ROBUST DOA ESTIMATION OF HEAVILY NOISY GUNSHOT SIGNALS

Ângelo M.C.R. Borzino†, *José A. Apolinário Jr.*†, *and Marcello L.R. de Campos*‡

† Military Institute of Engineering (IME)
Program of Defense Engineering
Rio de Janeiro, RJ – Brazil

‡ Federal University of Rio de Janeiro (UFRJ)
Program of Electrical Engineering (COPPE/PEE)
Rio de Janeiro, RJ – Brazil

E-mails: borzino@ime.eb.br, apolin@ime.eb.br, and mcampos@ieee.org

ABSTRACT

Direction of Arrival (DOA) estimation of gunshots is an important asset to law-enforcement agencies and defense forces, for shooter localization is key to improving public and troop safety. Solutions vary according to scenario, application, and available resources. As the distance between the firing position and the sensor array increases, as in a typical sniper scenario, signal to noise ratio decreases and the estimation degrades. This paper proposes a gunshot DoA estimation algorithm to be used with highly noisy signals. We combine exhaustive search for selecting pairs of microphones from the array to attain the best DOA estimation results and fast response time for different shooting scenarios. We are particularly interested in highly corrupted signals for which state-of-the-art algorithms fail. Experimental results from simulated and recorded gunshot signals were used to evaluate the performance of the proposed scheme.

***Index Terms*—** Gunshot signal, direction of arrival, iterative least squares, exhaustive search, consistent fundamental loop.

1. INTRODUCTION

The first work related to the propagation of ballistic waves [1] dates from 1946, in which experiments to ascertain the wave forms and laws of propagation and dissipation of ballistic shock waves were described. Issues related to the physics of the sound propagation [2], useful to gunshot signal analysis, appeared in 1971.

A typical signal from a rifle gunshot is impulsive and consists mostly of two characteristic waves: the muzzle blast (MB) and the shock wave (SW). The first component is a consequence of the explosion of the charge in the gun barrel, lasts 3 to 5 milliseconds, and propagates through the air at the speed of sound [3]. The latter is due to the dispersion of air molecules caused by motion of the projectile when traveling at supersonic speed [3] and usually arrives first in a microphone when it is located in the shock wave field of view. The "N" shape shockwave signal lasts typically 0.3 to 0.5 milliseconds. Fig. 1 shows a gunshot signal originated from a 7.62 mm M964 Light Automatic Rifle, at a distance around 300 meters from the recording position. The high frequency portion of the signal on the left side of this figure refers to the shock wave and its reflections, and the following one (lower frequency) corresponds to the muzzle blast and its reflections.

Assuming we are interested in the shooter direction, we estimate the DOA of the MB component (assumed detected [4]), which travels in a straight line from the shooter to the sensors. On the other hand, the SW component that arrives at the sensors originates in

Authors thank Brazilian Agencies CAPES (Project Prodefesa, Process no. 23038.009094/2013-83), CNPq, and FINEP for partial funding of this work.

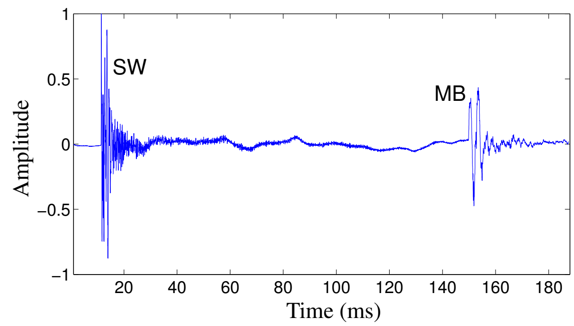

Fig. 1. Shockwave and muzzle blast components originated from a rifle 300 m from the recording microphone.

a point known as "detach point" [5] or "originating point" [6], located in the bullet trajectory. When the SW component is present, a coarse estimation of the shooter location can be obtained from a single microphone array [7]. When the gunshot is far from the recording position (say more than 400 meters), signal to noise ratio (SNR) decreases and DOA estimation degrades, as noise peaks may be confused with MB peaks, crucial to a reliable estimation.

