A Spectral Subtraction Based Algorithm for Real-time Noise Cancellation with Application to Gunshot Acoustics

António L. L. Ramos, Sverre Holm, Sigmund Gudvangen, and Ragnvald Otterlei

Abstract—**This paper introduces an improved spectral subtraction based algorithm for real-time noise cancellation, applied to gunshot acoustical signals. The derivation is based on the fact that, in practice, relatively long periods without gunshot signals occur and the background noise can be modeled as being short-time stationary and uncorrelated to the impulsive gunshot signals. Moreover, gunshot signals, in general, have a spiky autocorrelation while typical vehicle noise, or related, is periodic and exhibits a wider autocorrelation. The Spectral Suppression algorithm is applied using the pre-filtering approach, as opposed to post-filtering which requires *a priori* knowledge of the direction of arrival of the signals of interest, namely, the Muzzle blast and the Shockwave. The results presented in this work are based on a dataset generated by combining signals from real gunshots and real vehicle noise.**

Keywords—**gunshot acoustics, noise cancellation, spectral suppression, counter sniper systems.**

I. Introduction

GUNSHOT ACOUSTICS is a research topic that finds application in forensics and development of sniper positioning systems. These systems, as currently described in the literature, use a two-step algorithm to estimate the sniper's location. First, the shockwave and the muzzle blast acoustic signatures must be detected and recognized, followed by an estimation of their respective direction-of-arrival (DOA). Second, an estimate of the actual sniper's position is calculated based on the DOA obtained from the first step, via an iterative algorithm that varies from system to system.

The first comprehensive work on the acoustics of gunshots dates from 1946 by Jesse W. M. Du Mond *et al.* [1], where the authors address the issue of propagation and dissipation of ballistic shock waves. The physics of the sonic-boom propagation has been revisited in a broader context in [2]–[5]. In 1969, W. Snow [6] laid the foundations for the sniper position estimation algorithms largely used today in sniper positioning systems. Detection and DOA estimation algorithms applied to the shockwave signal have also received some attention

A. Ramos and S. Gudvangen are with Department of Technology, Buskerud University College, Kongsberg, Norway (e-mails: antonio.ramos@hibu.no; sigmund.gudvangen@hibu.no).

S. Holm is with the Department of Informatics, University of Oslo, 0373 Oslo, Norway (e-mail: sverre@ifi.uio.no).

R. Otterlei is at the head of POSICOM, Kongsberg, Norway (e-mail: ragnvald@posicom.no).

in relatively recent works [7]–[10] in the context of sniper localization systems.

Field tests have shown that detecting and estimating the DOA of the muzzle blast is a rather difficult task in real life situations. This is particularly true for long range detection in noisy environments and absorbing terrains, e.g., snow. At low SNR, the miss-detection rate increases and the system's performance is compromised. In [11], [12] the authors address, respectively, the DOA estimation and the background noise cancellation problems, making use of the beamforming, adaptive filtering, and spectral subtraction techniques in a unified approach.

This paper introduces a spectral subtraction based algorithm for real-time noise cancellation, applied to gunshot acoustical signals. This work has been partially presented in [13]. In our derivation, the background noise is modeled as being short-time stationary and uncorrelated with the impulsive gunshot signals. In the sniper positioning application, relatively long periods without signal occur and can be used to estimate the noise spectrum, as required in the spectral subtraction technique. Moreover, gunshot signals have a spiky autocorrelation in general, while typical vehicle noise, or related, is periodic and exhibits a wider autocorrelation. Nonetheless, sniper positioning systems require recognition and classification algorithms to be implemented for robustness sake, allowing them to discriminate gunshot signals from other possible sources of impulsive sounds, e.g., vehicle door slams, scrims, etc.

This work is organized as follows. Section II presents an overview of gunshots acoustics. The spectral subtraction technique is discussed in Section III in the context of gunshot acoustics, and the modified spectral suppression algorithm for gunshot signals is introduced. Application results using real data are presented in Section IV and, finally, conclusions are summarized in Section V.

II. Overview of the Acoustics of Gunshots

There are three different types of sound associated with the firing of a gun: the mechanical action sound, the muzzle blast, and the shockwave. The **mechanical action originated sounds**, useful in forensics [14], comprise three different sounds whose origin is associated with the trigger, the hammer, and the ejection of spent cartridges, respectively. These sounds are detectable only if a sensor is placed in the proximity of the gun.

