Acoustic Classification of Battlefield Transient Events Using Wavelet Sub-band Features

M. R. Azimi-Sadjadi*[a], Y. Jiang †[b] and S. Srinivasan [a]

[a]Information System Technologies, Inc.
Fort Collins, CO, 80521
[b]Department of Electrical and Computer Engineering
University of Colorado, Boulder, CO 80309

ABSTRACT

Detection, localization and classification of battlefield acoustic transient events are of great importance especially for military operations in urban terrain (MOUT). Generally, there can be a wide variety of battlefield acoustic transient events such as different caliber gunshots, artillery fires, and mortar fires. The discrimination of different types of transient sources is plagued by highly non-stationary nature of these signals, which makes the extraction of representative features a challenging task. This is compounded by the variations in the environmental and operating conditions and existence of a wide range of possible interference. This paper presents new approaches for transient signal estimation and feature extraction from acoustic signatures collected by several distributed sensor nodes. A maximum likelihood (ML)-based method is developed to remove noise/interference and fading effects and restore the acoustic transient signals. The estimated transient signals are then represented using wavelets. The multi-resolution property of the wavelets allows for capturing fine details in the transient signals that can be utilized to successfully classify them. Wavelet sub-band higher order moments and energy-based features are used to characterize the transient signals. The discrimination ability of the subband features for transient signal classification has been demonstrated on several different caliber gunshots. Important findings and observations on these results are also presented.

Keywords: Collaborative distributed sensors, Maximum Likelihood signal estimation, acoustic transient classification, wavelet subband features.

1. INTRODUCTION

The emergence of small, low-cost, and low-power sensor technologies with on-board signal processing and wireless communication capabilities has stimulated great interests in the utilization of distributed sensor networks in a wide variety of critical applications. One such application involves detection, classification and accurate localization of battlefield acoustic transient events namely gunshots, RPG, and artillery fires, especially in MOUT environments. In,[1] the use of distributed sensors has successfully been demonstrated for accurate localization of gunshots. The sensor nodes detect and transmit transient event time of arrivals (TOA's) to a base station where the localization takes place using an error correction process to mitigate for bad sensor reports and interference effects. Acoustic transient classification is also possible based upon the extracted parameters or attributes of the detected supersonic 'N-wave'.[1] However, if the supersonic shockwave does not exist or is not easily detectable/reliable, transient event classification cannot be accomplished.

In distributed sensor networks, transient events can be classified based upon muzzle blast signal captured by the closest sensor node. However, before the class of the detected acoustic transient event can be determined, it is necessary to estimate the waveform of the original transient source from noisy sensor recordings. More specifically, the goal here is to remove the effects of other sources of interference and restore the transient signature prior to feature extraction and transient classification. In this paper, a novel maximum likelihood (ML)-based signal waveform estimation method is developed that leads to a statistically optimal solution. This method fully exploits

*mo@infsyst.biz; phone: 1 970 224 2556; http://www.infsyst.com.
†Y. Jiang was a consultant at Information System Technologies, Inc.

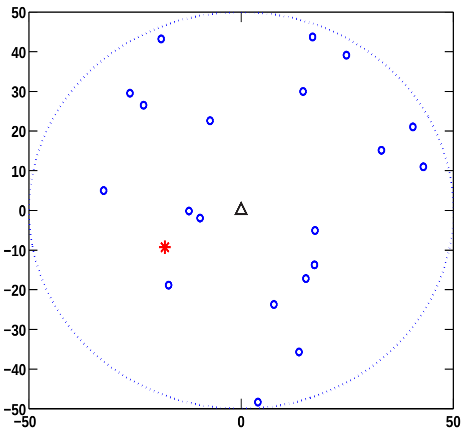
This work is funded by Army SBIR-Phase I contract # W15QKN-06-C-0058.

the inherent redundancy in the sensor network and hence is very robust. Once the acoustic transient signals are estimated/restored, a wavelet-based feature extraction scheme is applied to the estimated transient signal. This method provides multi-resolution representation of the estimated and restored transient signal, reduces the data dimensionality, and extracts a set of salient subband features that represent the characteristics of the transients. This reduced feature set is then used to classify the transients.