Fig. 2 shows a gunshot signal originated from a rifle 500 meters from the sensor. The MB component, highlighted by the red rectangle, has amplitudes comparable to the noise peaks, which hinders DOA estimation considerably.

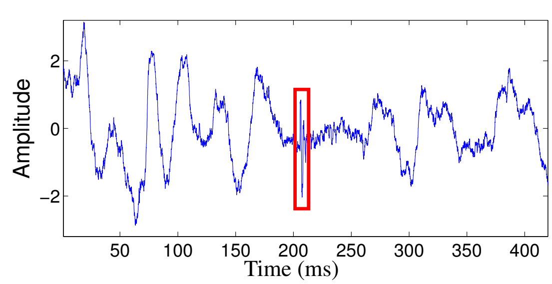

Fig. 2. MB signal of a rifle 500 meters away from the recording position.

A gunshot signal may also be degraded by reflection. For these cases, a deconvolution scheme was proposed in [8] which improves DOA estimation for SNRs between 3 dB and 12 dB. Furthermore, in [9], a refinement of the estimated DOA is attained by removing microphones that are not in the direction of the first estimate and is useful to eliminate reflections that reach those microphones with higher intensity. This method is applicable for SNR greater than

10 dB. In [10], a data selection algorithm, known as Iterative Least Squares (ILS) [11], [12], applied after spectral subtraction [5], [13] improves DOA estimation for gunshot signals with SNR between 2 and -3 dB. Nevertheless, in some cases, the distance between a sniper and its target is such that the SNR is lower than -3 dB. We propose, in this paper, a DOA estimation algorithm applicable to gunshot signals with SNR below these values.

The paper is organized as follows: Section 2 describes briefly the method of gunshot DOA estimation employed herein, as well as the ILS algorithm. Section 3 explains the exhaustive search (ES) approach and reviews the concept of consistent fundamental loop (cFL) applied to DOA estimation. Section 4 shows and discusses experimental results. Section 5 concludes the paper.

2. DOA ESTIMATION

There are several algorithms [14] that can be employed in DOA estimation, including "Delay-and-Sum beamforming," "Capon," "MUSIC," and "Generalized Cross Correlation (GCC)" [15]. The last one can be employed with wideband signals, such as a gunshot signal. A subclass of the GCC technique, used in this work, is the PHAT (Phase Transform) [15].

The direction of arrival can be characterized by two angles [16]: ϕ, azimuth, and θ, zenith. Fig. 3 shows the array with seven microphones used in our experiments, as well as angles ϕ and θ. The azimuth angle varies between zero and 360 degrees and the zenith varies between zero and 180 degrees. The unit vector in the direction of the wavefront propagation, $\mathbf{a}_{\theta,\phi}$, is given as

$$\mathbf{a}_{\theta,\phi} = \left[-\sin\theta\cos\phi \quad -\sin\theta\sin\phi \quad -\cos\theta\right]^{\mathrm{T}}. \qquad (1)$$

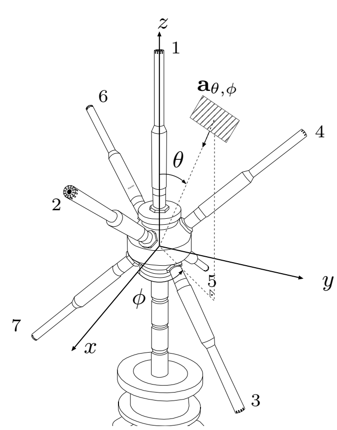

Fig. 3. Microphone array and angles of interest [11].

