

# Considerations on Developing a Chainsaw Intrusion Detection and Localization System for Preventing Unauthorized Logging

Valentin Andrei, Horia Cucu, Lucian Petrică

Speech and Dialogue Research Laboratory, University “Politehnica” of Bucharest, Bucharest, Romania

## Email address:

horia.cucu@upb.ro (H. Cucu)

## To cite this article:

Valentin Andrei, Horia Cucu, Lucian Petrică. Considerations on Developing a Chainsaw Intrusion Detection and Localization System for Preventing Unauthorized Logging. *Journal of Electrical and Electronic Engineering*. Vol. 3, No. 6, 2015, pp. 202-207.

doi: 10.11648/j.jeee.20150306.15

---

**Abstract:** This work presents a system designed to prevent unauthorized logging by detecting and locating chainsaw sound sources. We analyze the specifics of chainsaw related sounds and discuss about the possible approaches for classifying the input sounds. The work also highlights several approaches for sound source localization that can be used in wireless sensor network architecture for tracking the assumed intruders. Finally we describe the architecture of the system and discuss on how our approach is designed to be scalable, fail-safe and cost effective.

**Keywords:** Chainsaw Detection, Sound Source Localization, Sound Recognition

---

## 1. Introduction

Unauthorized logging is an issue that can have serious consequences over the surrounding environment and over the human settlements nearby an affected location. It has disastrous effects ranging from destroying the habitat of valuable wildlife to serious landslides that can result in human casualties. It is an issue that needs to be treated with maximum strictness. Often unauthorized logging is very difficult to prevent due to the lack of funding or inadequate surveillance methods. For example, using classical patrols is often no longer an option because the intruders that are not authorized to cut trees are very difficult to be tracked, especially in wide domains.

Over the years various automated surveillance solutions mostly based on Wireless Sensor Networks (WSNs) have been implemented to aid the authorities to detect intrusions in a forest domain and to track the intruders. The most efficient techniques rely on identifying a sound that is likely to be produced by a chainsaw and using microphone arrays for locating the source. For example, in [1] a light-weight approach to chainsaw identification is proposed, based on the auto-correlation of 256-sample blocks of audio signal. The authors refer to this method as Lightweight Acoustic Detection (LAC). Signal energy is utilized, as well as pitch

and pitch stability measures computed from the autocorrelation of the signal. These measures are chosen because chainsaws are much noisier than forest ambient sounds, have distinctive pitch signatures and have good pitch stability. LAC executes in real-time on a WSN sensor node based on the ATmega128RFA1 microcontroller. The reported accuracy of detection is up to 85%.

In [2] and [3] an algorithm called Normalized Peak Domination Ratio (NPDR) is utilized. Based solely on the signal spectrum and noise energy, NPDR looks for an overlap between the signal and pre-computed reference peaks in the spectrum, and also for a sufficiently high concentration of signal energy in the spectral vicinity of the reference peaks. Evaluation of NPDR on 1024-sample blocks of signal indicates over 99% accuracy in quiet room conditions on a range of sounds. NPDR runs on a PC and the authors indicate (but do not demonstrate) the possibility of executing NPDR on smartphones. Between the two algorithms, NPDR has better apparent accuracy and LAC a slightly simpler structure which may result in a more efficient implementation.

Other automatic surveillance solutions are strictly based on satellite image processing [4, 5] and local video surveillance [6, 7]. These solutions are focused and efficient on detecting forest fires or estimating fire risk, but they have important short-comings when it comes to (illegal) logging. The satellite signal based methods do not support real-time operation in

general, and also the spatial resolution is very limited. Conventional video surveillance suffers from the following:

- If the viewing angle of the lens is wide, then the resolution is low; otherwise the monitored area is too small.
- The analysis of data requires high complexity solutions.
- The required bandwidth for transmitting images is relatively high, while the energy consumption of such devices is also relatively large.

