SCIENCE AND TECHNOLOGY ORGANIZATION

COLLABORATION SUPPORT OFFICE

Cooperative Navigation
in GNSS Degraded and Denied Environments (SET 275)
29 - 30 September, 2021

Drone Detection and Classification Using an Acoustic Camera

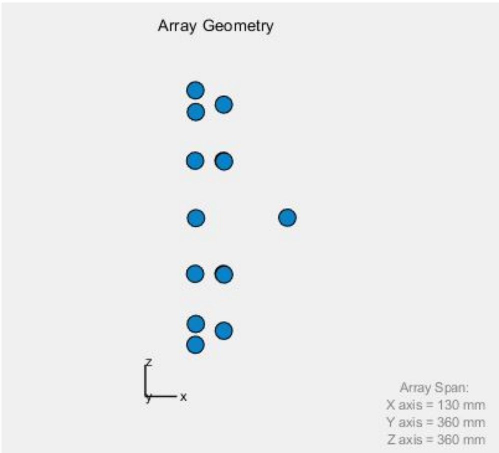
Sanja Grubeša[1], Jasna Stamać[2], Nikša Orlić[2], Tomislav Grubeša[2]

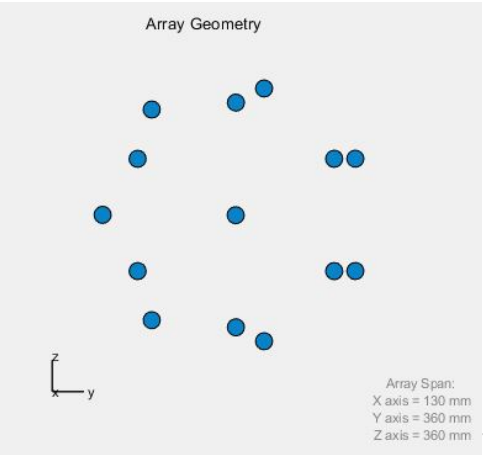
[1] University of Zagreb, Faculty of Electrical Engineering and Computing, Department of Electroacoustics, Zagreb, CROATIA, sanja.grubesa@fer.hr

[2] Geolux d.o.o., Ljudevita Gaja 62, Samobor, CROATIA

Design a broadband acoustic camera using MEMS microphones

- Choosing the shape of acoustic camera.
- Developed an algorithm called *Hemisphere* - written in MATLAB, to determine the optimal number of microphones and the position of microphones on the camera.
- Analysed the directivity patterns for frequencies f = 1 kHz, 2 kHz and 4kHz while maintaining the same number of microphones and their positions on the hemisphere.

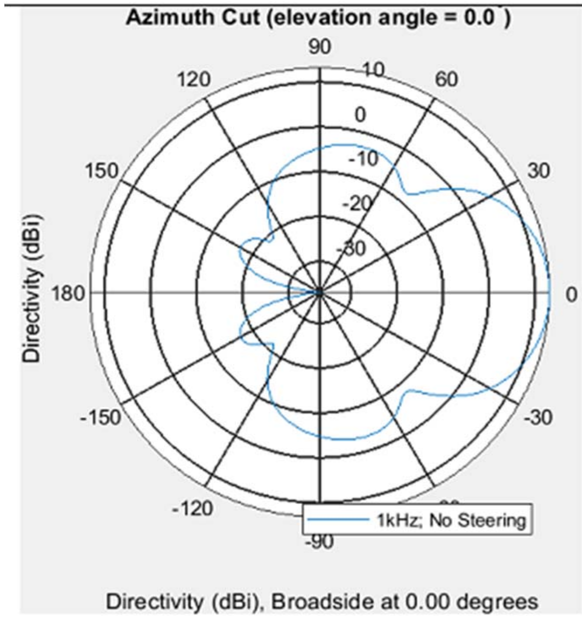




r_{max}	0.200 m
d	0.170 m
α_1	20°
r_1	0.188 m
n_1	7
α_2	35°
r_2	0.164 m
n_2	6
G	13.2 dBi
A	16.3 dBi
n	14

