Tracking of an Unmanned Aerial Vehicle with few sensors using time-frequency representation

Nathan Itare, Torea Blanchard, Jean-Hugh Thomas, Kosai Raoof

TRACKING OF AN UNMANNED AERIAL VEHICLE WITH FEW SENSORS USING TIME-FREQUENCY REPRESENTATION

N. Itare[1] **T. Blanchard**[1] **J-H. Thomas**[1,2] **K. Raoof**[1,2]

[1] Laboratoire d'Acoustique de l'Université du Mans,
LAUM UMR CNRS 6613, Université du Mans, Le Mans, 72000, France
[2] Ecole Nationale Supérieure d'Ingénieurs du Mans,
Université du Mans, Le Mans, 72000, France
nathan.itare.etu@univ-lemans.fr

ABSTRACT

Nowadays, UAV (Unmanned Aerial Vehicle) are increasingly used in many situations. It can be used for hobby, package deliveries, aerial photography, film making or surveillance. Detection and localization of UAV is necessary for example in sensitive areas. A lot of methods exist in literature to do so. This study deals with a particular use of traditional Delay and Sum Beamforming (DSB) method. Previous studies have shown that the acoustic signature of some UAV has an harmonic structure and that it was thus interesting to filter the signal recorded by the sensors taking into account this structure. This filter enables to select the fundamental frequency and some harmonics and consequently increases the signal-to-noise ratio. The DSB algorithm is then computed on the filtered signal. This method shows good results but can be difficult to implement in real time. The aim of our study is to compare this method with an approach based on the representation of the signal in the time-frequency domain. The output time signal provided by DSB is indeed represented in the time-frequency plane using a short time Fourier transform. Then only desirable frequencies, for instance the fundamental and the first three harmonics, are selected. The next step consists in calculating the energy in a time interval for the corresponding frequency bins chosen. This method enables to avoid the filtering processing. It could be more efficient in terms of computational time and could be implemented in real time.

1. INTRODUCTION

Acoustic source localization is an active research field and lots of methods exist for different types of sources. These methods are generally based on the exploitation of signals measured from a microphone array. Microphone array can be either used for acoustic source localization or parameter identification of a propagation medium. Depending on the context of the study, some of the localization or identification methods use a very few number of microphones. For example in [1] and [2], four microphones are used for estimating the source direction in a robotic context using TDOA (Time Difference of Arrival). Other methods use a large number of microphones as in [3] and [4], where two microphone arrays lead to identify the parameters of an ocean area from a double DSB method. Accuracy of the localization can be increased by the number of microphones but most of the methods only require a reasonable number of sensors. In this study, a ten microphone array is used with a different exploitation of DSB. Its geometry was designed in [5] for drone localization while two methods were used: DSB and acoustic goniometry which can characterize a propagation medium [6] or estimate sound source DOA (Direction of Arrival) [7]. The present study is only based on DSB to perform the localization. The work in [5] confirms that drone signals have an harmonic structure. Thus, this characteristic can be taken into account in the localization process. A filtering process only extracting harmonics in the bandwidth of the microphone array showed that it is possible to locate the position of the drone [5]. This process is based on bandpass filtering in the time domain by tracking the fundamental frequency and computes the DSB. It is interesting to compare this approach to one working in the time-frequency domain. The approach described here enables to directly select the frequency content. Thus, the time-frequency representation of the signal given allows us to only extract the information associated with some harmonics of the drone to calculate the energy. In part 2 the theoretical background and the proposed approach for the source localization are presented. In part 3 are shown the simulation results of the latter method. In part 4, the results obtained are summarized and some perspectives are discussed.

2. THEORETICAL BACKGROUND

2.1 Geometrical disposition (antenna and coordinate system)

The spherical coordinate system (O, r, φ, θ) for 3D source localization is shown in Figure 1. DSB applied to data acquired from a linear array enables to localize a source in a 2D plane. As this must be done in the 3D space, the array holds microphones on the 3 axis. The array has 10 microphones: 3 on each axis and 1 microphone at the center of the coordinate system. Figure 1 shows the geometry of

the microphone array used to record the broadband signal emitted by the drone. It has been designed with different spacings between microphones such as:

$$\begin{aligned} ||x_1|| = ||x_4|| = ||x_7|| = l_1, \\ ||x_2|| = ||x_5|| = ||x_8|| = l_2, \\ ||x_3|| = ||x_6|| = ||x_9|| = l_3, \end{aligned} \tag{1}$$

with $l_1 = 5\ cm$, $l_2 = 20\ cm$, and $l_3 = 110\ cm$. This arrangement gives the antenna a bandwidth of [220,5 3430] Hz.