The organization of this paper is as follows. Section 2, describes the transient event detection problem using the distributed sensor network framework. Section 3, presents the proposed ML-based signal estimation and restoration method together with the relevant simulation results. In Section 4, we develop the wavelet-based feature extraction method and present the feature separability for discriminating various transient events. Finally, Section 5 gives our concluding remarks and observations on this work as well as some important topics for future research.

2. DISTRIBUTED TRANSIENT EVENT DETECTION

In the distributed transient event detection the sensors need to determine and agree on the presence of a transient signature, its time duration, and more importantly its onset. This is crucial to the success of the subsequent operations including transient localization and classification. An illustration of transient signal detection using distributed sensor networks is shown in Figure 1(a). Several distributed sensors are coordinated by a gateway at the center of the cluster (denoted by "Δ") with certain radius (e.g. 50-100 meters). There could be multiple sensor clusters in a distributed sensor network. We assume that the sensors sample the received signal at a fixed sampling rate and each sensor clock is synchronized. The clock synchronization method in,[1] which is based upon message routing, can be used to synchronize the sensor nodes. The sensor positions are also assumed to be known to the base station. Transient signals, such as the flash of an explosion, the launch of an RPG, or machine gunfire are received by a subset of the sensors, especially the ones closer to the source. These transient signals are usually wideband and occupy only a small time duration.



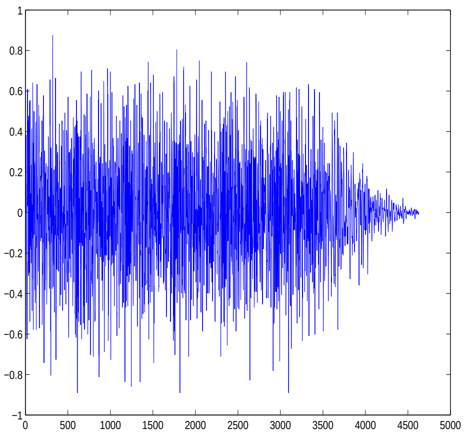


Figure 1. (a) The layout of randomly distributed sensors (denoted by "o") coordinated by a gateway (denoted by "Δ"). The transient acoustic source is denoted by "$*$"; (b) The original clean 20mm gun fire acoustic waveform.

The transient signal model for muzzle blast in presence of noise and multipath is as follows. Most of the time, the signal received by the k^{th} sensor is only background noise due to the transient nature of the signal of interest, i.e.,

$$x_k(m) = n_k(m), \qquad k = 1, 2 \ldots, K \tag{1}$$

where K is the number of the sensors. We model background noise $n_k(m)$ as zero-mean white Gaussian noise with variance σ_k^2. If at time T_0, a transient signal (e.g. muzzle blast) $s(m)$ is emitted from a wideband source, then the k^{th} sensor received signal is

$$x_k(m) = \begin{cases} \alpha_k s\left(m - \frac{d_{k,s}}{c}\right) + \sum_{j=1}^{J} \beta_{k,j} s\left(m - \delta_{k,j}\right) + n_k(m), & m = T_0, T_0 + 1, \ldots, T_1, \\ n_k(m), & 1 \leq m < T_0 \text{ and } T_1 < m \leq M, \end{cases} \tag{2}$$

where $\alpha_k = \frac{1}{d_{k,s}}$ represents the effect of signal attenuation which is proportional to its travel distance, $d_{k,s}$, between the source and the kth sensor, c is the speed of sound in air, and $T_1 - T_0$ is the time duration of the transient event. The second term on the right hand side of the first equation in (2) represents the collective effect of the multipath with unknown amplitudes $\beta_{k,j}$'s and time delays $\delta_{k,j}$'s.

The first step in the process is to detect the presence of the transient events in the signals recorded by all or a subset of the sensor nodes. A simple and yet very effective power-based algorithm for transient signal detection from the muzzle blast component has been developed, which determines not only the existence of the transient but also its onset and time duration. Our power-based detector is closely related to the Nuttall's maximum detector[2] with the difference that we also estimate the onset and duration of the transient. The statistical performance of this detector is also derived.