In the current work we describe implementation details of a system designed to detect and locate chainsaws in a forest domain under surveillance. The following section concentrates on discussing the specificities of the sounds produced by chainsaws and comments on the feature extraction approaches. Section 3 is dedicated to reviewing several sound localization techniques that can be used in the sensor elements of the proposed system. The architecture of the system is proposed in section 4. We discuss about the key features of the proposed system and about strategies to ensure a high level of quality of service. Finally the last section is dedicated to conclusions and future work.

## 2. Detection of Chainsaw Related Sounds

The first challenge to consider in developing the proposed system is the ability to recognize from a multitude of sounds, the specific ones produced by a chainsaw. In a forest there are numerous sound sources such as the noise caused by the tree leaves in windy conditions, the sounds caused by birds or animals, walking related sound, car related sounds, sounds produced by tourists who have authorized presence and plenty of other sources. The system should not fire the alarm for example when a car passes by. This would generate high usage costs and in the end would prove to be an unreliable solution.

Detecting a sound produced by a chainsaw is a task slightly more flexible than, for example, recognizing spoken words or speakers. A much greater level of detail needs to be captured by the sound analysis algorithms in order to recognize speech or speakers, because each voice is unique and each word produced by a certain voice can be considered an unrepeatable sample. We state that detecting a chainsaw is a flexible task because *we want to determine that a captured sound belongs to a certain class* – in our case the class of sounds produced by chainsaws. With the sounds being produced by a mechanical system the task becomes relatively easier. This is because mechanical systems like engines have an extremely good periodicity rate and therefore are likely to produce extremely repeatable sounds. Let's analyze how a signal produced by a chainsaw looks like. In Figure 1 we plotted the power spectral density computed using the Welch method for 4 signals captured from 4 different chainsaw types in different recording conditions. We want to observe if we detect a similarity pattern between the captured samples. If the pattern is detected, even in different recording conditions, we are more confident that detecting chainsaw presence based on the sound footprint is feasible.

We can observe that in the figure the spectrum shape obtained for the four types of chainsaws is extremely similar. In all quadrants the power per frequency step is higher towards the lower frequencies and decreases almost linearly towards the higher frequencies. This is an easily detectable pattern. We also speculate that using classical approaches for detecting if a sound belongs to a specific class can *fade the similarity* observed in Figure 1.

In [11], the authors collected a database of 10 chainsaw produced sounds. The results we obtained are in good correlation with data presented in [11].