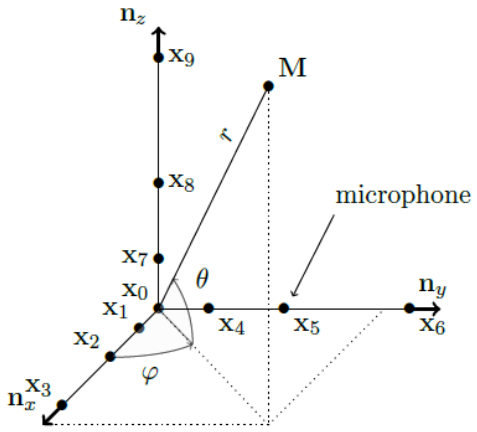


Figure 1: Microphone array used to localize the drone. [5]

2.2 DSB

Delay and Sum Beamforming consists in applying delays to the signals received by the microphone array. Indeed, each signal sent by the source arrives with a different time according to the microphone placed at the origin. Thus, by compensating these delays and summing the signals, an image of the source energy is reconstructed. There is one set of delays per angle $\Omega = (\varphi, \theta)$, and these delays are determined using a fictive source at $r = 1\ m$. Delay τ_i of the microphone i according to the origin microphone is simply the difference between the distance from the fictive source to the origin ($||OM||$) and the distance from the fictive source to the i^{th} microphone considered ($||M_iM||$, $i = 1, \ldots, N,\ N = 10$) divided by the speed of sound c. Equation (2) corresponds to a near-field source and could also be applied a fortiori to a far-field one.

$$\tau_i = \frac{|\ ||OM|| - ||M_iM||\ |}{c}. \tag{2}$$

This operation is realized for each angle Ω chosen in a spatial sampling grid.

2.3 Simulation of a drone signal

In order to test the proposed approach, the model of the drone's signal developed in [5] is used. This signal has an harmonic structure with weak and strong harmonics. Equation (3) shows the formulation of the simulated signal.

$$\begin{aligned} p_i(t) = \sum_{n=1}^{N_h} \underbrace{\beta \frac{cos(2\pi[2n-1]f_0(t - ||M_iM||(t)/c))}{4\pi||M_iM||(t)}}_{\text{weak harmonics}} \\ + \underbrace{\alpha_{f_0}(n) \frac{cos(2\pi[2n]f_0(t - ||M_iM||(t)/c))}{4\pi||M_iM||(t)}}_{\text{strong harmonics}}, \end{aligned} \tag{3}$$

with N_h the number of harmonics desired, f_0 the fundamental frequency of the drone, β is a gain for weak harmonics, and $\alpha_{f0}(n)$ is an attenuation factor for strong harmonics. Indeed, measured drone signals have shown that amplitudes of strong harmonics decrease with a logarithm function. The parameters for this study are $\beta = 10^{-1.5}$ and $\alpha_{f0}(n) = 10^{\frac{1}{20}(-11.6log_{10}(2nf_0)+65.4)}$. Figure 2 shows the Power Spectral Density (PSD) of the simulated drone signal used for the characterization of the proposed approach.

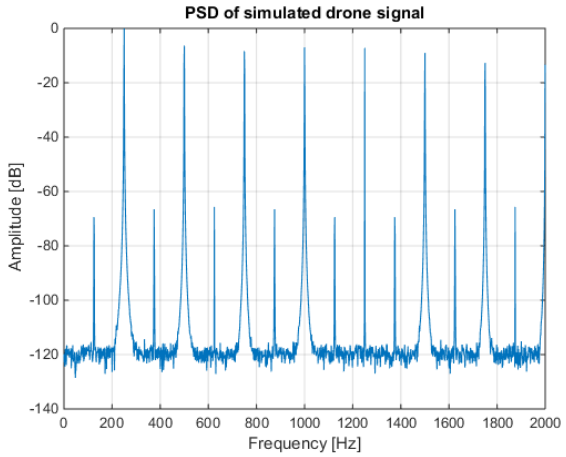


Figure 2: PSD of the simulated signal drone with $f_0 = 125$Hz.