Here, we present some results on a transient signal associated with 20mm gunfire. The noise-free gunfire acoustic waveform is shown in Figure 1(b). (The .wav files are downloaded from http://www.rcexchange.com/) The sample rate is $F_s = 4410$Hz. For a simulated distributed network, the signals received by four sensors at different distances from the source are shown in Figure 2. Note that here a simple distance-based signal attenuation model is imposed. As a result, the prominence of the received transient signal depends only on the distance between the acoustic source and the sensors and hence no phase perturbation, e.g. due to atmospheric effects, are considered. Figure 3 shows the plots of the energy in the snippets of the observed data in four sensors. In our simulation, we choose the snippet length of 42, which corresponds to a time window length of 9.5 millisecond. It can easily be seen from Figures 2 and 3 that different sensors may have different conclusions on the existence of a transient signal as well as on the duration of the signal.

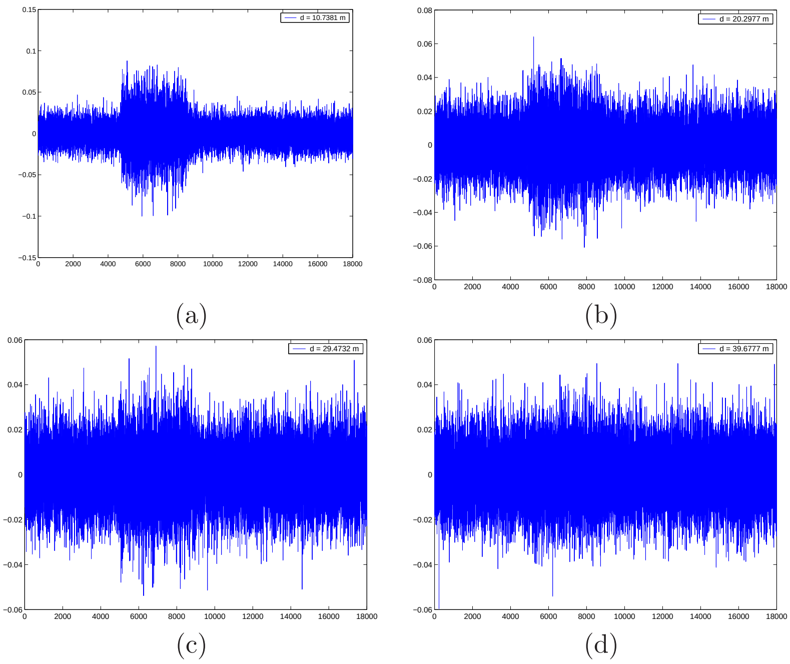


Figure 2. The received data containing the acoustic signal of 20mm gun fire. (a) $d_{1,s} = 10.7381$m (b) $d_{2,s} = 20.2977$m (c) $d_{3,s}$= 29.4732m (d) $d_{4,s}$=39.6777m.

