# 2019 IEEE Signal Processing Cup:
## *Search and Rescue with Drone-Embedded Sound Source Localization*
### - Competition Details –

MathWorks®     IEEE Signal Processing Society ®

*To ask questions and download competition data, please visit and enroll as a student on the competition's website: https://piazza.com/ieee_sps/other/spcup2019/*

Last update of this document: 11.19.2018

## 1. Introduction

Unmanned aerial vehicles (UAV), commonly referred to as *drones*, have been of increasing influence in recent years. **Search and rescue scenarios** where humans in emergency situations need to be quickly found in areas difficult to access constitute an important field of application for this technology. Drones have already been used in places like Haiti and the Philippines to map areas after a natural disaster, using high-resolution **embedded cameras** (see this recent United Nation report: http://tiny.cc/xmy0zy). On the other hand, whenever there is a lack of visual feedback due to bad lighting conditions (night, fog, etc.) or occlusions, **audio signals** could be of critical help to localize people in emergency. While **sound source localization** is a long-standing signal processing research topic, the **unique challenges** introduced by UAV search and rescue are at the heart of the IEEE signal processing cup 2019 (SPCup2019)



*Figure 1. Setup used for the SPCup2019: A Mikrokopter UAV equipped with an array of 8 microphones*

The main task is to **accurately estimate the direction of a sound source** (azimuth, elevation) using recordings from an 8-channel cube-shaped microphone array embedded in a flying UAV (Figure 1). This task is particularly challenging because of the high amount of **acoustic noise** produced by the UAV, resulting in negative signal-to-noise ratios (SNR). The noise consists of harmonics components related to the propellers speed as well as stationary structural noise and sometimes wind noise due to the UAV's movements and propellers rotations. Another difficulty

[1]A graduate student is a student having earned at least a 4-year univ. degree at the time of submission
[2]An undergraduate student is a student without a 4-year degree

comes from the fact that a UAV is **constantly moving**, sometimes with rapid changes of directions, resulting in complex relative source trajectories in the microphone array's reference frame.

# 2. Competition Rules

## 2.1. Organization

The competition will consist of two stages; an open competition that any eligible team can participate in and an invitation-only final competition. Teams participating in the open competition must submit their entries no later than **February 29, 2019**. The three teams with the highest performance in the open competition will be selected as finalists and will be invited to participate in the final competition. Finalists will be announced on **March 20, 2019**. The three teams invited to participate in the final competition will be judged at ICASSP 2019, which will be held **May 12-17, 2019**.

## 2.2. Formation of competition teams

Each team participating in the SPCup2019 should be composed of **one faculty member** or **someone with PhD degree employed by the university** (the Supervisor) at most **one graduate student**[1] (the Tutor), and at least **three** but no more than **ten undergraduate students**[2]. At least three of the undergraduate team members must hold either regular or student memberships of the IEEE Signal Processing Society. Undergraduate students who are in the first two years of their college studies as well as high school students who are capable to contribute are welcome to participate in a team. A participant should not be on more than one team.

## 2.4. Open competition submissions

In order to be considered for the final competition, teams that wish to participate in the open competition should submit the following material by **February 29, 2019** via the CMS system (details on the website)

1. A 2 to 4 pages report in the form of an IEEE conference paper describing the technical details of their sound source localization system.
2. Sound source localization results in .mat format, as described in Sections 3.1 and 3.2

In addition, extra points can be gained by submitting new audio recordings obtained from one or several microphones embedded in a flying UAV together with a report, as described in Section 4.

## 2.3. Prize for finalists

According to the guideline from the IEEE Signal Processing Society, a maximum of three members from each of the three finalist teams will receive travel support to attend ICASSP 2019 for the final competition (up to $1,200 for continental travel or $1,700 for intercontinental travel per

---

[1]A graduate student is a student having earned at least a 4-year univ. degree at the time of submission
[2]An undergraduate student is a student without a 4-year degree

member, and at most three people from each team will be supported). The participants can claim their travel expenses on a reimbursement basis.

More team members are also welcome to attend. Although these additional members will not be provided with any travel grant, those members who will not be presenting a paper at the conference will be offered a complimentary ICASSP registration. The finalist teams will also be invited to join the Conference Banquet as well as the Student Career Luncheon, so that they can meet and talk to SPS leaders and global experts.