2.4 Proposed approach

The proposed approach operates the traditional DSB but instead of showing a map of energy taking into account the whole energy of the signal, its time-frequency representation is used to calculate a partial energy. The representation of the output signal of the DSB is given by the Short Time Fourier Transform (STFT). Once the STFT is done, the time-frequency bins characterizing the drone can be selected and the corresponding energy is shown. Figure 3 shows the process of the proposed approach. First step is to choose the sampling space, which sets the resolution in azimuth and elevation. Second step is to realize the DSB for each angle Ω. In the third step, the STFT is applied to the beamformer's output. The fourth step consists in determining the fundamental frequency in order to select the desired harmonics in step five. The energy is then represented in step 6 in a map as a function of the steering angle Ω.

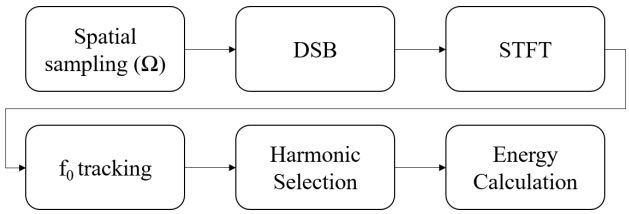


Figure 3: Synopsis of the proposed approach.

3. SIMULATION RESULTS

The first step of the process is to choose the resolution of the spatial sampling. This resolution will determine the accuracy of the localization and has also an influence on the computational time. Two resolutions of 4° in azimuth and 2° in elevation are chosen for the study. The first results are computed for a virtual source placed at two different locations ($r_1 = 10\ m,\ \varphi_1 = 0°,\ \theta_1 = 45°$), ($r_2 = 10\ m,\ \varphi_2 = 100°,\ \theta_2 = 70°$) with a Signal to Noise Ratio (SNR) of 20 dB and a fundamental frequency of 125 Hz. Figure 4 shows the energy of the output beamformer's for both locations. The localization shows a good result with an angular error of (4°, 1°) (corresponding respectively to azimuth and elevation errors) for the first position and (12°, 2°) for the second position. Figures 5, 6, and 7 show the energy obtained by taking only into account the frequencies chosen in the time-frequency plane for two positions. The frequencies chosen are harmonics of the fundamental frequency of the virtual source signal and this approach enables to see the contribution of each harmonic. Even harmonics are selected as they are predominant for a lower SNR and a greater distance. Energy maps show that the localization of the source using only one harmonic is possible. As the frequency increases, it appears that lobe width decreases and the number of secondary lobes increases. Thus with higher frequency harmonics, the localization of the source can be ambiguous.

Figures 8 and 9 show the influence of the noise on the accuracy of the localization for the classical DSB and the one based on the time-frequency representation. For each point of the figures, 50 virtual acoustic signals have been generated for one microphone and one SNR and the error shown is the average. This influence is described by taking only one harmonic in Figure 8 and the five first harmonics in Figure 9. For the first case, the azimuthal error is less than 5° until -10 dB for both representations whereas the elevation error is less than 2° until -10 dB where it begins to increase for both representations. For the second case, the azimuthal error is still less than 5° until -18 dB for the TF representation and until -16 dB for the classical approach. The elevation error is less than 2° until -18 dB for TF representation and until -16 dB for classical approach which is better than for the first case. Thus taking more harmonics into account could make the method more robust to a noisy environment.

Another simulation is done to evaluate the performance of this process with the presence of two sources having two different positions and spectral contents. The same positions as for the previous simulation [(0°, 45°) and (100°, 70°)] are taken. The fundamental frequencies of the two drone signals are $f_{01} = 150$ Hz and $f_{02} = 175$ Hz. The SNR is 20 dB for the same angular resolution. Figure 10 shows beamforming mappings using all signal energy (standard calculation) and the energy carried by two frequencies or a single frequency extracted from the time-frequency representation of the beamformer's output signal. The first map shows that both sources are located with a pretty good accuracy, the source is obtained with the search of only one maximum. The second map shows the energy taking into account the first harmonic of each source ($2f_{01}$ and $2f_{02}$). The first source is still located and there is more energy in the position of the second source but not enough to be located precisely. If each harmonic is taken separately which is shown in the last two maps, it appears that for each harmonic corresponding to the source, more precise localization of the source is obtained. Thus, it is possible to differentiate two sources with their spectral content and locate them using TF representation.