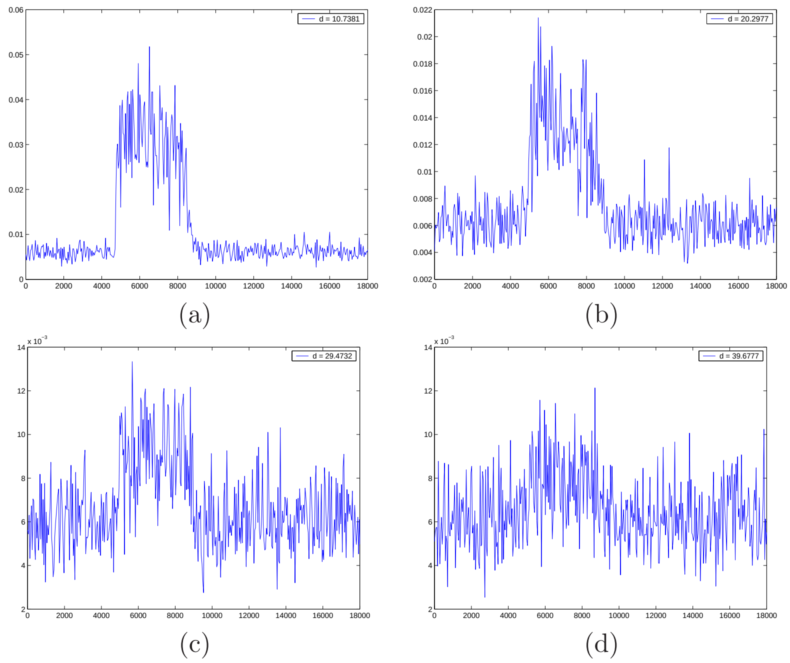


Figure 3. The output of the power detector applied to the signals in Figure 2.

As far as the supersonic projectile is concerned, one of its most observable properties is the shockwave it leaves in its wake. The shockwave is characterized by fast pressure rise times on the order of microseconds, a linear pressure drop to below the ambient pressure and an equally fast return to ambient pressure.[3] These waves are commonly known as N-waves based on their strong resemblance to the letter “N” and are only detectable by the sensors in the proximity of the traveling bullet. Detecting the presence of N-waves to localize snipers is preferred for many reasons. While the muzzle blast can be suppressed, the shockwave cannot. Moreover, multipath is not a major concern here as the shockwave travels directly from the bullet to the sensor. In,[1] a zero-crossing encoder and two state machines are used to detect the presence of N-wave as well as the muzzle blast. After a signal is sampled, it is zero encoded. Details of the signals rise time, fall time, duration, and amplitude are fed into a shockwave detection state machine and a filtered version is sent to a similar machine for detecting muzzle blast. These state machines search for patterns in the signal and are developed using zero-encoded recordings of different gunshots. If the signal has the right features at the right times, the machine confirms the presence of an N-wave, a muzzle blast or both. This method of N-wave detection could be used in conjunction with our power-based detector for the muzzle blast component.

3. ML-BASED JOINT SOURCE LOCALIZATION AND WAVEFORM ESTIMATION

A transient source in a 2-D space can be localized by estimating two time difference of arrivals (TDOA’s). On the other hand, we can estimate $\frac{K(K-1)}{2}$ TDOA’s using K spatially separate sensors to generate a very accurate localization by exploiting the redundancy in a network with moderate number of sensors. In,[4] the authors exploit this redundancy by combining weighted hyperbola to make more accurate source localization. However, their algorithm is not only computationally demanding but also ad-hoc. Our proposed approach is based upon searching over a grid on hypothesized source location to maximize the likelihood function. The problem of source localization using search over a hypothesized area has been applied before[5] for locating micro-seismic events. At each grid point, their algorithm calculates the time delays of the signal received by all the sensors. By compensating the time delays, one aligns and combines together the data sequences received by different sensors. They used the energy of the aligned and combined signal as a cost function for this search. The grid point associated with the largest signal energy is determined as the source location.

In this section, we present a novel ML-based joint source localization and signal waveform estimation method that leads to a statistically optimal solution to this problem. The waveform estimation of the transient signal is crucial for subsequent transient classification. There are numerous interfering sources on the battlefield and in MOUT environments that must be removed for successful transient event discrimination.

The proposed joint signal estimation and source localization method goes like this. We assume that the background noise $n_k(m)$ is zero-mean Gaussian with variance σ_k^2 and the transient signal $s(m)$ is an unknown deterministic signal. Denote $\mathbf{x}_k = [x_k(1), \ldots, x_k(M)]^T$, and $\mathbf{s} = [s(1) \ldots s(M)]^T$, where M is the data length, then the ML estimates of the source location $\mathbf{p}_s = [x_s \, y_s \, z_s]^T$ and $\mathbf{s}$ are

$$\hat{\mathbf{s}}, \hat{x}_s, \hat{y}_s, \hat{z}_s = \arg \max_{\mathbf{s}, x_s, y_s, z_s} f(\{\mathbf{x}_k\}_{k=1}^K | \mathbf{s}, \mathbf{p}_s) \tag{3}$$

