Automatic Forest Wood Logging Identification based on Acoustic Monitoring

Iosif Mporas
Artificial Intelligent Group, Wire Communication Laboratory, Department of Electrical and Computer Engineering, University of Patras, 26504, Rion Patras
Computer and Informatics Engineering Department, Technological Educational Institute of Western Greece, 30020 Antirio, Greece
imporas@upatras.gr

Michael Paraskevas
Computer and Informatics Engineering Department, Technological Educational Institute of Western Greece, 30020 Antirio, Greece
Computer Technology Institute and Press “Diophantus”
26504, Rion Patras
mparask@cti.gr

ABSTRACT
In this paper we describe a scheme for automatic identification of wood logging activity in forest based on acoustic surveillance. Specifically, we evaluate five machine learning classification algorithms using several audio descriptors for the identification of chainsaw wood logging sounds in the noisy environment of a forest. Different environmental noise interference levels, in terms of sound-to-noise ratio, were considered and the best performance was achieved using support vector machines.

CCS Concepts
• Information systems → Speech / audio search • Computing methodologies → Supervised learning by classification • Theory of computation → Support vector machines • Computing methodologies → Neural networks • Computing methodologies → Instance-based learning.

Keywords
biodiversity monitoring; audio based surveillance; audio processing; classification.

1. INTRODUCTION
Forests have an important role in the maintenance of the biodiversity and the earth's overall ecological balance. Global forest cover is a key indicator of the health of the planet. Forests purify air, preserve watersheds, improve water quality, prevent erosion and provide us with natural resources. Moreover they absorb a lot of carbon dioxide, a major greenhouse gas in global warming, and thus help to protect our planet from upcoming climate change. It is estimated that 1.6 billion people worldwide rely on forests for their livelihoods and 60 million indigenous people depend on forests for their subsistence [1].

Illegal logging causes unmanaged and often irreparable deforestation which is a great threat to maintaining biodiversity, as nearly ninety percent of terrestrial biodiversity is found in the forests [1]. Illegal logging poses a significant threat to the sustainability of forest ecosystems, resulting in large scale deforestation which has negative impact on the atmosphere resulting in global warming, flash floods, landslides, drought etc [2], as well as results in losses of government revenues, foster a vicious cycle of bad governance, and may contribute to increased poverty and social conflict [3]. The problem of illegal logging concerns not only forest rich countries but also countries that import and consume wood-based products from timber-producing countries with high levels of illegal logging, since the import of products without ensuring that they are legally sourced contributes to the problem [4].

The scale of illegal logging cannot accurately be estimated due to the nature of it. Globally, illegal forest activities are said to result in annual government revenue losses in the range of $10-15 billion USD [3, 5]. Illegal trade irregularities were estimated to be 15% of the total trade in the mid 1990s [6]. More than half of all logging activity in the most vulnerable forest regions is believed to be conducted illegally [7]. Despite the work of ecological movements, non-governmental organizations and existence of systems for tracking export timber products, the use of employed systems which would provide effective solution to the problem of illegal logging detection is essential [8].

Because of the adverse effects of illegal logging, forest management authorities take measures for surveillance, information collection and monitoring of the forest to prevent deforestation. Surveillance can be done either by ground-based techniques or using automatic sensor-based monitoring techniques exploiting the advances of existing technology [2]. The ground-based solutions typically include on-site surveillance by staff and mobile patrols for the monitoring of the forest by land, by water and/or by air [9]. Additionally, observation towers located at strategic points for visual detection of fire or other abnormal/illegal events by specialized personnel are often used. Such solutions are expensive, time consuming and require a lot of resources, thus technology-based solutions are needed.