The winner will be selected based on the obtained results during the final competition, the quality of the final presentation, and the capability to address questions. The champion team will receive a grand prize of $5,000. The first and the second runner-up will receive a prize of $2,500 and $1,500, respectively, in addition to the above mentioned travel grants and complimentary conference registrations.

# 3. Open Competition - Main Task: Sound Source Localization

## 3.1. Static source localization

For this task, 300 **8-channel** audio recordings at 44.1 kHz and of roughly 2 seconds each are provided in the form of wav files, named **1.wav** to **300.wav**. They can be found in the **static_task.zip** file, inside the **audio/** folder. All recordings were made with the 8-microphone array described in Section 3.5. Each were obtained by summing a clean recording of a static loudspeaker emitting speech from an unknown (azimuth, elevation) direction in the array's frame, and a recording of UAV noise of the same length in various flight conditions and using various signal-to-noise ratios.

The goal of this task is to retrieve the azimuth and elevation angles of the static speech source for each of the 300 recordings. To do so, a **mat file** (see Section 6.1.) called **SPCUP19_static.mat** must be provided, containing the following two **300x1** vectors:

- `static_azimuth` such that `static_azimuth(J)` is the estimated real-valued azimuth angle for recording **J.wav.**
- `static_elevation` such that `static_elevation(J)` is the estimated real-valued elevation angle for recording **J.wav.**

In addition to the 300 recordings, a mat file called **static_motor_speed.mat** is optionally provided. It contains a **300x4** matrix `static_motor_speed` such that `static_motor_speed(J,K)` contains the average angular velocity of propeller `K` during recording **J.wav**. It can be found in the **static_task.zip** file. Suggestions to use these data to reduce egonoise are given in Section 6.

## 3.2. In-flight source localization

For this task, 36 **8-channel** audio recordings at 44.1 kHz lastly precisely 4 seconds each are provided in the form of wav files, placed in the **audio/** folder of **flight_task.zip**. For recordings

[1]A graduate student is a student having earned at least a 4-year univ. degree at the time of submission
[2]An undergraduate student is a student without a 4-year degree

**audio/speech/1.wav** to **audio/speech/16.wav** the source is a loudspeaker emitting speech while for recordings **audio/broadband/1.wav** to **audio/broadband/20.wav** the source is a loudspeaker emitting white noise. All recordings were made **during flight** with the 8-microphone array described in Section 3.5 attached to a Mikrokopter UAV as showed in Figure 1. The flights were performed in a large room with moderate reverberation level. While the source (loudspeaker) is static during flights, the microphone array is moving with the drone and hence, the (azimuth, elevation) source direction in the array's frame is **time-varying**.

For each of these 4 seconds recordings **15 timestamps** are defined as follows, starting from the beginning of each file:

{0.25s, 0.5s, 0.75s, 1s, 1.25s, 1.5s, 1.75s, 2s, 2.25s, 2.5s, 2.75s, 3s, 3.25s, 3.5s, 3.75s}

The goal of this task is to retrieve the mean azimuth and elevation angles of the source within a **500ms** window centered on each of these timestamps, for each of the 32 recordings. To do so, a **mat file** (see Section 6.1.) called **SPCUP19_flight.mat** must be provided, containing the following four matrices:

- `speech_azimuth` (16x15) such that `speech_azimuth(J,T)` is the estimated real-valued azimuth angle at timestamp `T` for recording **J.wav.**
- `speech_elevation` (16x15) such that `speech_elevation(J,T)` is the estimated real-valued elevation angle at timestamp `T` for recording **J.wav.**
- `broadband_azimuth` (20x15) such that `broadband_azimuth(J,T)` is the estimated real-valued azimuth angle at timestamp `T` for recording **J.wav.**
- `broadband_elevation` (20x15) such that `broadband_elevation(J,T)` is the estimated real-valued elevation angle at timestamp `T` for recording **J.wav.**

In addition to the 300 recordings, a mat file called **flight_motor_speed.mat** is optionally provided. It contains the following two arrays (one for each type of source):

- `speech_flight_motor_speed` (16x15x4)
- `broadband_flight_motor_speed` (20x15x4)

such that the real value at (`J`,`T`,`K`) in each array is the average angular velocity of propeller `K` around timestamp `T` for recording **J.wav**. The averages are over **500ms** windows centered on timestamps. Suggestions to use these data to reduce egonoise are given in Section 6.