Suppose the noise received by the sensors are uncorrelated (multipath free environment). Then,

$$f(\{\mathbf{x}_k\}_{k=1}^K | \mathbf{s}, \mathbf{p}_s) = \prod_{k=1}^K \left(\frac{1}{\sqrt{2\pi}\sigma_k} \right)^K \exp \left\{ - \frac{\sum_{m=1}^M |x_k(m) - s(m - d_{k,s}/c)/d_{k,s}|^2}{2\sigma_k^2} \right\} \tag{4}$$

Note that $d_{k,s} = \sqrt{(x_k - x_s)^2 + (y_k - y_s)^2 + (z_k - z_s)^2}, \quad k = 2, \cdots, K$ is obviously a function of $\mathbf{p}_s$. We also note that σ_k of the ambient noise can be estimated by calculating the variance of received data during the period when signal is absent and hence is known *a priori*. The ML estimate given in (4) can be converted to

$$\hat{\mathbf{s}}, \hat{\mathbf{p}}_s = \arg \min_{\mathbf{s}, \mathbf{p}_s} \sum_{k=1}^K \frac{\sum_{m=1}^M |x_k(m) - s(m - d_{k,s}/c)/d_{k,s}|^2}{\sigma_k^2} \tag{5}$$

We can use either a grid search[5] or the robust source localization method in[6] for finding the source position. Then, the estimate of $\mathbf{s}$ is as follows.

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \sum_{m=1}^M \sum_{k=1}^K \left| \frac{x_k(m)}{\sigma_k} - \frac{s(m - \hat{d}_{k,s}/c)}{\sigma_k \hat{d}_{k,s}} \right|^2 \tag{6}$$

After some simple algebraic manipulations, the sample-wise estimate of $\mathbf{s}$ can be given by

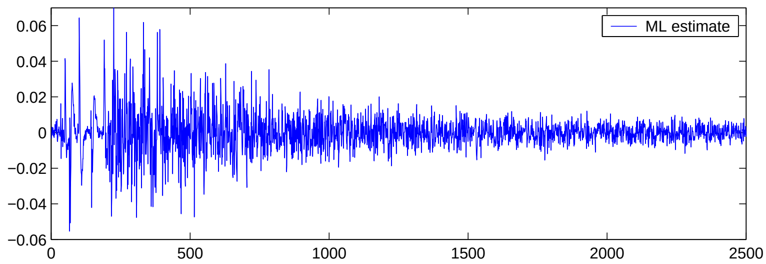
$$\hat{s}(m) = \frac{\sum_{k=1}^K \frac{x_k(m + d_{k,s}/c)}{\sigma_k^2 d_{k,s}}}{\sum_{k=1}^P \frac{1}{\sigma_k^2 d_{k,s}^2}} \tag{7}$$

Although $d_{k,s}/c$ is in general not an integer, we can round it to the nearest integer. Due to the edge effect, we need at least $M + \max\{d_{k,s}/c, \, k = 1, \ldots, K\}$ samples to estimate a signal waveform with length M. Inserting (7) into the cost function (4), yields the minimum of the cost function as

$$\sum_{m=1}^M \sum_{k=1}^K \left| \frac{x_k \left(m + \frac{\hat{d}_{k,s}}{c} \right)}{\sigma_k} \right|^2 - \sum_{m=1}^M \frac{\left| \sum_{k=1}^P \frac{x_k (m + \frac{\hat{d}_{k,s}}{c})}{\sigma_k^2 d_{k,s}} \right|^2}{\sum_{k=1}^P \frac{1}{\sigma_k^2 d_{k,s}^2}} \tag{8}$$

Thus, we can decouple the problem of jointly estimating the source location and waveform into two separate steps. The first step is to estimate the source location by minimizing the cost function given in (8) using an iterative search[5] or use the TDOA generated by the method.[6] The second step is to estimate the transient waveform using (7).

