24th IMEKO TC4 International Symposium
22nd International Workshop on ADC and DAC Modelling and Testing
IMEKO TC-4 2020
Palermo, Italy, September 14-16, 2020

System for Acoustic Detection

Jakub Svatos[1], Jan Belak[1], Jan Holub[1]

[1] *Department of Measurement, Czech Technical University in Prague, Prague 166 27, Czechia*
svatoja1@fel.cvut.cz

***Abstract* – At present, acoustic detection techniques of gunshots (gunshot detection) are increasingly being used not only for military applications but also for civilian purposes. Detection, localization, and classification of gunshots employing acoustic detection is perspective alternative to visual detection, which is commonly used. In certain situations, to detect a source of a gunshot, an automatic acoustic detection system may be preferable. This paper presents a system for acoustic detection, which can detect localize and classify acoustic events such as gunshots. Tested firearms are 9 mm short gun, 6.35 mm short gun, .22 short gun, and .22 rifle gun with various subsonic and supersonic ammunition. As "false alarms," sets of different impulse acoustic events like door slams, breaking glass, etc. have been used. To successfully classify the tested acoustic signals, Continuous Wavelet and Mel Frequency Transformation methods have been used for the signal processing, and the fully two-layer connected neural network has been implemented. The results show that the acoustic detector can be used for reliable gunshot detection, localization, and classification.**

I. INTRODUCTION

Acoustic detection (AD) of gunshots is a present topic which can help to detect hazardous and dangerous events in public areas. In recent days there is an increase in gunshot attacks in public areas such as schools, campuses, hospitals, and shopping centers. A fundamental goal of the AD is to record acoustic signals around the area of interest and to detect, localize and classify it into categories as an alert (gunshot) signals and standard, 'false alert' signals. Sometimes it is challenging to recognize a dangerous event from uncertain, inadequate received data by cameras or by security staff. The main asset of AD is based on the extraction of vital information from the recorded signal data and classify it due to the given issue (gunshot, a human scream, glass breaking, etc.). Due to this classification, AD can assist in Police, Law enforcement to better distinguish dangerous events and intervene in the going process on time and decrease the further casualties.

There are several experimental or commercial AD systems designed for gunshot events detection and localization available on the market [1] - [5]. These systems are intended to localize the source of the gunshot. The more sophisticated systems can even identify the particular firearm type using advanced classification methods. A drawback of the sophisticated systems is usually very high purchase and operational cost.

To successfully detect and classify a gunshot, essential characteristics of its complex physic have to be understood. A comprehensive explanation can be found in [6] and [7]. A gunshot is characterized by two phenomena, muzzle blast, and, if the bullet travels at supersonic speed, shock wave. Muzzle blast is caused by an explosive charge hot, high-pressured gases expanding as acoustic energy from the center of the barrel. The bullet with the supersonic speed generates a shockwave effect, which is propagating on the conic fashion behind the bullet trajectory. The shockwave is based on the combinations of compression and expansion shock.

These factors can include some valuable information that can be used for the detection capability of the detection system. Alongside this, the caliber of both bullet and barrel, the length of the latter, mechanical action caused by a gun itself, or even the chemical properties of the propellant cause different effects on the pattern of a gunshot. Last but not least, the temperature of the air, air humidity, wind speed, environment (e.g., foliage density, urban area) and soil characteristics have also impact on the resulting gunshot pattern. Considering these phenomena, to effectively detect and identify a gunshot, signal processing, including adaptive filtering and advanced data classification, have to be done [8] – [10]. An example of a typical pattern of a subsonic and a supersonic gunshot signal is in Figs. 1 and 2.

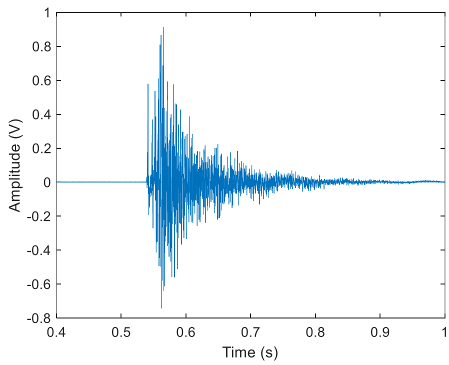

Fig. 1. Signal corresponding to 9 mm subsonic short gun shoot.

