Proceedings of the Annual Symposium of the Institute of Solid Mechanics
and Session of the Commission of Acoustics, SISOM 2018 Bucharest 24-25 May

ENVIRONMENTAL NOISE CLASSIFICATION USING THE GMM APPROACH

Elena OLTEANU[1,2], Svetlana SEGARCEANU[1,2], Inge GAVAT[2]

[1]Beia Consult International, [2]Department of applied Electronics and information engineering, Politehnica University of Bucharest, svet_segarceanu@yahoo.com, elena.olteanu@beia.ro

Environmental noise recognition is becoming doubtlessly an essential component of computer science and robotics as it simulates the important function of human hearing. It has applications in fields such as industry, urban or ecology monitoring solutions, allowing to identify flaws or threats in industrial processes or ecological aspects. We propose, and experiment several methods, based on the GMM approach, for sound recognition. The first approach applies the sheer GMM to several feature sets, including LPC and Mel-cepstral features, for closed-set identification. The second one is GMM-UBM, applied on the same types of features, to carry out open-set sound identification. It is intended to consolidate the recognition process by classifying uninteresting sounds outside the significant classes. We considered three types of noise: chainsaw, vehicles, and typical forest noise but intend to extend our research to more categories and to design a system meant to detect environmental threats.
Keywords: environmental sounds recognition, GMM, GMM-UBM, open-set sound identification

1. INTRODUCTION

Currently, environmental noise recognition is receiving more and more attention, as an essential component of computer science fields, or robotics, as it simulates the important function of human hearing, and moreover, is intended to overpower the human perception. Detecting footsteps behind you, rain drops through the window, creaking or slamming doors are some other examples of what means understanding the environmental sounds. For a robot, audio recordings may provide important information or hints about location, or direction of a moving vehicle, or environmental information, such as speed of wind. An important nowadays application concerns modern hearing aids [1, 2] which incorporate several programs that account for environmental models such as meeting rooms, auditoriums and other noisy environments. Automatic recognition of the surrounding environment allows devices to switch between tasks with minimum user interference. Another important application is the automatic detection of environment endangering factors. Such an example is illegal logging, which became a major global problem. Solutions based on monitoring systems represent an efficient means to cope with this issue and help protecting public and private forests at the request of the owner [3]. Transmission of data can be done within a short time interval and permit the authorities to intervene in a timely manner to reduce illegal deforestation.

Our paper investigates, in the context of a forest monitoring system, several approaches based on Gaussian Mixtures Modelling (GMM) to differentiate some classes of specific sounds. Investigation of these methods is accompanied by experimental results. Our goal is to devise an audio signal identification solution to be incorporated into a monitoring system. The proposed monitoring system has a built-in microphone for sound signal recording. The system is not completed but several components of the approach, such as acoustic signal processing, are already implemented. It should send the information as an alarm directly to the owners of forests with possible threats, giving them the opportunity to take appropriate measures in real time. The main obstacle in testing the proposed solution is the scarcity of relevant audio material.

The outline of the article is as follows. The next section presents the state-of-the art in environmental noise detection. Section 3 reviews several techniques based on the GMM methodology, to be applied in environmental noise recognition. Section 4 describes the experiments setup and results obtained by applying the methods described in Chapter 3. Section 5 presents the conclusions and future research directions.

2. STATE-OF-THE-ART IN ENVIRONMENTAL NOISE RECOGNITION

Several attempts to classify environment sounds have given rise to new sets of features and classification approaches. L. Grama, C. Rusu and E.R. Buhuș [4] investigated some classification algorithms to determine the effect of impulsive sounds (gunshots) and evaluate different types of features (such as LPC, prediction error variance). Two scenarios were conducted; the first used five different sound classes (birds, chainsaws, tractors, human voices, and gunshots) while scenario 2, was based on only the first four sound classes. The gunshot class was eliminated to check the influence of impulsive signals on the overall classifier accuracy, and it came out that the gunshot class alters the classification accuracy when increasing LPC order.

