Beamforming-Based Acoustic Source Localization and Enhancement for Multirotor UAVs

Daniele Salvati, Carlo Drioli, Giovanni Ferrin and Gian Luca Foresti
Department of Mathematics, Computer Science and Physics
University of Udine, Italy
Email: {daniele.salvati, carlo.drioli, giovanni.ferrin, gianluca.foresti}@uniud.it

***Abstract*—The problem of acoustic source localization and signal enhancement through beamforming techniques is especially challenging when the acoustic recording is performed using microphone arrays installed on multirotor unmanned aerial vehicles (UAVs). The principal source of disturbances in this class of devices is given by the high frequency, narrowband noise originated by the electrical engines, and by the broadband aerodynamic noise induced by the propellers. A solution to this problem is investigated, which employs an efficient beamforming-based spectral distance response algorithm for both localization and enhancement of the source. The beamforming relies on a diagonal unloading (DU) transformation. The proposed algorithm was tested on a multirotor micro aerial vehicle (MAV) equipped with a compact uniform linear array (ULA) of four microphones, perpendicular to the rear-front axis of the drone. The array is positioned slightly above the plane of propellers, and centered with respect to the drone body. The experimental results conducted in stable hovering conditions are illustrated, and the localization and signal enhancing performances are reported under various noise conditions and source characteristics.**

***Index Terms*—Acoustic source localization, signal enhancement, beamforming, spectral distance, unmanned aerial vehicles (UAV), micro aerial vehicle (MAV).**

I. Introduction

Acoustic source localization, an important task in microphone array processing since many decades [1]–[7], has interesting application perspectives in a number of scenarios involving mobile robotic devices. These include direction of arrival (DOA) estimation in single mobile robot [8] and in mobile robot sensors networks [9], or relative position estimation in a team of drones [10]. When the acoustic recording is performed using microphone arrays installed on multirotor unmanned aerial vehicles (UAVs), the localization and signal enhancement of acoustic sources of interest becomes especially challenging, due to the number and variety of acoustic disturbances generated by this class of devices. Moreover, in the case of micro aerial vehicles (MAVs) of small size, the consequent constraints on the size of the microphone array may lead to poor signal-to-noise enhancement and poor spatial resolution issues. As a matter of fact, attempts to tackle the acoustic related problems typical of multirotor aerial systems have been documented only recently [11]–[20]. Beamforming and blind source separation are typical signal processing techniques used to this aim. In [16], the problem of enhancing a target speech source located in front of the UAV is solved using a minimum variance distortionless response (MVDR) beamformer coupled to a Wiener postfiltering scheme. The beamforming approach is the method of choice in a number of other investigations, e.g. [18]. When acoustic source localization is the only concern, localization methods with high-robustness with respect to low signal-to-noise ratios (SNRs) have also been investigated. In [19], [20] methods derived from the multiple signal classification (MUSIC) [21] were assessed, and the reported results show good localization performances. However, the method described also requires the monitoring of UAV inertial sensors and motor controls, and the learning or monitoring of propellers noise signal. Supervised noise suppression techniques relying on the online monitoring of the propellers noise, or on the off-line learning of its spectral and statistical characteristics, were also proposed in [22]. When, on the other hand, signal enhancement is the principal concern, time-frequency spatial filtering or blind source separation techniques have been used to selectively attenuate the propeller noise component while preserving the target acoustic source component. In [11], [12], the time-frequency sparsity of specific target signals is exploited through the use of a time-frequency spatial filtering technique.

With respect to acoustic source localization on UAVs, MUSIC seems at today the best choice available due to its multisource localization capability and to noise robustness. However, its performance can be seriously degraded due to constraints imposed by the UAV hardware design. For example, if a small-size microphone array must be adopted to fit the size of a micro aerial device, the resulting spatial resolution would not be sufficient to discriminate between the source signal and the propeller noise. For these reason, a method which provides acceptable spatial resolution even with small arrays and at low SNR is desirable.