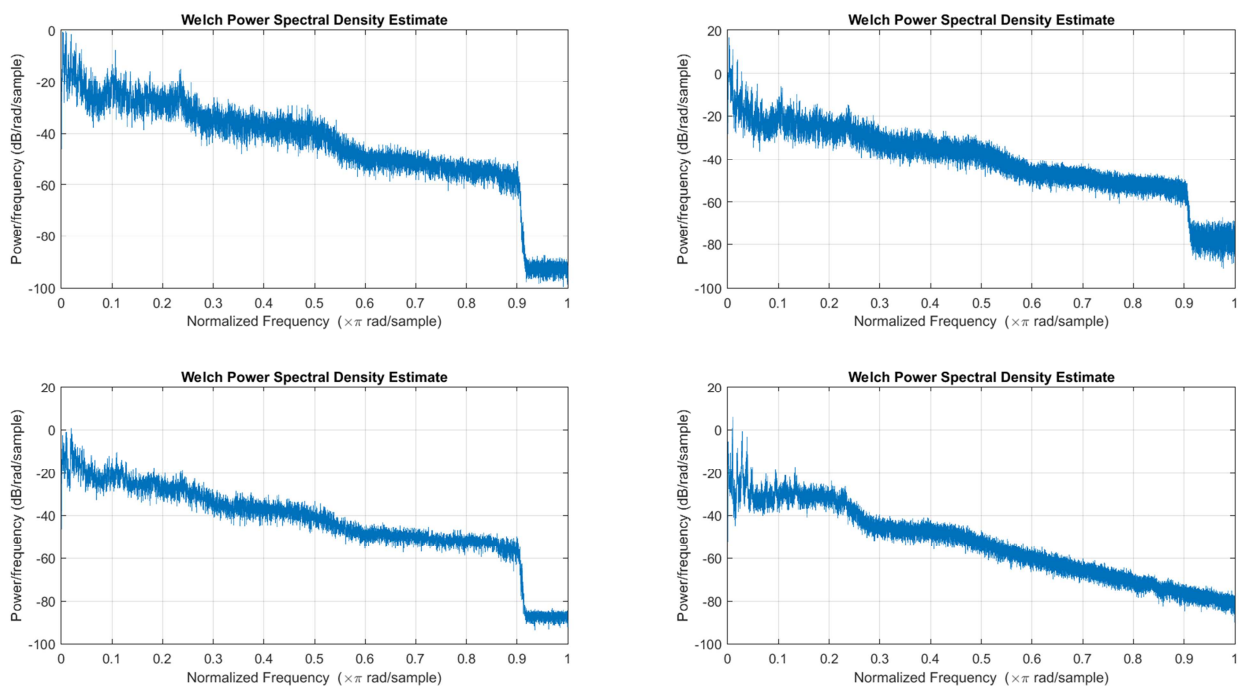
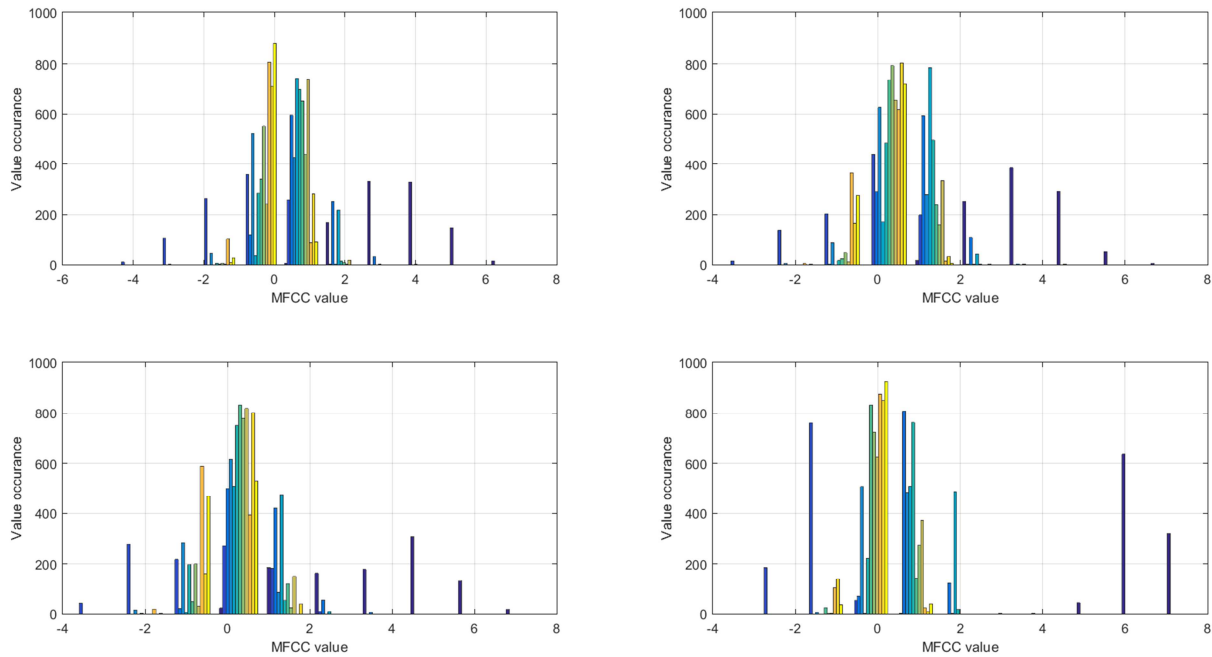
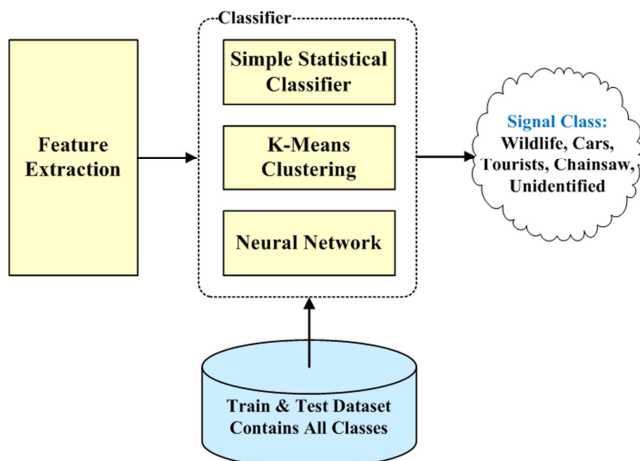


