World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:8, No:9, 2014

Low Cost Real Time Robust Identification of Impulsive Signals

R. Biondi, G. Dys, G. Ferone, T. Renard, M. Zysman

***Abstract*—**This paper describes an automated implementable system for impulsive signals detection and recognition. The system uses a Digital Signal Processing device for the detection and identification process. Here the system analyses the signals in real time in order to produce a particular response if needed. The system analyses the signals in real time in order to produce a specific output if needed. Detection is achieved through normalizing the inputs and comparing the read signals to a dynamic threshold and thus avoiding detections linked to loud or fluctuating environing noise. Identification is done through neuronal network algorithms. As a setup our system can receive signals to "learn" certain patterns. Through "learning" the system can recognize signals faster, inducing flexibility to new patterns similar to those known. Sound is captured through a simple jack input, and could be changed for an enhanced recording surface such as a wide-area recorder. Furthermore a communication module can be added to the apparatus to send alerts to another interface if needed.

***Keywords*—**Sound Detection, Impulsive Signal, Background Noise, Neural Network.

I. Introduction

AUTOMATIC sound and speech detection has received noticeable attention and achieves remarkable performance in everyday use of devices [1], [2]. However, the aspect of studying noise and extracting interesting samples from it is more abstract as it usually implies specific analysis [3]-[5]. The goal of present project is to detect and identify specific sounds that could be related to a dangerous situation. A dangerous situation is defined as an event that could bring harm to a nearby person in the area. If these situations accompanied by a sound, these events could be classified, such as a gunshot or someone screaming. The particularity of these singular events is their impulsive nature.

To study such signals several techniques are used to catch impulsive signals, and to compare them [6]. They are unable to detect precise characteristics in a noisy or urban environment. Sound classification has already been studied [7] and can be done using a combination of techniques to obtain efficient classifier (ICA, etc…). The proposed set of features has the ability to distinguish similar sounds whose sources have the same geometry and size but with different materials. Thus, considering that the same result can be achieved using different methods, present paper will focus on the technique used to detect and analyse impulsive signals and determine their nature. The systems recognition process evolved into the use of a neural network [8]. The complexity and needs for neural network vary in limitations and uses. Here a single layer neuron has been used [9] for the differentiation method explained in Part III. The ability to detect sounds related to dangerous situation could lead to a new type of security devices. This equipment could be added to existing video cameras and significantly improve the efficiency of security infrastructure. In order for the system to be considered as effective and trusty it must be reliable, real time and robust. These three objectives are kept in line during the entire process which was developed in present study.

Robin Biondi, Gareth Dys, Gilles Ferone, Thibault Renard and Morgan Zysman are Undergraduate Students of ECE Paris School of Engineering, 37 Quai de Grenelle, 75015 Paris, France (e-mail: biondi@ece.fr, dys@ece.fr, ferone@ece.fr, renard@ece.fr, zysman@ece.fr).

II. System Description

A. Preprocessing: Sound Feature Extraction

Through a mesh of constant background noise, the detection process must be able to detect an irregularity close to environment noise level. Furthermore complications are met with variation of background noise level.

To register environing events, a standard microphone is used to capture sound waves as electric readings. These analogic readings are converted to numeric values, by audio codec AD1836, from ADSP-BF533 evaluation board. The continuous noise input, converted as short integers (16 bits), is stored and segmented by a buffer of 512 bits. Once this buffer is full, the stream is recorded on another buffer and one can now process the current one.

Fast Fourier Transform (FFT) is applied on the recorded values, in order to bring out signal frequency dimension in the space of which frequency characteristics can be found. Each sound spectrum carries over characteristics from its source.