For sniper positioning applications, however, only the muzzle blast and the shockwave are the events of interest (see Fig. 1).

A. The Muzzle Blast

The **muzzle blast** results as a consequence of the sudden expansion of gas following the explosion in the gun barrel. It radiates in all directions propagating at the speed of sound (e.g., 343 m/s at 20 degrees Celsius) and lasts typically for less than 3 ms. Although the sound pressure is strongest in the direction the gun barrel is pointing to, the muzzle sound can be modeled as an acoustical monopole in the far-field. Furthermore, the sound pulse energy is directly proportional to the volume of gas flow rate (volume velocity) at the source and becomes very sensitive to background noise and other sources of interference with increasing range. Indeed, the source's sound pressure decreases exponentially with time and the rate of decrease depends on the flow characteristics of the source.

The muzzle blast is not always detectable, specially if sound suppressors are used. In some cases, the sensor, or array of sensors, can be several hundred meters away from the firing gun. In those cases, due to propagation losses, the muzzle sound pressure may drop below the noise floor and therefore become difficult to detect. Unfortunately, this issue has not received due attention in the context of sniper positioning systems, even though its correct detection and estimation is as important as the shockwave's in the sniper positioning systems context.

B. The Shockwave

A ballistic **shockwave**, present only when the bullet is moving at supersonic speed, is characterized by a sudden rise in pressure followed by an approximately linear decline to a value nearly as far below as the original rise and then an almost instant return to the atmospheric pressure [1]. The shape of the resulting wave form resembles that of an "N" (see Fig. 2a) and is sometimes referred to as the N-wave. In Figs. 2a and 2b,
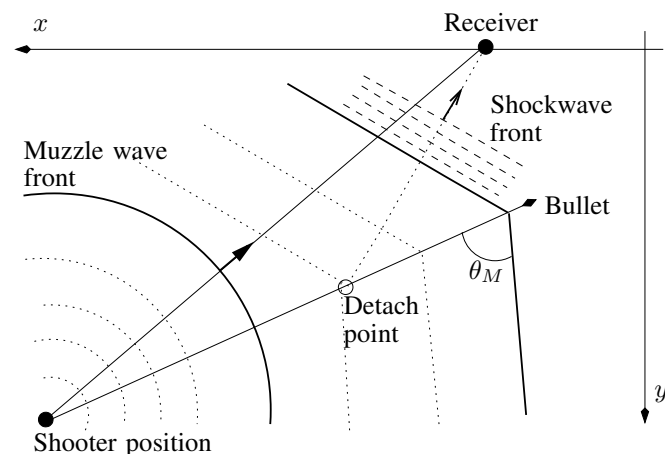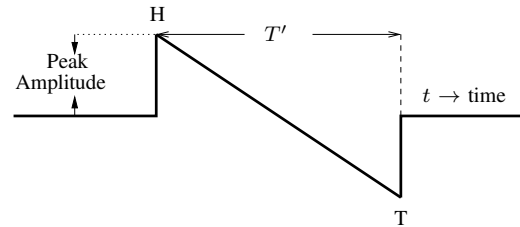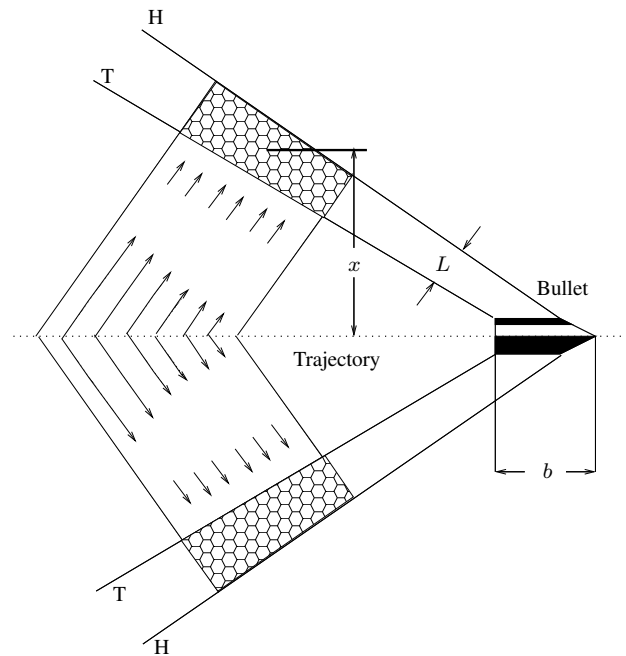
Fig. 1. A typical configuration for gunshot recording. θ_M is the Mach angle, and varies inverse proportionally to the bullet velocity. The detach point is the point on the bullet trajectory where the shockwave that reaches a given sensor is generated.