Figure 1. Example of PSD for chainsaw produced sounds.



**Figure 2.** The usage of MFC coefficients for detecting chainsaw related sounds.

In Figure 2 we computed the MFC coefficients for the four signals produced by chainsaws, also displayed in Figure 1. In order to verify if we can observe similar trends between the four sounds, when analyzing MFC coefficients, we computed for each sound a histogram that creates value bins and counts the occurrence frequencies for each bin. We can observe that while the values seem centered around zero for all the recordings, the similarity between the histograms is not as obvious as when analyzing the signals in frequency domain. Therefore we speculate that using complex techniques for extracting features from chainsaw produced sounds may not always yield promising results.



**Figure 3.** Chainsaw detection approach.

We used MFCC for this experiment as it is an extremely widely used technique adopted in systems for recognizing various sounds. Our results are somewhat in contrast with the work described by [8] and [9] which uses MFCC. However we firmly state that this approach is not necessarily suitable

for mechanical produced sounds as it was designed for speech related sounds. For example, in [10] the same authors propose a methodology of feature extraction, based on TESPAP which extracts simple metrics from the waveform and is stated to produce better classification results.

Considering the comments stated above, the general diagram for classifying sounds into classes and marking the ones that are likely to be produced by chainsaws, is presented in Figure 3. As stated we consider that the feature extraction algorithms should be fairly simple due to the specifics of the chainsaw produced sounds (e.g. compute a linear regression on the points that describe the spectrum and compute the correlation error as input parameter). A simple threshold base statistical classifier can be used to separate input sounds into classes or a more complex approach can be designed, using K-Means clustering. We recommend using a neural network only in conditions where chainsaw sounds are demonstrated as being rather complex therefore implying the use of feature extraction methods with wide feature sets. In this case a neural network is far better at learning the links between the feature sets than a simple statistical classifier and a K-Means Clustering algorithm. We estimate that the chainsaw detection approach should tag the recorded signals using a set of commonly met classes of signals for the desired use case. In a forest we can consider the following classes: wildlife, tourists, passing cars (in the case where there is a paved road through the forest) and finally we can consider a generic class: “unidentified”.

### 3. Sound Source Localization Challenges

Sound source localization in wireless sensor networks has been achieved for single-microphone sensor nodes in previous work through the use of distributed time difference

of arrival (D-TDOA). In D-TDOA, nodes synchronize their internal clocks and listen for a reference sound at a pre-determined moment in time. Because it has already been established through synchronization that the sampling is simultaneous, any difference in audio signal phase at the sensing nodes (SNs) is caused by differences in relative position of the sound source to the respective SNs. If the geometry of the WSN is known a priori, then the individual measurements are aggregated to determine the position of the sound source relative to the WSN. D-TDOA requires that time re-synchronization, hence radio communication, must occur before each SSL event to compensate for drift in SN timers. Following signal acquisition, the data must be aggregated for the SSL to be computed. The WSN incurs an energy consumption penalty for these communication events.