In this study, we illustrate the performance of a new beamforming-based spectral distance response algorithm, which is based on the diagonal unloading (DU) beamformer recently introduced in [23], when paired to a small-size and low-cost hardware configuration consisting in a four-sensor uniform linear array (ULA) mounted on a quadcopter MAV. The algorithm provides integrated source localization and signal enhancement capabilities, and provides superior spatial resolution properties when compared to conventional beamforming methods. The experimental results are conducted in stable hovering conditions and with different controlled SNRs.

ISBN 978-90-827970-1-5

II. Method

The frequency-domain model of a typical acoustic narrow-band beamformer, i.e. a spatial filter whose goal is to achieve directional signal reception, can be stated as

$$Y(f,\theta) = \mathbf{w}^H(f,\theta)\mathbf{x}(f), \tag{1}$$

where f is the frequency bin, $\mathbf{x}(f)$ is the vector of signals captured by the array, $\mathbf{w}(f,\theta)$ is the beamformer coefficients for time-shifting, weighting, and summing the data so to steer the array in the direction θ, $Y(f,\theta)$ is the output of the narrowband beamformer, and H denotes the Hermitian (complex conjugate) transpose. The power spectral density (PSD) of the spatially filtered signal is thus

$$P(f,\theta) = E\{|Y(f,\theta)|^2\} = \mathbf{w}^H(f,\theta)\mathbf{\Phi}(f)\mathbf{w}(f,\theta), \tag{2}$$

where $\mathbf{\Phi}(f) = E\{\mathbf{x}(f)\mathbf{x}^H(f)\}$ is the PSD matrix of the array signal ($E\{\cdot\}$ denotes mathematical expectation).

In the conventional steered response power (SRP) algorithms, the corresponding broadband SRP $P(\theta)$ is obtained by integrating these narrowband SRP components $P(f,\theta)$, over all frequencies. To increase the spatial resolution, the narrowband components are in general normalized with respect to some spectral characteristic, leading to the widely used phase transform (PHAT) [24], a pre-filter that uses the magnitude information of the PSD matrix to normalize the narrowband components in the SRP conventional beamforming, or to the incoherent frequency fusion [25]. The SRP $P(\theta) = \sum_f \frac{P(f,\theta)}{\max\limits_{\theta}[P(f,\theta)]}$ [25] of a beamformer conveys information on the acoustic energy coming from direction θ, thus it will be characterized by a maximum peak corresponding to the source direction $\hat{\theta}$. Therefore, the DOA estimate of the source is obtained by $\hat{\theta} = \underset{\theta}{\mathrm{argmax}}[P(\theta)]$. In the scenario considered in this paper, the propellers can be considered point-sources that make this case a complex multisource problem with four or more sources that have a very high signal pressure level and a target source with low signal-to-propeller noise ratio (SPNR). A compact array has thus some limitations in such scenario when using conventional high-resolution techniques such the MUSIC method.

We propose here the use of the DU beamformer and spectral dissimilarity between PSDs under the assumption that the DOAs of point-source propellers from the drone are known. The new approach consists in computing a beamforming in each DOA target direction and in measuring the spectral distance between a candidate source position and the propellers directions. The direction that maximizes the spectral distance response will be the target position, and the output from the corresponding beamformer will be the enhanced version of the target signal.

The beamforming method used in this paper is the low-complexity robust DU beamformer [23], which has been proposed to address the problem of acoustic source localization. Its design is based on a diagonal removal transformation of the covariance matrix involved in the conventional beamformer computation to achieve robust localization with low computational complexity. The method is illustrated in details in [23] and here we only briefly recall its principal characteristics, and demonstrate that it can be successfully used also to address the problem of target signal enhancement.

The transformation on which the DU method is based, is obtained by subtracting an opportune diagonal matrix from the PSD matrix $\mathbf{\Phi}(f)$ of the array output vector. As a result, the DU beamforming removes as much as possible the signal subspace from the PSD matrix and provides a high resolution beampattern by exploiting the orthogonality property between signal and noise subspaces. In practice, the design and implementation of the DU transformation is simple and effective, and is obtained by computing the matrix (un)loading factor.

