USE OF MODULATION SPECTRA FOR REPRESENTATION AND CLASSIFICATION OF ACOUSTIC TRANSIENTS FROM SNIPER FIRE

Lane Owsley, Les Atlas**, and Chad Heinemann***

*Applied Physics Laboratory, University of Washington, 1013 NE 40th St, Seattle, WA 98105
**Dept. of Elec. Engineering, University of Washington, Seattle, WA 98195

ABSTRACT

There are many applications for classification of acoustic transients produced by supersonic projectile fire. Analysis of existing models for such transients suggests they have properties which may be well-captured by a transform of a signal into joint acoustic and modulation frequency: a modulation spectral representation. Simple features are extracted from this representation which enables successful use in an important classification application.

1. INTRODUCTION

Automated or semi-automated identification and localization of snipers has obvious military and safety applications. Acoustic signatures from rifle shots are loud and distinctive, making them nominally easy to detect [1]. Passive acoustics sensing provides good range and non-line-of-sight capabilities [2].

Despite its many advantages, the performance of acoustic sensors can degrade under adverse propagation conditions. In particular, the direct weapon muzzle blast, which originates from the shooter location and the projectile shockwave, which does not relate to that location, overlap in frequency. Both signals are broadband and there is little, if any, reliable difference in their temporal and spectral patterns. Localization accuracy can be degraded by this overlap.

In this paper we add another signal dimension, modulation frequency, to potentially allow for increased discrimination between the muzzle blast and the shockwave. We show that this added dimension significantly increases discrimination accuracy over spectral features alone.

In the remainder of paper, we first describe and illustrate the nature of the muzzle blast transient. We then describe a system which jointly estimates standard acoustic frequencies along with modulation frequency. A simple model of the shockwave is then described, with an accompanying discussion as to how shockwaves produce higher modulation frequencies than a muzzle blast. This model case is then confirmed by examples on real rifle shot data. Overall system performance results are then estimated from a larger set of real rifle shot data. Our hypothesis that the added modulation frequency dimension improves the reliability of muzzle blast transient identification is supported by these experiments.

This work was supported by the Sensors & Electron Devices Directorate (Mr. Tien Pham) under the auspices of the U.S. Army Research Office Scientific Services Program administered by Battelle (Deliver Order 0228, Contract No. DAAD19-02-D-0001.) The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

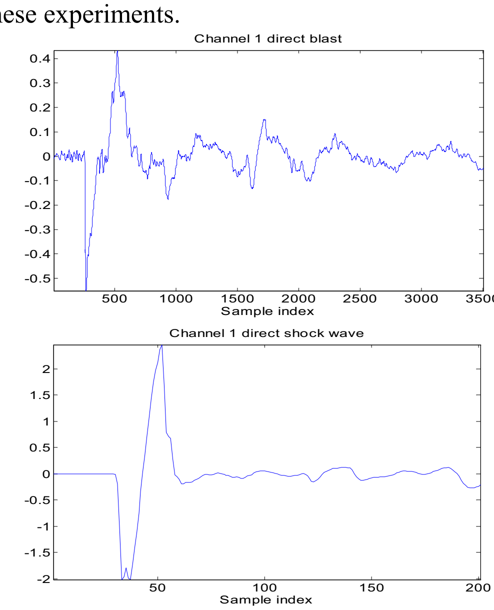

Figure 1. Acoustic transients from sniper fire: muzzle blast (above) and shockwave (below).

2. ACOUSTIC TRANSIENTS FROM SNIPER FIRE

Sniper fire in a battle situation can produce two types of acoustic transient events which can be received and used for detection and localization of the shooter. The first of these is the muzzle blast itself, emanated from the shooter's location. An example of this type of transient can be seen in Figure 1 (top). The second type of transient event is the shockwave produced by the projectile's supersonic travel; the shockwave takes the form of a cone with its apex at the current location of the projectile. An example of such a received transient can be found in Figure 1 (bottom). Further, both transient types may be

0-7803-8874-7/05/$20.00 ©2005 IEEE

received either directly through the air or after reflection off of the ground or other surfaces such as buildings or vehicles.

3. MODULATION SPECTRA

There is substantial evidence that many natural signals can be represented as low frequency modulators which modulate higher frequency carriers. Many researchers have observed that this concept, loosely called "modulation frequency," is useful for describing, representing, and modifying broadband acoustic signals such as speech and music. Modulation frequency representations usually consist of a transform of a one-dimensional broadband signal into a two dimensional joint frequency representation, where one dimension is typically standard Fourier frequency and the other dimension is a modulation frequency.