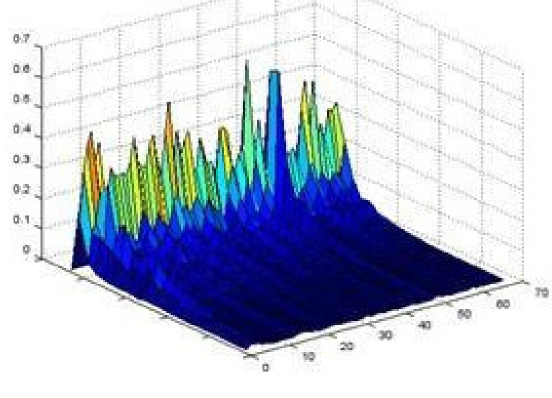

Fig. 1 Gunshot FFT Spectrum vsTtime

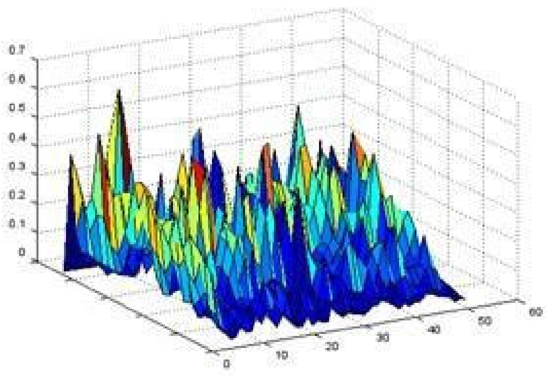

Fig. 2 Breaking Glass FFT Spectrum vs Time

Figs. 1 and 2 show that the spectra of the two different sounds are clearly differentiable. In the gunshot case, lowest and middle frequency amplitudes are very low whereas the breaking glass spectrum shows that almost every frequency has intense amplitudes. To obtain Figs. 1, 2, the inputs have to be normalised in order to produce a slightly less variable level of intensity. Furthermore, normalizing the signals enables to start detecting the particularities of sounds under reading. Present project is intended to find impulsive abnormality and characteristics of researched sounds inside collected noises. Over a 10-millisecond block of values, all amplitudes A(fn) for each frequency fn are summed up to form a single amplitude curve I(fn) given by.

$$I\_fn=\sum_{fn=fmin}^{fmax} A_{fn} \quad (1)$$

This curve is then normalised and one obtains a simple curve which will be used for next processing steps of the analysis.

The use of threshold mechanism through filtration and normalisation proves to be effective even in noisy backgrounds. The detection is still robust, to a certain degree, to an added artificial white noise background. Furthermore real-world background noises prove to be less critical, and similar detection systems have a higher percentage of success [10]. Figs. 3 and 4 correspond to different and comparable curves. Through differentiation or subtraction one can determine that they do not come from the same source.

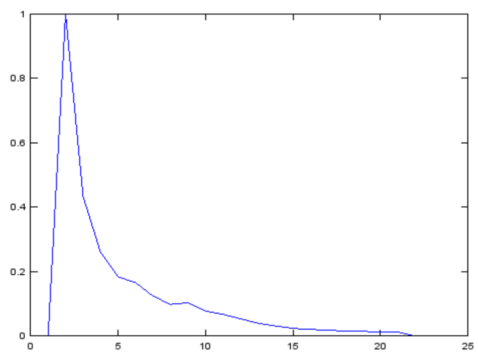

Fig. 3 Normalised Frequencies Sumfor a Gunshot

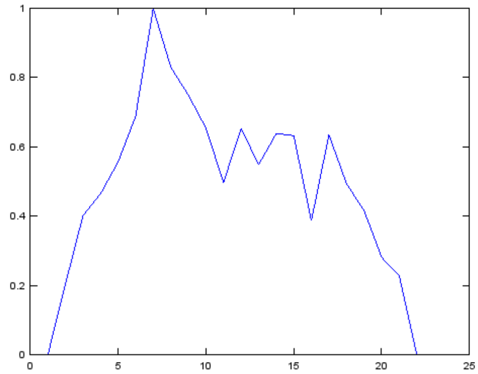

Fig. 4 Normalized Frequencies Sum for Breaking Glass

B. Classification

In order to classify the inputs a Feed Forward Neural Network (FFNN) is used [11]. The FFNN is divided into three parts, one input layer, one or more hidden layers and one output layer, see Fig. 5.These layers are connected together and process the input in order to draw conclusions on the nature of the sound. During the learning process the Neural Network is trained to recognize sounds such as gunshots, glass breaking and alarms. This learning process generates a weight matrix for every connection between two neurons which is then implemented in the system. Once the feature has been extracted, the data are used as input layer in the FFNN. Conclusions are drawn depending on the class of sound with best probability.