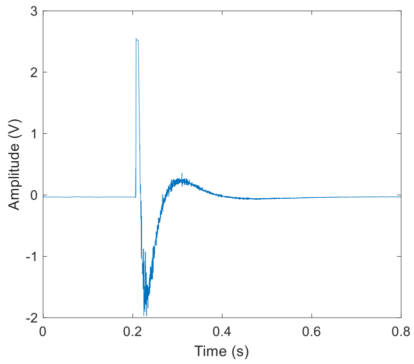

Fig. 2. Signal corresponding to 9 mm supersonic short gun shoot.

In Fig. 2, the shock arrival is not clearly visible due to the relatively low speed of the supersonic bullet (Mach number M = 1.09) and its proximity to the muzzle blast part of the signal.

In this article, a system for acoustic detection and classification is introduced. The system consists of sensor units which will be placed around monitored area in sufficient numbers and continuously monitor acoustic events around the unit and sends the data to the remote unit in case of possible gunshots detection. The central unit evaluates signals received from multiple sensor units and using advanced signal processing and classification methods determines the location of gunshot and the probable type of used firearm caliber. The localization accuracy of the system depends on the density and number of sensor units. In comparison with other existing available systems (e.g. [3] or [5]), the presented system has a novel modular flexible structure. It can be deployed on the building or moving object (car, person, animal) while the central unit can be installed in a distant protected place.

In the future, the presented acoustic detector can be used in public areas like schools, campuses, shopping centers to detect and localize gunshot.

The paper is organized as follows. In Section II, the sensor units, detection algorithm and signal processing are introduced and described. In Section III, the experimental measurements and results are presented. The conclusion and future work directions are stated in section IV.

II. METHODS

The presented system for acoustic detection and identification consists of multiple (at least four to estimate the correct localization of the event) sensor unit and one central unit. Such a topology enables additional analyses at the central unit, i.e., triangulation localization of the acoustic event position using timestamped data from multiple sensor units receiving the acoustic signal related to the shooting. Each sensor unit has to cover analog signal pre-processing, digitization, simple detection algorithm and simple evaluation according to Fig. 3.

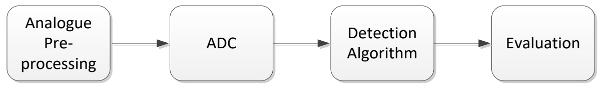

Fig. 3. Unit function requirements.

The stand-alone sensor working on this principle has been introduced by the authors in [8]. It works on the principle of dividing the recorded signal to the optimal number of time windows, which are fed to the median filter and the resulting signal from the median filter is then filter the analyzed signal from the background noise. Thereafter the detection algorithm based on multiple thresholds distinguish actual gunshot events from other ‘false alarms’. More details about the algorithm can be found in [11].

Requirements for flexible modular sensor units are even more demanding. Since the sensor units can be deployed everywhere, it has to be able to send their exact position. Moreover, all the sensor units have to by precisely time-synchronized to get timestamp about the detected event properly and send it to the central unit wirelessly. To fulfill all these criteria, the sensor unit has been designed according to Fig. 4.

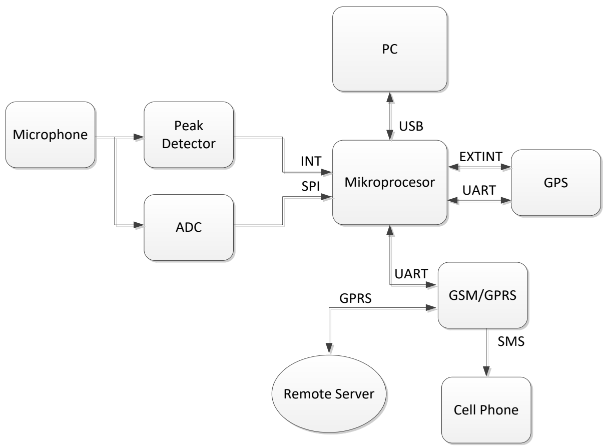

Fig. 4. Block diagam of the sensor unit.

Every sensor unit uses a pre-polarized, electret condenser microphone that has a flat frequency response from 20 Hz to 20 kHz and omnidirectional sensitivity in all directions, see Fig. 5 and 6.

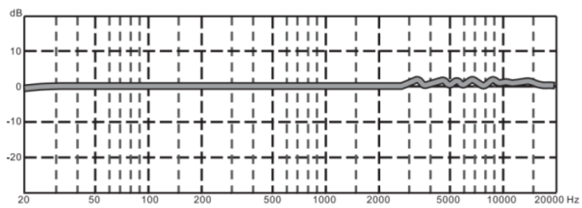

Fig. 5. Frequency response of the microphone.