Fig. 2. The Shockwave. (a) The N-wave. (b) Path of Energy transfer from bullet to shock wave. x: miss-distance; L: wavelength.

H and T stand for head and tail, respectively, and represent the sudden rise and decline in pressure caused by the head and tail of the bullet as it travels through the air at supersonic speed. These two discontinuities, usually referred to as Bow-shock and Trailing-shock, respectively, are separated by a time interval T' known as the period of the N-wave.

The period T' varies proportionally to the bullet length and the propagation time. The last holds because the two discontinuities H and T travel at different speeds. Moreover, the head discontinuity travels at a velocity slightly higher than the speed of sound whereas the tail discontinuity travels at a slightly slower velocity than the speed of sound. These facts are depicted in Figure 2b, where the increase of the time interval, T', between the bow and trailing shocks as they propagate outwards from the bullet trajectory is evident. Figure 2b also illustrates how the energy is transferred from the bullet to the shockwave. The average energy decreases as the shockwave propagates due to three main factors [1]: 1) increase in the mean radius x, the distance from the measured shockwave pulse to the bullet trajectory, also known as the miss-distance; 2) increase in the wave-length L; and 3) dissipation into heat. In between the two main shock waves there can also be secondary shockwaves which are commonly referred to as
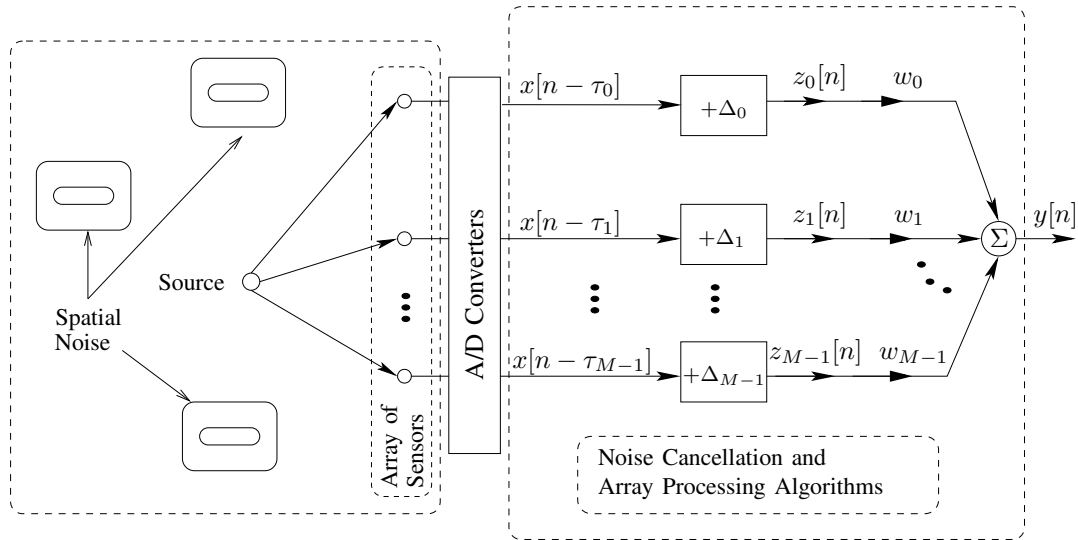
Fig. 3.   Signal model.

compression and expansion waves. As they propagate, they can merge into the main shock waves, or give rise to new shock waves if they have enough power.

By extrapolating the sketch in Fig. 2b to three-dimensional space, it is evident that the shockwave exhibits a cone shape involving the bullet trajectory and, therefore, it is not detectable when the bullet is moving away from the sensor. The angle of propagation θ_M with regard to the bullet trajectory is referred to as Mach angle and is given by

$$\theta_M = \sin^{-1}\left(\frac{1}{M}\right), \tag{1}$$

where $M = v/c$ is the Mach number, with v and c being the instantaneous bullet velocity and the speed of sound, respectively. The sound velocity in air is related to the air temperature according to [8]

$$c = c_0\sqrt{1 + \frac{T}{273}}, \tag{2}$$

where T is the air temperature in degree Celsius and $c_0 = 331$ m/s is the sound velocity at $T = 0$ ºC.