The same authors present an audio signal classification system, involving low computational costs, using a signal classifier based on LPC and Random Forests. The signals that can be detected include: birds, gunshots, chainsaws, human voices, tractors. The identification rate reported by the paper is 99.25% [5].

In [6] the authors make an exhaustive evaluation of several techniques to be used at feature extraction, modelling and classification levels. The paper considers time-domain and frequency-domain types of features, perceptual or not, ranging from zero-crossing rate to features derived from LPC, Mel scale analysis or Bark scale analysis, these features are rated by their mutual similarity and redundancy. Most of the features are the same as those used in speech processing, apart from the amplitude descriptor, applied mainly in animal sound detection. As classification approaches, the authors have tested the k-NN, LVQ and SV|M methods. The feature vectors are combinations of some of the above features. The best results are obtained using the k-NN technique. Another approach [1] proposes a combination of simple time and frequency features modelled by Neural Networks and HMM. The process includes two stages: A first sound classification into one of the sound categories such as collision sounds, friction sounds, vibration sounds, electric sound, and other noises. Next, a second classification process takes place inside the category previously detected.

Concerning monitoring solutions, most of the developed monitoring systems use wireless sensor network. As Forest Guardian, a system meant to spot specific sounds of logging [7]. J. Papán, M. Jurecka, J. Púchyová also propose a wireless sensor network to prevent illegal logging using acoustic signal evaluation and the principles of network nodes communication [8]

3. ACOUSTICAL SOUND PROCESSING

The final goal of the audio signal processing is the correct identification of an unknown acoustic sound. The correct identification process involves either the proper assignation of the unknown sound into one of the considered classes of sounds or possibly placing it outside any of the considered classes. In the first case sound identification can be regarded as closed-set identification, in the second case as open-set identification process. We remind the fact that Open-Set Identification is based on the result of Closed Set Sound Identification, where the identified (type of) sound is further verified by comparing its score value with a specific threshold. Figure 1 presents the open-set identification process.

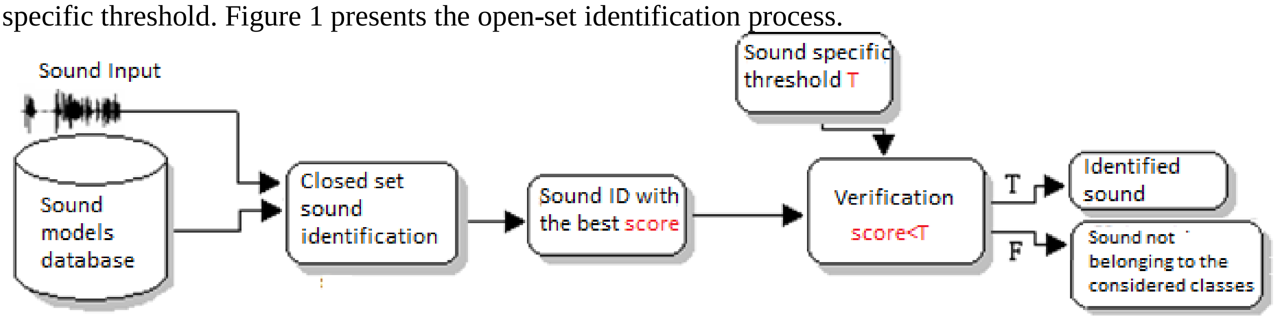

Figure 1 - Open-Set Sound Identification decision logic

Closed-Set sound identification itself involves acoustic signal processing in the training and recognition phases, feature space modelling at training, and classification at recognition stage. Next we will detail the procedures applied in these phases.