This algorithm was applied to the corrupted and faded signals received by the neighboring sensors to estimate the transient signal in Section 2. Figure 4 compares the ML estimate of the transient signal and the raw signal received by the sensor closest to the source, which has the highest confidence level of detection. Comparing the tails of the two signals, one can conclude that ML estimate has significantly higher SNR than the raw data. This SNR improvement is measured to be around 3 dB. Most importantly, we will show in the next section that the features extracted from the estimated transient waveforms are very close to those of the clean signature leading to very good transient discrimination.



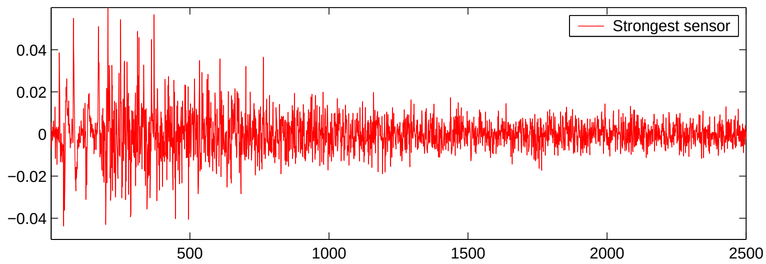


Figure 4. Comparison of ML estimate of waveform estimate and the signal received by the strong sensor.

4. WAVELET-BASED FEATURE EXTRACTION FOR TRANSIENT EVENT CLASSIFICATION

In this section, we develop a wavelet-based feature extraction scheme for transient signal representation. The potentially large categories of transient events on the battlefield and in MOUT makes it nearly impossible to identify any one particular type of feature that captures all relevant information for transient classification.[7,8] For instance, the transient events associated with different weapons have widely different characteristics. We represent each wavelet subband signal using a compact statistical set of features that capture the behavior of each subband signal. The wavelet subband high order moments and energy-based features[7,9] are used to characterize the transient signals. Let $\mathbf{a}_J$ denote the lowest order approximation and $\mathbf{d}_1, \cdots \mathbf{d}_J$ be the J added details in a J-level DWT decomposition[10] of a transient signal. The total energy of the wavelet coefficients is,

$$E_{total} = \|\mathbf{a}_J\|^2 + \sum_{j=1}^{J} \|\mathbf{d}_j\|^2 \tag{9}$$

where $\|\mathbf{a}_J\|^2 = \frac{1}{N_{a_J}} \sum_{k=1}^{N_{a_J}} |a_J(k)|^2$ and $\|\mathbf{d}_j\|^2 = \frac{1}{N_{d_j}} \sum_{k=1}^{N_{d_j}} |d_j(k)|^2$, for $j = 1, \cdots, J$. N_{a_J} and N_{d_j} are the lengths of the lowest order approximation and j^{th} level added detail subband signals, respectively. The feature vector consists of the normalized energy of the wavelet signals in different subbands,

$$\mathbf{f} = [\frac{\|\mathbf{d}_1\|^2}{E_{total}}, \cdots, \frac{\|\mathbf{d}_J\|^2}{E_{total}}, \frac{\|\mathbf{a}_J\|^2}{E_{total}}]^T \tag{10}$$

Energy normalization is performed to ensure that the feature vectors are invariant to the strength of the observations. This property is highly desirable as it makes the feature vector invariant to the distance between the source and the sensors.

Wavelet subband moment features could also aid in transient event discrimination. Define $a'_J(k) = \frac{a_J^2(k)}{\|\mathbf{a}_J\|^2}$ and $d'_j(k) = \frac{d_j^2(k)}{\|\mathbf{d}_j\|^2}, j \in [1, J]$ that are the normalized squared lowest order approximation and added detail wavelet coefficients in different subbands, respectively. Then, the p^{th} order (p = 2,3 and 4) central moment corresponding to j^{th} wavelet subband is obtained by treating $d'_j(k),\ \ j \in [1, J]$ and $a'_J(k)$ as probability density functions,

$$\mu_p = \sum_k (k - \mu_1)^p d'_j(k) \tag{11}$$

where $\mu_1 = \sum_k k d'_j(k)$ is the mean in this subband. Similar moments are also extracted from the lowest order approximation subband. The dimensionless moments that describe the average, dispersion, skew and kurtosis are, $M_1 = \mu_1$, $M_2 = \sqrt{\mu_2}/\mu_1$, $M_3 = \mu_3/M_2^{3/2}$ and $M_4 = \mu_4/M_2^2 - 3$, respectively. These wavelet subband moment features provide excellent representation of the transient events.