III. THE SPECTRAL SUBTRACTION METHOD

The spectral suppression algorithm was introduced in [15] in the context of noise reduction in speech signals. Since then, the algorithm has been widely studied and various improvements or modifications have been proposed for many different applications, e.g., speech enhancement [16]–[18], and automotive [19] and biomedical engineering [20]. This paper discusses the application of the spectral subtraction method for denoising of acoustic gunshot signals in the context of sniper positioning systems. Usually, these systems make use of an array of microphones spatially distributed in a 3-D geometry and various signal processing algorithms are needed either for preprocessing, e.g., noise reduction, or for estimation of parameters of interest such as the direction-of-arrival of the incoming signals. A complete signal model is illustrated

in Fig. 3 above, where the effect of spatially distributed noise sources is also accounted for. The Spectral Suppression algorithm is applied using a pre-filtering approach, as opposed to post-filtering which requires *a priori* knowledge of the direction of arrival of the signals of interest, namely, the Muzzle blast and the Shockwave.

A. The Signal Model for Noise Reduction

Consider a 3-D geometry array of microphones and a signal $s(t)$ propagating from a source located in the far field. The signal received at the mth microphone is given by

$$\begin{aligned} x_m[n] &= x_m(nT) \\ &= s[n - \tau_m(\phi,\theta)] + v_m(n), \end{aligned} \tag{3}$$

where $x_m[n]$ denotes a discrete-time version of the propagating signal received at the mth sensor and sampled at a sampling period T and ϕ and θ are the azimuth and elevation angles, respectively. The quantity $\tau_m(\phi,\theta)$ represent the time-delay measured at microphone m with respect to a predefined origin, whereas $v_m[n]$ is the additive noise component measured at the mth microphone. Furthermore, $s[n - \tau_m(\phi,\theta)]$ represents a discrete-time version of the signal of interest, $s(t)$. The model as given by (3) assumes that all microphones are properly calibrated.

B. The Spectral Subtraction Algorithm Revisited

Taking the Fourier transform of both sides of (3) yields

$$X_m(e^{jw}) = S(e^{jw})e^{-j\tau_m(\phi,\theta)} + V_m(e^{jw}). \tag{4}$$

The spectral subtraction algorithm is applied to each channel before the beamforming [11], [12] using same spectrum estimate and runs independently of the time delay, τ_m. We can therefore simplify the notation by dropping the subscript and disregarding τ_m in (4), yielding

$$X(e^{jw}) = S(e^{jw}) + V(e^{jw}). \tag{5}$$

The power spectrum of the signal of interest, $|S(e^{jw})|^2$, can be estimated according to

$$|\hat{S}(e^{jw})|^2 = |X(e^{jw})|^2 - [\overline{V}(e^{jw})]^2, \qquad (6)$$

and its respective Fourier transform can be expressed as

$$\hat{S}(e^{jw}) = \left[|X(e^{jw})|^2 - [\overline{V}(e^{jw})]^2\right]^{\frac{1}{2}} e^{j\sphericalangle X(e^{jw})}, \quad (7)$$

where $\sphericalangle X(e^{jw})$ denotes the phase of $X(e^{jw})$ and $[\overline{V}(e^{jw})]^2$ is the average value of the background noise power, $|V(e^{jw})|^2$, estimated using successive overlapping frames.

Equation (7) can be rearranged into the following form

$$\begin{aligned}
\hat{S}(e^{jw}) &= \left[1 - \frac{|\overline{V}(e^{jw})|^2}{|X(e^{jw})|^2}\right]^{\frac{1}{2}} X(e^{jw}) \\
&= H(e^{jw})X(e^{jw}), \qquad (8)
\end{aligned}$$

where $H(e^{jw})$ is the so called spectral subtraction filter.

The term $\left[|X(e^{jw})|^2 - [\overline{V}(e^{jw})]^2\right]$ in (7) can exhibit negative values at frequencies where there is no or little signal energy in $|X(e^{jw})|$, and the additive noise component is less than the average value, $\overline{V}(e^{jw})$. To overcome this problem one can use the so called nonlinear rectification which consists of replacing the possible negative values in $\left[|X(e^{jw})| - \overline{V}(e^{jw})\right]$ by zero [15], [16], [18]. This approach, however, can introduce a specific type of undesirable disturbance known as "musical noise" that has a significant impact in the case of speech enhancement applications.

