Techniques for Classifying Acoustic Resonant Spectra*

Randy S. Roberts Paul S. Lewis Jerome T. Chen Octavio A. Vela

Group ESA-MT, Mail Stop J580

Los Alamos National Laboratory

Los Alamos, New Mexico, 87545 USA

Abstract

A second-generation nondestructive evaluation (NDE) system that discriminates between different types of chemical munitions is under development. The NDE system extracts features from the acoustic spectra of known munitions, builds templates from these features, and performs classification by comparing features extracted from an unknown munition to a template library. Improvements over first-generation feature extraction, template construction and classification algorithms are reported. Results are presented on the performance of the system on a large data set collected from surrogate-filled munitions.

1. Introduction

Los Alamos National Laboratory is developing a nondestructive evaluation (NDE) system that discriminates between different types of chemical munitions [1]. The system is based on Acoustic Resonance Spectroscopy (ARS), a technique for studying the resonance structure of mechanical objects [2]. The system allows on-site inspectors to rapidly verify items declared under the Chemical Weapons Convention, a multilateral chemical weapons agreement scheduled to take effect next year [3]. It also provides an alternative to the time consuming and hazardous drill-and-sample technique of munition verification. (See reference [4] for an overview of verification technology.) This paper reports on improvements to the feature extraction, template construction and classification algorithms previously reported in [1]. Additionally, results are presented on the algorithm's performance with a large data set.

Algorithms for the NDE system are based on the premise that chemical munitions filled with the same agent have similar acoustic spectra, and that munitions filled with different agents have significantly different spectra. The system classifies chemical munitions by first constructing a library of templates built from munitions of known fill. An individual template is built by extracting and clustering spectral features (resonant frequencies and associated parameters) from the acoustic spectra of munitions filled with the same agent. An unknown munition is classified by comparing its spectral features to all templates in the library. The munition is declared to belong to the class associated with the template that produces the closest match.

The algorithms presented in this paper have been tested on a large data set. The data set is composed of acoustic spectra obtained from munitions filled with a variety of surrogate compounds. (The surrogate compounds have been designed to mimic the acoustic properties of various chemical agents.) Results are presented that illustrate the potential of this classification technique.

2. Feature Extraction, Clustering and Classification Algorithms

The measurement system, feature extraction, template construction and classification algorithms for the NDE system are described in [1]. Here, we briefly review the measurement system and then focus on modifications made to the algorithms reported in [1] as a result of extensive testing.

The measurement system is designed to collect acoustic spectra from munitions under inspection. It consists of a signal synthesizer, two piezoelectric transducers mounted in a small handset, a receiver and support equipment. The signal synthesizer generates a stepped-frequency sinusoid that is coupled into the munition by one of the piezoelectric transducers. The output signal is received by the other transducer, and is quadrature demodulated and digitized by the receiver. The start and stop frequencies of the sweep are adjust-

* This work was supported by the Defense Nuclear Agency, IACRO 94-840.

1058-6393/95 $4.00 © 1995 IEEE

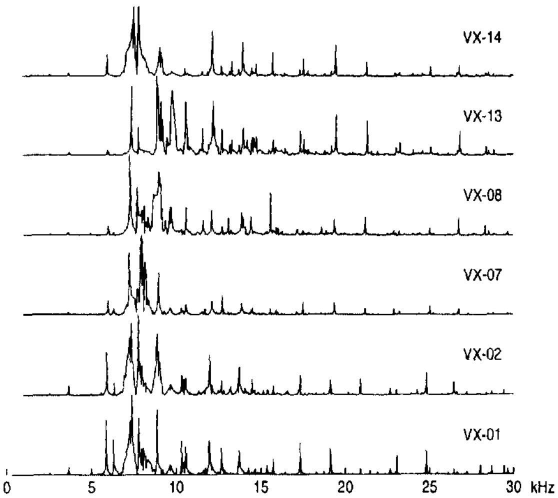

FIGURE 1: Magnitude plots of acoustic spectra, two measurements each from three surrogate-filled munitions.

