Outdoor field trials for the measurement of the acoustic signals of mini UAVs

Pierre Naz, Sébastien Hengy, Aro Ramamonjy, Oussama Rassy, Eric Bavu

Outdoor field trials for the measurement of the acoustic signals of mini UAVs

NAZ Pierre[1] HENGY Sebastien[1] RAMAMONJY Aro[2]
RASSY Oussama[1] BAVU Eric[3]

[1] ISL, French-German Research Institute of Saint-Louis (ISL), Saint-Louis, France
[2] Acoustics Laboratory, Le Mans University, Le Mans, France
[3] Laboratoire de Mécanique des Structures et des Systèmes Couplés, CNAM (LMSSC), Paris, France
Correspondence: pierre.naz@isl.eu

ABSTRACT

Acoustic detection and tracking of UAVs is considered by means of Unattended Ground Sensors equipped with microphonic sensors. Experimental campaigns were conducted with flying drones (DJI, Parrot…) in an anechoic chamber and in countryside. The acoustic database includes various scenario such as hovering flight, translation flight, etc. At the same time, "disturbing noises" have been recorded: ambient noises including birds, insects, people speaking, detonations and fire shot noises have been recorded to feed our database.

A part of the recorded database has been used to train a classifier (learning phase). Then another part of the dataset was used to estimate the F-score to evaluate both the precision and recall of the classifier. Adding artificial noise to the data, and selecting acoustic features with evolutionary programming enabled the detection of an unknown drone in an unknown soundscape within 200 meters with the JRip classifier (Fscore of 0.88 for distances between 0 and 100 m, and 0.56 between 100 and 200 m).

Main results obtained during the signature analysis and the classifier assessment will be presented and the perspectives in terms of performance improvement with the use of MEMS multi-microphones array.

1. INTRODUCTION

The large diffusion of small unmanned aerial vehicles (sUAV) for both civil and military applications is a major concern for security problems as well as for protection of private life of the citizen. Network of detectors of various types (Acoustic, Optical, Radio waves, Radar…) are envisaged for surveillance system against UAVs. Within the envisaged technologies of detection, acoustics is a passive technique interesting for the detection, localization and reconnaissance of small UAVs.

For this purpose, we have used some background knowledge gained in the years 1990-2000 on helicopter detection and tracking. The main results gained by the ISL team on this topic will be discussed in the next paragraph.

Then we will present the experimental measurements done in order to have a physical insight of sUAV noises, the signal processing technique for the localization and tracking of the flight, the reconnaissance and classification of such a "threat". This paper is based on the work done by the CNAM/LMSSC [1], by the ISL [2] and with a prominent contribution of a common PhD study [3].

2. BACKGROUNDS

2.1 Helicopter noises

Helicopter noise has been studied intensively in the 90's for environmental purposes (diminution of the noise pollution generated by heliports) or for military purposes (detection of attacking helicopters flying at low altitude in small valleys in order to be undetectable by classical air-control radars [4 and 5]). This noise is mainly of aerodynamic origin and is situated in a frequency band at the lower limit of the audible frequencies and even slightly lower for some types of helicopters. The acoustic energy results largely from the main rotor (with a fundamental frequency between ten and thirty hertz), from the tail rotor (its fundamental frequency is close to hundred Hertz for a classical rotor tail and to 1 kHz for a "fenestron"), from the gear box and the turbine (some kilohertz), (Fig. 1). The most intense noise is created at very low frequencies by the main rotor, which is favorable for long range detection, because these low frequencies propagate very well and are only very little attenuated during their travel through the atmosphere.

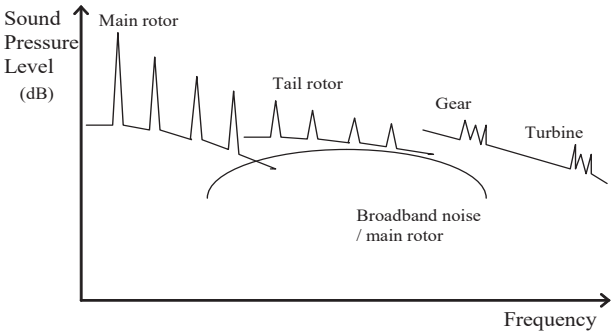


Figure 1. Typical spectrum of helicopter noise

The noise generated by the main rotor is composed by 3 main sources: a monopole sound caused by the volume of air displaced by the movement of the blades; a dipole sound caused by the fluctuations in pressure impinging on the surface of the blades and a quadrupole sound caused

by shearing phenomena within the fluid streams disturbed by the passage of the blades. This latter noise dominates when the part of the blade in question is moving at a transonic or supersonic speed; this is the case with the advancing tip of a blade when the helicopter is in high-speed forward flight, or when turbulence created by a blade meets another blade (e.g. during rapid descent).

Since the acoustic level is very high, the propagation in the atmosphere has a highly non-linear behavior, with the creation of important harmonics and a tendency of acoustic waves to evolve towards a form of shock (nested "N" waves, Fig. 2). The frequency analysis of these signals shows that the fundamental frequency and the different harmonics of the main rotor and tail rotor form two "combs" of nested lines (Fig. 1 and 4).