C. The Modified Spectral Subtraction Algorithm

Let $P_v(w, \lambda)$ denote the smoothed power spectrum of the background noise, $v(n)$, at frequency bin w, and frame index λ of length L samples. The background noise is assumed to be short-time stationary and uncorrelated to the gunshot signals. The smoothed power spectrum can then be estimated recursively during periods of noise only according to

$$P_v(w, \lambda) = \alpha P_v(w, \lambda - 1) + (1 - \alpha)X(w, \lambda), \quad (9)$$

where $0 < \alpha < 1$ is the so called smoothing parameter. It was shown in [21] that, for the case of speech enhancement applications, the background noise can be considered as non-stationary and, therefore, the smoothing of (9) with a fixed parameter α might lead to estimation inaccuracies. When denoising short-time or impulsive events, however, the background noise can be considered as being short-time stationary and (9) provides good estimates for $0.94 < \lambda < 0.98$. The background noise power estimate is updated during noise only periods and frozen during the occurrence of a impulsive event. This is easily accomplished using a simple impulse detector [22], [23]. In a broader framework, recognition is also required in order to select the impulsive events of interest for further processing, e.g., DOA and caliber estimation.

The algorithm can be summarized as follows. During the initialization stage, a few frames are used to generate an initial spectral power estimate. The power spectral subtraction is not applied at this stage. After the algorithm is initialized, the power spectral subtraction is then applied for each subsequent frame. The output signal, $\hat{s}(t)$, is then obtained using the
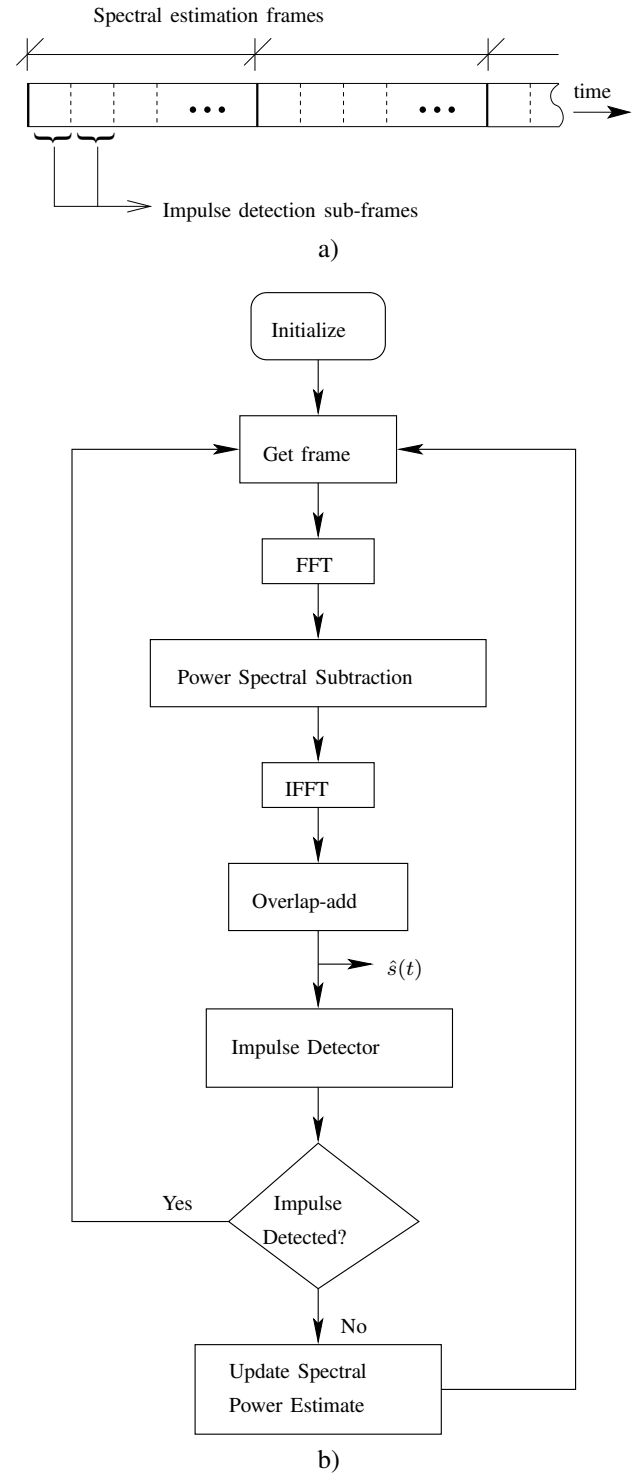