An invertible modulation transform was proposed in [3] where standard Fourier frequency, referred to as "acoustic frequency" in this paper, is represented along one dimension, and modulation frequency, representing the spectral analysis of the temporal envelope, is represented along a second dimension. The two dimensions are computed through a separable two-dimensional transform. The process first segments the source signal into consecutive overlapping frames which are transformed to the frequency domain and aligned to form a time-frequency representation. A detection operation (i.e., magnitude) is performed on each frequency sub-band to extract the envelope information of the signal. A second transform is then performed across time in each frequency sub-band to form the acoustic-modulation frequency representation.

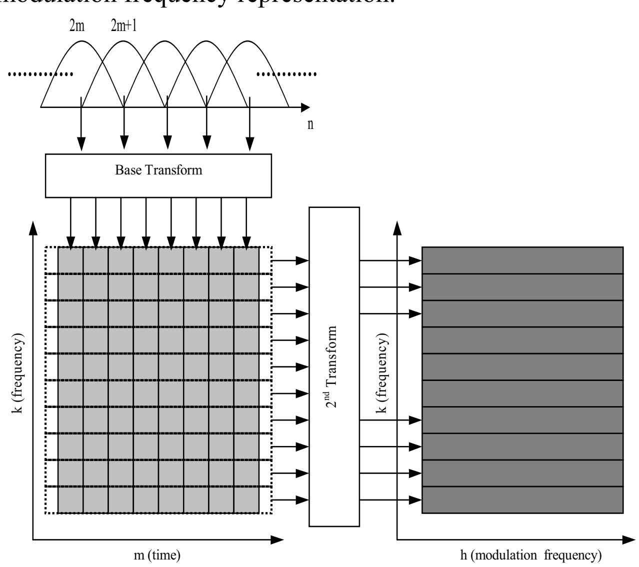

Figure 2. Simplified structure of the two-dimensional transform.

Figure 2 depicts a simplified overview of the modulation spectral transform. This modulation spectral transform used a short first (or "base") transform window size of 130 microseconds in duration. Since we are most concerned with short duration events, this duration is significantly shorter than usual frame sizes for audio and speech applications. The second transform size was about 15 ms.

4. USE OF MODULATION SPECTRA TO REPRESENT SYNTHETIC SHOCKWAVES

The shockwave produced by the supersonic travel of a projectile has a well-known idealized form [1]: the N-wave:

$$f(t)=1-2t/t_0\,; \quad 0 \le t \le t_0\,.$$

This form can be generalized slightly by adding onset and offset regions:

$$f(t)=t/t_{on}+1\,; \quad -t_{on}\le t<0$$
$$f(t)=(t-t_0)/t_{on}-1\,; \quad -t_0<t\le t_0+t_{off}$$

The result is a signal which appears as shown in Figure 3:

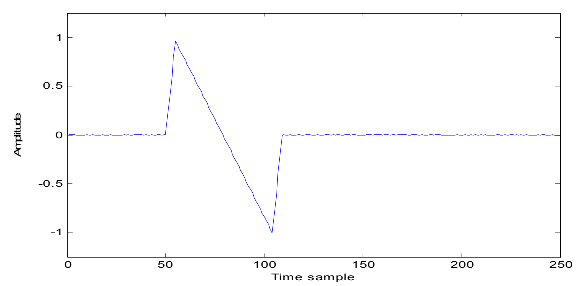

Figure 3. A synthesized shockwave, modeled as an N-wave, with t_{on}=5 samples, t_0=50 samples, and t_{off}=5 samples

Use of synthesized signals enables us to experiment with the effect of varying the N-wave structure on the modulation spectral representation. The primary effect is clear: the harmonic structure in modulation frequency is a function of the distance between the peaks. The reason for this is that the N-wave can be viewed as two occurrences of a single transient, one of which has been shifted and flipped:

$$g(t)=1-2t/t_0\,; \quad 0\le t\le t_0/2\,.$$
$$f(t)=g(t)-g(t_0-t)$$

As a result, since each spectral component in the first instance of $g(t)$ also appears in the second instance of $g(t)$, we get the harmonics associated with the transform of these repeated signals.

The second parameter we can control is the onset time (for this simple investigation we have also set the decay time to the same value.) The result of this variation is primarily manifested in the relative power of the harmonics rather than in the frequency (though that too is affected as we are in effect increasing the length of the entire wave as we change the onset). Our interpretation is that the increase in onset time is in fact a spreading of the energy at each individual frequency band and, as such, the higher harmonics are not as strong.