3.1. Acoustic Signal pre-processing

In the pre-processing phase the signal is divided into frames of around 25ms, with a 10ms overlap. This means, for an audio file at a sampling frequency of 44.1 kHz, about 1102 samples per frame. The length of the FFT analysis was the set as the closest power of 2 exceeding the number of frame samples, so in the case above it was set to 2048. Signal pre-emphasis was applied using a pre-emphasis filter $H(z) = 1 - a * z^{-1}$ with α = 0.97. No frames were eliminated. Although the length of the analysis frame is a theme of debate and polemic, most approaches to environmental sound processing use signal frames of lengths between 11.5ms and 25ms. A more thorough study of the spectral properties of this type of signals would make a consistent basis for setting the length of the analysis frame. Figures 2a and 2b present the overall image of spectra for chainsaw, and vehicle signals respectively. As can be inferred from these images the useful frequency domain is situated beneath 8kHz, while the Fourier transform generates a much wider domain for the given sampling frequency. Figures 3a and 3b present the sub-spectra of the same two signals for frequency values under 8kHz.

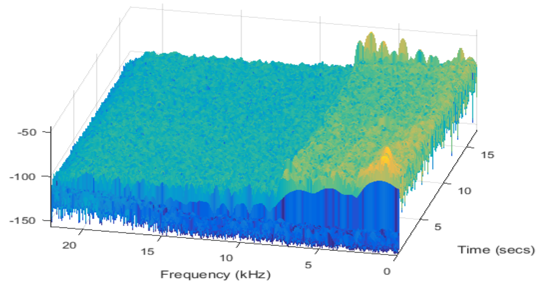

Figure 2a - Spectrum of a chainsaw recording of 15s

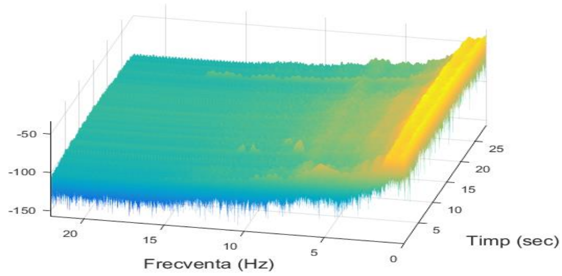

Figure 2b - Spectrum of a vehicle sound of 25s

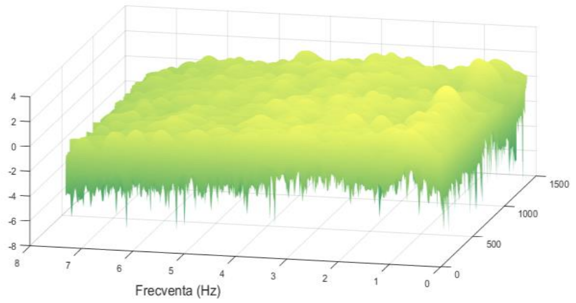

Figure 3a - Sub-Spectrum of a chainsaw sound for frequencies under 8kHz

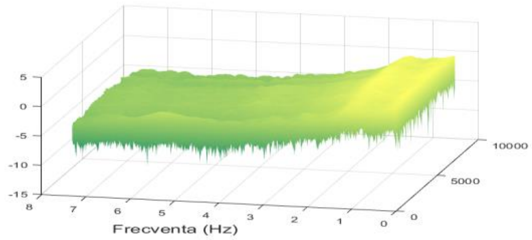

Figure 2b - Sub-Spectrum of a vehicle sound for frequencies under 8kHz

3.2. Feature space generation

For feature space generation we have considered the following types of features:

- 16 LPC order cepstral features [9]
- 14 Mel cepstral features [10]
- Zero crossing rate [6]

For the Mel-scale analysis the frequency domain was set to [0.3kHz, 3.7kHz].

3.3. Feature space modelling using GMM

GMM is a typical variant of probabilistic clustering, which generates a covering of the feature space by K Gaussian components f_k, expressed by [11]:

$$f(x) = \sum_{k=1}^{K} c_k f_k(x) = \sum_{k=1}^{K} c_k \phi(x/\mu_k, \Sigma_k) \tag{1}$$

$$f_k(x) = \phi(x/\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{d}{2}} \Sigma_k} \exp\left[-\frac{1}{2}(x-\Sigma_k)^T \Sigma^{-1} (x-\Sigma_k)\right], \quad \sum_{k=1}^{K} c_k = 1, \quad 0 \le c_k \le 1, \quad 1 \le k \le K \tag{2}$$