Fig. 4. Summary of the modified spectral subtraction algorithm. (a) Signal framing: 40% overlapping samples for spectral estimation; and smaller non-overlapping sub-frames for impulse detection. (b) Algorithm flow-cart.

overlap-add method. This generates a pre-processed block of samples in the time-domain at each iteration. This larger block of samples is then subdivided into smaller frames using a rectangular window (see Fig. 4a) in order to search for the occurrence of impulsive events using the impulse detector sub-routine. The updating of $P_v(w, \lambda)$ will depend on whether an impulse signal is detected or not.
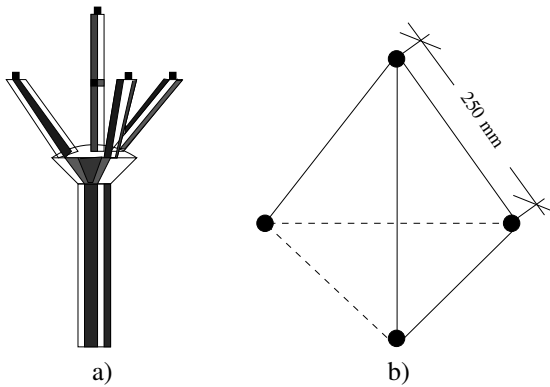
Fig. 5. Layout of the current prototype's array. Sensors are placed 250 mm apart from each other. Actual sensors' xyz-coordinates are listed in Table I. (a) Array layout. (b) Sensors geometry.

Fig. 6. Impulse detection.

The advantage of this approach for denoising gunshot signals is two-fold: 1) Having the impulse detector operating on a signal with higher SNR, improving thereby the detection rate; and 2) Providing an enhanced output signal that can be used at the recognition and classification stage. The algorithm is summarized in Fig. 4b.

D. Implementation Issues

In [15], a number of other key issues regarding the implementation of the spectral subtraction algorithm, e.g, residual noise reduction, additional noise reduction during noise only periods, magnitude averaging, input-output data buffering, have been discussed, in the context of a speech enhancement application. Although most of these ideas can also be implemented here in a straightforward manner, we highlight the fact that there are two fundamental differences in connection with the nature of speech and gunshot signals. The first one has to do with the windowing method as illustrated in Fig. 4. The second one is regarding magnitude averaging using neighboring frames that must be avoided when denoising gunshot signals using the spectral suppression method, given their impulsive nature.

TABLE I
ACTUAL POSITION OF THE SENSORS ON THE CARTESIAN PLANE

Mic Number	Mic position (m)		
	x	y	z
1	0.144	0	1.5
2	-0.072	0.125	1.5
3	-0.072	-0.125	1.5
4	0	0	1.704

IV. APPLICATION RESULTS ON REAL DATA

This section discusses the results of applying the modified spectral subtraction algorithm presented in this work on real data.

A. Data Acquisition

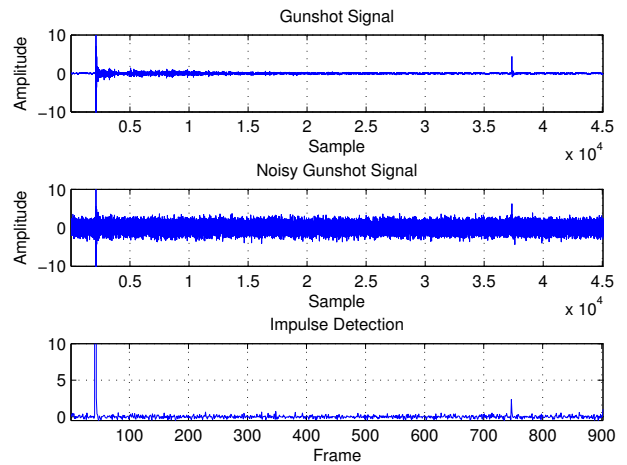The experiment was carried out using real gunshot acoustical data acquired using a 4-microphone array that is currently being used in our sniper positioning system prototype. Our data base comprises shooting samples from various ranges ($60 - 800$ meters), and miss-distances. The array layout is as depicted in Fig. 5, and the actual position of the sensors are as given in Table I. The signals were recorded as cleanly as possible, with both shockwave and the muzzle blast signatures present, and using a sampling rate of 96 kHz. The noise was added during the experiment in order to have better control over the input SNR. The noise component was generated using a steel tracked military tank running on asphalt and includes, therefore, the sound from the vehicle engine, which varies slightly in frequency over time according to the engine's rpm, and the sound from the steel tracks as the vehicle moves.