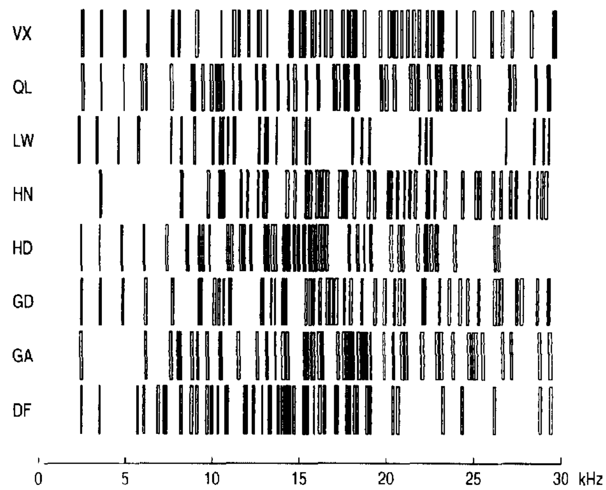

FIGURE 2: Templates constructed from eighteen measurements per surrogate agent. (Six measurements per munition and three munitions per surrogate agent.)

able (a typical sweep is from 1 to 30 kHz), as is the frequency increment (typically set from 15 to 25 Hz). When the system is operated with nominal sweep parameters approximately 2000 complex-valued spectral samples are collected. Figure 1 illustrates magnitude plots of the acoustic spectra of three munitions, with two measurements per munition

The feature extraction algorithm locates peaks in the spectral data and estimates their relative heights. The algorithm begins by computing the magnitude of the spectrum, then normalizes and smooths the spectrum with an $M = 8$ point unity-area Hamming window. Peaks in the smoothed spectrum are located using the first-order differencing procedure described in [1]. The relative height of each peak is then estimated using the following procedure. The i^{th} peak, located at frequency f_i, is surrounded by upper and lower valleys located at frequencies f_u and f_l respectively. The relative height of the i^{th} peak is $H_i = S(f_i) - S(f_v)$ where f_v is the frequency f_u or f_l that satisfies

$$\min(|f_i - f_u|, |f_i - f_l|)$$

and $S(f)$ is the smoothed spectrum. The frequencies of the $N_q = 100$ peaks with the largest relative heights are selected as spectral features.

After features have been extracted from a representative class of munitions, a template for the class is constructed. A template is built by first clustering spectral features and then selecting appropriate clusters. Clustering is performed using a procedure described in [1]. In essence, a frequency window is set so that is spans a fraction of the bandwidth of the collected spectra. Features within the window are grouped into a cluster, and the cluster that has the smallest variance is selected. At least $F_m = 80\%$ of the measurements must contribute to the cluster, or the cluster is abandoned. The frequency window is shifted to the next position and the process is repeated. The output of the clustering algorithm is a table of clusters, where an entry in the table consists of the frequency of the cluster (the mean value of the frequencies that compose the cluster), and the width of the cluster (the standard deviation of the frequencies that compose the cluster).

After all clusters have been formed a selection criteria is applied to limit the number of clusters used in the template [1]. The selection criteria sorts the clusters in order of increasing width, and computes a running mean of the first-order difference of the widths. If a difference exceeds four times the running average, clusters past that point are declared to be artifacts of the clustering algorithm and are discarded. If no large jump occurs in the magnitudes of the cluster widths, then the $N_c = 30$ clusters with the smallest widths are selected for the template. If the cutoff point occurs beyond $N_m = 50$ clusters, then the N_m clusters with the smallest widths are selected. Figure 2 illustrates a set of templates built for eight agents. Each template was constructed from 18 measurements collected from three munitions.