The technological achievements in wireless communications (e.g. WiFi, Bluetooth and ZigBee technologies) as well as in statistical signal modeling and pattern recognition allow the development of cost effective solutions for forest monitoring, against

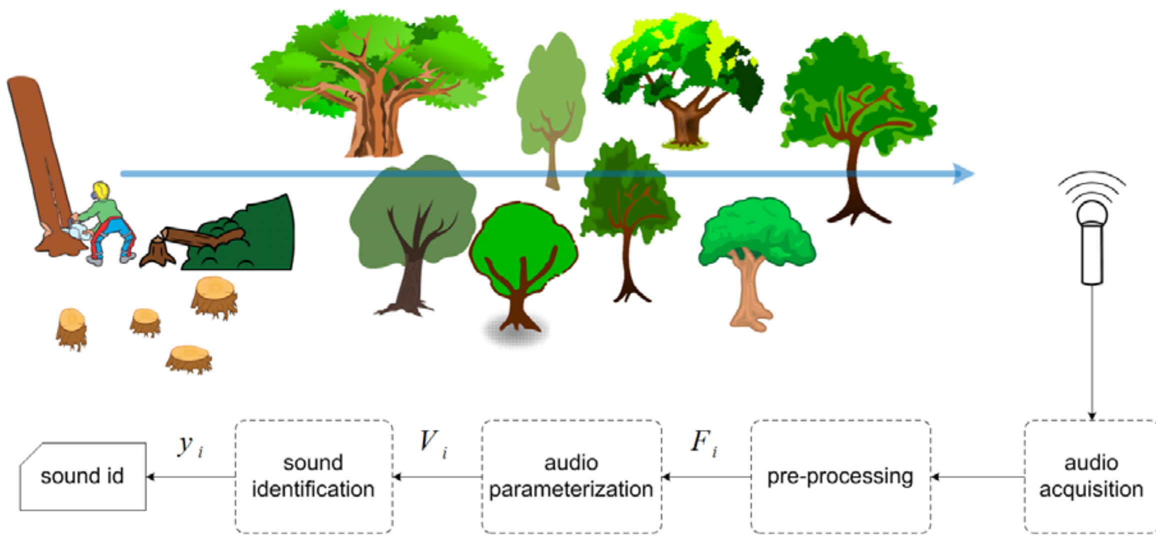

Figure 1. Block diagram of the acoustic monitoring scheme for forest logging identification.

deforestation. Aerial or satellite image/video surveillanceis very costly for the monitoring of illegal activities in forests, while in ground solutions image/video surveillance is difficult due to the typically dense and high vegetation of forests. Moreover, image/video surveillance performance will drop when illumination is low or with the existence of fog. Thus, audio-based surveillance is advantageous.

Several acoustic monitoring approaches for the detection of logging have been proposed in the literature. In [10] evaluation of chainsaw logging sound using spectral audio characteristics and several decision trees was applied. In [11] the authors present a ZigBee based solution for the detection of wood logging based on sound recognition, while in [12] a sound similarity between wood logging chainsaw and forest interference sounds is studied.

In this article we evaluate several classification algorithms using a short-time analysis scheme and well known audio descriptors in the task of logging identification considering different levels of forest noise in terms of sound to noise ratio. The remainder of this article is organized as follows: In Section 2 we present the acoustic monitoring scheme for the identification of wood logging sounds. In Sections 3 and 4 we describe the experimental setup and the evaluation results, while in Section 5 conclusions are provided.

2. ACOUSTIC FOREST MONITORING

Automatic forest monitoringusing audio data offers 24/7 surveillance of specific habitats and provides the information needed for the protection of their biodiversity. The wood logging identification task falls in the category of audio pattern recognition, and thus can briefly be structured in the audio acquisition stage, the audio preprocessing and parameterization stage and finally the identification (classification) stage. The identification of wood logging sounds in the forest includes interferences that are additive to the sound source (e.g. a chainsaw). Such interferences are the wind, the rain, the sum of the leaves, vocalizations from animal species of the habitat, sounds produced by human activities, etc. The concept of audio-based wood logging identification is illustrated in Figure 1.