A different approach to sound source localization has been explored in [12] utilizing microphone arrays. We refer to this method as Array TDOA (A-TDOA). In A-TDOA, each SN is equipped with an array of microphones arranged in a fixed geometry. To obtain positioning in a 2D plane, a planar symmetrical geometry is best suited, such as placing the microphones on a circle. Delay-and-Sum (DS) [13] is the algorithm utilized in [12] for sound direction estimation, although more sophisticated algorithms exist [14]. The multi-microphone data gathering and processing in A-TDOA is above the capabilities of a microprocessor and needs either a

DSP or a FPGA to be added to the system.

Two or more A-TDOA direction measurements by separate WSN nodes may be aggregated as in [12] in order to increase the localization accuracy if needed. A single SN can only determine through A-TDOA the probability that a sound source is located in a certain direction relative to the SN. Adjacent SNs with overlapping sensing ranges may superimpose their localization, resulting in a more accurate localization. The cost of the increased localization accuracy is energy expended for SN communication.

It must be noted that A-TDOA can only determine the direction of a sound source relative to the SNs microphone array. Therefore, for A-TDOA there must also be some way for the SN to determine the absolute spatial orientation of the microphone array.

Localization by A-TDOA avoids the need for WSN time synchronization. Conversely, on-node computation is required for A-TDOA to determine the sound direction. Both A- and D-TDOA require data aggregation between nodes to determine the sound source accurately, however A-TDOA is capable of obtaining a certain measure of localization without SN communication. In some cases, knowing the general direction of the sound source is sufficient to discriminate between legal and illegal deforestation activity (e.g., at the edges of a forest or natural reserve).