These two basic frequencies are related to certain characteristics of the helicopter (number of blades, rotation speed of rotors, mass, speed, etc.). At first approximation, the fundamental rotor frequency depends only on the rotation speed (Ω) and the number of blades (n). The following simple formulae as in Eqn. (1) and (2) make it possible to calculate it in the case of a hovering helicopter (f_0) and in the case of a steady-speed (V) advance flight (f_V), with C_0: speed of sound. In some publication f_0 is also called Blade Passing Frequency (BPF).

$$f_0 = \mathrm{n}\,\Omega \tag{1}$$

$$f_V = \frac{f_0}{1 - V/C_0} \tag{2}$$

Some other parameters complicate this characterization: the acoustic emission of a helicopter is directive in site and field, it also depends on the flight configuration (speed, load, maneuver), and will be perceived differently depending on the distance and altitude of the helicopter.

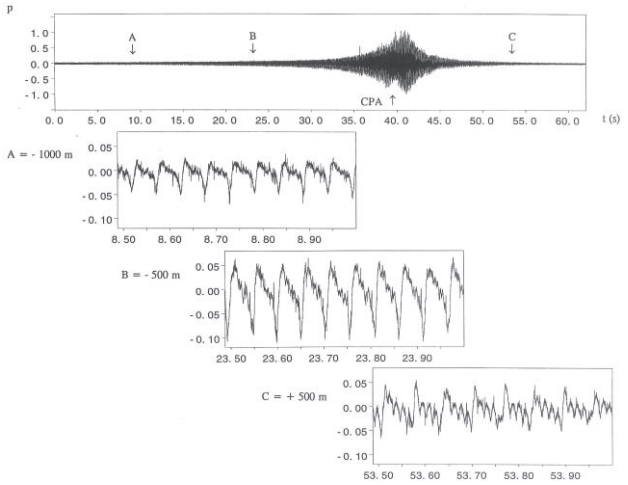


Figure 2. Time signals of a translation flight (global and zoom for 3 distances)

Fig. 3 and Fig. 4 show that the frequency line family's characteristics of the main rotor and the rear rotor are clearly visible. On Fig. 3 the frequency inflection due to the Doppler Effect when the helicopter passes through the microphone is visible. In more detail, we observe that when the helicopter is very close, a broadband noise superimposes itself on the noise of lines. At greater distances, only a few pure frequencies emerge from the background noise. After the helicopter passes, the spectral richness and the acoustic level drop very quickly.

In comparison with the problem of the study of noise nuisances, only low frequency contributions are to be taken into account for remote sensing applications.

The next graph (Fig. 5) illustrates the fact that an helicopter is well heard, even if it is masked behind a hill, which is an important feature in military context [6]. The value of the attenuation due to the hill is relatively low for the main rotor frequency and its first harmonics (as expected the attenuation is greater for higher frequencies).

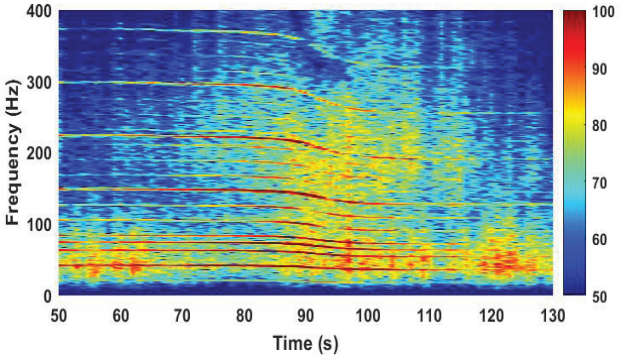


Figure 3. Spectrogram of a helicopter noise (translation flight)

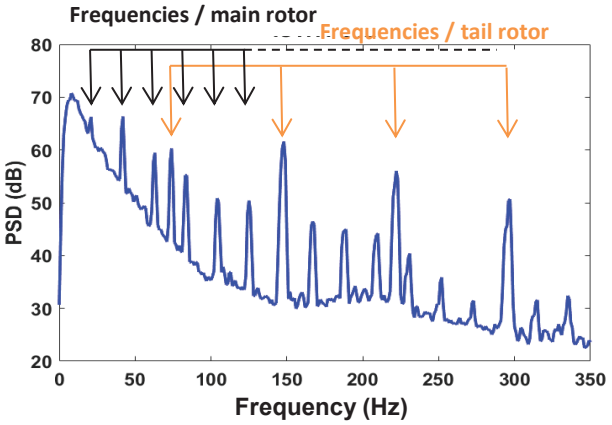


Figure 4. Example of spectrum of a helicopter noise

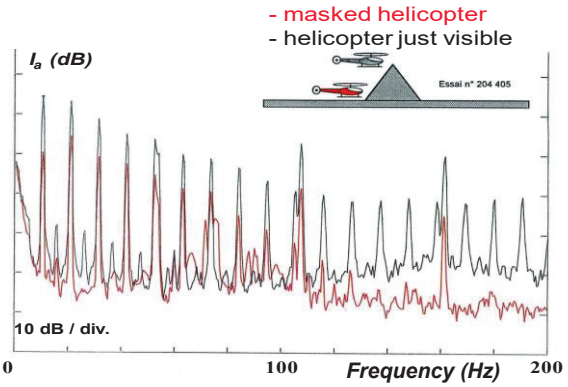


Figure 5. Spectrum for a masked and a visible helicopter [6]

2.2 Acoustic antenna for helicopter detection

The tasks to be performed by the surveillance systems studied can be defined into three main items: detection,

recognition, localization of the "threat". These functions correspond to the three basic questions: is there an abnormal (noise) event? What is it? Where is it?