Munition classification is performed by extracting features from the acoustic spectrum of the munition under inspection, and counting the number of features that fall within the clusters of templates in the template library. For the comparison, the cluster widths are scaled by a factor $W_s = 1.6$, and features that fall within a cluster of the template are counted as a "hit." A hitratio for a particular template is computed by dividing the number of hits by the number of clusters in the template. The munition is declared to belong to the class associated with the template with the largest hitratio.

3. Experimental Results

The algorithms were tested on a large set of data collected at the U. S. Army Dugway Proving Ground during August 1994. Two instruments, which are designated as "A" and "B", were used to collect data from 72 155mm munitions. (From here on data collected with the "A" instrument will be designated A data, templates built from data collected with the "B" instrument as B templates, etc.) The munitions were filled with surrogates for the following agents, denoted by their common name and NATO designation: blistering agents distilled mustard (HD), nitrogen mustard three (HN-3), and Lewisite (L); nerve agents Tabun (GA), Soman (GD), and VX (VX); and binary nerve agent precursors DF (DF) and QL (QL). (In this paper we substitute the designators LW for Lewisite and HN for nitrogen mustard three.) Each surrogate agent was contained in nine munitions, three 100% full, three 75% full, and three 50% full for a total of 72 munitions. Three to six measurements per munition were collected, where all measurements were taken from different locations on the upper half of the munition. In all the data set consisted of 674 measurements.

Three types of experiments were performed on the data. A self consistency test was performed by constructing templates from a set of measurements that represented a class, e.g, all A measurements on all agents in munitions at 100% fill level. After the templates were constructed each measurement was classified against all templates. An error occurred if the measurement was misclassified, for example, if a measurement from a VX munition was classified as DF. In the second test, templates were built by holding out one measurement and then classifying the held out measurement against all templates. This test provided an indication of how classification would extrapolate to new measurements. The third test consisted of classifying A data with templates built from B data, and vice versa.

As a statistical measure of the test results, the reliability of the test at 90% confidence was computed. The following tables represent a subset of the tests we performed on the data, and the results presented here are typical the of results we obtained on other tests.

Table 1 presents results for the consistency test on the 100% filled munitions. Templates were constructed from all A (or B) measurements on munitions containing a particular agent. Since there are 3 munitions per agent and 6 measurements per munition, 18 measurements were used to construct each template. (The templates in Figure 2 are the A templates of this test.) Table 2 illustrates the results of the hold-out test for the same set of templates. Tables 1 and 2 are arranged to show the number of misclassifications per munition. In the consistency test, one measurement was misclassified in both the A and B data. The resulting success rate was 99.3% and at 90% confidence the reliability is 97.3%. For the hold-out test, the A data was misclassified in eight measurements for a success rate of 94.4%, while the B data was misclassified in seven measurements for a success rate of 95.1%. At 90% confidence the reliabilities of the tests are 91.2% and 92.0% respectively.

Table 3 illustrates results obtained by building templates with A data from 100% filled munitions and classifying the B data from 100% filled munitions, and vice versa. Again, the table is arranged to show the number of misclassifications per munition. (Like the previous test, we have six measurements per munition and three munitions per agent.) The success rate of classifying B data with A templates was 88.2% (17 classification errors), and the rate for classifying A data with B templates was 81.3% (27 errors). The reliability at 90% confidence was 84.1% and 76.6% respectively.

Finally, the A and B measurements for each 100% filled munition were combined to build templates for individual munitions. In all there were 24 classes (three munitions per agent and eight agents) and 288 measurements. Consistency and hold-out tests were performed on all measurements. The results are illustrated in Table 4. In the consistency test no misclassifications occurred for a 100% success rate. The associated reliability at 90% confidence is 99.2%. The hold-out test had three misclassifications for a success rate of 99.0%. The associated reliability for this test at 90% confidence is 97.7%.