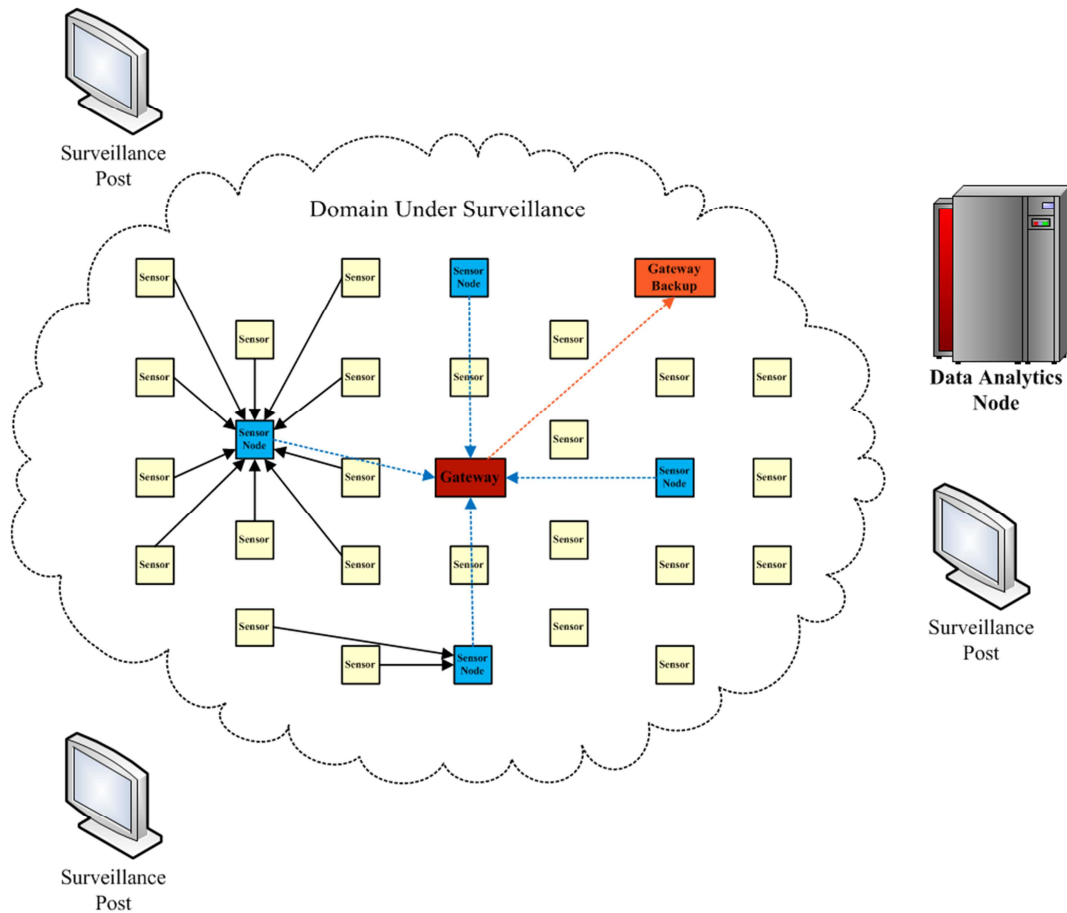


Figure 4. Intrusion localization system diagram.

## 4. System Design and Reliability Considerations

We propose a flexible design for detecting chainsaw related sounds in a monitored area. We consider that such a system needs to have the below features:

- *Scalability* – Extending the monitored domain needs to be done without involving redesign costs.
- *Fail-Safe* – Because it contains multiple components it should expose redundancy to ensure that in the case one component fails, the system will not fail completely.
- *Cost-Effective* – We state that the investment needed for such a system needs to have a reasonable return. Technically this involves using generally available hardware with low maintenance and support costs. Additionally the amount of energy consumed by the system needs to be kept at a reasonably low level.

The diagram of the proposed system is presented in Figure 4. It can be viewed as a distributed architecture, similar with the architecture of a wide area cellular network. Let's describe the features of each component.

### 4.1. Sensors

The basic element of the design is the sensor. However this component is more complex than a plain sensor. In our case, the sensor can have the logic for capturing the audio signal, for estimating the location of the source and for performing basic processing over the collected input. In Figure 3 we illustrated the architecture of the sound recognition system. We estimate that basic recognition using either a simple set of statistical thresholds or even a k-means clustering approach can be performed at the sensor level. Of course even more processing can be done, but that would greatly increase the energy footprint of the proposed system.

### 4.2. Sensor Node

The sensor node acts like a hub, connecting multiple simple sensors. Its main functionality is to route the acquisition data towards the gateway. Additionally it can be designed to contribute to the processing especially in the case where the simple sensors do not contain complex processing logic. Such enhancements could be related to the preprocessing of the captured sound signal, like filtering or even feature extraction.

### 4.3. The Gateways

The gateways have a central role in the architecture of the proposed system. First, we propose at least 2 gateways. If one should fail there should always exist a backup. The gateway number can increase as the domain under surveillance grows larger.

At the gateway level we can consider more complex

processing, for example even adding a neural network for labeling sounds. Unlike the sensors and the sensors nodes, the gateway needs to have a more powerful processor because it is responsible for processing the information from a high number of sensors. The computational power demand is expected to grow seriously if the lower level components do not contain complex processing.

### 4.4. Data Analytics and Surveillance Posts

The final level is represented by the data-analytics node. This component will provide the ultimate decision to the surveillance post. For example the lower levels like the sensor, the sensor node or the gateway can predict the occurrence of an intruder, but the data analytics node has the final decision of raising the intrusion alarm. This can be accomplished by storing information related to the terrain where the chainsaw sound source was located. If the data analytics node has a detailed map of the domain under surveillance it can assign intrusion likelihoods to each spatial division of the domain and therefore if the intrusion is estimated to be in a highly inaccessible area (without having detected „unknown” sounds a priori) the alarm will be raised with a lower probability. The data analytics level needs to be powered by a server system.

## 5. Conclusions

We discussed in this paper about several aspects that need to be considered when designing a WSN solution for chainsaw intrusion detection. We analyzed that the spectrum of a chainsaw produced sound has a shape that can be extracted with fairly simple feature extraction techniques and we also illustrated that there is a high degree of similarity between spectra associated to different types of chainsaws. Nevertheless because we are talking about a mechanically produced sound, it has a high degree of periodicity. We also illustrated that the usage of complex feature extractions, like MFCC can visually lower the similarity between feature sets. We also mentioned several sound localization techniques, like A-TDOA and D-TDOA that can be used in our setup. Finally we proposed a WSN architecture for chainsaw intrusion detection that can be scaled and is equipped with fail safe mechanisms. We also discussed about techniques to distribute the processing load in order to minimize the energy footprint of the system.