To show the robustness of the wavelet subband energy and moment features, as well as the performance of the ML-based signal estimation method in Section 3, a preliminary study is conducted here. In this experiment, three acoustic transients namely AK-47, 75mm and 20mm gun fires were considered. The sensor network consisted of three simulated vehicle-mounted systems with 24 acoustic sensors. Each vehicle contains a 3-D volume array of eight microphones mounted on the upper and lower corners of a HMMWV vehicle. The structure of this volume array is based upon the specifications of the MCFS vehicle-mounted system developed by AAI Corporation. The signals received by the sensors on different vehicles are modeled using the model in (2) without considering the multipath effects. However, coherence loss due to signal attenuation or fading as a function of distance is considered in the model. The three vehicles were randomly placed in a coverage area of radius 100m. The source was also randomly placed within a 3-D volume of 100m radius at every Monte Carlo simulation. A total of 20 trials were tried where at every trial a different distributed sensor network was adaptively formed from sensors mounted on these vehicles depending on the relative locations of the vehicles and the source. The transient waveforms were estimated from these observations using the ML-based method described in Section 3, assuming that only a few sensors contain strong direct path (line-of-sight) signals. The wavelet subband energy and wavelet subband moment features were then extracted from the estimated signal associated with the closest sensor to the source using 8-level and 3-level DWT, respectively. Figure 5 (a) shows the 2-D scatter plot of two (4th and 5th out of 9 subband energy features) wavelet subband energy components; while Figure 5 (b) shows the 2-D scatter plot of two (out of 16 moments from 4 subbands) wavelet subband moments, which correspond to the first order moment of the third level detail and the lowest order approximation coefficients. These features were selected since they exhibited high discrimination ability out of all 25 features. The samples that are indicated by solid diamond, circle and cross correspond to the features extracted from the original (clean) transient signals while those that are hollow correspond to the estimated signal waveforms. The clustering of the features for these three different transient sources is a clear indication that the feature vectors are unaffected by the variation in distance between the source and the sensors. These results also indicate that the feature vectors of the estimated waveforms closely resemble those of the original (clean) sources. Moreover, as far as AK-47, 20mm and 75mm gunshots are concerned, the feature space is linearly separable in the wavelet domain using our feature extraction method. Thus, a simple classifier can successfully discriminate these three types of gunshots with high classification accuracy.

5. CONCLUSIONS

Sparse distributed sensor networks that consist of several single low-cost low-power sensor nodes (acoustic, seismic, etc.) offer numerous important benefits including: simplicity and ease in deployment, large coverage area, better spatial resolution for separating multiple closely spaced targets, less hardware complexity and hence significantly lower costs, more flexibility in configuring different dynamic sensor array configurations, and potential widespread applications in urban warfare, homeland security, border patrol, industrial monitoring, etc. For transient event (e.g. gunshots) detection, localization, and classification problems, distributed sensor processing also allows for better detection and capture of the supersonic shockwave in multipath rich environments as well as better shooter localization accuracy owing to larger coverage area.