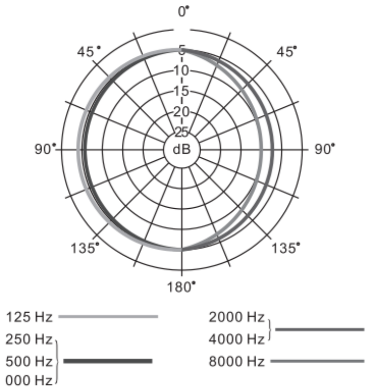

Fig. 6. Omnidirectional sensitivity of the microphone.

The heart of the acoustic sensor unit is low consumption 32-bit microprocessor LPC 1837, which processes the recorded data from a microphone digitized by a 16-bit Analogue-to-Digital Converter (ADC) ADS8866. The ADC samples the recorded signal with sampling frequency f_s = 44 kHz. The peak detector provides an interrupt for the microprocessor if an acoustic even is triggered. After the interrupt, the data stored in a circular buffer as a 16-bit number from the ADC are ready for processing.

If a trigger from the peak detector occurs, the median filter algorithm described above evaluates if there is the possibility of a gunshot. If the possible gunshot is detected, low consumption GSM chip GL865-QUAD V3 then sends the recorded data together with the position and a timestamp for further analysis to the remote central unit PC. The localization of the sensor unit position and synchronization with the other sensor modules secures GPS module LEA-6T by adding the precise timestamp to the recorded event. For synchronization, NMEA protocol [12] is used. The UCT timestamp and exact position of each unit are determined. The time accuracy of the LEA-6T module, for synchronization of all units, is 0.1 ms. The whole sensor unit is power supplied from U_{cc} = +5V.

The central remote unit process the data using advanced signal processing and classification to state if there is a gunshot detected and identify the caliber of the firearm or if its only 'false alarm'. The whole sensor unit with a testing microphone is in Fig. 7.

The Continuous Wavelet (CW) and Mel Frequency Transformation (MFT) methods are used for the processing of received data to obtain features for the further classification. MFT algorithm uses band-pass filters to get the energy of the signal for each defined band. Then the algorithm uses frequency distribution to create Mel Frequency Coefficients (MFC). MFC is computed using cosine transformation on the logarithm of bank energies. In this case, 20 MFC has been used with filter bands of 500 Hz to 5 kHz. More details about MFC can be found in [13] or [14].

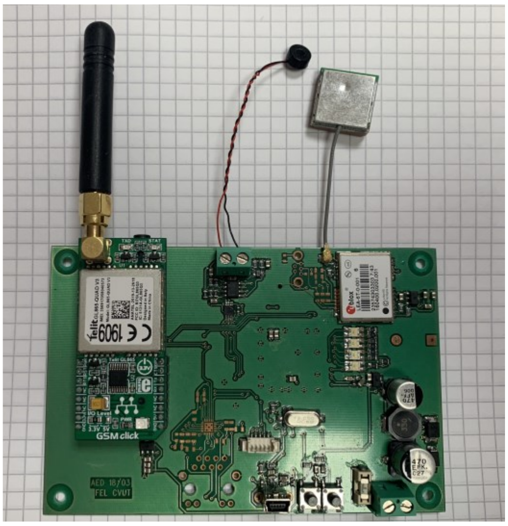

Fig. 7. The sensor unit for acoustic detection.

To limit the influence of an echo on a classification, the CW algorithm was considered. Unlike FFT, CW transform uses defined waves to create frequency spectrum with significant time resolution. Time resolution allows us to limit the influence of echo, due to the fact that the echo does not usually interact with the beginning of the shot sound. As a mother wavelet, Bump wavelet [15] and [16], which is defined by (1), has been used, Fig. 8.

$$\Psi(s\omega) = e^{1-\left(\frac{1}{1-(s\omega-\mu)^2/\sigma^2}\right)} 1_{\left[\frac{\mu-\sigma}{s},\frac{\mu+\sigma}{s}\right]} \tag{1}$$

, where $1_{[(\mu-\sigma)/s,\ (\mu+\sigma)/s]}$ is the indicator function, s is the scale, ω is angular frequency, σ is standard deviation and μ is mean value.