As can be seen in Figure 1, the audio signal is captured by a microphone installed in the forest field. The audio signal is typically amplified and sampled before sent for further processing. At the preprocessing stage the signal is low-pass filtered and frame blocked by a sliding time window with constant length of N samples, resulting to a sequence of audio frames $F_i \in \Re^N$ with $i = 1, 2, \ldots$. After preprocessing, each of the audio frames, F_i, is forwarded to the audio parameterization stage, where d audio parameters (audio descriptors) are computed for each frame, thus decomposing the audio input to a sequence of audio feature vectors $V_i \in \Re^d$ with $i = 1, 2, \ldots$. The identification of the sounds of interest (in our case wood logging sounds produced by chainsaws) is performed in frame level. Specifically, each of the computed feature vectors is processed by a classification algorithm, f, which labels the corresponding feature vector as a wood logging sound or not, i.e.

$$y_i = f\left(V_i, M\right) \quad (1)$$

where M is the classification model and y_i is the label (i.e. the sound id) assigned to the i-th frame.

Based on the identified labels, y_i, further post-processing can be applied for improving the identification performance. Post-processing typically includes the exploitation of the identified labels of the w preceding and w successive audio frames for the fine-tuning (re-estimation) of the y_i label.

The presence of non-stationary noises originating from the environment makes the wood logging identification task more challenging. The degree of interference of the environmental noises and the actual sound-to-noise ratio are crucial for the identification of wood logging sounds. In this work we focus on the effect of the distance of the sound source (produced by chainsaws) from the monitoring station (field microphone) as expressed by different signal-to-noise ratios, to the logging identification performance.

3. EXPERIMENTAL SETUP

This section provides a description about the audio data used in the present evaluation, the audio parameterization algorithms used, the machine learning classification algorithms that were tested and the experimental protocol that was followed.

3.1 Data

The dataset used in the present research consists of audio recordings of eleven different types of chainsaws of total duration of approximately 5 minutes. Together with the wood logging sounds we used recordings from forest which include several background sounds such as wind, rain, sum of leaves and bird vocalizations. All data were collected from online and freely available audio data repositories and were down-sampled at 8 kHz.

In order to evaluate the wood logging identification performance at different signal-to-noise ratios randomly selected forest recordings were interfered to the wood logging sounds as additional noise.

3.2 Audio Parameterization

For the parameterization of the audio signals we used a diverse set of audio parameters. Specifically, the audio signal was frame blocked by a time sliding window of 20 msec length and 10 msec step. For audio parameters we used two temporal audio descriptors and sixteen spectral audio descriptors. The temporal audio descriptors which were used are the frame intensity and the zero crossing rate. As considers the spectral audio parametersthese are: the 12 first Mel frequency cepstral coefficients (MFCCs), the root mean square energy of the frame, the voicing probability, the harmonics-to-noise ratioby autocorrelation function and the dominant frequency. All audio parameters were computed using the openSMILE acoustic parameterization tool [13]. The computed audio parameters were post-processedby applying dynamic range normalizationin order to equalize the range of their numerical values.

3.3 Classification Algorithms

For the evaluation of the wood loggingsound identification we relied ona number of different machine learning algorithms for classification. These are:

- the k-nearest neighbors classifier with linear search of the nearest neighbor and without weighting of the distance – here referred as instance based classifier (IBk) [14].
- a 3-layer Multilayer perceptron (MLP) neural network with architecture 18–10–1 neurons (all sigmoid) trained with 50 000 iterations [15].
- the support vector machines (SVM) utilizing the sequential minimal optimization algorithm with a radial basis function kernel [16].
- the pruned C4.5 decision tree (J48), with 3 folds for pruning and 7 for growing the tree [17].
- the Bayes network learning (BN) using a simple data-based estimator for finding the conditional probability table of the network and hill climbing for searching network structures [18].

For the implementation of these algorithms we relied on the Weka [18] machine learning toolkit. For all of the evaluated algorithms the parameters the values of which are not been defined have been set equal to the default ones.