The elements of a Gaussian Mixture Model, for an established number of components, the means, standard deviations and weights are calculated using the EM algorithm. An important step of this algorithm is the initialization step, where the values of these parameters are set to some initial values. The system performance depends strongly on these values, moreover certain components cannot be generated due to unproperly chosen initial values. An efficient and well-balanced initial configuration is given by applying in the initialization step certain hierarchical agglomerative clustering approaches. Such approaches are [12, 13]:

- Complete linkage
- Average linkage
- Simple linkage
- Weighted linkage – variant of Average linkage
- Pairwise Nearest Neighbours (PNN)

Applying such a method generates a tree called dendrogram, whose branches gather representatives belonging to the same class. Figure 4 presents a dendrogram obtained by applying one of the above methods on a feature space We have used the GMM approach in several variants:

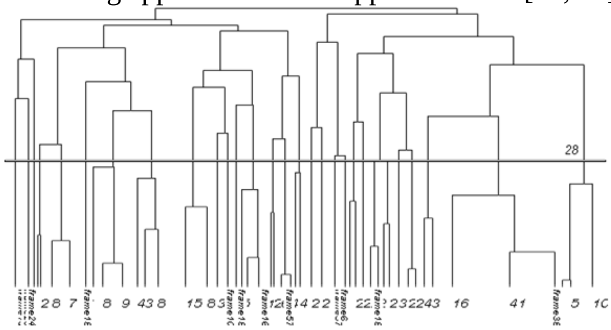

Figure 4 - Dendrogram obtained by applying a hierarchical clustering method

First, in the case of closed set sound identification, the membership of an unknown sample sequence X = {x_1, x_2,., ,x_T} to a certain class of sounds λ, among C classes $\lambda \in \{\lambda_1, \lambda_2, .., \lambda_C\}$ is established by applying the formula:

$$\lambda_S = \underset{\lambda \in \{\lambda 1, \lambda 2, \ldots, \lambda C\}}{\operatorname{argmax}} \left(\log(p(X, \lambda))\right) \tag{3}$$

where

$$\log(p(X,\lambda)) = \sum_{t=1}^{T} \log(p(x_t,\lambda)) \quad \text{and} \quad p(x_t,\lambda) = \sum_{k=1}^{K} c_k \phi(x/\mu_k, \Sigma_k) \tag{4}$$

For open-set sound identification two approaches were investigated:

The first approach considers a new class, besides the C classes mentioned before, accounting for the sounds that do not belong to any of the C classes. The method is based on the same formula as above, except for the fact that C+1 acoustic models are considered.

The second method uses the GMM-UBM approach where besides the sound models, a model of alternative classes, or a universal alternative class model are considered. The alternative model of a class or for all the classes, denoted by λUBM, is also a mixture of Gaussian models, but with a specific number of components, and trained on alternative sounds to those belonging to the identified class model, λS. After applying the closed set identification as in relations (3) the system verifies if the score calculated in (5) is positive.

$$\log(p(X, \lambda_S)) - \log(p(X, \lambda_{UBM})) - \theta \tag{5}$$

3.4. System Performance Evaluation

Evaluation of a closed-set Identification system is usually expressed by the identification rate (IR), which counts the percent of correctly classified sounds. It can be more thoroughly evaluated by the confusion errors, placed in confusion matrices, accounting for all the cases where the sound is identified as belonging to each class of sounds.

Upon the evaluation of an open set identification system three types of errors may be encountered:

- *Identification rate,* as above is expected to be smaller than the one achieved closed-set identification

- *False acceptance* (FA) occurs when an audio signal outside the specified classes is wrongly identified as belonging to one type of signals in the system
- *False rejection* (FR) is met when an audio signal is erroneously classified as not belonging to the group of target sound classes.
- *Confusion errors* count the cases where the sound is wrongly identified is as belonging to other classes.