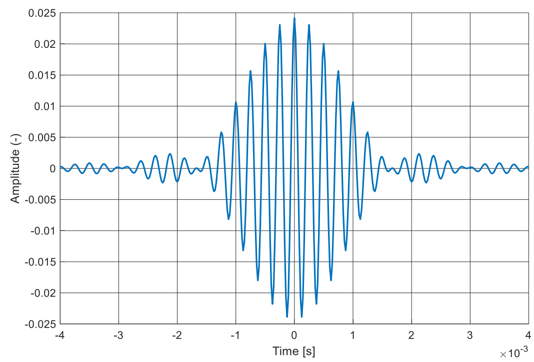

Fig. 8. Bump mother wavelet.

For classification, a set of gunshots were used as well as several false signals similar to a gunshot signal. Each false signal has been chosen to be challenging to differentiate from real gunshots by a human operator.

III. RESULTS

To acquire the test data, multiple gunshot measurements in different shooting ranges were done. For increasing the diversity of gunshots, three different calibers were measured. Tested firearms has been 9 mm short gun, 6.35 mm short gun, .22 short gun, .22 rifle gun. A various subsonic and supersonic (up to Mach number of M = 1.1) ammunition has been used with the 9 mm short gun. Each type of gun has been measured at least 60 times. In total, approx. 250 samples corresponding to a gunshot have been taken and recorded at least by three sensor units described in *Section II*. For false signals, impulse acoustic events similar to gunshots were measured. As an example of a false signal, glass breaking, doors slam, handclaps, or close bubble wrap popping have been recorded. All the tested signals has been recorded by at least three sensor units placed in nearby locations.

To accurately localize the source of the detected gunshot, a timestamp of the detected event and exact location of at least three acoustic units, which detected the event, have to be known. The relationship between the source of the gunshot and three sensor units describe equations (2) to (4)

$$(t_1 - t_0) * c = (x - a_1)^2 + (y - b_1)^2 \tag{2}$$

$$(t_2 - t_0) * c = (x - a_2)^2 + (y - b_2)^2 \tag{3}$$

$$(t_3 - t_0) * c = (x - a_3)^2 + (y - b_3)^2 \tag{4}$$

, where t_0 is the time where the gunshot has occurred, t_1, t_2 and t_3 are the times where the gunshot has been detected by unit 1, 2 and 3, (a_x, b_x) are the coordinates of the relevant unit and (x, y) are the coordinates of the gunshot. The resulted coordinates are necessary to recalculate the Cartesian system. The unit which detects the gunshot first (t_1) is considered as placed in point (0, 0). The coordinates in meters of other units are then calculated by (5).

$$d = acos(sin(\Phi_1)\, sin(\Phi_2)) + \\ + cos(\Phi_1)\, cos(\Phi_2)\, cos(\delta\lambda) R \tag{5}$$

, where Φ_1 and Φ_2 are the coordinates in meters and $\delta\lambda$ is the difference of the longitudes and R is the mean Earth radius (6378 km). Once the coordinates of the Cartesian system are calculated, it is necessary to recalculate it back to the latitude and the longitude coordinates. Since the distance of one degree of longitude is different at the North Pole and the Equator, it is necessary to know the relationship between degree and meter at a given latitude [17]. This is done by the simplified formula (6) and (7), which recalculates meters to one degree of latitude of longitude, resulting in inaccuracies in the order of centimeters.

$$x_{latitude} = 111123.92 - 559.82\, cos(2\lambda) + \\ +1.175\, cos(4\lambda) - 0.0023\, cos(6\lambda) \tag{6}$$

$$x_{longitude} = 111412.84\, cos(\Phi) - \\ -93.5 cos(3\Phi) - 1.175\, cos(5\Phi) \tag{7}$$

An example of the sent timestamp with the exact position in longitude and latitude coordinates and the timestamp by the three sensor units (AED 1 – AED 3) and calculated source of a detected acoustic event is in Fig. 9.

```
--- AED 1 ---
Latitude: 50.0
Longitude: 14.00
Timemark [ms]: 5193
Last captured: 26 Dec 18 02:01:12. PM
-------------
--- AED 2 ---
Latitude: 50.01
Longitude: 14.0025
Timemark [ms]: 5193
Last captured: 26 Dec 18 02:02:01. PM
-------------
--- AED 3 ---
Latitude: 50.0
Longitude: 14.005
Timemark [ms]: 5203
Last captured: 26 Dec 18 02:00:54. PM
-------------
--- Source of acustic event ---
Latitude: 50.004866764791
Longitude: 14.002504288864
```

Fig. 9. An example of a message with calculated source of an acoustic event sent by the sensor unit.

