Animal Sound Activity Detection Using Multi-Class Support Vector Machines

W. Astuti, A.M. Aibinu, M. J. E Salami, R. Akmelawati, and Asan G.A Muthalif
Intelligent Mechatronics System Research Group
Department of Mechatronics Engineering, International Islamic University Malaysia.
P.O. Box 10. 50728, Kuala Lumpur, Malaysia.
maibinu @iium.edu.my, Winda1977@gmail.com

Abstract- **On March 11th 2011, the whole world was taken aback by another tragic experience of Tsunami triggered by a magnitude 9.8 earthquake in Japan. Just few days after that, on March 25th 2011, another earthquake of magnitude 6.8 hit Myanmar deaths and destructions. Despite the loss incurred on properties and human being, available data show that relatively few numbers of animals died during most natural disasters. Prior to the occurrence of these disasters, available reports shows that animals do migrate to higher level or leave the areas en masse ahead of the event. Other related account show that animal sometimes behaves in unusual ways prior to the occurrence of these natural disasters. These overwhelming evidences point to the fact that animals might have the ability to sense impending natural disaster precursor signals ahead of time. This paper discusses the preliminary results obtained from the use of support vector machine (SVM) and Mel-frequency cepstral coefficients (MFCC) in the development of animal sound activity detection (ASAD) which is an integral part in the development of earthquake and natural disaster prediction using unusual animal behavior. The use of MFCC has been proposed for the features extraction stage while SVM has been proposed for classification of the extracted features. Preliminary results obtained shows that the MFCC and SVM can be used for features extraction and features classification respectively.**

Keywords-Earthquake, Mel-frequency cepstral coefficients (MFCC)Natural disaster;Animal Sounds Activity detector;Support Vector Machine (SVM)

I. INTRODUCTION

Earthquake, landslide, Tsunami occurrences are some of the significant tragic natural disaster which cause irretrievable financial, human and physical losses in various parts of the world. The 1973 Hawaii earthquake , the 22 December 1996 earthquake in Oman, the Bam Earthquake in Iran in 2003[3], the Sumatra Earthquake in 2004, the Japan earthquake in March 2011, and the Myanmar Earthquake also in March 2011 are some of the calamities that have besieged man in the recent past.

On March 11th 2011, the whole world was taken aback by another tragic experience of Tsunami triggered by a magnitude 9.8 earthquake in Japan. As finishing touches are being put to this paper, on March 25th 2011, another earthquake of magnitude 6.8 hit Myanmar leaving behind the death of over 75 people and rendering hundreds homeless. Aside from these major catastrophes in 2011, similar tragic events were reported across Asia in 2004, 2005, 2008 and 2009, with lots of occurrence of landslide in various parts of Malaysia.

Despite the loss incurred on properties and human being, available data show that relatively few numbers of animals died during most natural disasters. Prior to the occurrence of these disasters, available reports shows that animals do migrate to higher level or leave the areas en masse ahead of the event. Other related account show that animal sometimes behaves in unusual ways prior to the occurrence of these natural disasters. These overwhelming events point to the fact that animals might have the ability to sense impending natural disaster precursor signals ahead of time.

In a related case, the very big tsunami triggered by a magnitude 9 of northern Sumatra Island on December 26, 2005 rolled through the Indian Ocean, killing more than 150,000 people. It was reported that relatively few animals were found dead [1]. Similarly, in Colombo Wildlife, officials in Sri Lanka expressed surprise that they found no evidence of large-scale animal deaths from the tsunamis, indicating that animals might have sensed the wave coming and fled to higher ground [1-9].

There have been lots of concerted efforts in reducing the catastrophic effects of earthquake and other natural disaster in recent time [2]. Efforts are being made in accurately predicting the incoming earthquake far ahead of time. Earthquake predictions involve forecasting the occurrence of this unwanted natural disaster of specific magnitude, time and likely region of occurrence. Thus, the term "earthquake prediction" refers to the knowledge of earthquake prognostic parameters which include earthquake epicenter, time of occurrence and the magnitude of the expected earthquake [3-9] It can be divided into three different types, namely long term, short-term and intermediate prediction [2]. Long term prediction, which is not so accurate, is rarely used for public evacuation while; intermediate prediction, which consists of prediction over years to weeks, similar to the long term prediction is also of less usage. Short term predictions involve forecasting the likely occurrence of earthquake within months, weeks, and days from the time of prediction.

Various methods have been proposed for natural disaster predictions, these include: the use of satellite image, Radio communication technique, Prediction using earth precursor geophysical phenomena and signals, and unusual animal behavior just to mention but a few [1-9].

The animal sound activity detection (ASAD) scheme as a preliminary stage in the use of unusual animal behavior for the prediction of impending disaster is discussed in this paper. The remaining part of this paper is organized as follows: section II discusses some of the processes involve in ASAD while the proposed ASAD system is presented in section III. Preliminary results and discussions are contained in the last section.