Conclusions

Test results demonstrate the potential use of the ARS system to classify chemical munitions by agent. Table 1 illustrates that the system produces consistent classification on data collected with the same instrument. Table 2 indicates the ability of the system to extrapolate classification to unknown munitions. On the other hand, the minor degradations found in Table 3 could have been caused by several factors including instrument variations as well as temperature differences in the munitions. Both possibilities are under investigation. Along the same lines, variability of spectral features as a function of munition age, fill level and concentration of agent are also topics of investigation. While the system performs reasonably well at classifying munitions according to agent, it performs very well at classifying individual munitions (cf. Table 4). Thus, the system appears to be well suited for classifying objects that have acoustic spectra with sharp resonances.

References

[1] R. S. Roberts, J. T. Chen, O. V. Vela and P. S. Lewis, "Munitions Classification Using an Acoustic Resonance Spectroscopic Technique," *Proceedings of the Twenty-Seventh Annual Asilomar Conference on Signals, Systems, and Computers,* Pacific Grove, CA. Nov. 1-3, 1993.

[2] D. N. Sinha "Acoustic Resonance Spectroscopy," *IEEE Potentials*, April 1992, vol. 11, No. 2, pg 10.

[3] Special Issue on Chemical Weapons Verification, *Verification Technologies*, U. S. Department of Energy, Office of Arms Control and Nonproliferation, First/Second Quarters 1992.

[4] Proceedings of the Defense Nuclear Agency Conference on Arms Control and Verification Technology (ACT) 1-4 June 1992, prepared for the Defense Nuclear Agency by the Kaman Sciences Corp. under contract DNA 001-88-C-0025, December 1993.

Agent	100% fill A			100% fill B		
DF	0	0	0	0	0	0
GA	0	0	0	0	0	0
GD	0	0	0	0	0	0
HD	0	0	0	0	0	0
HN	0	0	0	0	0	0
LW	0	1	0	0	0	0
QL	0	0	0	1	0	0
VX	0	0	0	0	0	0
Success rate	143/144 = 99.3%			143/144 = 99.3%		
% reliability	97.3%			97.3%		

TABLE 1: Consistency test for 100% fill munitions. Eight classes and 144 measurements. Reliability is computed for 90% confidence

Agent	100% fill A			100% fill B		
DF	0	1	0	0	0	0
GA	0	0	0	0	0	0
GD	0	0	0	0	0	0
HD	0	1	0	0	0	1
HN	0	0	0	0	0	0
LW	0	3	0	0	2	0
QL	2	0	0	1	0	1
VX	0	1	0	1	1	0
Success rate	136/144 = 94.4%			137/144 = 95.1%		
% reliability	91.2%			92.0%		

TABLE 2: Hold-out test for 100% fill munitions. Eight classes and 144 measurements. Reliability is computed for 90% confidence.

Agent	A Template			B Template		
DF	3	0	4	0	1	0
GA	0	0	0	0	0	1
GD	0	0	0	0	3	4
HD	0	4	0	2	6	3
HN	2	0	0	0	0	2
LW	0	2	2	0	2	0
QL	0	0	0	1	0	0
VX	0	0	0	0	0	1
Success rate	127/144 = 88.2%			117/144 = 81.3%		
% reliability	84.1%			76.6%		

TABLE 3: A(B) data applied to B(A) templates. Eight classes and 144 measurements. Reliability is computed for 90% confidence.

Agent	Consistency test			Hold-out test		
DF	0	0	0	0	0	0
GA	0	0	0	0	0	0
GD	0	0	0	0	0	0
HD	0	0	0	0	0	0
HN	0	0	0	0	0	0
LW	0	0	0	2	0	0
QL	0	0	0	1	0	0
VX	0	0	0	0	0	0
Success rate	288/288 = 100%			285/288 = 99.0%		
% reliability	99.2%			97.7%		

TABLE 4: Each 100% filled munition is a class. Twenty four classes and 288 measurements. Reliability is computed for 90% confidence.