The main advantages of using acoustics in such a system are: Passive and discrete means; Low cost sensors; Non line of sight detection capability.

The main disadvantages are: Variable performances with environmental conditions (including false alarm management); Time delay due to the acoustic wave propagation.

A very simple algorithm may be used to calculate the azimuth of the acoustic events (more sophisticated and more robust algorithm was presented in previous publications to be used on an operational sensor network [5]). The main equation uses the differences in the time of arrival of the waves between the pairs of microphones (eqn. 3 and Fig. 6). In the presence of a good signal to noise ratio, the previous simple formula gives good results. It is more difficult for weak signals in adverse conditions. Example of results of the angular localization of a helicopter versus time is given in Fig. 7. These results correspond to instantaneous detections given by two close arrays of 1 m (array A) and 0.5 m (array B) width.

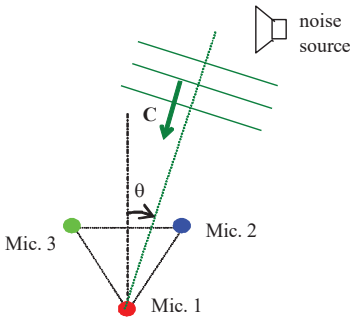


Figure 6. Spectrum for a masked and a visible helicopter

$$tg(\theta) = \sqrt{3} \cdot \frac{\Delta t_{12} - \Delta t_{13}}{\Delta t_{12} + \Delta t_{13}} \qquad (3)$$

with :

Δt_{ij} : time delay between the signals of microphones i et j

θ : azimuth of the source

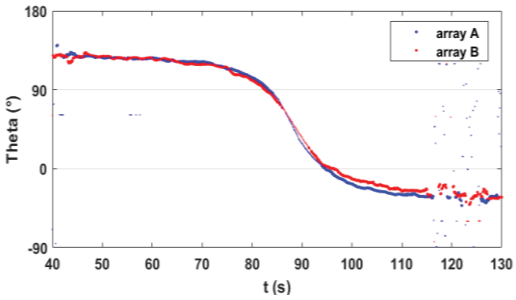


Figure 7. Angular localization of a helicopter versus time (bold points: validated localization, small points: unvalidated localization)

3. MEASUREMENT OF UAV NOISES IN AN ANECHOIC CHAMBER

3.1 Bibliography

Recently, some research on the noise generated by sUAV has been published. In [7], measurements have been carried-out in an anechoic chamber with a DJI Phantom 2 and 4 types of commercial blades. The authors show clearly the contribution of the Blade Passing Frequency (BPF, 128 Hz) and harmonics, then the contribution of the electric motors at mid frequency (600 – 6000 Hz). The individual contribution of 1, 2 and 4 rotors at nominal rotational speed (around 6000 RPM) was also characterized.

Another study investigates more in details the aerodynamic origin of these noises [8]. The tonal noise is associated with the dynamic loading of the rotor blade. The broadband noise level is largely determined by the complex turbulent flow structures and shear layers. Effects of rotor-to-rotor interactions on the aerodynamics and aeroacoustics performances of sUAV was also noticed.

The NASA Langley Research Center [9] has published a study with 4 main areas: outdoor measurements of noise generated by flyovers of sUAV, measurements in controlled test facilities to understand the generated noise, computational predictions, and psychoacoustic tests to assess human annoyance from these noises. Published results confirm and complete the ones obtained by other authors. In particular they noticed that the under-airframe configuration results in greater tonal noise compared to placing the rotors above the airframe. They noticed also that, at an equal sound exposure level, test subjects rated the sUAV as more annoying than road vehicles.

3.2 ISL measurements

Unlike most of the bibliographical results reported in the previous paragraph where the drones were fixed on a stand, we have first measured the acoustical signals in an anechoic chamber in "free flight conditions" (Fig. 8). We have recorded the noise generated by a DJI Phantom 3 (quadcopter, 2 blades per propeller), a Parrot Bebop (quadcopter, 3 blades per propeller) and a Parrot Mars (quadcopter, 2 blades per propeller).

Figure 8. UAV flying in the ISL anechoic chamber (rotating device for the microphonic array on the right)

Before take-off, the rotational speed of the motors and blades are very stable, the spectrogram of the acoustic data exhibit a great number of horizontal spectral lines with quite constant frequencies from 1 to 1.8 kHz (Fig. 9). After

take-off, we can observe the emergence of lower frequency lines. As the UAV needs to do small corrections to maintain a stationary flight, the rotation speeds of the quadcopter's rotors are slightly different which generates frequency variations around the observed fundamental frequency and its harmonics. (Fig. 10 and 11). The main contribution of the aerodynamic noises and of the motor noises are clearly shown in Fig. 12.