B. Experimental Results

The spectral suppression algorithm was run using a (Hamming window) frame of 500 samples to estimate the power spectrum and a (rectangular window) sub-frame of 100 samples for the impulse detector, yielding a total of 5 sub-frames. In our implementation, we used the robust impulse detector of [23]. Its application is illustrated in Fig. 6 under additive white Gaussian noise, with both shockwave and muzzle blast included. The final experimental results are plotted in Fig. 7. Here we highlight the muzzle sound which is most affected by additive noise. Figure 7a illustrates the result of applying the impulse detector to the noisy muzzle blast signal whereas Fig. 7b illustrates the result of applying the modified spectral suppression algorithm to the same signal. The input SNR was set to about 0 dB and the measured SNR at the output was roughly 5 dB. As can be seen from the comparison of the spectrograms of the noisy and denoised signals, Figs. 7a and 7b, respectively, the algorithm performed very well.

V. CONCLUSION

This paper presented an improved spectral subtraction based algorithm for real-time noise cancellation, with application to gunshot acoustical signals. The experimental results using real
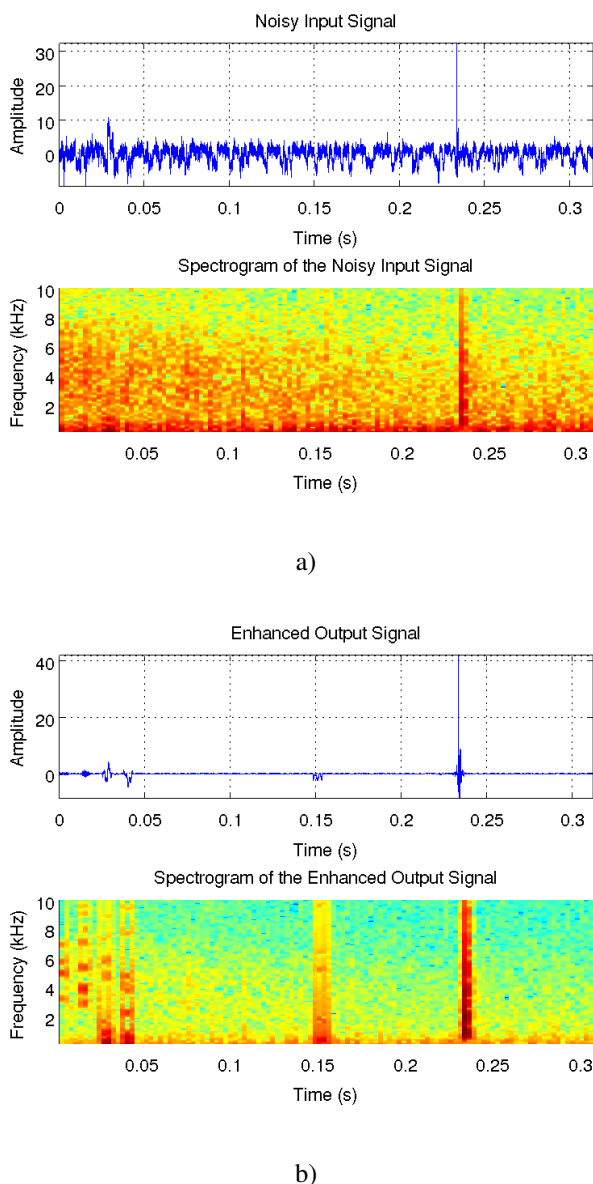
a)

b)

Fig. 7. Simulation results: denoised output emphasizing the muzzle blast signal. (a) Input signal. (b) Output signal.

data are very promising and show that the algorithm can be used effectively in order to improve the reliability of sniper positioning systems, specially in low SNR environments such as a battle field. The algorithm can potentially be used in other applications, such as seismic exploration and underwater acoustics.

REFERENCES

[1] J. W. M. DuMond, E. R. Cohen, W. K. H. Panofsky, and E. Deeds, "A determination of the wave forms and laws of propagation and dissipation of ballistic shock waves," *The Journal of the Acoustical Society of America*, vol. 18, no. 1, pp. 97–118, 1946.