4. EXPERIMENTS ORGANIZATION AND RESULTS

4.1. Acoustic material

Currently, we dispose of several recordings representing forest sounds:

- 18 chainsaw recordings (10 used for training, 8 for testing)
- 28 various vehicle sounds (12 for training, 16 for testing)
- 30 forest specific sounds (10 for training, 20 for testing)
- 28 recordings of several environmental sounds, many of them taken from urban environment; a part of these are supposed to contribute to set the background model; the rest will be used to evaluate the system, especially for assessing the False Acceptance Rate. Among them “Horse gallop on dirt constantly then stop”, “Door-Wood-Sauna-Open-Confident”, “Footsteps on hardwood in bare feet with shifting and steps”, “Wooden creaky door”, “Wiretap”.

All the recordings are acquired from Internet. Most of these recordings are sampled at 44.1 kHz and recorded using 2 channels. All of them use a 16 bits representation. Other sapling frequencies are 48kHz, or 96kHz. All the recordings last more than 5 s. As most of the recording contained two versions recorded on two channels we used separately each of them for training and testing the system as well. So, for instance the chainsaw test material, contained 3 files recorded on 1 channel and 5 on 2 channels, we have tested the chainsaw case on 13 recordings.

4.2. Training

For acoustic classes modelling we used the following alternatives concerning frame features

- 14 cepstral features based on 16 order LPC
- 14 Mel Cepstral coefficients and the zero-crossing rate on each frame
- 14 Mel cepstral coefficients.

In the modelling stage we applied the GMM modelling for each type of sounds. For closed-set identification we considered 3 classes of sounds:

- Unaltered typical forest noise (birds, wind, falling leaves, leaves rustle);
- Vehicles: tractor, ATV, other utility vehicles;
- Specific grubbing tools: chainsaw, saw, ax, hatchet;

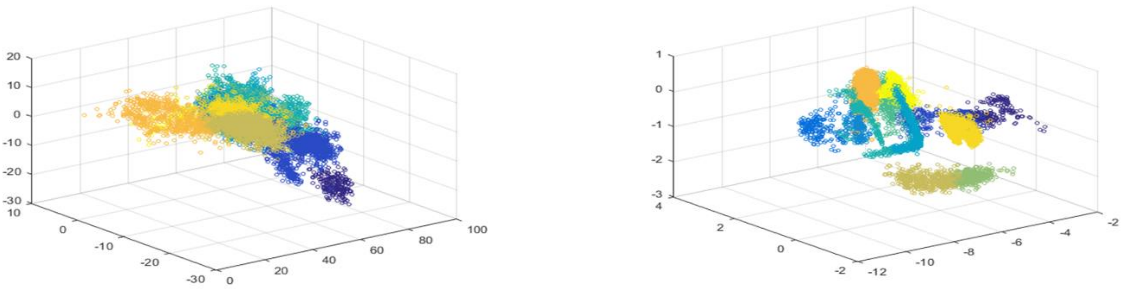

Figure 5 - Initial partition of the Mel-cepstral feature space (left) and LPC cepstral feature space (right) obtained by applying PNN

The number of components varied from 8 to 14. These methods were implemented in Matlab. We used the hierarchical clustering approaches provided by Matlab [13], among them those mentioned in Subsection

3.2. Figures 5 present initial partitions of the Mel-cepstral feature space and LPC cepstral feature space respectively, dimensions 3, 4, 5, obtained by applying PNN. Moreover, each of these techniques can be configured to use different distance measures, apart from the usual Euclidian distance:

- Minkovski distance, parametrized by *p*, where the default power is *p=2*
- Seuclidian distance - Standardized Euclidean distance. Each coordinate difference between observations is scaled by dividing by the corresponding element of the standard deviation,
- Chebychev distance (maximum coordinate difference).
- Mahalanobis distance using the sample covariance of the sequence.