Given the matrix $\mathbf{\Phi}(f)$ which represents the array output vector covariance, the DU transformed matrix can be written as

$$\mathbf{\Phi}_{\mathrm{DU}}(f) = \mathbf{\Phi}(f) - \mu\mathbf{I}, \tag{3}$$

where μ is a real-valued, positive scalar, selected in such a way that the resulting PSD matrix is negative semidefinite. In the single source case with spatially white noise, the parameter μ is obtain by imposing in $\mathbf{\Phi}(f)$ that the eigenvalue corresponding to the signal subspace is null, and that the eigenvalues corresponding to the noise subspace are non-zero. The value of μ that verifies such constraints can be shown to be

$$\mu = \mathrm{tr}(\mathbf{\Phi}(f)) - (N-1)\sigma^2, \tag{4}$$

where $\mathrm{tr}(\cdot)$ is the operator that computes the trace of a matrix, N is the number of microphones, and σ^2 is the noise variance for all sensors. Theoretically, this solution totally removes the signal subspace from the PSD matrix and the beamformer output is therefore null in the source direction. However, if also signal enhancement is pursued, it is necessary that the beamformer output signal is non null, and thus part of the signal subspace should be retained in the PSD matrix. Therefore, a suboptimal DU procedure is used here:

$$\mu = \mathrm{tr}(\mathbf{\Phi}(f)). \tag{5}$$

This solution guarantees that the resulting PSD matrix $\mathbf{\Phi}_{\mathrm{DU}}(f)$ is negative semidefinite and that the signal is present at the beamformer output. The DU procedure provides a solution for the beamforming coefficients $\mathbf{w}$:

$$\mathbf{w}_{\mathrm{DU}}(f,\theta) = \frac{\mathbf{\Phi}_{\mathrm{DU}}(f)\mathbf{a}(f,\theta)}{\mathbf{a}^H(f,\theta)\mathbf{\Phi}_{\mathrm{DU}}(f)\mathbf{a}(f,\theta)}, \tag{6}$$

where $\mathbf{a}(f,\theta)$ is the array steering vector. Given the beamforming coefficients, it is thus possible to also compute the spatially filtered signal with respect to the estimated DOA, which will provide an enhanced version of the target source signal.

Let θ_p (p=1,2 and $\theta_1 < \theta_2$) denote the two DOAs of the four propellers since the microphone array is located on the top of the MAV, centered with respect to the four propellers that thus are symmetric with respect to the array.

ISBN 978-90-827970-1-5 © EURASIP 2018

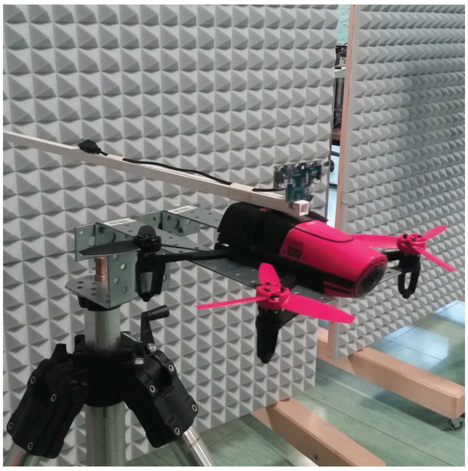

Fig. 1. The MAV and the microphone array used in the experiments.

The spectral dissimilarity is calculated with the Itakura-Saito distance measure [26]:

$$S(\theta) = \sum_{p=1}^{2} \sum_{f=0}^{L-1} \left| \frac{P(f,\theta)}{P(f,\theta_p)} - \log \frac{P(f,\theta)}{P(f,\theta_p)} - 1 \right|, \tag{7}$$

where L is the size of discrete Fourier transform. Let Δ denote the desired spatial resolution, the DOA estimate of the source is obtained by solving the following constrained problem

$$\begin{aligned} &\hat{\theta} = \underset{\theta}{\operatorname{argmax}}[S(\theta)], \\ &S(\theta - \Delta) < S(\theta) < S(\theta + \Delta), \\ &\theta_1 \leq \theta \leq \theta_2. \end{aligned} \tag{8}$$

The beamforming-based spectral distance response algorithm is summarized as follow:

1) For all θ in the range $\theta_1 \leq \theta \leq \theta_2$ with spatial resolution Δ, calculate the output signal beamformer $Y(f,\theta) = \mathbf{w}_{\mathrm{DU}}^H(f,\theta)\mathbf{x}(f)$ implemented with the overlap-add technique and calculate its average power $P_m(f,\theta) = \frac{1}{N}\sum E\{|Y(f,\theta)|^2\}$ for a signal segment composed of N frames.
2) Measure the spectral dissimilarity using equation (7) substituting $P_m(f,\theta)$ to $P(f,\theta)$, obtaining the steered spectral distance response $S(\theta)$.
3) Calculate the source position $\hat{\theta}$ using (8) and recover the signal enhanced version of the source $Y(f,\hat{\theta})$.

III. Experimental Setup

The DU beamforming-based spectral distance response algorithm is applied to the task of localizing a source and enhancing its signal by processing the data recorded by a MAV device equipped with a microphone array and conducted in stable hovering conditions. The hardware components were selected so to obtain a small-size, low-cost device relying on a four-sensor ULA. The MAV used was a Parrot Bebop 1 quadcopter, with a 250 mm frame type, 400 g weight, and overall dimensions of 280 $\times$ 320 mm. The microphone array

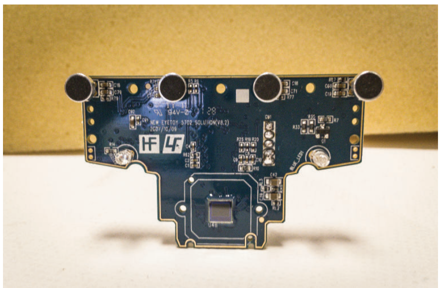

Fig. 2. The compact microphone array used in the experiments.

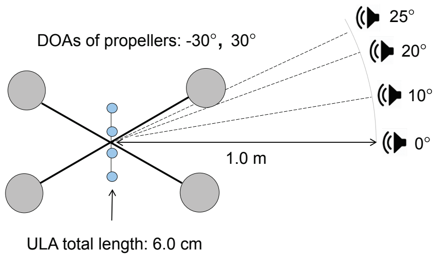

Fig. 3. Recording configuration. The microphone array is located on the top of the MAV, centered with respect to the four propellers. The target acoustic source was positioned at a distance of 1 m from the array, and at different angles.

was adapted from the acoustic sensor components embedded in the PlayStation Eye digital camera device, which provides a four-microphone uniformly spaced linear array with total size of 6 cm (the inter-microphone spacing being of 2 cm). The microphone array was located on the top of the MAV, slightly above the plane of propellers and centered with respect to the four propellers. This choice, which is not necessarily the best one in terms of acoustic properties, is however a good compromise when attempting at: 1) obtaining a compact MAV-based sensing system; 2) providing a sensed acoustic region located in front of the MAV coherent with the video camera. The resulting DOAs of the four propellers are θ_1 =-30 degrees and $\theta_2 = 30$ degrees. Note that at this stage we decided not to deal with MAV on-board digital audio recording issues. The recording was thus performed by a dedicated computer positioned next to the MAV. The array was connected to the recording computer through its USB interface. A picture of the MAV setup is provided in Figure 1. The compact microphone array is showed in Figure 2. The audio sampling frequency was 16 kHz, and the block size L was 2048 samples with an overlap of 256 samples. A Hann window was used. The spatial resolution Δ was set to 1 degree. The signal segment duration for the estimation of the averaging $P_m(f,\theta)$ and of the steered spectral distance response was 5 s.

IV. Experimental Results

Several recording sessions were conducted to build a database featuring different target acoustic sources at different positions with respect to the MAV, and corrupted by the propeller noise in realistic acoustic conditions. Numerical

ISBN 978-90-827970-1-5

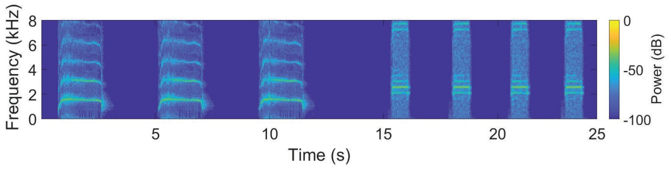

Fig. 4. The spectrogram of the entire clean audio sequence composed by three female screams and four emergency whistle sounds.