Directivity pattern for the camera before optimization

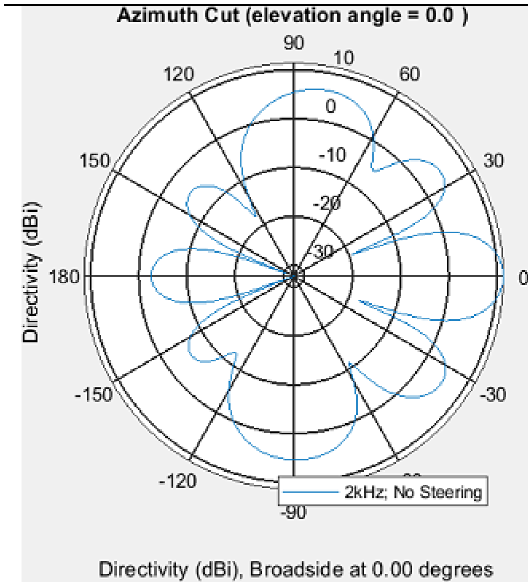
1 kHz



G = 13.2 dBi

A = 16.3 dBi

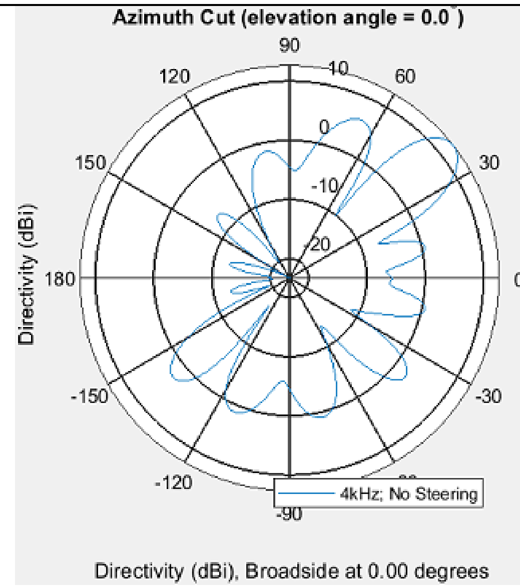
2 kHz



G = 11.4 dBi

A = 4.9 dBi

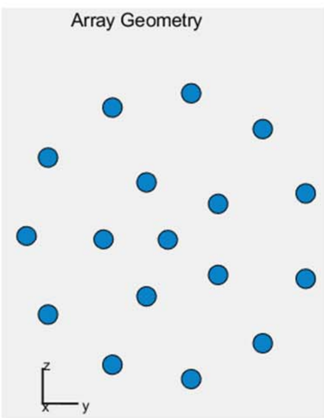