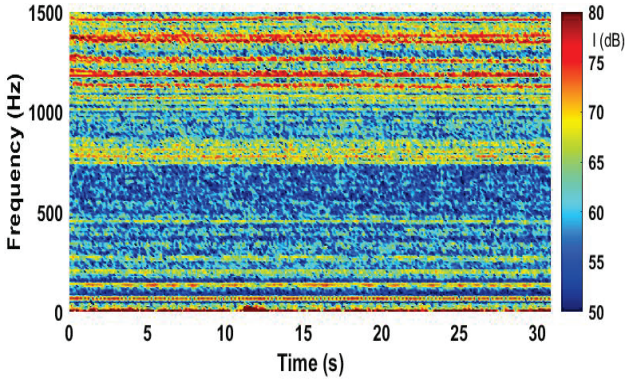


Figure 9. Spectrogram of a DJI Phantom 3 (before take-off)

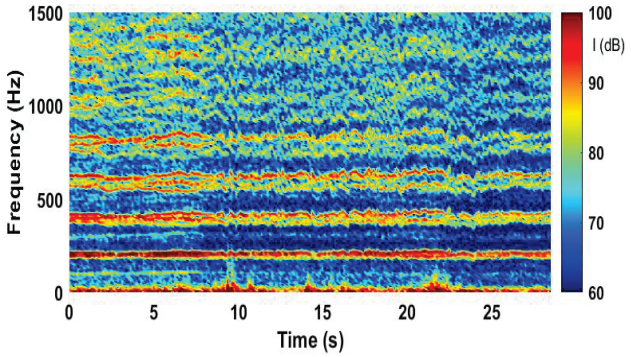


Figure 10. Spectrogram of a DJI Phantom 3 (hovering)

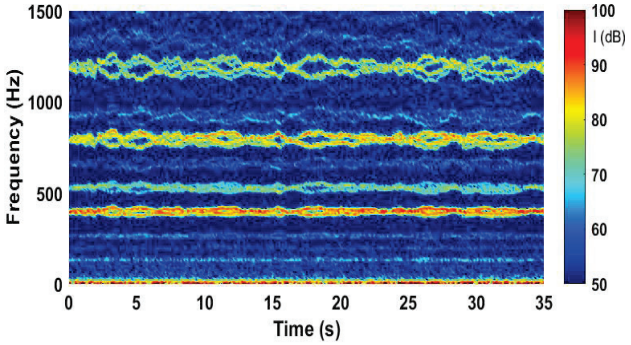


Figure 11. Spectrogram of a Parrot Bebop-2 (hovering)

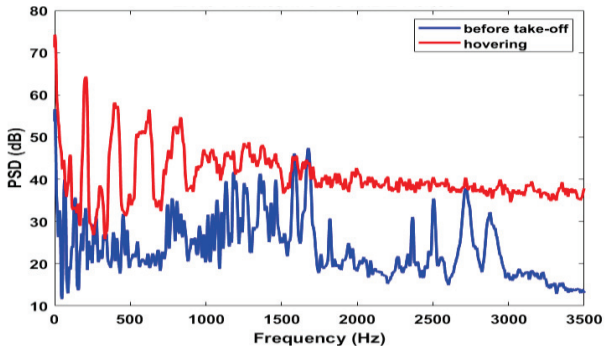


Figure 12. Spectrum of a DJI Phantom 3

Compared with the helicopter noises, the spectrograms exhibit similar frequency lines, but at higher frequencies due to the higher rotation speed of the rotors (main frequency in the range 200-400 Hz for UAVs compared to 10-20 Hz for helicopters). Due to the absence of tail rotor, no corresponding frequency lines may be observed. The other main difference is the frequency variations of these lines of the multi-rotors UAVs, visible in the spectrograms but also in the spectrum were consequently some frequency contributions are less visible.

4. FREE FIELD TEST

ISL organized free field experiments with various acoustic arrays and UAVs in order to acquire acoustic signals in realistic conditions (real flights and various environmental conditions). This section gives a short overview of the measurement systems deployed and of the results obtained.

First, a Real Time Kinematics (RTK) system based on GNSS receivers was implemented on a Base Station and an extra GNSS receiver on the UAV. This RTK System is used to collect ground truth points in a local "East North Up" (ENU) coordinate System. It was used in order to have precise localization of the UAV for post-processing analysis of the acoustic data. To compensate the extra weight of the GNSS/Receiver device we have removed the camera and corresponding gimbal (Fig. 13).

Acoustic arrays using conventional metrological microphones as well as prototypes of new arrays using MEMs microphones have been deployed at different points of the site (for more details see paragraph 5).

Specific microphone(s) of these arrays have been used in order to build the acoustic database that will be used for the classification study (for more details see paragraph 6).

Recordings have been done with various types of flights: stationary flight (hovering), translation flight S-N and E-W at a constant height, circular or square flight pattern around the sensors … Most of the tests have been done with one drone flying at a time, some have been done with 2 drones having a common or a different trajectory. In this context, the RTK system is of a great help for the post-processing of the data.

The drones tested outdoor are quadrotors, same as previously used in the anechoic chamber, to which a new one was added: a flying wing (Parrot Disco).

Figure 13. GNSS/RTK receiver mounted on a DJI Phantom 3

The spectrograms of the signals measured outdoor exhibit similar features as seen previously, but the spectral lines are less pronounced due to the greater distance

between the sUAV and the microphones and to the frequency variability due to variations of the rotors speed (Fig. 14 and Fig. 15). The variability is still greater in case of the flying wing (Fig. 16, Fig. 17).