A custom software packet for the remote unit PC to communicate with acoustic units and process sent data has been developed as a part of the acoustic detection system (Fig. 10).

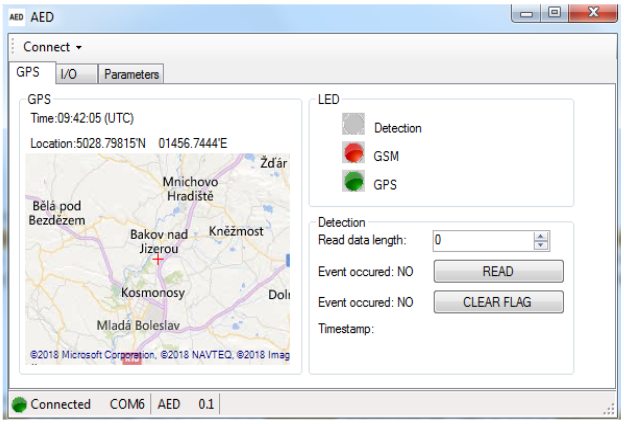

Fig. 10. PC application.

The application has been developed in C# programmable language. The program shows a map with the exact location of the chosen acoustic unit and precise/synchronized UTC time. In the case of detection, the application signalizes the event, displays an exact timestamp in UTC and stores the data for further processing. The remote central unit then applies the CW and MFC algorithms described in S*ection II* to obtain the features for classification. The application can also set the sampling frequency, parameters for the detection algorithm based on the median filter described in

S*ection II* [8], and the connection of additional detection units.

For classification, two independent Neural Networks (NN) was created. First, NN was designed to classify signals based on MFC. Considering the low dimensionality of these coefficients, a fully two-layer connected neural network has been used. For the CW algorithm, a convolution neural network has been used. As the spectrum is a two-dimensional matrix, the convolution network is significantly better than a fully connected system. Convolution network allows finding local features in multidimensional, location-dependent input data. Spectrogram features are location-dependent, and therefore, the convolution network has better results than a typical fully connected network. Also, the convolution network is typically much smaller in dimension, which significantly reducing the computation time. A single network, combining both designs, has been implemented and used for the classification. The network had two inputs, one for the MFC and one for the CW algorithm.

In this work, classification into 'false alarms' and into an individual caliber is presented. To train the classifiers in an optimal way, the recorded acoustic events data have to be divided into the two groups – to training and to validation sets. Limitations of overfitting problems, validation and training datasets have to be the same size. Recorded 'false alarms' and gunshots samples have been randomly chosen into the training and the validation sets. Each set had approx. 140 independent measurements. The experiment results for the validation data congaing 141 measured 'false alarms' and gunshots shows Tab. 1.

Table 1. Validation data classification results.

original/ predicted	false alarm	6.35 mm caliber	9 mm caliber	.22 caliber
false alarm	10	0	5	2
6.35 mm caliber	0	25	6	1
9 mm caliber	0	0	44	3
.22 caliber	0	1	0	44

The result shows the validation success rate of classifying into the false alarms and the gunshots classes, together with an exact caliber of a firearm class caliber, was nearly 90%. The system can classify the gunshot from the 'false alarm' with a 100% success rate but, on the other hand, approx. 40% of the false alarms are identified as a gunshot. The main reason for this is the limited number of signals corresponding to the false alarms. A significant number of impulse acoustic events similar to gunshots have to be recorded and used for classifier training.

From the tested measurements took in different shooting ranges, with at least three acoustic units to record the impulse acoustic even, the presented acoustic detector system, operating on the principle of the modified median filter, CW and MFC, can successfully detect and classify an acoustic event. Used NN classifier can classify an individual caliber of a used firearm with a 90 % success rate. As a next step, the System for Acoustic Detection will be deployed in an environment (residential area) to test it in real conditions.

IV. CONCLUSION

The system for acoustic detection, localization, and classification of a gun caliber has been presented. The system consists of sensor units that continuously monitor acoustic events around the unit and the remote unit. Sensor units uses a modified median filter algorithm to state if there is a possibility of a gunshot. The remote PC unit then evaluates the signal by advanced signal processing and classification in case of detection. Continuous Wavelet and Mel Frequency Transformation methods are used to get features for a neural network classifier.