5. CLASSIFICATION SYSTEM

During this feasibility study, we investigated a wide variety of possible applications for modulation spectral representations in aiding the processing of sniper transients, including automatic difference-of-arrival calculations, classification of direct-path versus reflected-path transients, and separation of shockwaves and muzzle blasts. The last was identified as a key goal by the sponsor and is summarized here. To facilitate this work, the sponsor provided a dataset collected in a MOUT (Military Operations in Urban Terrain) environment. The recordings were taken at a variety of different shooter positions with respect to the recording microphones. Each gunshot was recorded with four microphones. The first channel is an open-field B&K microphone, channels two and three are B&K microphones mounted on an HMMWV (humvee), and the fourth channel is a low-end production microphone mounted on the humvee.

In Figure 4, the modulation spectrum is shown for typical shock wave and muzzle blasts. Darker regions correspond to the highest and most significant energy. Looking first at the shock wave modulation spectral structure, we see that there is a harmonic between 5 and 10 kHz in modulation frequency. This is due to the time difference between the leading and trailing edges of the shock wave. Also, the structure is high in energy out to relatively high frequencies along both axes. This structure does not exist in the case of the muzzle blast, where most of the energy is contained in lower acoustic and modulation frequencies. The increased range in both acoustic and modulation frequency in the case of the shock wave lead us to investigate features that take this into account as a means for discriminating between the signal types.

A basic transient detector is used to identify the transients of interest. This block detects significant increases in signal power with respect to the established background power. Modulation spectral analysis is applied to each normalized snippet of data. When comparing the shock waves and muzzle blasts, it was apparent that the muzzle blasts died off more quickly along both the acoustic and modulation frequency axes. To capture this behavior, we created two parameters: maximum extent over some threshold in acoustic frequency, and maximum extent over some threshold in modulation frequency. These are simply the maximum frequencies at which the modulation spectral representations of the normalized signals exceeded an experimentally determined threshold. These features and the classifier derived from them are not presented as a robust classifier but merely to enable demonstration of the utility of the representation and the feasibility of a classifier based on that representation.

Experiments were conducted across a variety of noise conditions, shooter/target geometries, microphone locations, and microphone types.

6. RESULTS

In this section we present results illustrating our classification algorithm. We show that the shock wave and muzzle blast can be separated in the chosen 2-D feature space by linear thresholding, and that the frequency extent features are robust in moderate noise levels.

To test our classification algorithm in a noisy environment we added military environmental noise to each transient such that the resulting signal-to-noise ratio (SNR) was 15 dB. The noise sample was taken from a B&K microphone recording of the humvee engine. The modulation spectra of the noisy shock wave and muzzle blast are shown in Figure 4.

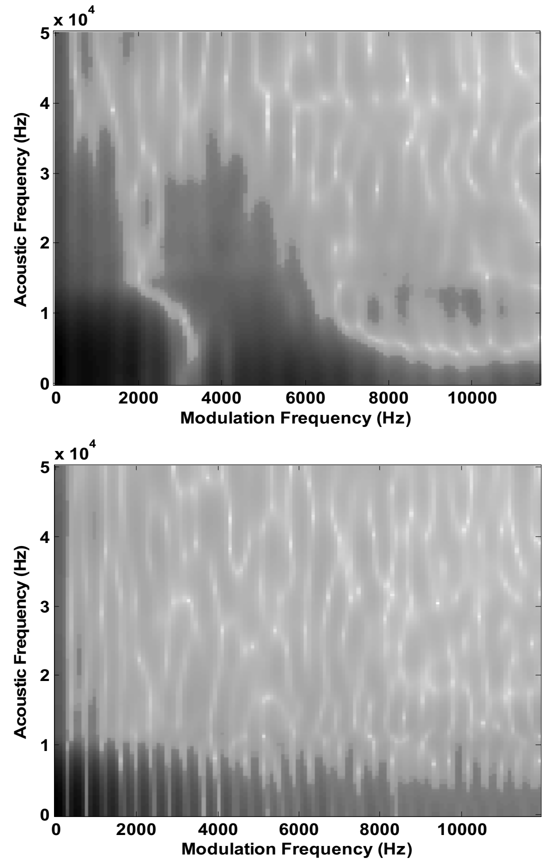

Figure 4. Modulation spectrum representations of shockwave (top) and muzzle blast (bottom).

Prior to extracting the frequency extent features from the modulation spectrum we perform the following preprocessing steps:

1. Compute the modulation spectra in dB and set to zero all locations which fall 80 dB below the maximal spectral value.
2. Map the resulting spectral range to the interval [0,1].
3. Multiply by a 0-1 mask, zeroing all values less than an empirically-determined threshold.