This paper develops a novel algorithm to estimate and restore the transient signal waveform (muzzle blast) by removing the fading effect (e.g. due to acoustic pressure loss) as well as uncorrelated interference (ambient noise). Using this ML-based method source localization is also possible using a search process. The estimated waveform is then used to develop and test subband-based feature extraction methods using wavelet transform to

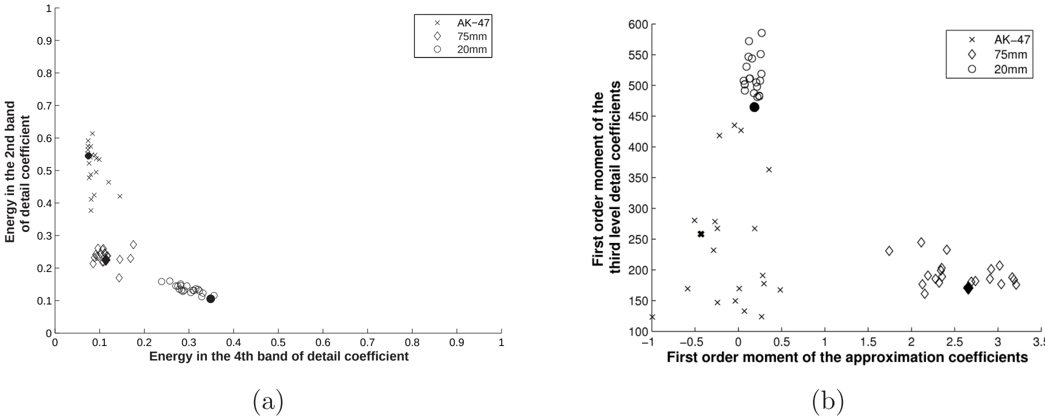


Figure 5. Separability of 2-D feature space of AK-47, 75mm and 20mm gun fires.

extract a set of salient features that represent the characteristics of the transient events. The combined feature set is then used to discriminate different caliber gunshots. The results attest to the success of the ML-based signal estimation method as well as to the effectiveness of the subband wavelet features developed in this paper.

Among the future goals is to extend the transient waveform estimation method for multipath and complex urban environments and when non-line-of-sight problems exist. Moreover, the features extracted from the restored muzzle blast signature should be augmented with the N-wave parameters[3] extracted from the detected supersonic shockwave for a more accurate classification.

REFERENCES

1. A. Lédeczi, A. Nádas, P. Vëlgyesi, G. Balogh, B. Kusy, J. Sallai, G. Pap, S. Dóra, and G. Simon, "Countersniper system for urban warfare," *ACM Transactions on Sensor Networks*, vol. 1, pp. 153–177, November 2005.
2. A. Nuttall, "Detection capability of linear-and-power processors for random burst signals with unknown location," *NUWC-NPT Tech Report 10 822*, August 1997.
3. B. Sadler, T. Pham, and L. Sadler, "Optimal and wavelet-based shockwave detection and estimation," *J. Acoust. Soc. Amer.*, vol. 104, pp. 955–963, August 1998.
4. D. P. Young, C. Keller, D. W. Bliss, and K. W. Forsythe, "Ultra-wideband (UWB) transmitter location using time difference of arrival (TDOA) techniques," *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1225 – 1229, November 2003.
5. B. Daku, J. Salt, and L. Sha, "An algorithm for locating microseismic events," *Conference on Electrical and Computer Engineering*, vol. 4, pp. 2311 – 231, 2-5 May 2004.
6. Y. Jiang and M. R. Azimi-Sadjadi, "A robust source localization algorithm for distributed acoustic sensor networks," *to appear Proc. of the IEEE Internatinal Conference on Acoustics, Speech and Signal Processing (ICASSP'2007)*, April 2007.
7. A. Kundu, G. C. Chen, and C. E. Persons, "Transient sonar signal classification using hidden Markov models and neural nets," *Journal of Ocean Engineering*, vol. 19, pp. 87–99, Jan 1994.
8. J. Ghosh, L. M. Deuser., and S. D. Beck, "A neural network based hybrid system for detection, characterization, and classification of short-duration oceanic signals," *IEEE Journal of Ocean Engineering*, vol. 17, pp. 351–363, Oct 1992.
9. S. A. Probert and Y. H. Song, "Detection and classification of high frequency transients using wavelet analysis," *IEEE Power Engineering Society Summer Meeting*, vol. 2, pp. 801 – 806, July 2002.
10. S. Malat, *A wavelet tour of signal processing.* Academic Press, 1998.