The main objectives of the tests were to

- Evaluate several initialization methods using various distances
- Evaluate several combinations of features
- Assess the optimum number of Gaussian components

Table 1 - Closed-set identification rates using 8 components GMM for modelling the feature space, consisting of vectors of 14 Mel features and zero-crossing rate, initialized by several combinations of hierarchical clustering methods and distance measures

		Distance				
		euclidian	minkovsky	chebychev	seuclidian	mahalanobis
Hierarchical clustering method	Ward (PNN)	65.88	65.88	72.94	72.94	
	Complete	64.70		63.52		-
	Average	63.52	63.52	65.88	-	-
	Weighted			61.17		

Closed-set identification rates obtained using 8 Gaussian components for modelling the feature space consisting of 14 Mel features and the zero-crossing rate are presented in table I. Several combinations of initializing hierarchical methods and distances were operated. The cells where no results are displayed indicate that the Gaussian Mixture was not generated because of scarce data.

		Distance			
		euclidian	minkovsky	chebychev	seuclidian
Number of Gaussian components	12	71.76	71.76	72.94	72.94
	10	72.94	72.94	70.58	70.58
	14	69.41	71.76	69.41	71.76
	8	65.88		71.76	70.58

Table 2 - Closed set identification rates (IR), using GMM with varying number of components for modelling the of 14-dimensional Mel features space, using PNN and several measures for initialization

Table 3 - Confusion matric for closed set identification using GMM for modelling the Mel - cepstral feature space

	Chainsaw	**Vehicle**	**Genuine forest**
Chainsaw	53.84	16.16	0
Vehicle	15.625	50	34.375
Genuine fprest	2.5	0	97.5

Table 4 - Confusion matric for closed set identification using GMM for modelling the LPC- cepstral feature space

	Chainsaw	**Vehicle**	**Genuine forest**
Chainsaw	61.53	7.69	30.76
Vehicle	9.375	31.25	59.37
Genuine forest	0	12.5	87.5

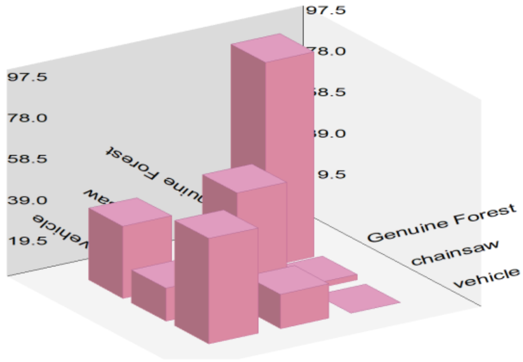

Figure 6a - Visual representation of confusion errors, obtained using Mel-cepstral features showing a significant confusion between vehicles and the other two classes and poor identification rates for chainsaw and vehicles

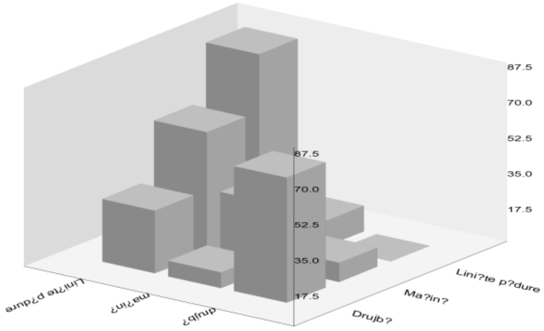

Figure 6b -Visual representation of confusion errors, obtained using LPC-cepstral features, showing a confusion rate between vehicles and the other two classes more significant that the identification rate for vehicles

As the most stable and best results were obtained using PNN as initialisation approach, we further investigated the impact of varying the number of Gaussian components on the identification rate, using PNN in the initialisation step of EM. Closed set identification rates using a varying number of Gaussian components in modelling the feature space consisting of only 14-dimensional Mel feature space and the PNN initialization method with several distances are presented in table II. The zero-crossing rate was excluded from the test to evaluate its impact on system accuracy. The best identification rate, 73% was obtained using 10 or 12 Gaussian components, the same as using the zero-crossing rate in addition to the Mel-cepstral coefficients and 8 Gaussian components. The benefit of using the zero-crossing rate for 8components was only of 1 percent. However, although the identification rate was not very bad the confusion matrices in table III show that the performance is very poor in what concerns the recognition of chainsaw and vehicle sounds. Table IV presents the confusion matrix obtained in similar experiments using the LPC cepstral features. The results are even worse, and the identification rate using LPC approach is about 65%.