2.1. The GCC PHAT

Considering an array with *M* microphones, we have a total of N $= \frac{M(M-1)}{2}$ possible pairs of microphones and an equal number of cross-correlations which will be used by GCC PHAT in the DOA estimation procedure. Let τ_{ij} be the time difference of arrival (TDOA), in samples, between microphones i and j; it may be estimated from the peaks of the cross-correlation between their signals as in

$$\tau_{ij} = arg\, max\, r_{x_i x_j}(\tau), \qquad (2)$$

where $x_i(n)$ and $x_j(n)$ are the signals arriving at the i^{th} and j^{th} microphones, respectively, and $r_{x_i x_j}(\tau) = \mathrm{E}\left[x_i(n)x_j(n-\tau)\right]$ their cross-correlation, estimated as

$$\hat{r}_{x_i x_j}(\tau) = \frac{1}{L}\sum_{n=0}^{L-1} x_i(n)x_j(n-\tau), \qquad (3)$$

for $x_i(n)$ with a sample support of size L. In order to provide more accurate results, we can interpolate the correlations around the peak [16].

This cross-correlation estimate corresponds to the convolution $\hat{r}_{x_i x_j}(\tau) = \frac{1}{L}\left[x_i(\tau) * x_j(-\tau)\right]$, and is usually computed as the inverse Fourier transform of the cross-power spectrum density (CPSD) $\frac{1}{L}\left[X_i(e^{j\omega})X_j(e^{-j\omega})\right]$. In the case of GCC PHAT, the CPSD is normalized by its absolute value prior to the inverse Fourier transform [15].

Defining the TDOA in time unit as $\overline{\tau}_{ij} = \frac{\tau_{ij}}{f_s}$, f_s being the sampling frequency, we know that it corresponds to the time the sound travels distance d_{ij}, from microphone i to microphone j, which corresponds to

$$d_{ij} = \mathbf{a}_{\theta,\phi}^{\mathrm{T}}(\mathbf{p}_i - \mathbf{p}_j), \qquad (4)$$

where $\mathbf{p}_i$ and $\mathbf{p}_j$ are the microphone coordinates. Therefore, we can also write

$$\overline{\tau}_{ij} = \frac{d_{ij}}{v_{sound}} = \mathbf{a}_{\theta,\phi}^{\mathrm{T}}\Delta\overline{\mathbf{p}}_{ij}, \qquad (5)$$

where $\Delta\overline{\mathbf{p}}_{ij} = \frac{\mathbf{p}_i - \mathbf{p}_j}{v_{sound}}$. If we define the least squares (LS) cost function as

$$\xi_{\theta,\phi} = (\overline{\tau}_{12} - \Delta\overline{\mathbf{p}}_{12}^{\mathrm{T}}\mathbf{a}_{\theta,\phi})^2 + ... + (\overline{\tau}_{(M\text{-}1)M} - \Delta\overline{\mathbf{p}}_{(M\text{-}1)M}^{\mathrm{T}}\mathbf{a}_{\theta,\phi})^2, \qquad (6)$$

we can find a DOA estimation by taking the gradient of this function with respect to $\mathbf{a}_{\theta,\phi}$ and making the result equal to zero. The resulting estimate is given by $\mathbf{a}_{DOA} = \mathbf{R}^{-1}\mathbf{p}$, where

$$\mathbf{R} = \Delta\overline{\mathbf{p}}_{12}\Delta\overline{\mathbf{p}}_{12}^{\mathrm{T}} + ... + \Delta\overline{\mathbf{p}}_{(M\text{-}1)M}\Delta\overline{\mathbf{p}}_{(M\text{-}1)M}^{\mathrm{T}}, \text{ and} \qquad (7)$$

$$\mathbf{p} = \overline{\tau}_{12}\Delta\overline{\mathbf{p}}_{12} + ... + \overline{\tau}_{(M\text{-}1)M}\Delta\overline{\mathbf{p}}_{(M\text{-}1)M}. \qquad (8)$$

Finally, from the elements of $\mathbf{a}_{DOA} = [a_x \quad a_y \quad a_z]$, the horizontal angle (azimuth) is given by $\phi = tan^{-1}\frac{a_y}{a_x}$ and the vertical angle (zenith) is given by $\theta = cos^{-1}(-a_z)$. This LS-based TDOA estimation method was proposed in [17] and applied in [18] to determine which technique among three closed-form localization techniques provides best results.