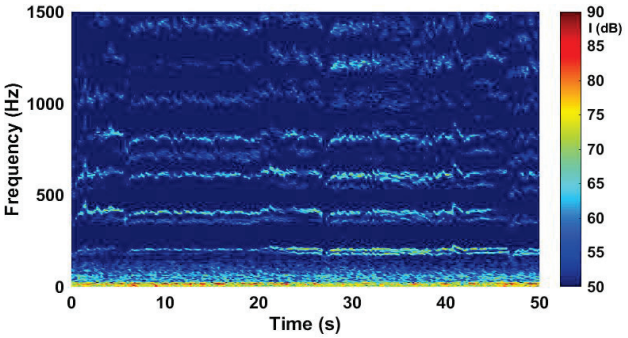


Figure 14. Spectrogram of a DJI Phantom 3 (outdoor)

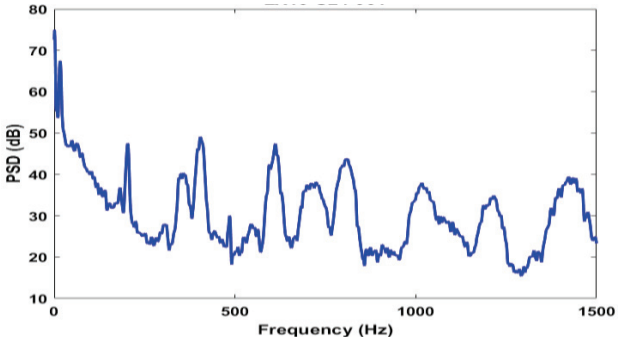


Figure 15. Spectrum of a DJI Phantom 3 (outdoor)

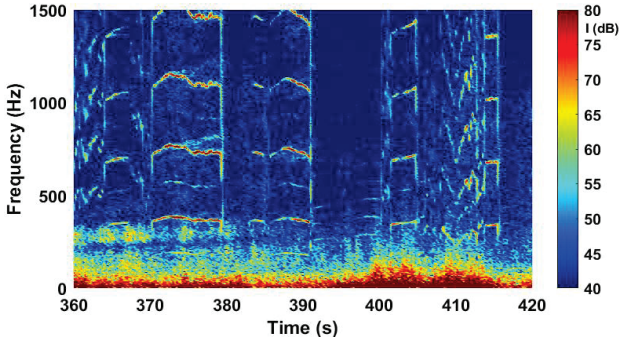


Figure 16. Spectrogram of a Parrot Disco (outdoor)

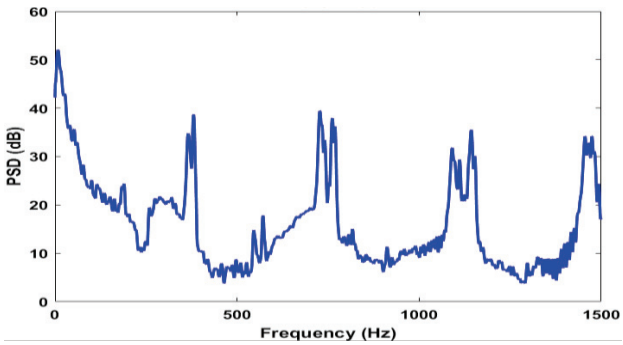
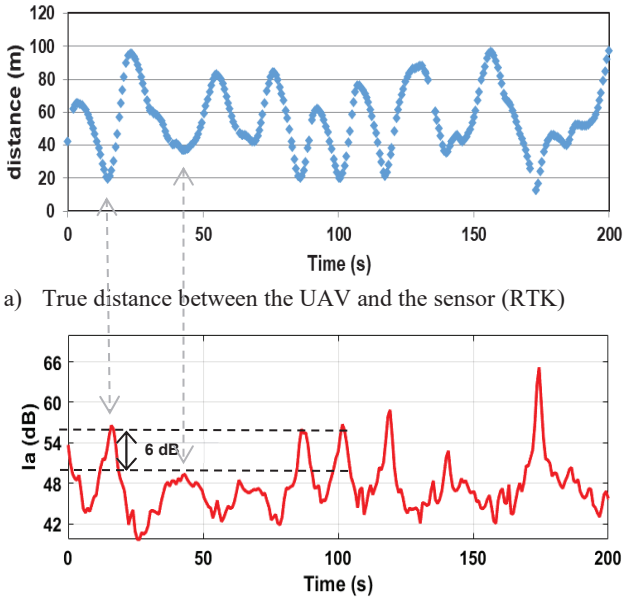


Figure 17. Spectrum of a Parrot Disco (outdoor)

Fig. 18 illustrates the evolution of the Sound Pressure Level (calculated on the frequency band [80-3500 Hz]) during E-W and N-S paths. The comparison with the RTK data shows that the local maximum of the pressure corresponds to the local closest points of approach (CPA) of each segment of the trajectory. A "classical" 6 dB difference is observed when the distance of the closest points of approach (d_{CPA}) is varying from 20 to 40 m.



a) True distance between the UAV and the sensor (RTK)

b) SPL_rel 20 µPa

Figure 18. Results with DJI Phantom 3 (N-S and E-W flights)

5. LOCALIZATION ANTENNAS

5.1 Arrays with metrological microphones

During the field experiment, three seven-microphones acoustic array have been deployed with pre-polarized metrological microphones (2 with B&K 4189 microphones and 1 with G.R.A.S. 46AE). These arrays are constituted with two interlocked tetrahedrons respectively of 10 and 2.5 centimeters in width (Fig. 19). The measured acoustic signatures are low-pass filtered at a cut-off frequency of 6.8 kHz adapted to the array's shape before the localization process is started.

