in figure 3. In figure 5, the feature extracted from each data set is represented by a vector that belongs to the four-dimensional real space $\mathbb{R}^4$ because $|\Sigma| = 4$ and $D = 1$. Since the elements of each feature vector are stationary probabilities of the four states of a PFSA, the sum of the (positive) elements of each feature vector is unity (i.e., belonging to the three-dimensional simplex). Therefore, only three elements of the feature vector are linearly independent, implying that the decision space is three-dimensional in this setting. For each of the three plots in figure 5, the two-dimensional hyperplane unambiguously separates the patterns of Rifle 2 from those of Rifle 1. In this way, a time series of rifle data is reliably tested for identification of the unknown rifle type, namely, Rifle 1 or Rifle 2 by observing to which side of the hyperplane the feature vector belongs.



(*a*) Rifle 1 (*b*) Rifle 2

Figure 4. Acoustic signals and respective feature vectors.



(*a*) Aim-point 1

(*b*) Aim-point 2

(*c*) Aim-point 3

Figure 5. Rifle classification at aim-points 1, 2 and 3 (Microphone 4).

The execution time of the algorithm for each plot in figure 5 is less than 1.5 s on a desktop computer, which demonstrates its real-time execution capability.

3.1. Channel effects on classification results

This section presents a summary of channel effects (i.e. the degradation of acoustic information due to propagation through the atmosphere) on the classification results, which are attributed to: (i) variations in the recorded acoustic signals due to relative positions of microphones and calibration settings; (ii) environmental conditions such as vegetation, terrain and urban buildings that influence the channel. While the first cause is mainly a hardware issue that can be mitigated by appropriate calibration, the second cause (i.e., environmental effects) is the focus of this section.

In real-world applications (e.g., an urban environment with tall buildings and various other objects), the shock wave and the muzzle blast from sniper firing are subject to reflection, attenuation, absorption, diffraction and other wave modifications as they propagate through the atmosphere. A microphone placed in the vicinity of an aim-point receives pressure waves arriving directly from the source and waves arriving later from other directions due to reflections and scattering. At distances far from the rifle shot trajectory, the shock wave is expected to disperse sufficiently by spatial spreading such that it may no longer be detectable compared to the ambient noise [1]. The situation becomes much more complicated if the surroundings include obstacles and reflecting surfaces so that the received acoustic signal contains multipath interference, diffraction effects and other propagation-related flaws. In essence, the environmental effects could be totally different at different sensor locations. An ideal weapon classification system should be independent of the channel effects due to environmental variations.

The results presented earlier in this section make use of the acoustic time series data from a single microphone, namely, Microphone 4 for weapon classification. This subsection reports the results obtained based on the data from all eight microphones, including Microphone 4, to investigate the impact of environmental effects on the weapon classification. The microphones are placed in different locations and have varying levels of echo/reflection due to their slightly different environment. Table 1 summarizes the classification results obtained from all eight microphones for each of the three aim-points. The total number of tests (i.e., rifle shots) for each aim-point is 24, and the classification success rate is calculated by subtracting the ratio of the number of false classifications over the total test number from 1. It is seen in table 1 that a majority of the microphones have a very high classification success rate. The classification success in Microphone 1 is slightly lower than that in other microphones due to improper calibration as recorded in the original log of

Table 1. Rifle classification results of all microphones.

Microphone no.	Aim-point 1	Aim-point 2	Aim-point 3
1	95.83%	100.00%	83.33%
2	100.00%	100.00%	100.00%
3	100.00%	100.00%	95.83%
4	100.00%	100.00%	100.00%
5	91.67%	100.00%	100.00%
6	91.67%	100.00%	100.00%
7	91.67%	100.00%	100.00%
8	91.67%	91.67%	100.00%
		Average	97.22%

the field test. The average rate of successful classification is 97.22%, which is a clear indication of its reliability and effectiveness. Trade-off between probabilities of successful and false classifications is a topic of future research and is not addressed in this communication.

The quality of classification could be further improved by (i) increasing the number of states in the PFSA and (ii) converting the linear hyperplane in the SVM into a hypersurface by an appropriate nonlinear transformation.