2.2. Other Algorithms

Steering Response Power (SRP) [19] is another algorithm that can be employed in DOA estimation. It is based on spatial spectral estimation [20], making use of the spatial correlation matrix. In [21], the equations governing a 3-dimensional geometry array are described, and the equation that corresponds to DOA estimation is presented as

$$\hat{\Theta}(\theta,\phi) = \arg\max_{\theta,\phi}\mathbf{1}^{\mathrm{T}}\hat{\mathrm{R}}_{\theta,\phi}\mathbf{1}, \qquad (9)$$

where $\hat{\mathrm{R}}_{\theta,\phi}$ denotes an estimate of the spatial correlation matrix obtained by averaging the samples that arrive at the array. In other words, SRP searches (in a grid of possible values) for the set of angles (θ,ϕ) that maximizes the array's output power.

The ILS algorithm tries to eliminate TDOAs subject to errors due to spurious signals [12] that produce undesirable peaks in the cross-correlations between signals arriving at the microphones. Our array has seven microphones, leading to a total of N $= 21$ possible pairs and an equal number of cross-correlations. This algorithm takes into account the fact that there may be cases in which not all estimated delays contribute to a good outcome [11]. The ILS algorithm eliminates correlations (among 21 possibilities) that contribute less to the minimization of the cost function in Eq. (6). In other words, the pair (i, j) of microphones that originates the highest value $(\overline{\tau}_{ij} - \Delta\overline{\mathbf{p}}_{ij}^{\mathrm{T}}\mathbf{a}_{\theta,\phi})^2$ is removed and the cost function is evaluated again, normalized by the number of terms. This procedure is repeated until only six (or five) terms remain in the cost function [11], [12].

3. EXHAUSTIVE SEARCH

Although effective in many cases, the ILS algorithm leads to a sub-optimal solution. We maintain that better results can be obtained with exhaustive search without adding unreasonable computational complexity. In the case of an array with a small number of microphones ($M = 7$ in our case), typically employed in sniper location devices, evaluating the LS cost function for all possible combinations of $n < N$ microphone pairs, improves performance considerably. Moreover, if n is not close to $N/2$, the number of combinations to be evaluated, $\begin{pmatrix} N \\ n \end{pmatrix}$, is small and would cause no noticeable delay. The number of pairs of microphones n ranges from 3 to $N = 21$ while all possible combinations can be found by incrementing a binary counter from 1 to 2^N and separating those whose sum of bits equals n. The position of each bit corresponds to a specific pair of microphones. In the following, after defining ES(n), the exhaustive search of n pairs, we see that a small value of n presents the best results, in average, for heavily noisy signals.

At first, a pseudo-code for ES(n) for an array of M microphones is described:

Set (binary counter) BC $= 1$ and N $= \frac{M(M-1)}{2}$
for BC $= 1$ to 2^{N}
  if sum of bits of BC $= n$
    Evaluate cost function $\xi_{\theta,\phi}$, in Eq. (6)
    Store $\xi_{\theta,\phi}(BC)$
  end
end
Find BC with the lowest $\xi_{\theta,\phi}(BC)$
Choose the n mics pairs from positions of bits "1"
Estimate DOA with the n corresponding $\overline{\tau}_{ij}$

As before mentioned, when the SNR decreases, noise peaks tend to increase. In these cases, several peaks appear in the cross-correlations between signals arriving at a pair of microphones, and there may be cases in which the highest peaks are due to noise, instead of gunshot signal. For very low SNRs, among all $N = 21$ possible cross-correlations, there may be cases where only a few ones have the correct highest peak due to signal, not to noise. That is the reason why we obtain more accurate results with fewer data. We might even think that the probability of having 3 correct peaks would be larger than those of having 4, 5, 6 or more correct peaks. Nevertheless, for low SNR, ES(3) does not result, in average, to better results than ES(n), $n > 3$. In Fig. 4, we take as an example the case of $n = 3$ pairs from only three microphones—(2,4), (4,6) and (6,2)—which define a plane.