Figure 19. Microphonic array (7 microphones)

5.2 Localization and tracking

The localization process is based on the MUltiple SIgnal Classification (MUSIC) algorithm [10]. This high-resolution localization algorithm allows localizing multiple acoustic sources synchronously. It is based on the

decomposition of the measured data s(Θ,f) into its signal and noise subspaces, where Θ represents the position of the source (azimuth and elevation information), and f the frequency.

In order to minimize the number of outliers in the estimated values, two predictive filtering techniques have been applied: Kalman filtering and particle filtering [11]. In the context of our data, it has been shown that the particle filter provides more stable results and a better rejection of the artefacts (Fig. 20 and 21). This filtering step is applied first for each array on the estimated azimuth and elevation values (live streamed data). In a second step a fusion of the predicted data sent by the 3 arrays is processed, providing estimated latitude and longitude coordinates. Finally, these latitude-longitude coordinates are processed again using a particle filter to smooth the estimated trajectory estimates.

The Fig. 22 shows the data fusion resulting from the processing of the angular data of the three acoustic antennas. The obtained filtered results are compared to the ground truth data recorded with the RTK GPS System.

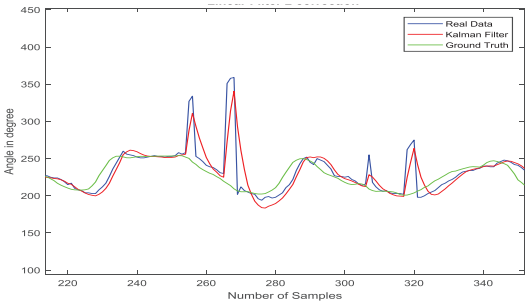


Figure 20. Kalman filter data compared to raw, and ground truth data for a measured azimuth angle over time

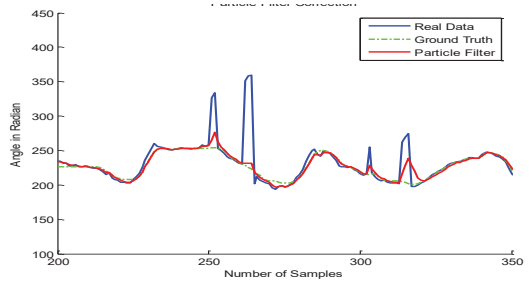


Figure 21. Particle filter data compared to raw data for a measured azimuth angle over time with the array "BK1"

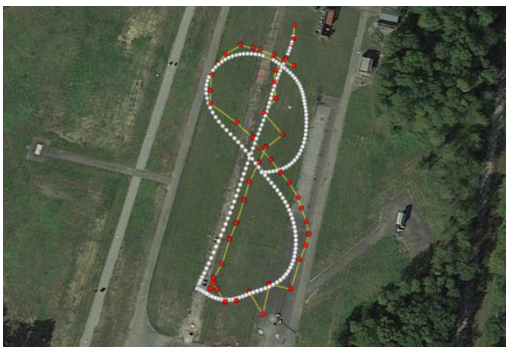


Figure 22. Ground truth (white dots) and estimated drone positions using data of three acoustic arrays (red dots)

This filtering process leads to sensible increase in performance as the estimation error standard deviation drops from 50% on the East-West axis, while it drops of 33% on the North-South axis.

5.3 MEMS array

Prototypes of compact microphone arrays (CMA) using 16 and 32 digital MEMS microphones have been designed for the localization and tracking of the UAVs acoustical sources. These compact arrays consist in two orthogonal lines of 8 or 16 MEMS microphones which are placed in the horizontal plane. Multiple microphone pairs are used to estimate the pressure (P) and the particle velocity components (V_X, V_Y) on two orthogonal axes at the center of the array, i.e. at the crossing of the two lines of microphones. Different spacing between the microphones are used either separately to measure the acoustic field at different frequencies (in this case decreasing spacing is used for increasing frequencies) or together to obtain a more accurate estimate (higher order estimations). The use of logarithmic spacing between the microphones allows to perform localization in log scaled frequency bands with a limited number of microphones (Fig. 23).