4. EXPERIMENTAL RESULTS

In all experiments we followed a common experimental protocol as described in Section 3. Ten-fold cross validation experiments were performed on the audio data described in the previous section, thus resulting to non-overlapping training and test data subsets. The performance of the binary classification algorithms (i.e. logging sound vs. not) in frame level for various signal-to-noise ratios is shown in Table 1, in percentages.

Table 1. Acoustic wood logging classification accuracy (in percentages) for different signal-to-noise ratios

SNR	IBk	MLP	SVM	J48	BN
-6 dB	73.90	76.37	77.04	77.19	73.03
0 dB	76.69	79.26	81.65	80.75	75.80
6 dB	78.03	82.40	84.32	81.20	77.44
12 dB	80.61	83.64	88.11	81.58	80.03
16 dB	81.91	85.80	89.45	83.33	81.59
20 dB	84.78	87.83	91.07	86.02	84.20

As can be seen in Table 1, the best performing algorithm for all evaluated SNR values is the support vector machines. Specifically, SVM algorithm achieved classification accuracy equal to 81.65% for sound-to-noise ratio equal to 0 dB, while the performance increased to approximately 91% for sound-to-noise ratio equal to 20 dB.

The two discriminative algorithms, i.e. the SVM and MLP neural network, demonstrated the highest classification accuracy for almost all SNR levels, followed by the C4.5 decision tree (J48). The IBk and Bayes network (BN) algorithms did not achieve competitive performance.

It is worth mentioning that in noisy environments, such as SNR equal to 0 dB and -6 dB, the C4.5 decision tree (J48) performs equally well with the support vector machine algorithm, which is in agreement with [10], where J48 also proved to achieve robust classification accuracy performance. However, in the present evaluation SVM algorithm achieved superior classification accuracy scores independently of the SNR value, which is an indication of the advantage that SVM can offer in real-field environments, where the presence of non-stationary interfering noises is frequent. Moreover, low signal-to-noise ratio issues are met in audio acquisition when the sound source (in our case wood logging sounds) is not close to the microphones installed in the field.

In a next step, we applied a post-processing sliding window filter to the recognized labels of each frame. The purpose of the post-processing is to eliminate sporadic erroneous labeling of the current audio frame, e.g. due to momentary burst of interference, and thus contributes for improving the overall classification accuracy. Specifically, during post-processing we apply a decision-smoothing rule to each frame, F_i, i.e. when the k preceding and the k successive audio frames are classified to one class (wood logging sound or not) then the current frame is also (re)labeled as of this sound class. The length, L, of the smoothing window is subject to investigation and in the general case is equal to $L = 2 \cdot k + 1$. The case $L = 1$ corresponds to baseline setup without post-processing of the classified labels. The effect of the smoothing window in the wood logging sound classification performance for the best performing algorithm (i.e. the support vector machines) and for several SNR values is tabulated in Table 2, in percentages.

Table 2. Acoustic wood logging classification accuracy (in percentages) using post-processing

SNR	L = 1	L = 3	L = 5
-6 dB	77.04	78.93	78.46
0 dB	81.65	82.89	82.09
6 dB	84.32	85.85	84.78
12 dB	88.11	89.34	89.20
16 dB	89.45	90.65	89.90
20 dB	91.07	92.30	92.04

As can be seen in Table 2,the effect of the post-processing stage is significant for all signal-to-noise ratios and especially in the case of noisy environment, i.e. for low signal-to-noise ratios. Specifically, the window length equal to three offers the best performance for all the evaluated signal-to-noise ratio values. After the application of the post-processing with $L = 3$, the achieved classification accuracy was improved by approximately 1% in terms of absolute improvement for all signal-to-noise ratio values, while for the case of noisy environment (i.e. for SNR value equal to -6dB) the improvement was almost 2% comparing to the case where no post-processing was applied ($L = 1$).

5. CONCLUSION

We evaluated of five classification algorithms, employed in a typical-standard scheme for audio processing, on the wood logging identification task. In our setup this evaluation involves identification of chainsaw sound when logging in the forest. The best classification accuracy was achieved by the support vector machine algorithm.