Notice in Fig. 4 that, for this choice of pairs, unit vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ have the same horizontal projection (on plane $x'y'$) and therefore the same components, $\mathbf{a}_{1x} = \mathbf{a}_{2x}$ and $\mathbf{a}_{1y} = \mathbf{a}_{2y}$. Also, their

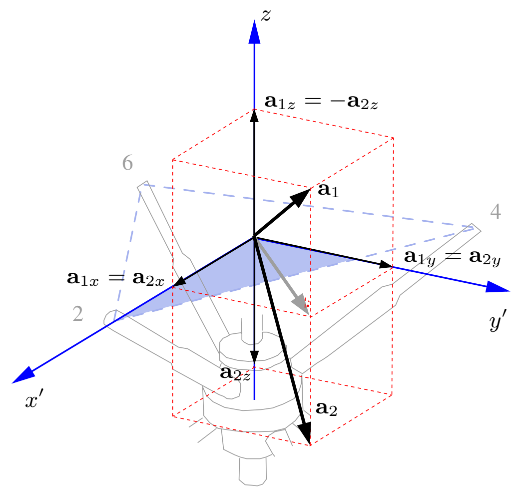

Fig. 4. Ambiguity in ES(3) when the three pairs define a plane.

vertical components are opposite, $\mathbf{a}_{1z} = -\mathbf{a}_{2z}$, which, being perpendicular to plane $x'y'$, have no influence in the delays. Hence, the azimuths of signals arriving from the directions of the unit vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are the same ($\phi_1 = \phi_2$), and their zenithal angles are symmetrical ($\theta_1 = -\theta_2$). As both vectors cause the same delay when captured by the three microphones, it becomes impossible to determine, from the three TDOAs, whether the correct DOA is $\mathbf{a}_1$ or $\mathbf{a}_2$. Also note that, in these cases, matrix $\mathbf{R}$ in Eq. (7) is singular. Because of this *ambiguity*, whenever 3 pairs define a plane, we do not have a single solution for ES(3). On the other hand, when three pairs do not define a plane—(2,4), (4,6) and (3,5), for instance—this problem does not occur and ES(3) presents a single solution.

We decided to choose the value of n which was most likely to give the best results for simulated signals. In order to assess the performance of ES(n) for different values of n, we synthesized gunshot signals using as clean reference a muzzle blast component originated from a rifle recorded at a distance $236m$ from the shooting position (SNR greater than $20dB$). We then chose randomly directions (azimuth and zenith) and gave the corresponding delays to the reference signal in order to simulate signals arriving at each microphone of the array depicted in Fig. 3. Afterwards, we added noise also recorded by the microphone array (noise only portions of the recorded signal), such that we were able to control the SNR[1]. In our experiment, for each simulated gunshot, we varied n in ES(n) from 3 to 21, and observed for which n ES(n) offered the best result, for SNR from -3 dB to -8 dB. For SNR below -8 dB, current state-of-the-art algorithms fail if denoising schemes are not applied. For SNR above -3 dB, ES(4) provides results comparable to ES(5), ES(6) or others.

Table 1 shows the relative frequency of best results, with an ensemble of 200 independent runs, for ES(3) to ES(7) with simulated signals having a SNR of -8 dB. For each run, the best result corresponds to the ES(n), n varying from 3 to 7, that provides the lowest error defined as $\mathrm{e}_\theta^2 + \mathrm{e}_\phi^2$, where $\mathrm{e}_\theta = \theta - \hat{\theta}$ is the error between simulated and estimated zenithal angles, whereas $\mathrm{e}_\phi = \phi - \hat{\phi}$ is the error between simulated and estimated azimuthal angles.

ES(4) provided the lowest (average) error for the largest number of runs for SNR equal to -8 dB. We have observed that the same happens for different SNR in the range of interest, -8 dB to -3 dB.