4 kHz

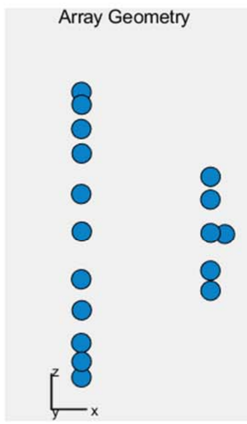


G = -5.7 dBi

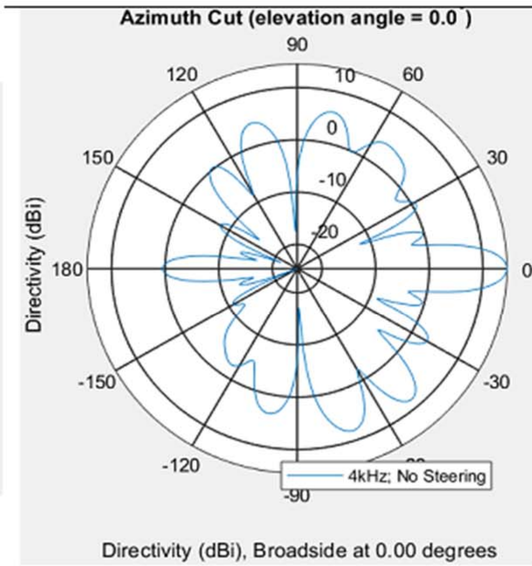
A = -11.8 dBi

Hemispherical acoustic camera optimized for frequency 4 kHz



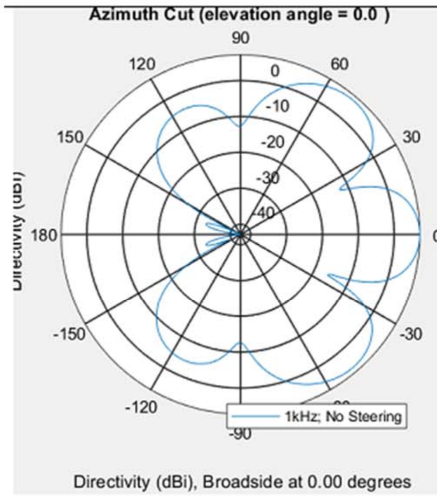


4 kHz