## 3.3. Evaluation metric

A source localization is considered correct if and only if **both** the estimated azimuth and elevation angles are **within 10°** of the true azimuth and elevation angles. Each correct localization is worth **1 point**. Hence, **300 points** can be earned from the static source localization task, and **36x15 = 540 points** can be earned from the in-flight source localization task, for a total of **840 points.**

## 3.4. Development data

In addition to the challenge data, some development data are provided to help the participants.

---

[1]A graduate student is a student having earned at least a 4-year univ. degree at the time of submission
[2]An undergraduate student is a student without a 4-year degree

- **dev_static.zip** contains three 2 seconds clean recordings of a static loudspeaker emitting white noise (broadband source). They are named **1.wav**, **2.wav** and **3.wav**. The true (azimuth, elevation) directions of arrival for these recordings are given in **SPCUP19_dev_static.mat**, using Section 3.1 instructions.
- **dev_flight.zip** contains three 4 seconds clean recordings of a static loudspeaker emitting white noise (broadband source) made while the UAV was moved around in a room with its **motors switched-off**. They are named **1.wav**, … , **5.wav**. The true (azimuth, elevation) directions of arrival for these recordings are given in **SPCUP19_dev_flight.mat**, using Section 3.2 instructions.
- **individual_motor_recordings.zip** contains 21 **8-channel** .wav files at 44.1 kHz corresponding to audio recordings of each of the 4 **individual propellers** of the UAV at different speed (angular velocity). For this, the UAV equipped with the microphone array was fixed on a nylon string hanging freely in a room and held still by a human operator. This was done in order to record pure motor noise at a chosen speed without having the influence of balancing and stabilizing movements and with limited amount of in-flight wind noise. The .wav files are named according to the convention **MotorX_YY.wav** where X is the propeller's number from 1 to 4 and YY is the motor speed in rotation/second from 50 to 90. An additional recording allMotors_70.wav in which all 4 propellers runs at the same time is also provided.

## 3.5 Array geometry

The positions of the eight microphones in the array's coordinate frame centered at the barycenter of the array are given in the file **micPos.txt**. The structure holding the microphones is cubic, while the microphone constellation forms two parallel horizontal squares, each being twisted in opposite directions in the azimuthal plane.

## 3.6. Baseline

A baseline correlation-based sound-source localization method implemented in both Python and Matlab will be provided on the website to help participants.

# 4. Open Competition - Bonus Task: Data Collection

On top of the **840 points** that can be gained by correctly localizing all the sources (Section 3.3), participating teams that wish to do so can send their own audio recordings obtained from one or several microphones embedded in a flying UAV. The recordings should be made either outdoor or in an environment with mild reverberation, and should not feature other sound sources than the UAV's noise. A report thoroughly detailing the microphones and UAV used, the recording conditions and including pictures of the setup and experiments should be given with the audio files. **Up to 60 extra points** will be granted to teams submitting audio files, based on the quality of the report.

---

[1]A graduate student is a student having earned at least a 4-year univ. degree at the time of submission
[2]An undergraduate student is a student without a 4-year degree

After the end of the SPCup2019, a new UAV egonoise dataset will be created from the teams' contributions and will be made freely available online for research and education purpose.

# 5. Final Competition at ICASSP 2019

The three best competing teams will be selected according to their submitted results. These teams will be invited to the final at ICASSP 2019. There, the teams will be given new recordings made with the UAV, with increasing levels of difficulty. They will have a **limited amount of time to run their methods** on each of them. The estimated trajectories will be evaluated on the spot, using the same metrics as in the open competition. The teams will also give a detailed presentation of their work. The winner will be selected by a Jury based on the obtained results, the quality of the final presentation, and the capability to address questions.

# 6. Technical Background

## 6.1. Mat files

Files with a **.mat** extension contain MATLAB formatted data and are extensively used in the SPCUp2019 to store data and results. They can contain any number of vectors and arrays as well as other variable types that are not used in this challenge. They can be created, read and written using both Python and Matlab programming languages a like.

- For Matlab users see https://fr.mathworks.com/help/matlab/matlab_env/save-load-and-delete-workspace-variables.html
- For Python users see  https://docs.scipy.org/doc/scipy/reference/tutorial/io.html

*Additional technical background in sound source localization and noise reduction will be added to this section during the competition.*

[1]A graduate student is a student having earned at least a 4-year univ. degree at the time of submission
[2]An undergraduate student is a student without a 4-year degree