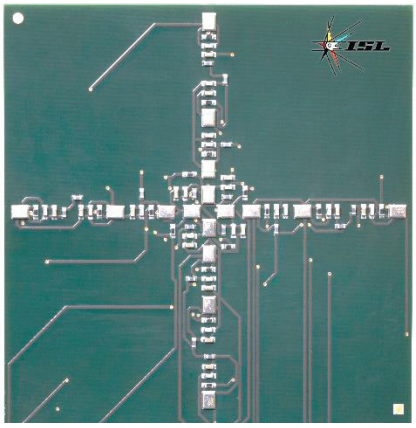


Figure 23. View of an array (16 MEMS microphones)

An original Direction Of Arrival (DOA) estimation algorithm is proposed. This real-time time-domain method uses the RANSAC algorithm (RANdom SAmple Consensus) on broadband [100 Hz - 10 kHz] pressure and particle velocity data which are estimated using finite differences and sums of signals of microphone pairs with frequency-dependent inter-microphone spacing (Eqn. 4 to 6).

$$\theta = atan2\{-(\rho_0 c_0 V_Y/P),\ -(\rho_0 c_0 V_X/P)\} \quad (4)$$

$$V_X \cong \rho_0 \int_0^t \frac{p_{1x}-p_{2x}}{d} d\tau \quad (5)$$

$$V_Y \cong \rho_0 \int_0^t \frac{p_{1y}-p_{2y}}{d} d\tau \quad (6)$$

The principle of the DOA estimation and tracking of a trajectory has been first tested in the CNAM/LMSCC laboratory, using their spatialization sphere. This sphere is able to generate moving sound sources in the Ambisonic domain. An average DOA estimation error of 4 degrees

has been observed (Fig. 24). The use of such a sphere may be very useful in the future to extend our database with synthetic acoustic data generated in flight configurations not yet experienced. Outdoor experiments with the MEMS CMA will be organized in the next months.

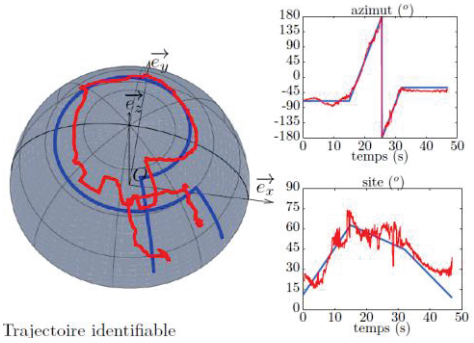


Figure 24. Filtered latitude and longitude coordinates

6. UAVS CLASSIFICATION

A global approach has been suggested [1]. This original approach consists in an initial detection of a potential target, followed by a DOA estimation and tracking process, along with a refined detection, facilitated by adaptive spatial filtering (Fig. 25).

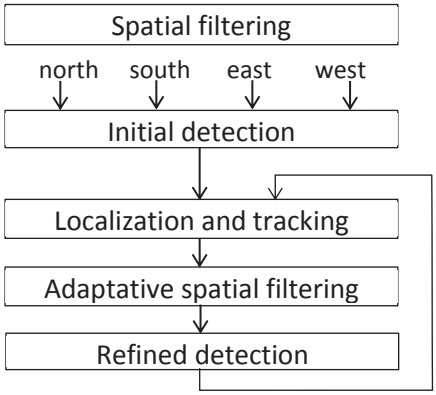


Figure 25. Global schematic of the signal processing tasks

The initial phase of drone detection was addressed as a quick and binary classification problem of presence/no-presence of drone in a noise sound recording. Few computer resources must be used for continuous on-site use, and lead to a low rate of false negatives, even if it may have a relatively high rate of false positives at first, which can be reduced in a second time during a later refined detection phase.

6.1 Detection / classification

For this initial phase, the signals are processed to obtain acoustic descriptors for each signal frame (duration of 20 ms) and are averaged by groups of 5 successive observations, giving a probability of presence between 0 and 1 on which a detection threshold is set to obtain a cost-sensitive classifier. A more detailed analysis is proposed in [3]. In order to facilitate the detection and extend its range, a spatial filtering by differential channel formation in four main directions (North, South, East, West) [12] has been implemented. The directivity diagram obtained by each of the microphone lines is quasi-constant with the frequency, thus not altering the frequency signature of the sources to be detected.

Standard audio descriptors have been selected and associated to the Jrip classifier [13] from the WEKA library (Waikato Environment for Knowledge Analysis). The Mel-frequency cepstral coefficients (MFCCS) are coefficients that together form a particular representation of the signal. MFCC coefficients give a compressed representation of the signal, which is commonly used as a set of descriptors in machine learning from audio data [14, 15]. In order to minimize the computation load, a set of nine optimized descriptors have been selected ([3]) and five additional descriptors from the MIR toolbox [16], leading to a total of 14 descriptors used in the presented classification process.

In a second step called “refined detection/classification” machine learning tools will be used to detect the presence or absence of drones from the calculated descriptors. The refined detection stage is preceded by MVDR (Minimum Variance Distortionless Response) beamforming informed by the target’s DOA. For this second step, the signal is processed on a greater number of successive frames (up to one second).

6.2 Construction of the database

Using the experimental data described in the paragraph 4 (Baldersheim tests), a database has been set up, using segments of 20 ms of the acoustic signals of the flying UAVs, with the knowledge of the distance between the drone and the sensor. It has been noticed in the literature [16] that the robustness of real outdoor sound classification may be improved or at minimum “strengthened” by the addition of various artificial noises. Consequently, the database has been extended by mixing the short duration signals from 4 drones with 4 environmental sounds from the TUT/DCASE 2016 residential sounds data base (Detection and Classification of Acoustic Scenes and Events) [17]. Drones flying from 4 categories of distances have been considered [0 to 50 meters], [50 to 100 meters], [100 to 200 meters], [200 to 400] meters. The 2/3 first samples of both Baldersheim and DCASE sounds are dedicated to the training database, while the 1/3 last samples are dedicated to the test database.