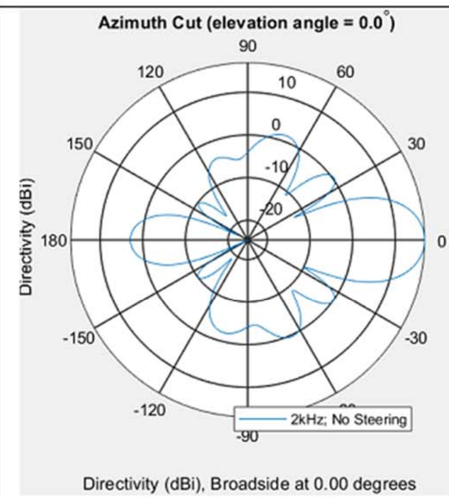
$G = 14.5$ dBi
$A = 6.2$ dBi

1 kHz



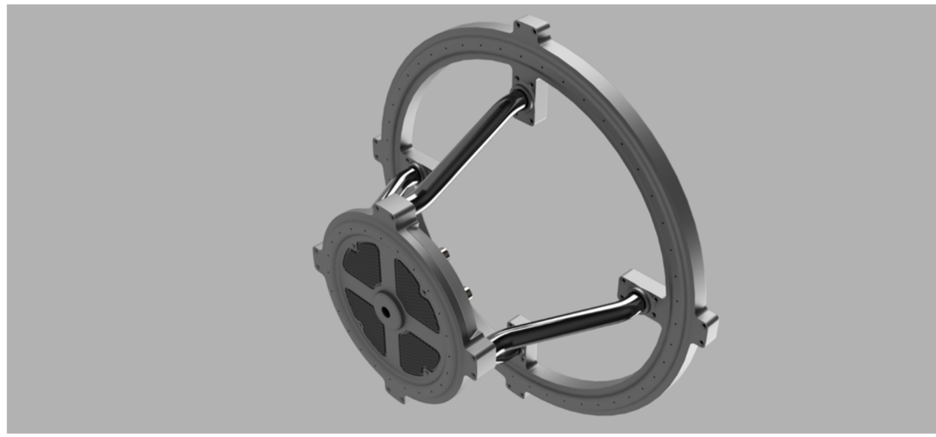
$G = 7.2$ dBi
$A = 1.2$ dBi

2 kHz



$G = 16.9$ dBi
$A = 15.5$ dBi

Prototype of the Acoustic camera