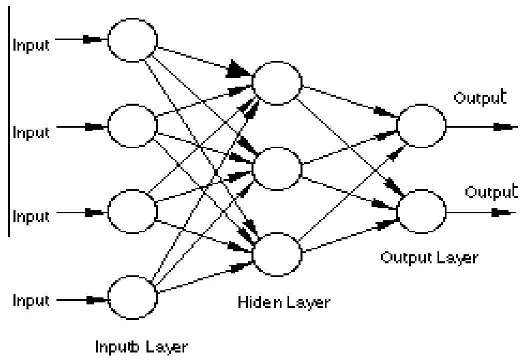

Fig. 5 Neural Network Structure

C. System Architecture

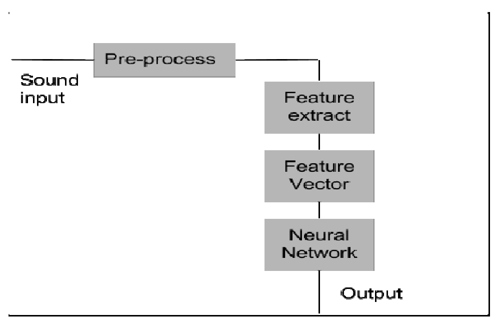

Fig. 6 Diagram Structure of Processing Steps

In order to conduct the feature extraction and the classification a Digital Signal Processor ADSP-BF533 has been used. The sound input is provided by a ".wav" file. The output is in the form of an integer; each type of sound has an ID (ex. 1: Gunshot, 2: Glass Breaking, 3: Alarm). Once the sound has been identified the system can communicate the information to the relevant authorities.

III. Results

Three classes of sound have been tested directly on the system, gunshots (1), glass breaking (2) and alarm signal (3). Each class includes ten samples of different sounds. First test consisted on playing the sample without background noise.

TABLE I
Test Results without Background Noise

Class	Recognition accuracy (%)
Gunshot	86
Glass breaking	93
Alarm signal	95

The second conducted test consisted in mixing the samples with background noises extracted from a public area.

TABLE II
Test Results with Background Noise

Class	Recognition accuracy (%)
Gunshot	69
Glass breaking	90
Alarm signal	92

It can be seen from Tables I and II that the recognition accuracy for gunshot class (1) decreases while it stays almost unchanged for glass breaking (2) and alarm signal (3) classes. This can be explained by the fact that gunshots are more likely to merge with background noises as their frequency is very low whereas the two other classes operate in a different frequency range.

In this sense, present system brings a major difference with similar systems able for instance to detect gunshots [12], [13]. It has the ability to differentiate several sounds from one another by finer analysis of their frequency spectra with and without background noise as illustrated by results on Tables I and II. It can thus provide relevant results for use in security domains for instance. It should be borne in mind that present set-up is not intended as in other approaches [14], [15] to recognize any sound which usually requires large Neural Network and long training period with sometimes weak determination. Here the system has been specialized to the specific class of very typical sounds related to dangerous situations. It is then possible as a consequence of this restriction to focus on their characteristic features with higher efficiency even with modest equipment.

IV. System Properties

A. Low Cost Alternative

The development of present project was undertaken with evaluation board from Analog Devices, the ez-kit associated with ADSP-BF533 board which, as well as the software associated with it, is relatively expensive. However system architecture portrays is quite simple. For the experiment, a direct analog input was used for sounds submission into the system. However for live capture, a microphone and an amplification circuit will be needed. Once the input is dealt with, the analytical part of the system requires tailoring for the specific process. A tailored circuit reduces components waste and surplus expenses.

Information processing can be done through the use of converters, amplifiers, arithmetic logic units and memory space. These circuit necessities can be found in digital signal processors. To achieve a low-cost design the DSPic, by microchip, comes equipped with the basic necessities.

B. Robustness

The analysis applied to captured inputs, focuses on precise variables and excludes others, somewhat similar to a white list. The raw sound is broken down into analogical variables, and over a period of time the range of possibilities and differences are practically innumerable. This is especially the case with a particularly noisy background or with particularly loud recorded sounds.

As involved process does not trigger differences through a threshold mechanism, unimportant or any loud artefacts are not flagged in system records. The process breaks down the variations of a recording time period and bases the analysis upon signal frequency signature.

As seen in Figs. 3 and 4, a certain sound has its own signature in the frequency domain. This signature can be considered as print or archetype for its respective class of sounds. Though two sounds of same nature might vary in intensity or in time, the frequencies they generate are unique and intrinsic to their nature.

Sounds of same nature should present the same frequency pattern, it is a possibility that this pattern can vary depending on the composition or the event triggering the sound. Indeed the event of glass shattering might present different frequency patterns depending on its molecular composition, and this is also true depending on the surface condition the glass is shattered with. Nevertheless if the variations are small and as the system compares pattern similarities, it is possible to define an exactness threshold below which the sounds can be considered as similar.

The system manages to differentiate two sounds effectively through the frequency analysis and, as shown in previous example, a slight variation in the signature proves that the sound might be similar but its nature is different. So the system proves to be robust in the differentiation of two sounds.