[1] We define SNR herein from a 7.5 ms window of clean signal containing the MB component with variance (σ_s^2). The variance of the noise (σ_n^2) is chosen such that a desired SNR (in dB) is obtained as $10 \log \left(\sigma_s^2/\sigma_n^2\right)$.

Table 1. Relative frequency of lowest error (RFLE) for ES(n)

n	3	4	5	6	7
RFLE	0.10	**0.36**	0.26	0.14	0.14

Values of n from 8 to N $= 21$ were also tested for highly noisy signals, but did not lead to better results than those with lower n.

For ES(3), the exhaustive search was carried out in a sub-space of possible combinations; those in which the three pairs of microphones defined a plane were not considered. In the cases where the 3 pairs of microphones form a plane, a Fundamental Loop (FL) [22] is formed. We say that an FL with n pairs of microphones is formed when there is a cyclic path. For example, when $n = 3$ and the microphone pairs are (i,j), (j,k), and (k,i). An FL is said to be consistent (cFL) if the sum of TDOAs in this loop is zero ($\tau_{ij} + \tau_{jk} + \tau_{ki} = 0$). This was defined in [22] as "zero cyclic sum (ZCS) condition." If all n TDOAs are correctly estimated, then ZCS condition is satisfied. However, converse is not true: if ZCS is satisfied, we cannot guarantee that all TDOAs are correct.

The concept of FL is also valid for more than three pairs of microphones and could be taken into account in the DOA estimation of gunshot signals. Whenever the DOA is perfectly estimated, all TDOAs are correctly estimated such that the cost function in Eq. (6) equals zero. In a cFL where $\sum_{ij} \tau_{ij} = 0, ij \in$ FL, we claim that the cost function $\xi_{\theta,\phi}$ formed by the terms corresponding to that loop is also zero. Herein, we chose to minimize (in the least squares sense) the cost function $\xi_{\theta,\phi}$ instead of searching for cFL solutions motivated by the fact that these criteria are equivalent when dealing with non-degraded signals. For highly noisy signals, optimizing $\xi_{\theta,\phi}$ is still effective while searching for cFLs is not.

4. EXPERIMENTAL RESULTS

In order to verify which algorithm performs better, either for simulated directions or for recorded gunshot signals, we varied the SNR starting from -3 dB and below and compared the results for four algorithms here designated as SE (standard GCC PHAT estimation with 21 pairs), ILS [11, 12], SRP [19, 21], and the proposed ES(4).

In all tests, the ES(4) algorithm produced similar or better results than the others. Fig. 5 compares SE, ILS, SRP, and ES(4) in 25 gunshots fired with random DOAs, SNR equal to -8 dB. We can observe that ES(4) performs much better than the current state-of-the-art ILS and SRP algorithm for SNR= -8 dB.

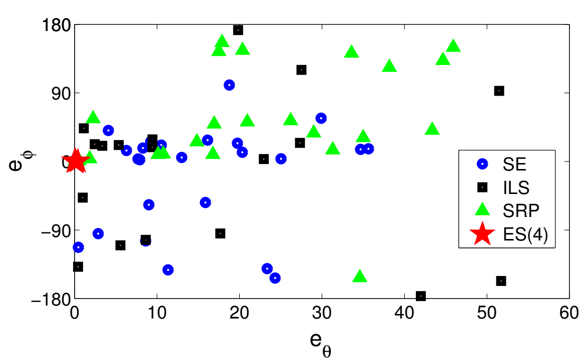

Fig. 5. Comparing ES(4).

We ran 200 random DOAs in the same manner as the simulations that originated Table 1, for each integer SNR from -8 dB to -2 dB, and computed the average error, in dB, as $10 \log \left(e_\theta^2 + e_\phi^2\right)$ for each algorithm (SE, ILS, SRP, and ES(4)). Fig. 6 shows the results. Here, again, the performance of the proposed algorithm was better than those of competing state-of-the-art algorithms.