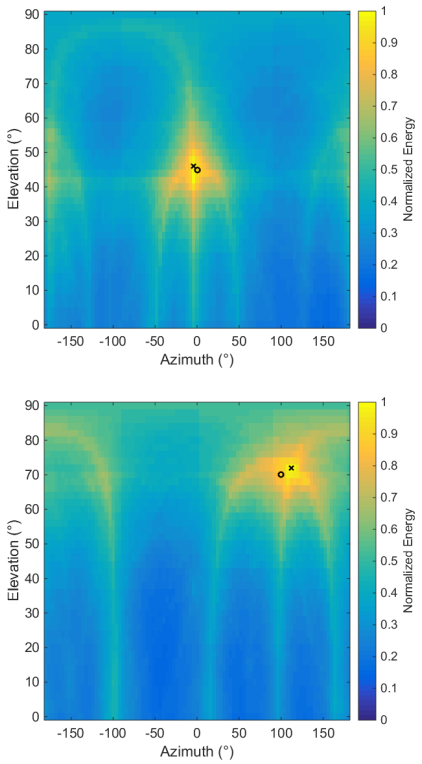


Figure 4: Energy of the classical DSB in the azimuth/elevation plane (the circle represents the actual position of the virtual source, the cross represents the position of the maximum of energy) for the (0°, 45°) source (top) and the (100°, 70°) source (bottom).

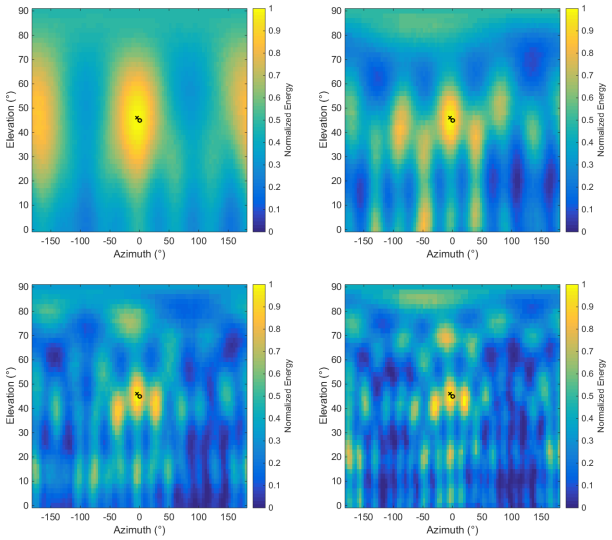


Figure 5: Energy maps of the energy of the corresponding chosen even frequencies (from the second harmonic to the eighth harmonic) for the (0°, 45°) source.

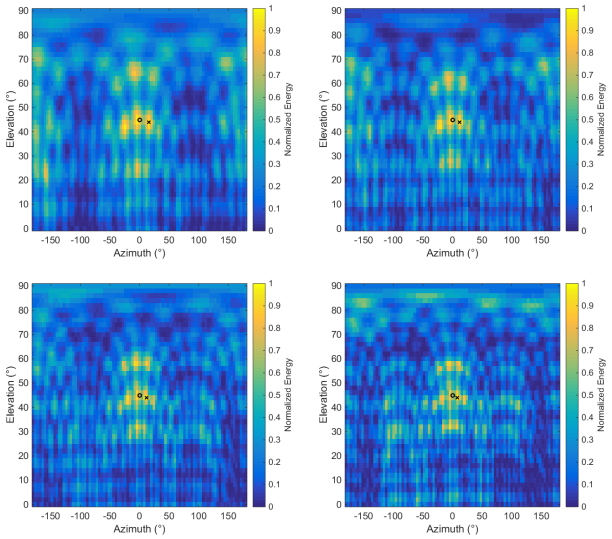


Figure 6: Energy maps of the corresponding chosen even frequencies (from the tenth harmonic to the sixteenth harmonic) for the (0°, 45°) source.