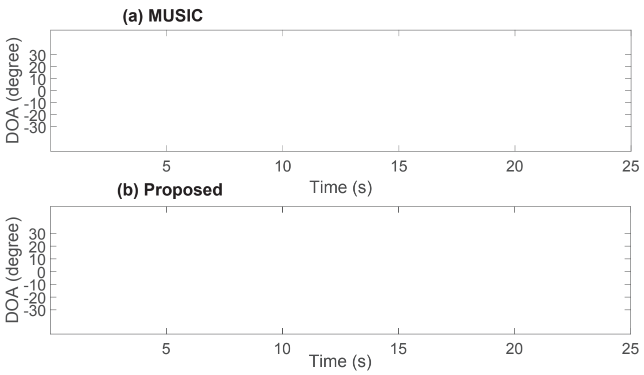

Fig. 5. MUSIC based (a) versus the proposed DU beamforming-based spectral distance technique (b) sound source localization. The MUSIC algorithms fails to distinguish the target sound source form the propeller noise, whereas the proposed approach correctly estimates the target source DOA around 20 degrees.

TABLE I
THE RMSE (DEGREE) OF THE LOCALIZATION PERFORMANCE OBTAINED WITH THE PROPOSED ALGORITHM.

	DOA Source			
SPNR	0°	10°	20°	25°
-3 dB	3.708	5.253	3.354	34.519
-13 dB	6.614	8.071	5.590	21.354
-23 dB	12.503	32.824	30.004	25.375

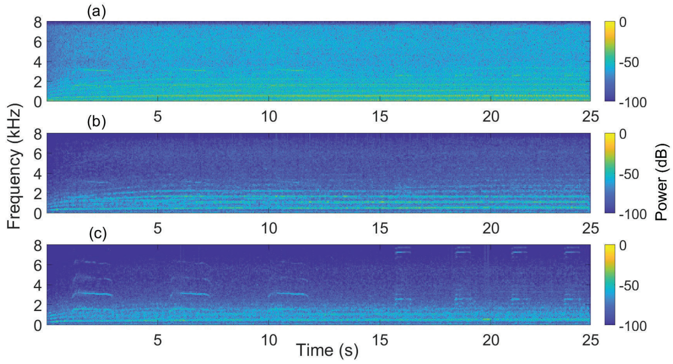

Fig. 6. The spectrograms of the channel 1 with the propeller noise and the source (a), of the DU beamformer in the propeller noise direction (b) (-30 degrees), and of the signal enhancement DU beamformer in the source direction (c) (0 degrees). The SPNR is -13 dB.

experiments were conducted on these data to assess both the localization and the enhancement performance with the algorithm described in Section II. The assessment was conducted on a set of two sound sources: a loud screaming female voice and an emergency whistle. The target acoustic source was generated by a loudspeaker positioned at a distance of 1 m from the array, and at different angles within the range $[0°, 30°]$, as illustrated in Figure 3, namely at 0, 10, 20, and 25 degrees. The sequence was repeated at each different position with three different SPNR ratios, i.e. -3 dB, -13 dB, and -23 dB. The measured propellers noise loudness at the array was 93 dB, and the mean loudness of the source signal at the array (with no propeller noise) was 90 dB, 80 dB and 70 dB on average, for the three different intensities respectively. The spectrogram of the entire clean audio sequence is shown in Figure 4. The sequence is composed by three female screams and four emergency whistle sounds. A first attempt of adopting a conventional localization method based on the MUSIC algorithm failed, due to the low resolution provided by the given microphone array and to the low SPNR. The MUSIC algorithm was implemented assuming three sources and searching for three peaks in the spatial spectrum. The performance of source DOA estimation when the source is located at an angle of 20 degrees, and SPNR is -3 dB, is reported in Figure 5 for the MUSIC approach and the proposed DU beamforming-based spectral distance response algorithm. As we can observe, the MUSIC algorithm estimates only the propeller noise directions and fails to distinguish the target sound source since no peak appears in the spatial spectrum for the source direction. Table I reports the localization performance using the root mean square error (RMSE) for the different positions of the source and the three SPNR considered. We can observe the good performance of the proposed approach when SPNR is -3 dB and -13 dB for the DOAs in the range [0,20] degrees with an average RMSE of 5.431 degree. The localization instead fails when SPNR is -23 dB. The spectrograms of the channel 1, of the DU beamformer in the propeller noise direction (-30 degrees), and of the signal enhancement DU beamformer in the source direction (0 degrees) when the SPNR is -13 dB is showed in Figure 6. A spectral distance response $S(\theta)$ for a source direction of 20 degrees with an SPNR of -3 dB is depicted in Figure 7. In order to quantify the level of the signal enhancement, we report the signal-distortion index (SDI) defined as