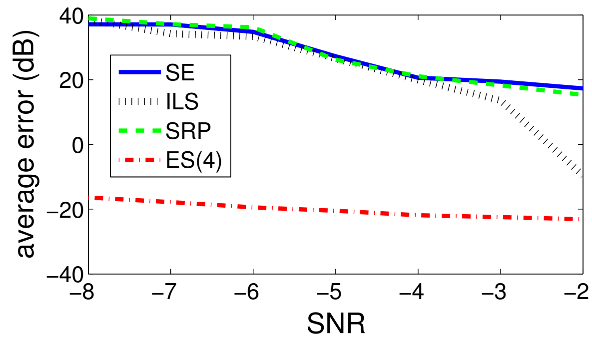

Fig. 6. Performance of each algorithm with different SNRs.

In order to verify that the results obtained with simulated signals are equivalent when dealing with recorded signals, we applied the same algorithms to 8 gunshot signals. The weapons used were the 7.62 mm M964 Light Automatic Rifle and the 5.56 mm Tavor Assault Rifle; the distance between the shooter and the microphone array was 1,062 m. The recordings took place at a Brazilian Army[2] unit responsible for the evaluation of ordnance. In these shots, the lowest SNR is equal to -8.13 dB. The array was positioned in a place so that there was a direct line of sight between the microphones and the shooter. The correct azimuthal and vertical angles were measured with a "TOPCON" topography station, model CTS3000. Fig. 7 shows the errors, and we can observe that ES(4) outperforms the other algorithms.

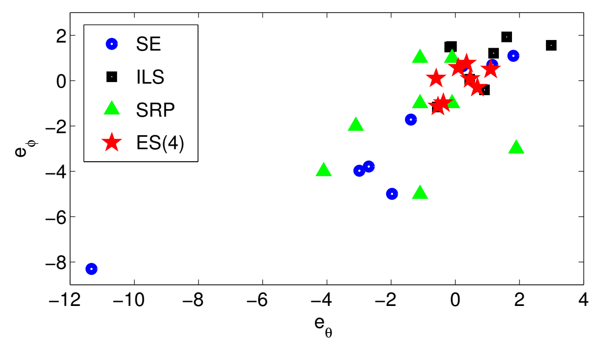

Fig. 7. Performance of each algorithm with real signals.

5. CONCLUSION

We employed exhaustive search to DOA estimation of gunshot signals and showed that the use of 4 pairs leads to better results when SNR is low. Combining ES(n) for different values of n is currently under investigation. We also explained why ES(3) does not work as well as ES(4), relating this fact to the concept of consistent fundamental loop. We compared the performance of SE, ILS, SRP, and ES(4) with simulated and real signals, and verified that, in both cases, ES(4) performs better than the others when the SNR is low, more precisely from -8 dB to -3 dB. For SNR between -3 dB and 2 dB, the algorithm also works well, but the technique presented in [10] produces satisfactory results. Above 2 dB, the ILS algorithm works very well, while for clean signals (SNR greater than 20 dB), the standard estimation is as robust as the others.

[2] *Centro de Avaliações do Exército* (Army Evaluation Center).

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(1), pp. 93–118, 1946.

[2] A. D. Pierce, "Statistical theory of atmospheric turbulance effects on sonic-boom rise times," *The Journal of the Acoustical Society of America*, vol. 49(3B), pp. 906–924, 1971.

[3] R. C. Maher, "Modeling and signal processing of acoustic gunshot recordings," in *Digital Signal Processing Workshop*. IEEE, September 2006, vol. 59, pp. 257–261.

[4] U. Libal and K. Spyra, "Wavelet based shock wave and muzzle blast classification for different supersonic projectiles," *Elsevier Expert Systems with Applications*, vol. 41, pp. 5097–5104, September 2014.

[5] A. L. L. Ramos, S. Holm, S. Gudvangen, and R. Otterlei, "A spectral subtraction based algorithm for real-time noise cancellation with application to gunshot acoustics," *International Journal of Electronics and Telecommunications*, vol. 59, pp. 93–98, April 2013.