The open-set identification, as presented before, was accomplished, as described previously in two ways. In the first approach we used an additional class of sounds representing various environmental sounds or noises and used it as an alternative class. So, we brought the problem to a closed-set identification issue. Identification rates and error rates for sound identification using an additional complementary class, to model the alternative class of noises are presented in table V. As previously we applied the PNN algorithm with various distances in the EM initialisation stage. This method generates high FAR rates.

Table 5 - Open-set identification rates (IR), false acceptance and false rejection rates using GMM with varying number of components for modelling the of 14-dimensional Mel features space, and PNN with several measures for initialization

		Distance											
		euclidian			minkovsky			chebychev			Seuclidian		
		IR	FRR	FAR	IR	FRR	FAR	IR	FRR	FAR	IR	FRR	FAR
Number of Gauss components	10	65.88	34.12	45.45	65.88	34.12	45.45	63.52	36.48	45.45	63.52	36.48	45.45
	14							65.88	34.12	45.45	65.88	34.12	45.45

In the second range of experiments we applied the GMM-UBM, where we rather defined a specific background model for each of the three classes of sounds we want to deal with (chainsaw, vehicle and genuine forest) than create a universal or general background model. We chose to do so because the number of classes is low, so it does not entail high computational or storage burden, and it should guarantee better performance. More exactly the background model for each class was defined as the union of the models of the alternative classes models (in fact the union of all the Gaussian components, with recalculated weights). This process is regarded as a closed-set identification followed by a verification, as presented in Figure 1. The sound identified at the closed-identification step is checked using (5) against a specific threshold θ, at the verification step. To define the threshold θ for a specific class, or a general threshold, for a sequence of θ values used as in (5) the FAR and FRR rates are determined using a validation or training set of environmental audio signals. The θ value were the two errors are approximatively equal (and equal to a such called Equal Error Rate EER) is recommended to be used, but it also depends on the choice of the specialist or user. Table VI presents the FAR and FRR rates for several values of threholds ranging from -1.8*0-15 to 1.8*10-15 for verification of the genuine forest sounds. The threshold value is usually chosen where the two rates are approximatively equal, in our case this value is close 0.3*10-15.

Using the threshold values established as above, set to 0 for chainsaw and vehicle classes and to 2.4 for forest sounds, we performed the open-set identification, with. The results are presented in table VII.

Table 6 - Closed set identification rates (IR), using GMM with varying number of components for modelling the of 14-dimensional Mel features space, using PNN and several measures for initialization

		-1.8	**-1.5**	**-1.2**	**-0.9**	**-0.6**	**-0.3**	**0**	**0.3**	**0.6**	**0.9**	**1.2**	**1.5**	**1.8**
FRR		0	0	0	0	0	0	0	**0.325**	0.725	0.875	0.9	0.9	0.9
FAR		1	1	1	1	1	0.964	0.625	**0.25**	0.107	0	0	0	0

Table 7 - Open-set identification rates, FAR and FRR for the experiments using the GMM-UBM approach, and PNN as initialisation method with Chebychev și seuclidean distanceas

		Classes of audio Signals								
		chainsaw			Genuine Forest			vehicle		
		IR	FR	FA	IR	FR	FA	IR	FR	FA
Distance	chebychev	53.84	47.16	9.63	87.5	12.5	19.64	50	50	9.375
	seuclidian	53.84	47.16	9.63	87.5	12.5	19.64	50	50	10.93

5. CONCLUSIONS

The paper intended to evaluate some environmental sounds recognition methods based on GMM. We considered three classes of sounds and assessed the identification rates, confusion errors and false acceptance/ rejection rates. We used different types of features and feature combinations, such as LPC or Mel-scale cepstral features and the zero crossing rates. The best results were obtained using the Mel-cepstral features. But as one can deduct from the Signal Pre-processing subsection the most significant part of the spectrum to be explored is situated under 8kHz, fact that was speculated in Mel-cepstral analysis, by limiting the frequency domain to the interval [0.3kHz, 3.7kHz]. The algorithm to calculate the LPC coefficients was not adapted to deal with a restrained frequency domain, so the analysis took place on the whole spectral domain [0kHz, 22kHz]. In addition, we remarked the not very significant effect of using the zero-crossing rate.