Experiments with additive noise using typical acoustic background from forest, for several sound-to-noise values, demonstrated the robustness of the wood logging identifier in noisy conditions, such as the ones found in the real field. Finally, the use of post-processing on decision level per frame demonstrated an improvement of more than 1% and especially for low sound-to-noise ratios.

6. REFERENCES

[1] Harvanova V., Vojtko M., Babis M., Duricek M., Pohronska M. (2011), "Detection of wood logging based on sound recognition using zigbee sensor network", In Proc. of International Conference on Design and Architectures for Signal and Image Processing.

[2] Chethan K.P., Srinivasan J., Kriti K. and Sivaji K. (2012), "Sustainable Forest Management Techniques", Deforestation Around the World, Ed: Paulo Moutinho, InTech, 2012.

[3] Tacconi L., Boscolo M., Brack D. (2003), "National and International Policies to Control Illegal Forest Activities", A report prepared for the Ministry of Foreign Affairs of the Government of Japan, July, 2003.

[4] Hoare A. (2014), "Illegal Logging and Related Trade - The Response in Ghana", Energy, Environment and Resources, A Chatham House Assessment.

[5] Brack D. (2006), "Briefing Paper: Illegal Logging", Chatham House.

[6] Brack D., Hayman G. (2001), "Intergovernmental Actions on Illegal Logging, Options for Intergovernmental Action to Help Combat Illegal Logging and Illegal Trade in Timber and Forest Products", The Royal Institute of International Affairs, London.

[7] Lawson S. and MacFaul L. (2010), "Illegal Logging and Related Trade - Indicators of the Global Response", Chatham House, July 2010.

[8] Babis M., Duricek M., Harvanova V. , Vojtko M. (2011), "Forest Guardian – Monitoring System for Detecting Logging Activities Based on Sound Recognition", Researching Solutions in Artificial Intelligence, Computer Graphics and Multimedia, IIT.SRC 2011, pp. 1–6.

[9] Magrath W.B., Grandalski R.L., Stuckey G.L., Vikanes G.B., Wilkinson G.R. (2007), "Timber Theft Prevention: Introduction to Security for Forest Managers", Sustainable Development-East Asia and Pacific Region, Discussion Papers, The World Bank Publication.

[10] Czúni L., Varga P.Z., "Lightweight Acoustic Detection of Logging in Wireless Sensor Networks", 2014, International Conference on Digital Information, Networking, and Wireless Communications (DINWC2014), pp. 120-125.

[11] Harvanová, V., Vojtko, M., Babiš, M., Ďuríček, M., Pohronská, M. (2011). Detection of Wood Logging Based on Sound Recognition Using Zigbee Sensor Network. International Conference on Design and Architectures for Signal and Image Processing.

[12] Tang Y., Han P., Wang Z., Hu L., Gao Y., Li H., "Based on intelligent voice recognition of forest illegal felling of detecting methods", 2012, 2nd International Conference onCloud Computing and Intelligent Systems.

[13] F. Eyben, M. Wollmer, and B. Schuller, "OpenEAR - introducing the Munich open-source emotion and affect recognition toolkit," In Proc. of the 4th International HUMAINE Association Conference on Affective Computing and Intelligent Interaction (ACII 2009).

[14] D. Aha, D. Kibler, "Instance-based learning algorithms", Machine Learning, 6 (1991), pp. 37–66.

[15] T.M. Mitchell, Machine Learning, McGraw-Hill International Editions (1997).

[16] S.S. KeerthiS.S., S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy, "Improvements to Platt's SMO algorithm for SVM classifier design, Neural Computation, 13 (3) (2001), pp. 637–649.

[17] R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, San Mateo, CA (1993).

[18] H.I. Witten, and E. Frank. Data Mining: practical machine learning tools and techniques. Morgan Kaufmann Publishing.