[6] T. Mäkinen and P. Pertilä, "Shooter localization and bullet trajectory, caliber, and speed estimation based on detected firing sounds," *Elsevier Applied Acoustics*, vol. 71, pp. 902–913, October 2010.

[7] I. L. Freire and J. A. Apolinário Jr., "Localização de atirador por arranjo de microfones (in Portuguese)," in *Simpósio Brasileiro de Automação Inteligente*. SBAI, September 2011, vol. X, pp. 1049–1053.

[8] A. M. C. R. Borzino, J. A. Apolinário Jr., Marcello L. R. de Campos, and C. L. Pagliari, "Gunshot signal enhancement for DOA estimation and weapon recognition," in *European Signal processing Conference (EUSIPCO)*. EURASIP, September 2014.

[9] D. M. Paredes-Calderón and J. A. Apolinário Jr., "Estimación de la posición de un francotirador usando arreglos de micrófonos en plataformas elevadas (in Spanish)," in *Convención de Centro América y Panamá*. CONCAPAN, November 2014.

[10] A. M. C. R. Borzino, J. A. Apolinário Jr., and Marcello L. R. de Campos, "Estimating direction of arrival of long range gunshot signals," in *International Telecommunications Symposium (ITS)*. IEEE and SBrT, 2014.

[11] I. L. Freire, P. C. Prandel, and J. A. Apolinário Jr., "Sobre a escolha de sinais em arranjos de microfones estimando DOA com GCC-PHAT (in Portuguese)," in *Simpósio Brasileiro de Telecomunicações*. SBrT, September 2012.

[12] I. L. Freire, "Robust direction-of-arrival by matched-lags, applied to gunshots," *Journal of the Acoustical Society of America*, vol. 135, pp. 246–251, May 2014.

[13] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," in *Transactions on Acoustics, Speech, and Signal Processing*. IEEE, April 1979, vol. 27, pp. 113–120.

[14] H. L. V. Trees, "Optimum array processing - part iv - detection, estimation, and modulation theory," *Wiley-Interscience, New York, United States of America*, 2002.

[15] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," in *Transactions on Acoustics, Speech, and Signal Processing*. IEEE, August 1976, vol. 24, pp. 320–327.

[16] I. L. Freire and J. A. Apolinário Jr., "DOA of gunshot signals in a spatial microphone array: performance of the interpolated generalized cross-correlation method," in *Argentine School of Micro-Nanoelectronics Technology and Applications*. EAMTA, August 2011, pp. 1–6.

[17] Y. Chan, R. Hattin, and J. B. Plant, "The least squares estimation of time delay and its use in signal detection," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, April 1978, vol. 3, pp. 665–669.

[18] J. O. Smith and J. S. Abel, "Closed-form least-squares loaction estimation from range-difference measurements," in *Transactions on Acoustics, Speech, and Signal Processing*. IEEE, December 1987, vol. 35, pp. 1661–1669.

[19] J. H. DiBiase, H. F. Silverman, and M. S. Brandstein, "Microphone arrays: Digital signal processing: Robust localization in reverberant rooms," *Springer-Verlag, Berlin, Germany*, pp. 157–180, 2001.

[20] J. Dmochowski, J. Benesty, and S. Affes, "Direction of arrival estimation using the parameterized spatial correlation matrix," in *Transactions on Audio, Speech, and Language Processing*. IEEE, May 2007, vol. 15, pp. 1327–1339.

[21] António L. L. Ramos, Sverre Holm, Sigmund Gudvangen, and Ragnvald Otterlei, "Delay-and-sum beamforming for direction of arrival estimation applied to gunshot acoustics," in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X*. SPIE, May 2011, vol. 8019.

[22] J. Scheuing and B. Yang, "Disambiguation of TDOA estimation for multiple sources in reverberant environments," in *Transactions on Audio, Speech, and Language Processing*. IEEE, November 2008, vol. 16, pp. 1479–1489.