During the learning phase the tested drone and the environment corresponding to the recording have been excluded.

6.3 Main results

The F-score is used to express the global performance of the classifier. Both the *precision* and the *recall* are considered in order to estimate the score. For the 4 tested drones, we observed that this score is globally decreasing

for increasing distances and decreasing SNR values. Globally the detection of an unknown drone in an unknown soundscape is very good within 100 meters and good within 200 meters (Fig. 26).

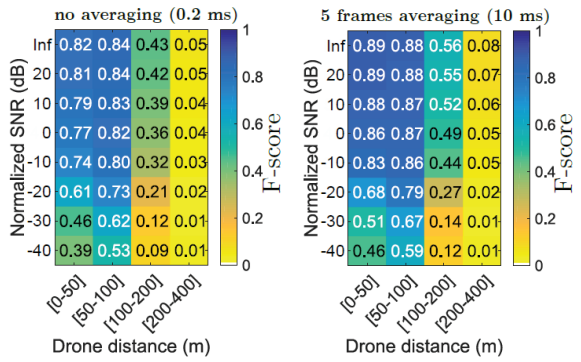


Figure 26. Results of classification F-score (JRip)

7. CONCLUSIONS

We have presented an overall view of the ISL & CNAM work done on the topic of acoustic detection and tracking of small UAVs. This study includes an experimental part, some electronic developments and a signal processing part. Original aspects are related to the pressure/velocity array and on the detection/recognition part combining two successive phases of detection/localization.

The next phase of our work will further investigate the use of deep-learning techniques [18].

8. ACKNOWLEDGMENT

The ISL studies have been initiated in the context of its program of work and have been supported by governmental contracts from the DGA and the BAAINBw. The ISL and the CNAM/LMSSC are grateful for the DGA PhD grant of A. Ramamonjy and the ASTRID contract DEEPLOMATICS.

9. REFERENCES

[1] A. Ramamonjy et al.: "Source localization and identification with a compact array of digital MEMS microphones", *25th. ICSV*, 2018.

[2] S. Hengy et al.: "Multimodal UAV detection: study of various intrusion scenarios", *Proc. of SPIE Vol. 10434,* Oct. 5, 2017.

[3] A. Ramamonjy: "Développement de nouvelles méthodes de classification/localisation de signaux acoustiques appliquées aux véhicules aériens", *PhD thesis*, CNAM Paris, 28 Mai 2019.

[4] P. Naz: "Detection and Acoustic Signatures of Helicopters", *Defense & Technologie Int.*, Sept. 1992.

[5] A. Lemer et al.: "Acoustic/Seismic Ground Sensors for Detection, Localization and Classification on the Battlefield", *NATO Meeting Proc.,* RTO-MP-SET-107, pp. 17-1 – 17-12, 2006.

[6] P. Naz: "Influence of terrain masking on the acoustic propagation of helicopter noise", Joint Acoustic Propagation Experiment (JAPE-91) Workshop, *NASA Conf. Publication* 3231, April 28, 1993.

[7] N. Intaratep et al.: "Experimental Study of Quadcopter Acoustics and Performance at Static Thrust Conditions", *AIAA Conf.*, May 30-31, 2016.

[8] W. Zhou et al.: "An Experimental Investigation on Rotor-to-Rotor Interactions of Small UAV", *35th AIAA Applied Aerodynamics Conf.*, 5-9 June, 2017.

[9] N.S. Zawodny et al.: "A Summary of NASA Research Exploring the Acoustics of Small Unmanned Aerial Systems", *NASA*, 2018.

[10] R.O. Schmidt: "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. Antennas Propagation*, Vil. AP-34, pp. 276-280, March 1986.

[11] N.J. Gordon et al.: "Novel approach to nonlinear/non Gaussian Bayesian state estimation", *IEEE proceedings*, 140(2), pp.107-113, 1993.

[12] J. Chen et al.: "On the design and implementation of linear differential microphone arrays," *The Journal of the Acoustical Society of America*, 136, 3097, 2014.

[13] W.W. Cohen: "Fast effective rule induction", *Machine Learning Proceedings,* pp. 115–123, 1995.

[14] J. Vavrek et al.: "Acoustic events detection with support vector machines". In *Faculty of Electrical Engineering and Informatics of the Technical University of Košice Proc.*, pp. 796–801, 2010.

[15] M. Müller: "Information retrieval for music and motion", vol. 2. Springer, 2007.

[16] O. Lartillot, et al.: "A Matlab toolbox for music information retrieval". *Data analysis, machine learning and applications*, pp. 261–268. Springer, 2008.

[17] S. Mukkamala et al.: "Feature selection for intrusion detection with neural networks and support vector machines". *Journal of the Transportation Research Board*, (1822): 33–39, 2003.

[18] A. Mesaros et al.: "Tut database for acoustic scene classification and sound event detection", *24th European Signal Processing Conf.*,1128–1132, 2016.

[19] E. Bavu et al.: "TimeScaleNet: A Multiresolution Approach for Raw Audio Recognition Using Learnable Biquadratic IIR Filters and Residual Networks of Depthwise-Separable One-Dimensional Atrous Convolutions", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 13, No. 2, pp. 220-235, May 2019.