$$\mathrm{SDI} = \frac{1}{L}\sum_{f=0}^{L-1}\frac{|P(f,\theta_s) - P_s(f)|}{P_s(f)}, \tag{9}$$

where $P_s(f)$ is the signal power of the clean signal and θ_s is the DOA of the source. Table II shows the average SDI for the channel 1, for the signal enhancement with the conventional delay and sum (DS) beamformer [27], and for the DU beamformer. We can note the better capability of the DU beamformer to reduce noise since it has the lower SDI for both source characteristics.

V. CONCLUSIONS

In this study, we discussed the problem of acoustic source localization and signal enhancement using a compact micro-

ISBN 978-90-827970-1-5

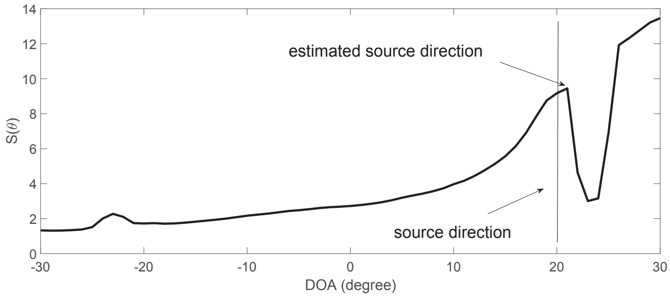

Fig. 7. The spectral distance response $S(\theta)$ for a source direction of 20 degrees with an SPNR of -3 dB.

TABLE II
THE AVERAGE SDI OF THE SIGNAL ENHANCEMENT PERFORMANCE FOR THE CONVENTIONAL DS BEAMFORMER AND THE DU BEAMFORMER COMPARED TO THE CHANNEL 1 (CH1) SDI.

Source	CH1	DS	DU
Scream	49.02	43.41	12.23
Whistle	16.56	10.96	1.94

phone array installed on a quadcopter MAV. We presented a beamforming-based spectral distance response algorithm for both localization and enhancement of the source. The new approach consists in computing a DU beamformer for each DOA target direction and measuring the spectral distance between a candidate source position and the propellers directions using the Itakura-Saito distance measure. The DU beamformer provides high-resolution without adding significant computational load. Experimental results conducted in stable hovering conditions showed that the proposed approach can localize successfully a source with an SPNR up to -13 dB. We have also reported the better signal enhancement capability of the DU beamformer if compared to the conventional one.

REFERENCES

[1] D. Salvati, C. Drioli, and G. L. Foresti, "Frequency map selection using a RBFN-based classifier in the MVDR beamformer for speaker localization in reverberant rooms," in *Proceedings of the Conference of the International Speech Communication Association*, 2015, pp. 3298–3301.

[2] X. Li, L. Girin, R. Horaud, and S. Gannot, "Estimation of the direct-path relative transfer function for supervised sound-source localization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 11, pp. 2171–2186, 2016.

[3] D. Salvati, C. Drioli, and G. L. Foresti, "A weighted MVDR beamformer based on SVM learning for sound source localization," *Pattern Recognition Letters*, vol. 84, pp. 15–21, 2016.

[4] L. Kumar and R. M. Hegde, "Near-field acoustic source localization and beamforming in spherical harmonics domain," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3351–3361, 2016.

[5] J. Traa, D. Wingate, N. D. Stein, and P. Smaragdis, "Robust source localization and enhancement with a probabilistic steered response power model," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 3, pp. 493–503, 2016.