The results are worse than those obtained using the SVM methodology. But it is known that SVM uses a binary classification, as it was conceived especially for chainsaw detection, which makes the comparison a bit inappropriate.

As future directions we will continue of course try to improve the identification rate, by exploring other types of features, for instance spectral features, and other modelling strategies. Another important immediate direction is to acquire a more proper acoustic database.

Acknowledgment

This work has been supported in part by UEFISCDI Romania through projects SeaForest, CitiSim, SenSyStar and WINS@HI, and funded in part by European Union’s Horizon 2020 research and innovation program under grant agreement No. 777996 (SealedGRID project) and No. 643963 (SWITCH project).

6. REFERENCES

1. XUAN GUO, YOSHIYUKI TOYODA, HUANKANG LI, JIE HUANG, SHUXUE DING, YONG LIU *“Environmental Sound Recognition Using Time-Frequency Intersection Patterns”,* Hindawi Publishing Corporatio. Applied Computational Intelligence and Soft Computing, Volume 2012, Article ID 650818, 6 pages doi:10.1155/2012/650818,

2. SUNIT SIVASANKARAN, K.M.M. PRABHU, "*Robust features for Environmental Sound, Conference: Electronics*", Computing and Communication Technologies (CONECCT), January 2013 IEEE, available at https://www.researchgate.net/publication/261489370_Robust_features_for_environmental_sound_classification
3. WWF: Deforestation available at https://www.worldwildlife.org/threats/deforestation
4. D. MITROVIC, M. ZEPPELZAUER, H. EIDENBERGER,"*On Feature Selection In Environmental Sound Recognition*", IEEE, ELMAR International Symposium, 2009., 28-30 September 2009, Zadar, Croatia, available at https://www.researchgate.net/publication/224085432_On_feature_selection_in_environmental_sound_recognition .
5. R. BUHUȘ, L. GRAMA, C. RUSU, "*Several Classifiers for Intruder Detection Applications*", 9th International Conference on Speech Technology and Human-Computer Dialogue, SpeD 2017 , 2017
6. L. GRAMA, C.RUSU, "*Audio Signal Classification Using Linear Predictive Coding And Random Forest",* 9th International Conference on Speech Technology and Human-Computer Dialogue, SpeD 2017, 2017
7. M. BABIŠ, M. ĎURÍČEK, V. HARVANOVÁ, M. VOJTKO, "*Forest Guardian –Monitoring System for Detecting Logging Activities Based on Sound Recognition*", IIT.SRC, pp. 1-6, 4 May, 2011.
8. J. PAPÁN, M. JURECKA, J. PÚCHYOVÁ, "*WSN for Forest Monitoring to Prevent Illegal Logging*", IEEE, Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 809-812, 2012
9. MARKEL, J. D., GRAY JR, A. H., *Linear Prediction of Speech*, Springer Verlag - New York, 1976.
10. DAVIS, S. B., AND MERMELSTEIN P., "*Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences*", IEEE Trans. on ASSP 28, pp 357–366, 1980.
11. X. HUANG, A. ACERO, H. W. HON. "*Spoken Language Processing:A Guide to Theory, Algorithm, and System Development*", Prentice Hall, 2001.
12. PIER LUCA LANZI:, "*A turorial on Clustering*": available at http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/
13. "*Agglomerative hierarchical cluster tree - MATLAB linkage*" available at https://www.mathworks.com/help/stats/linkage.html