Figure 5 shows the results of these preprocessing steps for the modulation spectra shown in Figure 4. This figure clearly shows the shock wave modulation spectra having greater extent in both acoustic and modulation frequency than the muzzle blast modulation spectra.

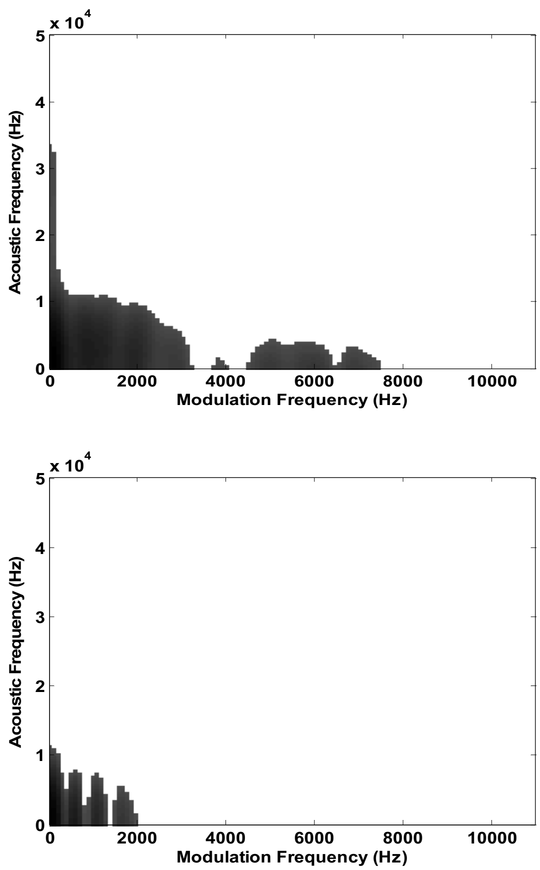

Figure 5. Preprocessed MS corresponding to those shown in Figure 4. The MS for the shock wave (top) clearly extends further in both acoustic and modulation frequency than that for the muzzle blast (bottom).

We computed the modulation spectra for 123 shock wave/muzzle blast transient pairs, extracting the frequency extent features from each modulation spectra. Figure 6 shows the extracted features in the 2-D feature space. Here we see that the features for shock waves are well separated in the feature space from those of the muzzle blast, and that a simple linear threshold may be used to distinguish the two transient types.

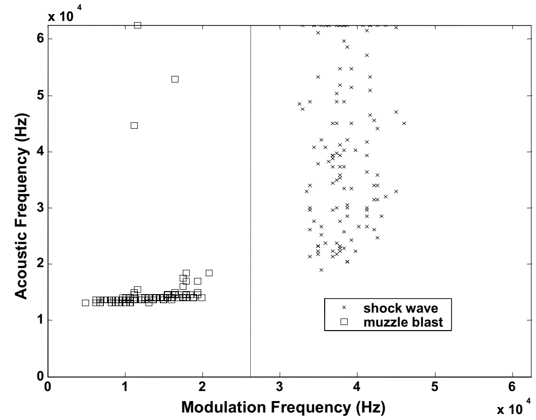

Figure 6. Spectral extent features from the shock waves (crosses) and muzzle blasts (squares). A simple linear threshold may be used to distinguish the transient classes in the feature space.

7. FUTURE WORK

The results presented in this paper support the analysis of the theoretical model of sniper fire acoustic transients which suggests that modulation spectral representations may be of use in representing the signals for classification. Current work centers on a robust system which can handle noisy signals, including other high-energy transient events, and which performs detection, noise reduction, and false alarm rejection as well as classification. Previous work has shown that many types of noise which may be encountered in military environments manifest themselves primarily in regions of the modulation spectral plane which are disjoint from those of the signals of interest. Long-term goals include expansion to classification of other types of military acoustic signals.

We acknowledge the help of Keith Davidson of UW Applied Physics Lab for his assistance in some of the experiments.

8. REFERENCES

[1] "Optimal and wavelet-based shock wave detection and estimation," B.M. Sadler, T. Pham, and L.C. Sadler, *J. Acoust. Soc. Am.* 104 (2), Pt. 1, August 1998, 955-963.

[2] Gervasio Prado, Hardave Dhaliwal, and Philip O. Martel, "Acoustic sniper localization system," *Proceedings of SPIE -- Volume 2938, Command, Control, Communications, and Intelligence Systems for Law Enforcement*, Edward M. Carapezza, Donald Spector, (Eds.), February 1997, pp. 318-325.

[3] M. S. Vinton and L. E. Atlas, "Scalable and progressive audio codec," *Proceedings ICASSP '01*, May 2001, pp. 3277 - 3280.