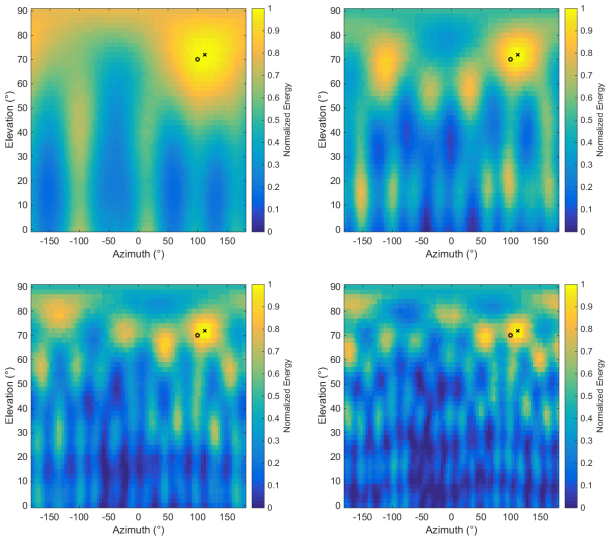


Figure 7: Energy maps of the corresponding chosen even frequencies (from the second harmonic to the eighth harmonic) for the (100°, 70°) source.

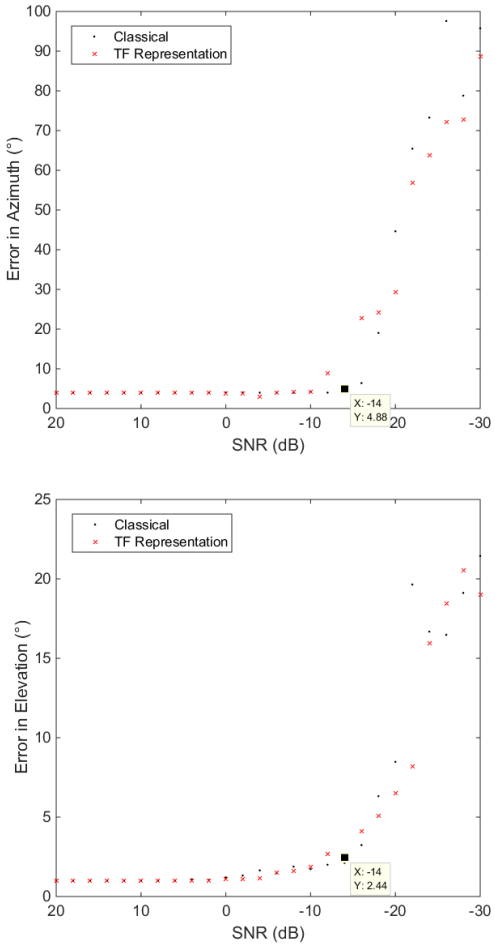


Figure 8: Evolution of the angle errors as a function of the SNR for the (0°, 45°) source with one harmonic ($2f_0$).