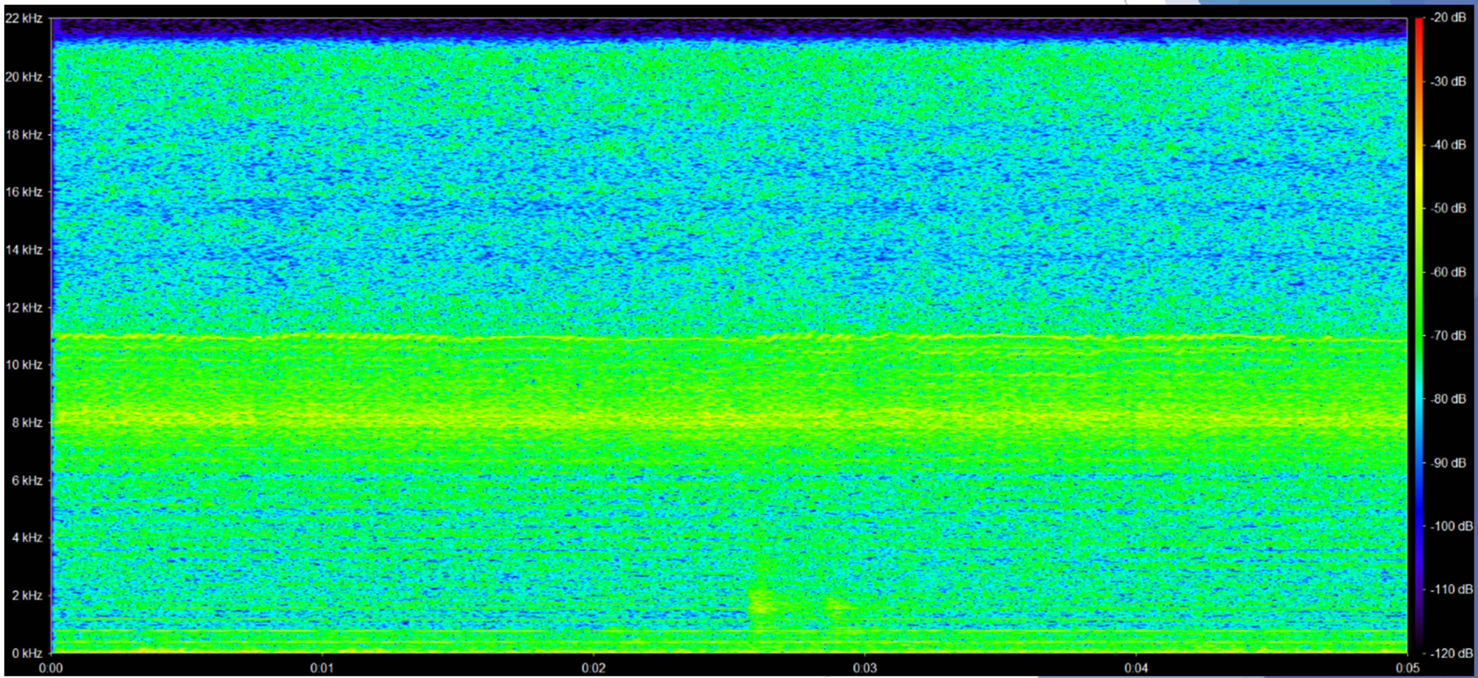


Create a database of different sounds

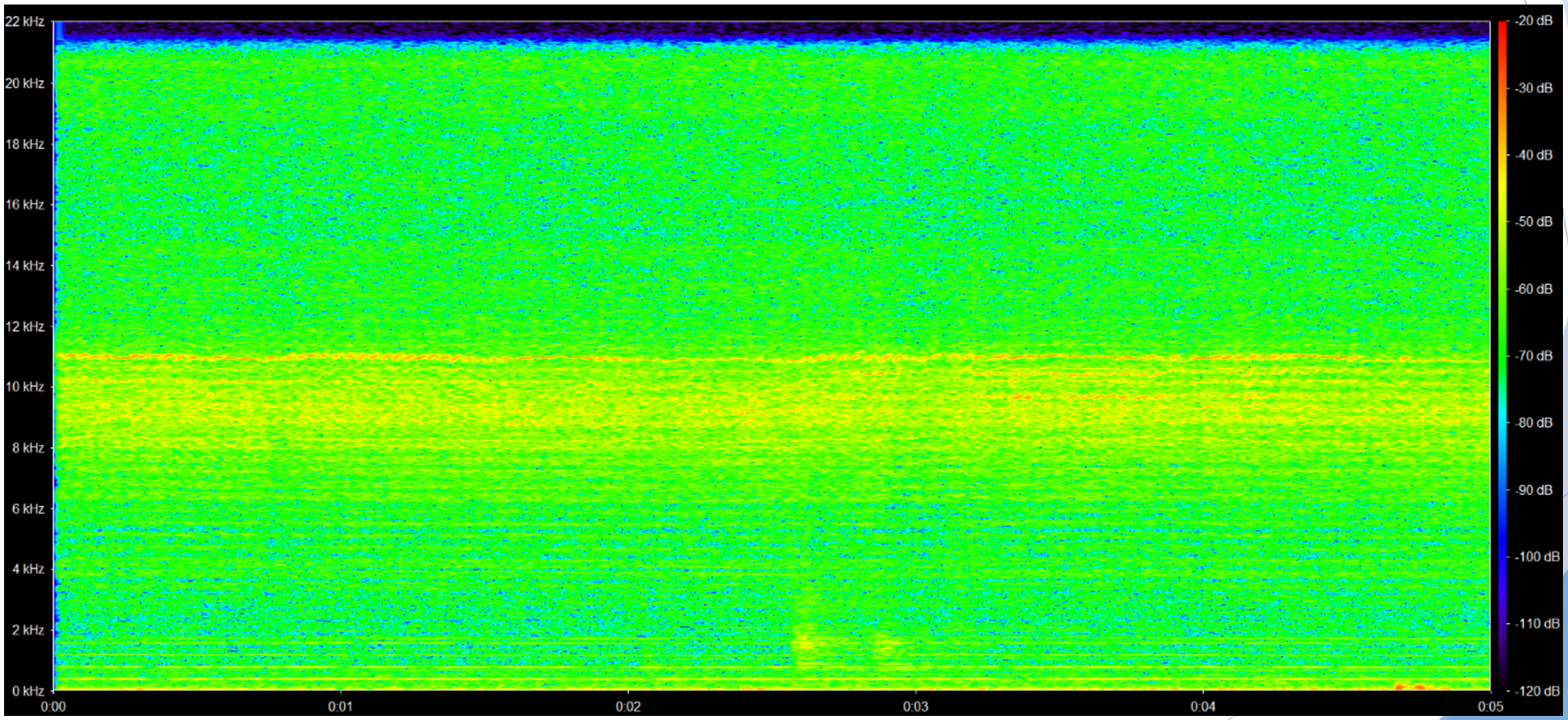
First step of classification was to create a database of different sounds.

Samples are divided into the following four categories:

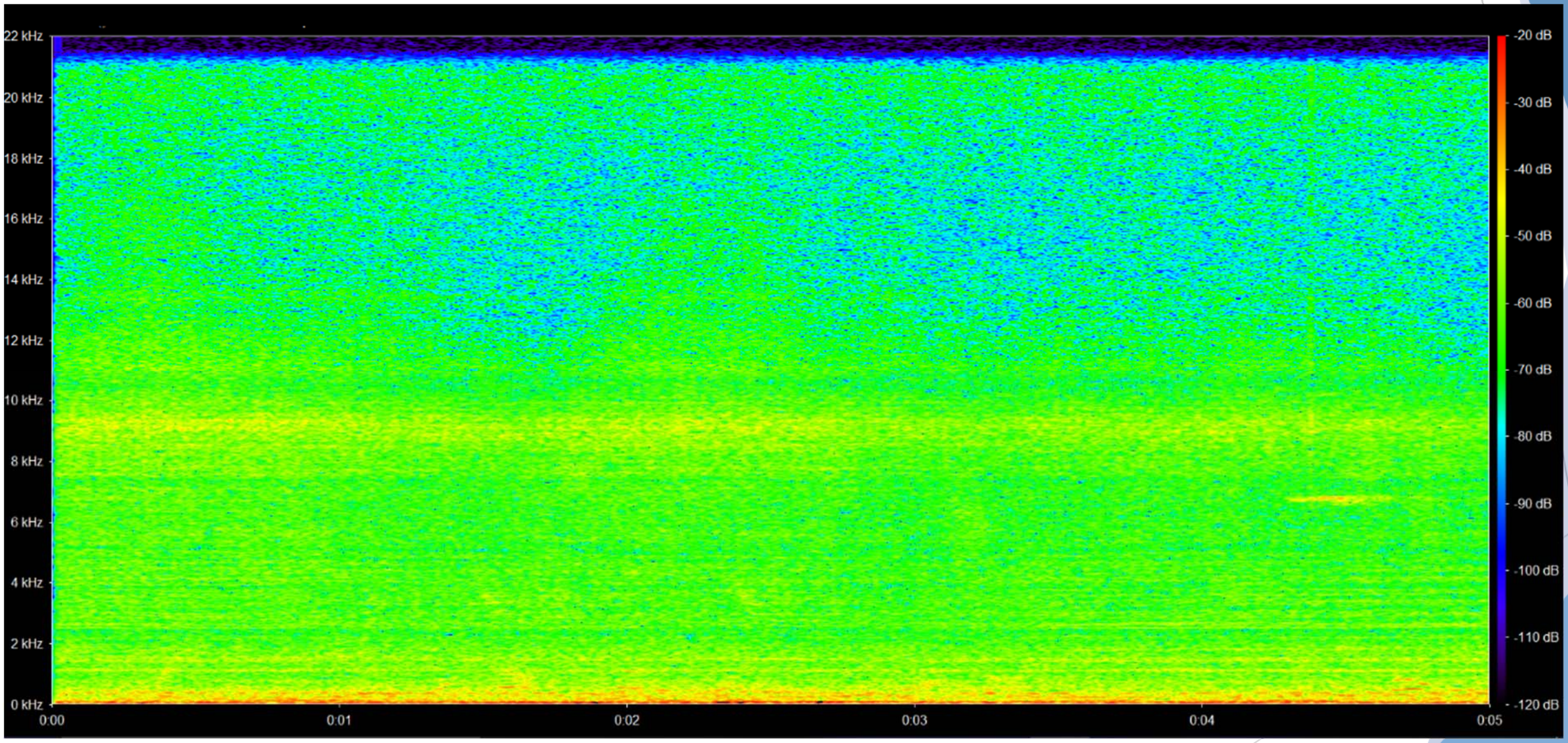
- noise,
- walking,
- speech,
- drone flight.



Sound Spectrogram Generated by a Drone 5 m Away From the Acoustic Camera



Sound Spectrogram Generated by a Drone 50 m Away From The Acoustic Camera



Implementation of a Convolutional Neural Network for Acoustic Camera

TensorFlow tool easy to experiment with:

- different architectures of convolutional neural networks
- different sets of recorded samples (training sets)

to select the optimal network architecture and check if the training set contains enough recordings.

In the classification of acoustic samples, the input to the first layer of the neural network is a 2D spectrogram matrix, where each element of the matrix contains the intensity of the spectrogram at a given point.

Implementation of a Convolutional Neural Network for Acoustic Camera

Neural network architecture that looks as follows:

- 1. Input layer
- 2. Max-pooling layer
- 3. 2D convolution 3x3, 64 elements
- 4. Max-pooling layer
- 5. Dropout 50%
- 6. 2D convolution 2x2, 64
- 7. Max-pooling layer
- 8. Dropout 50%
- 9. 2D convolution 2x2, 64
- 10. Flatten layer, 128 elements
- 11. Dropout 50%

Samples divided into the following four categories:

- Noise,
- Walking,
- Speech,
- Drone flight.

Final results on the validation set

		Real samples for validation			
		Noise	Walking	Speech	Drone flight
Classification results	Noise	87%	7%	7%	4%
	Walking	2%	89%	5%	1%
	Speech	3%	3%	85%	3%
	Drone flight	8%	1%	3%	92%

CONCLUSIONS

- The architecture of convolutional neural networks has proven to be extremely effective in working with images and recognizing features from them.
- Since convolutional neural networks are frequently used to classify images, we can imagine the input to the convolutional network as a grayscale image which represents the spectrogram.
- Preliminary classification results show a high classification accuracy of the samples, i.e. spectrograms, which shows that the method and architecture of the selected neural network are appropriate.
- Although the number of collected samples used as the training set for the convolutional neural network was relatively small, the achieved classification accuracy proved to be high, hence it will not be necessary to modify the neural network architecture.