[6] M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, and B. Lee, "A survey of sound source localization methods in wireless acoustic sensor networks," *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–24, 2017.

[7] D. Salvati, C. Drioli, and G. L. Foresti, "Exploiting a geometrically sampled grid in the steered response power algorithm for localization improvement," *Journal of the Acoustical Society of America*, vol. 141, no. 1, pp. 586–601, 2017.

[8] J. M. Valin, F. Michaud, J. Rouat, and D. Letourneau, "Robust sound source localization using a microphone array on a mobile robot," in *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003, vol. 2, pp. 1228–1233.

[9] R. Levorato and E. Pagello, "DOA acoustic source localization in mobile robot sensor networks," in *Proceeding of the IEEE International Conference on Autonomous Robot Systems and Competitions*, 2015, pp. 71–76.

[10] M. Basiri, F. Schill, P. Lima, and D. Floreano, "On-board relative bearing estimation for teams of drones using sound," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 820–827, 2016.

[11] L. Wang and A. Cavallaro, "Time-frequency processing for sound source localization from a micro aerial vehicle," in *Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 496–500.

[12] L. Wang and A. Cavallaro, "Microphone-array ego-noise reduction algorithms for auditory micro aerial vehicles," *IEEE Sensors Journal*, vol. 17, no. 8, pp. 2447–2455, 2017.

[13] P. Nguyen, H. Truong, M. Ravindranathan, A. Nguyen, R. Han, and T. Vu, "Matthan: Drone presence detection by identifying physical signatures in the drone's RF communication," in *Proceeding of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, 2017, pp. 211–224.

[14] K. Hoshiba, K. Washizaki, M. Wakabayashi, T. Ishiki, M. Kumon, Y. Bando, D. Gabriel, K. Nakadai, and H. G. Okuno, "Design of UAV-embedded microphone array system for sound source localization in outdoor environments," *Sensors*, vol. 17, no. 11, 2017.

[15] S. Oh, Y.-J. Go, J. Lee, and J.-S. Choi, "Sound source positioning using microphone array installed on a flying drone," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 3422–3422, 2016.

[16] Y. Hioka, M. Kingan, G. Schmid, and K. A. Stol, "Speech enhancement using a microphone array mounted on an unmanned aerial vehicle," in *Proceeding of the IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2016, pp. 1–5.

[17] S. Argentieri, P. Danes, and P. Soueres, "A survey on sound source localization in robotics: from binaural to array processing methods," *Computer Speech and Language*, vol. 34, no. 1, pp. 87–112, 2015.

[18] T. Ishiki and M. Kumon, "Design model of microphone arrays for multirotor helicopters," in *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 6143–6148.

[19] K. Furukawa, K. Okutani, K. Nagira, T. Otsuka, K. Itoyama, K. Nakadai, and H. G. Okuno, "Noise correlation matrix estimation for improving sound source localization by multirotor UAV," in *Proceeding of the International Conference on Intelligent Robots and Systems*, 2013, pp. 3943–3948.

[20] K. Okutani, T. Yoshida, K. Nakamura, and K. Nakadai, "Outdoor auditory scene analysis using a moving microphone array embedded in a quadrocopter," in *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 3288–3293.

[21] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[22] Sojeong Yoon, S. Park, Youmie Eom, and Sunggeun Yoo, "Advanced sound capturing method with adaptive noise reduction system for broadcasting multicopters," in *Proceeding of the IEEE International Conference on Consumer Electronics (ICCE)*, 2015, pp. 26–29.

[23] D. Salvati, C. Drioli, and G. L. Foresti, "A low-complexity robust beamforming using diagonal unloading for acoustic source localization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 609–622, 2018.

[24] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[25] D. Salvati, C. Drioli, and G. L. Foresti, "Incoherent frequency fusion for broadband steered response power algorithms in noisy environments," *IEEE Signal Processing Letters*, vol. 21, no. 5, pp. 581–585, 2014.

[26] R. J. McAulay, "Maximum likelihood spectral estimation and its application to narrow-band speech coding," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 2, pp. 243–251, 1984.

[27] B.D. Van Veen and K.M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.

ISBN 978-90-827970-1-5