[2] A. D. Pierce, "Statistical theory of atmospheric turbulence effects on sonic-boom rise times," *The Journal of the Acoustical Society of America*, vol. 49, no. 3B, pp. 906–924, 1971.

[3] H. E. Bass, B. A. Layton, L. N. Bolen, and R. Raspet, "Propagation of medium strength shock waves through the atmosphere," *The Journal of the Acoustical Society of America*, vol. 82, no. 1, pp. 306–310, 1987.

[4] R. Raspet, H. E. Bass, L. Yao, P. Boulanger, and W. E. McBride, "Statistical and numerical study of the relationship between turbulence and sonic boom characteristics," *The Journal of the Acoustical Society of America*, vol. 96, no. 6, pp. 3621–3626, 1994.

[5] P. Boulanger, R. Raspet, and H. E. Bass, "Sonic boom propagation through a realistic turbulent atmosphere," *The Journal of the Acoustical Society of America*, vol. 98, no. 6, pp. 3412–3417, 1995.

[6] W. B. Snow, "Survey of acoustic characteristics of bullet shock waves," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 4, pp. 161–176, 1967.

[7] B. M. Sadler, T. Pham, and L. C. Sadler, "Optimal and wavelet-based shock wave detection and estimation," *The Journal of the Acoustical Society of America*, vol. 104, no. 2, pp. 955–963, 1998.

[8] R. C. Maher, "Acoustical characterization of gunshots," *SAFE 2007: Workshop on Signal Processing Applications for Public Security and Forensics*, pp. 109–113, 11–13 April 2007, Washington D.C., USA.

[9] A. Lédeczi, P. Völgyesi, M. Maróti, G. Simon, G. Balogh, A. Nádas, B. Kusy, S. Dóra, and G. Pap, "Multiple simultaneous acoustic source localization in urban terrain," in *IPSN '05: Proceedings of the 4th international symposium on Information processing in sensor networks*. Piscataway, NJ, USA: IEEE Press, 2005, p. 69.

[10] J. R. Aguilar, R. A. Salinas, and M. A. Abidi, "Acoustical model of small calibre ballistic shock waves in air for automatic sniper localization applications," in *Poceedings of SPIE, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VI*, 2007.

[11] A. L. L. Ramos, S. Holm, S. Gudvangen, and R. Otterlei, "Delay-and-sum beamforming for direction of arrival estimation applied to gunshot acoustics," *Poceedings of SPIE, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X*, vol. 8019, May 2011.

[12] ——, "Real-time vehicle noise cancellation techniques for gunshot acoustics," *Poceedings of SPIE, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X*, vol. 8359, May 2012.

[13] ——, "A modified spectral subtraction algorithm for real-time noise reduction applied to gunshot acoustics," in *Poceedings of the International Conference on Signals and Electronic Systems, ICSES 2012*, Wroclaw, Poland, September 2012.

[14] B. M. Brustad and J. C. Freytag, "A survey of audio forensic gunshot investigations," in *Audio Engineering Society Conference: 26th International Conference: Audio Forensics in the Digital Age*, July 2005.

[15] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, no. 2, pp. 113–120, April 1979.

[16] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, December 1984.

[17] M. Mizumachi and M. Akagi, "Noise reduction by paired-microphones using spectral subtraction," *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 1001–1004, May 1998.

[18] H. Gustafsson, S. Nordholm, and I. Claesson, "Spectral subtraction using reduced delay convolution and adaptive averaging," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 8, pp. 799–807, November 2001.

[19] J. Meyer and K. Simmer, "Multi-channel speech enhancement in a car environment using wiener filtering and spectral subtraction," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 1167–1170, April 1997.

[20] T. Tosanguan, R. Dickinson, and E. Drakakis, "Modified spectral subtraction for de-noising heart sounds: Interference suppression via spectral comparison," in *IEEE Biomedical Circuits and Systems Conference. BioCAS 2008*, November 2008, pp. 29–32.

[21] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, July 2001.

[22] T. Kasparis and J. Lane, "Suppression of impulsive disturbances from audio signals," *Electronics Letters*, vol. 29, no. 22, pp. 1926–1927, October 1993.

[23] A. Dufaux, "Detection and recognition of impulsive sound signals," Ph.D. dissertation, University of Neuchâtel, Switzerland, 2001.