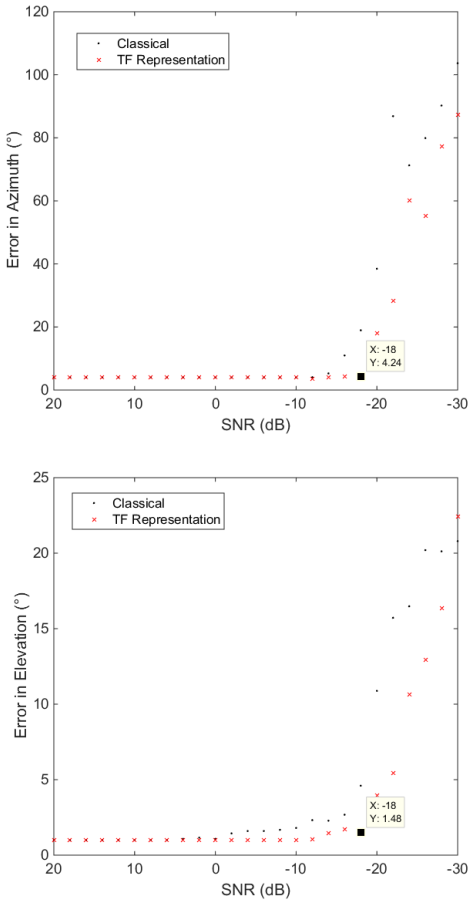


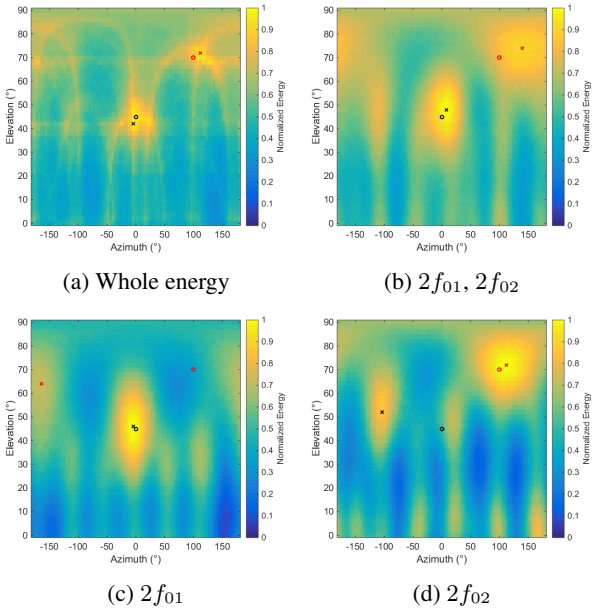
Figure 9: Evolution of the angle errors as a function of the SNR for the (0°, 45°) source with five harmonics ($2f_0$, $4f_0$, $6f_0$, $8f_0$, $10f_0$).

4. CONCLUSION AND PERSPECTIVES

This paper presents a study on the localization of a drone in the 3D space. Two approaches have been compared. The first approach is the classical DSB and the second approach uses a TF representation to select desirable frequency content. A designed antenna for drone localization in a particular band pass, and a simulated drone signal have been used. Simulation results show that using the TF representation enables to have an accurate localization of the source using only a few spectral components. This approach has also shown its ability to differentiate two sources with two different spectral contents and positions in space. The results have been compared with the classical DSB algorithm. Also, a comparison of the noise robustness of the two approaches has been shown. This study is a preliminary work using TF representation of beamformer's output and it was tested in simulation. Thus it could be interesting to evaluate the performance of this process with real measurements.

5. ACKNOWLEDGMENTS

This PhD is financed by the General Direction of Army (DGA).



(a) Whole energy (b) $2f_{01}$, $2f_{02}$

(c) $2f_{01}$ (d) $2f_{02}$

Figure 10: Energy maps in the case of the presence of two sources (the circles represent the actual positions of the virtual sources, the crosses represent the positions of the two energy maxima).

6. REFERENCES

[1] F. Ruffier, S. Benacchio, F. Expert, and E. Ogam, "A tiny directional sound sensor inspired by crickets designed for micro-air vehicles," in *2011 IEEE SENSORS Proceedings*, (Limerick, Ireland), pp. 970–973, IEEE, Oct. 2011.

[2] S. Lee, Y. Park, and J.-S. Choi, "Estimation of multiple sound source directions using artificial robot ears," *Applied Acoustics*, vol. 77, pp. 49–58, Mar. 2014.

[3] F. Martellotta, "On the use of microphone arrays to visualize spatial sound field information," *Applied Acoustics*, vol. 74, pp. 987–1000, Aug. 2013.

[4] B. Nicolas, I. Iturbe, and P. Roux, "Double Beamforming for wave separation and identification : robustness against noise and application on FAF03 experiment," p. 21, 2008.

[5] T. Blanchard, *Caractérisation de drones en vue de leur localisation et de leur suivi à partir d'une antenne de microphones*, PhD thesis, Le Mans Université, 2019.

[6] R. Boulandet, X. Falourd, M. Rossi, and H. Lissek, "Localisation des premières réflexions dans une salle par chrono-goniométrie acoustique," *Congrès Français d'Acoustique 2010, Lyon*, p. 6.

[7] E. V. Lancker, "Goniométrie acoustique et estimation des différences de temps de propagation," *Congrès Français d'Acoustique 2000, Lausanne*, p. 4.