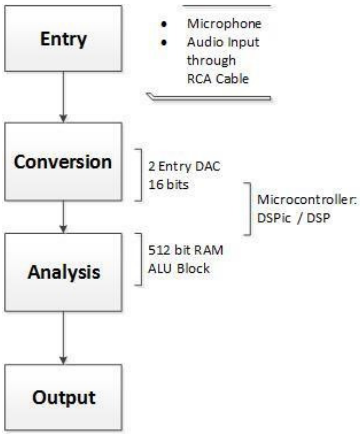

Fig. 7 Sketch of Sound Analysis Process

C. Limits

Despite the ability of present system to detect disruptive sounds with high frequency components, sounds with low frequency are more likely to merge with background noises. In order to build an effective security system more classes of sounds, or noises, should be covered as well as effective pre-processing and filtering techniques in order to submit more relevant information to FFNN. Other disruptive sounds like human screams are very complicated to identify as a dangerous event, because screams can be used in various safe situations as well.

If kept in consideration a big difference between two sounds, initially defined as of same nature, brings out questions on testing conditions. Two events of “glass shattering” are indeed of same nature, but glass shattering against a hard surface and glass shattering against a soft surface might have to be considered as of two different natures.

V. Conclusion

With low cost equipment it has been possible to detect three classes of sounds related to dangerous situations. Experimental results show reliability and robustness of proposed approach against disturbance from background noises. Additional filtering and pre-processing could be used to increase system detection rate. Shot detection equipment has already been adopted in some places like banks or public areas. Present design could lead to a new way of considering public safety by allowing security systems to detect dangerous events and communicate with relevant authorities in complete autonomy.

Acknowledgment

The authors are very much indebted to ECE for having provided the environment where the work has been developed and Pr. M. Cotsaftis for help in preparing the manuscript.

References

[1] D. Hoiem, Y. Ke, and R. Sukthankar, "SOLAR: Sound Object Localization and Retrieval in Complex Audio Environments", *ICASSP* 2005

[2] K.D. Martin: Sound-Source Recognition: a Theory and Computational Method, PhD Thesis, EE and Computer Science Dept., MIT, 1999.

[3] C. Clavel, T. Ehrette, G. Richard: Events Detection for an Audio-Based Surveillance System,, Proc.2005 IEEE Intern. Conf. on Multimedia and Expo, ICME 2005, July 6-9, 2005, Amsterdam, the Netherlands, pp. 1306–1309, 2005.

[4] BR. Stiefelhagen, R. Bowers, J. Fiscus, eds. : Multimodal Technologies for Perception of Humans, International Evaluation Workshops CLEAR 2007 and RT 2007, Berlin, Heidelberg: Springer-Verlag, 2008.

[5] T. Heittola, A. Mesaros, T. Virtanen, A. Eronen : Sound Event Detection in Multisource Environments Using Source Separation, CHiME 2011 Workshop on Machine Listening in Multisource Environments, 2011.

[6] A. Dufaux : Detection and Recognition of Impulsive Sounds Signals, PH.D Thesis, Faculté des Sciences, Université de Neuchatel, 2001.

[7] S. Cavaco, J. Santos Rodeia : Classification of Similar Impact Sounds, Proc. ICISP 2010, Editors: A. Elmoataz et al., Series: LNCS, Number: 6134, Springer- Verlag, pp. 307-314, 2010.

[8] S. Haykin : Neural Networks - A Comprehensive Foundation, Macmillan, 1994.

[9] M.T. Hagan, H.B. Demuth, M. Beale: Neural Network Design, PWS Publishing Company, Boston, USA, 1995.

[10] C. deGroot, D. Wurtz: Plain Back-Propagation and Advanced Optimisation Algorithms: a Comparative Study, Neurocomputing, Vol.6, pp.153-161, 1994.

[11] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford Univ. Press, Oxford, 1995.

[12] S. Lecomte, R. Lengellé, C. Richard, F. Capman, B. Ravera :Abormal Events Detection Using Unsupervised One-Class SVM - Application to Audio Surveillance and Evaluation, 8th IEEE Intern. Conf. on Advanced Video and Signal-Based Surveillance, 2011.

[13] A. Dufaux, L. Besacier, M. Ansorge, F. Pellandini : Automatic Sound Detection and Recognition for Noisy Environment, Proc. EUSIPCO 2000, European Signal Processing Conference 2000, pp. 1033-1036, Tampere, FI, September 5-8, 2000.

[14] T. Zhang and C. Kuo : Hierarchical System for Content-based Audio Classification and Retrieval, Proc. Int. Conf. on Acoustics, Speech, and Signal Processing, 1999.

[15] B. Feiten, S. Gunzel : Automatic Indexing of a Sound Database Using Self-organizing Neural Nets, Computer Music Journal, Vol.18(3), pp.53-65, 1994