## Acknowledgement

This work was supported in part by the Sectoral Operational Programme "Human Resources Development" 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU /159/1.5/S/132397 and in part by the PN II Programme "Partnerships in priority areas" of MEN - UEFISCDI, through project no. 32/2014.

---

## References

- [1] L. Czúni, P. Varga, "Lightweight acoustic detection of logging in wireless sensor networks", in Proc. DINWC 2014, pp. 120-125, 2014.
- [2] M. Babis, M. Duricek, V. Harvanova, M. Vojtko, "Forest guardian – monitoring system for detecting logging activities based on sound recognition", Proc. of IIT. SRC, pp. 1-6, 2011.
- [3] I. C. Yoo, D. Yook, "Automatic Sound Recognition for the Hearing Impaired", IEEE Transactions on Consumer Electronics, vol. 54, no. 4, pp. 2029-2036, 2008.
- [4] Q. Guangmeng, Z. Mei, "Using MODIS land surface temperature to evaluate forest fire risk of northeast China", IEEE Geoscience and Remote Sensing Letters, vol. 1, issue 2, pp. 98-100, 2004.
- [5] K. Kyzirakos, M. Karpathiotakis, G. Garbis, C. Nikolaou, K. Bereta, I. Papoutsis, C. Kontoes, "Wildfire monitoring using satellite images, ontologies and linked geospatial data", Web Semantics: Science, Services and Agents on the World Wide Web.
- [6] Z. Xiong, R. Caballero, H. Wang, A. M. Finn, M. A. Lelic, P. Y. Peng, "Video-based smoke detection: possibilities, techniques, and challenges", in IFPA, fire suppression and detection research and applications - a technical working conference (SUPDET), Orlando, USA.
- [7] Bo Lei; Zhijie Zhang; Song Yue and Chensheng Wang, "Forest smoke detection in video images based on constant speed rotating platform," in Proc. ISPD 2013 - Fifth International Symposium on Photoelectronic Detection and Imaging, 2013.
- [8] G. Oltean, L. Grama, L. Ivanciu, C. Rusu, "Alarming events detection based on audio signals recognition", Proc. of 8<sup>th</sup> Conference on Speech Technology and Human Computer Dialogue, pp. 83-90, 2015.
- [9] M. Ghiurcau, C. Rusu, R. C. Bilcu, J. Astola, "Audio based solutions for detecting intruders in wild areas", Journal of Signal Processing, vol. 92, Elsevier, pp. 829-840, 2012.
- [10] L. Grama, C. Rusu, G. Oltean, L. Ivanciu, "Quantization effects on audio signals for detecting intruders in wild areas using TESPAS S-Matrix and Artificial Neural Networks", Proc. of 8<sup>th</sup> Conference on Speech Technology and Human Computer Dialogue, pp. 1-6, 2015.
- [11] L. Tudor, V. Zoicas, L. Grama, C. Rusu, "The SPG (Signal Processing Group) Sound Database", Novice Insights in Electronics, Communications and Information Technology, Issue 15, pp. 62-65, 2014.
- [12] J. Tiete, F. Dominguez, et. al. "Sound Compass: a distributed MEMS microphone array-based sensor for sound source localization", MDPI Sensors, vol. 14, no. 2, pp. 1918-1949, 2014.
- [13] A. Greensted, "Delay Sum Beamforming", The Lab Book Pages, October 2012.
- [14] H. Do, H. Silverman, Y. Yu, "A real-time SPR-PHAT source location implementation using stochastic region contraction (src) on a large-aperture microphone array", in Proc. ICASSP, pp. 121-124, 2007.