Gunshots of different calibers and various false alarms similar to gunshots have been recorded on a shooting ranges to test the system. At least three sensor units recorded the acoustic events.

The system shows the ability to detect the gunshot with a 100% accuracy and to correct classify the caliber of a gun with approx. 90% accuracy. Considering the limited size of training dataset, such results are impressive. However, all measurements were measured in a similar environment, and a significantly larger dataset in a real environment, such urban areas, should be examined for future tests and unit improvements.

In the future, presented Systems for Acoustic Detection can be used as a standalone unit placed in schools, campuses, shopping centers or other public areas in general, to detect and to localize gunshot events and to increase the safety for the civil population.

ACKNOWLEDGEMENT

This research was supported by the "Energy for Smart Objects" grant provided by Electronic Components and Systems for European Leadership Joint Undertaking in collaboration with the European Union's H2020 Framework Programme (H2020/2014-2020) and National Authorities, under grant agreement n° 692482.

REFERENCES

[1] B. Kaushik, D. Nance, K. K. Ahuja "A Review of the Role of Acoustic Sensors in the Modern Battlefield", 11th AIAA/CEAS Aeroacoustics Conference, 2005.

[2] H. E. Bree, "The microflown. Journal of Clinical Monitoring and Computing" - J CLIN MONITOR COMPUT, dissertation thesis, 1997.

[3] SHOTSPOTTER, available at: http://www.shotspotter.com/, 2018.

[4] J. Millet and B. Baligand, "Latest Achievements in Gunfire Detection Systems", Battlefield Acoustic Sensing for ISR Applications (pp. 26-1–26-14)

NATO RTO-MP-SET-107, 2006.

[5] G. L. Duckworth, D.C. Gilbert, J.E. Barger, “Acoustic Counter-sniper System”, Command Control, Communications, and Intelligent Systems for Law Enforcement, SPIE Proceedings, Vol. 2938, 1996.

[6] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," 2006 IEEE 12th Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop, Teton National Park, WY, pp. 257-261, 2006.

[7] R. C. Maher, "Acoustical Characterization of Gunshots," 2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics, Washington, DC, USA, pp. 1-5, 2007.

[8] K. Łopatka, J. Kotus, A. Czyżewski, “Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations,” Multimedia Tools and Applications, vol. 429, 2015.

[9] J. Sallai, W. Hedgecock, P. Völgyesi, et al, “Weapon classification and shooter localization using distributed multichannel acoustic sensors,” Journal of Systems Architecture - Embedded Systems Design. Vol. 57, 2011.

[10] M. Lojka, M. Pleva, et al, “Efficient acoustic detector of gunshots and glass breaking,” Multimedia Tools and Application, vol. 75, 2016.

[11] J. Svatos, J. Holub, "Smart Acoustic Sensor," 5th International Forum on Research and Technologies for Society and Industry: Innovation to Shape the Future, pp. 161-165, 2019.

[12] K. Betke, “The NMEA 0183 Protocol,” URL http://www.tronico.fi/OH6NT/docs/NMEA0183.pdf, 2010.

[13] V. Fernandes, L. Mascarehnas, et al, "Speech Emotion Recognition using Mel Frequency Cepstral Coefficient and SVM Classifier," 2018 International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, pp. 200-204, 2018.

[14] A. D. P. Ramirez, J. I. de la Rosa Vargas, et al, "A comparative between Mel Frequency Cepstral Coefficients (MFCC) and Inverse Mel Frequency Cepstral Coefficients (IMFCC) features for an Automatic Bird Species Recognition System," 2018 IEEE Latin American Conference on Computational Intelligence (LA-CCI), Gudalajara, Mexico, pp. 1-4, 2018.

[15] Mathworks, “Matlab documentation,” URL https://www.mathworks.com/help/wavelet/ref/cwtft.html, 2020.

[16] M. F. Ghazali, A. M. Hamat, et al, “Irregularity detection in artificial signal using time-frequency analysis,” Journal of Engineering and Applied Sciences, vol. 11, pp. 3593-3597, 2016.

[17] F. Ivis, “Calculating Geographic Distance: Concepts and Methods”, NESUG 2006, Data Manipulation, Canada, 2006. *Available at* https://www.lexjansen.com/nesug/nesug06/dm/da15.pdf