4. Summary, conclusions and future work

This communication presents a data-driven method of real-time weapon classification based on time series data collected from asynchronous microphones on a sensor network. The pattern classification algorithm is feature-based in the sense that it first converts the raw time series data into a feature vector of significantly lower dimension, and then pattern classification is performed based on the extracted feature vectors. The feature-based approach is well suited to the weapon classification problem because direct usage of a large volume of raw data with unknown or partially known noise statistics is inefficient from both analytical and computational perspectives. Based on the data set analyzed in this study, the proposed classification algorithm appears to be nearly channel independent, which renders it potentially reliable and effective in the real-world applications.

The discipline of data-driven weapon classification is relatively new and requires further theoretical and experimental research. In this context, the following topics are recommended for future research before execution of a field application of the proposed weapon classification algorithm.

- Extension of the current algorithm to multi-class pattern classification with advanced SVM tools.
- Enhancement of classification performance through usage of multi-sensor information fusion.
- Investigation of the effects of the signal-to-noise ratio and clutter parameters for automatic target recognition (ATR) [17]. In this context, while making a trade-off between probability of false alarms and probability of successful detection, additional costs related to weapon localization could be included in the composite cost functional, which will augment the standard receiver operating characteristics (ROC) curve to a higher dimensional Pareto surface.

Acknowledgments

This work was supported in part by the US Army Research Laboratory and the US Army Research Office under Grant No. W911NF-07-1-0376, by the US Office of Naval Research under Grant No. N00014-09-1-0688, and by NASA under Cooperative Agreement No. NNX07AK49A. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

References

[1] Maher R C 2007 Acoustical characterization of gunshots *Proc. SAFE 2007 (Washington, DC, IEEE Signal Processing Society, 11–13 April 2007)* pp 109–113

[2] Lédeczi A, Nádas A, Völgyesi P, Balogh G, Kusy B, Sallai J, Pap G, Dóra S and Molnár K 2005 Countersniper system for urban warfare *ACM Trans. Sensor Netw.* **1** 153–77

[3] Stoughton R B 1997 SAIC Sentinel acoustic counter-sniper system *Proc. SPIE* **2938** 276–84

[4] Sadler B M, Pham T and Sadler L C 1998 Optimal and wavelet-based shock wave detection and estimation *J. Acoust. Soc. Am.* **104** 955–63

[5] Duckworth G, Barger J, Carlson S, Gilbert D, Knack M, Korn J and Mullen R 1999 Fixed and wearable acoustic counter-sniper systems for law enforcement *Proc. SPIE* **3577** 210–30

[6] Volgyesi P, Balogh G, Nadas A, Nash C and Ledeczi A 2007 Shooter localization and weapon classification with soldier-wearable networked sensors *Proc. MobiSys '07 (11–14 June 2007, San Juan, Puerto Rico)* pp 113–26

[7] Damarla T, Whipps G and Kaplan L 2008 Sniper localization for asynchronous sensors *Proc. 26th Army Science Conf. (1–4 December 2008, Orlando, FL)* Paper No. AP-03

[8] Ray A 2004 Symbolic dynamic analysis of complex systems for anomaly detection *Signal Process.* **84** 1115–30

[9] Gupta S and Ray A 2009 Symbolic dynamic filtering for data-driven pattern recognition *Pattern Recognition: Theory and Application* (Hauppage, NY: Nova Science) chapter 5, pp 17–71

[10] Ray A 2005 Signed real measure of regular languages for discrete-event supervisory control *Int. J. Control* **78** 949–67

[11] Chattopadhyay I and Ray A 2006 Renormalized measure of regular languages *Int. J. Control* **79** 1107–17

[12] Bishop C M 2006 *Pattern Recognition and Machine Learning* (Berlin: Springer)

[13] Vapnik V 1998 *Statistical Learning Theory* (New York: Wiley)

[14] Subbu A and Ray A 2008 Space partitioning via Hilbert transform for symbolic time series analysis *Appl. Phys. Lett.* **92** 084107

[15] Bapat R B and Raghavan T E S 1997 *Nonnegative Matrices and Applications* (Cambridge: Cambridge University Press)

[16] Cortes C and Vapnik V 1995 Support-vector networks *Mach. Learn.* **20** 273–97

[17] Ratches J, Walters C, Buser R and Guenther B 1997 Aided and automatic target recognition based upon sensory inputs from image forming systems *IEEE Trans. Pattern Anal. Mach. Intell.* **19** 1004–19