II. REVIEW OF RELATED PROCESSES

Despite the advancement in human technology, its ability to detect and accurately predict the occurrence of these calamities still trail behind that of animal because some of the presently in-use methods suffer from one defect or the other thus limiting man's ability to master the art of disaster prediction. The accuracy of satellite technique depends on the satellite imagery received and the correlation between the acquired images and previous images. The artificial intelligent system based technique suffers from training data because unknown cases cannot be accurately predicted based on the past available data . Hence, the ability to predict unknown signals depends on the occurrence of such event in the past.

In corroborating this age long unproven belief on animal ability to sense impending disaster, various eyewitness accounts shows that elephants screamed and ran for higher ground, dogs refused to go outdoors, flamingos abandoned their low-lying breeding areas and zoo animals rushed into their shelters and could not be enticed to come back out some hours to the event and fish migrated to deepest path of the sea [1-9].

Several problems in the past mitigate the analysis of reports concerning anomalous animal behavior prior to earthquakes. Three categories of unusual animal response prior to earthquake have been identified [10]. Firstly, most but not all of the unusual animal behavior precursors occur close to the epicenter within 1 or 2 days before earthquake. Secondly, some but not all of the unusual animal behavior occur within a few minutes before earthquake. Lastly, a few of the unusual animal behaviors occur in days to weeks before earthquake [10]. Based on these characteristics, it suffices to use some of the animals' behaviors' can be regarded as precursor phenomenon prior to the earthquake occurrence. Hence, for this preliminary work, the use of animal sound has been explored.

There are two important procedures in the identification of ASAD, namely feature extraction and classification algorithm. Features are some quantities, which are extracted from preprocessed signals and can be used to represent the animal sounds. Many features have been proposed in the literature such as sinusoidal modeling, which based on the tonal quality of many birds song, MFCC and wavelet transformation [11-14]. The second process termed features classification algorithm, is a process to determine the animal identity based on the feature extracted from the recorded sound. Also, many techniques for animal feature classification have been proposed in the literature, these include linear discriminant analysis, and decision tree, SVM and artificial neural network (ANN)[13-17]. The fully automatic system for bird species recognition has been developed based on bird sound recognition of the syllables that are building block of the bird song and call [16]. The pattern recognition technique is used for continual bird sound monitoring. The technique is based on event detection and repetition rate estimation of the bird song element [17].

The choice of SVM is based on its better performance as compared to other Artificial Intelligent techniques. The advantages of SVM over other artificial intelligent techniques are given in Table 1[18]. It has faster learning time, robust in learning process and relatively good optimization capability. SVM as a machine learning technique is based on statistical learning theory. This method is mathematically simple and avoid over-fitting associated with other technique such as ANN, hence reason for its usage in this preliminary work..

TABLE I. Formula for kernel function

No	Criteria	ANN	GA	SVM	Evaluation
1	mathematical model	complicated	complicated	simple	SVM
2	Learning time	fast	medium	faster than ANN	SVM
3	robustness in learning process	robust	slightly robust	more robust than ANN	SVM
4	optimization	slightly good	good	the best	SVM

.

The basic idea of SVM is mapping non linearly training data into higher-dimensional feature space through the kernel function.

Three typical kernel functions used in SVM are listed in Table II.

Table II. Formula for kernel function

Kernel	$K(x,x_i)$
Linear	$x^T \cdot x_i$
Polynomial	$(x^T \cdot x_i +1)^d$
Gaussian RBF	$\exp(-\|x-x_i\|2/2\sigma^2)$

Though SVM was originally designed for binary classification, however, it can be extended to multi-class classification [18]. The extension of the SVM to solve problem containing more that two classes is important. The use of multi-class classification SVM, not only decrease the error rate for the learning and testing set data but also results in faster adaptation and classification times [16]. Several methods have been proposed to solve multi-classification problem, among which include the use of one-against-one approach, which is considered in this paper [18-20].

III. PROPOSED SVM-BASED ASAD SYSTEM

The overall model of the proposed SVMs-based ASAD system is shown in Fig. 1. A digital speech signal is passed through a processing stage from which MFCC features are extracted and passed through the SVMs learning algorithm for both training and testing. The detailed analysis of each stage is as subsequently discussed.

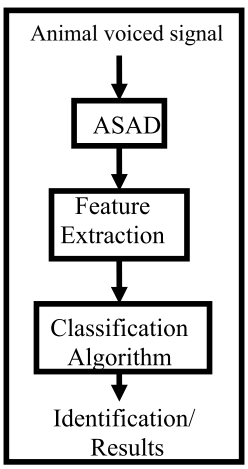

Figure 1. The proposed ASAD system

A. *Animal Sound Activity Detection (ASAD)*

ASAD is used to separate animal sound (voiced), unvoiced and silence frames from recorded animal sounds. Generally, voiced sound shows periodicity in time and frequency domain, with power at low to middle frequency. Unvoiced sound exhibits stronger power at higher frequency than lower frequency. The ratio of sound power in high frequency to that in lower frequency is an important feature to differentiate unvoiced sound from voiced and silence sounds [21].

B. *Feature Extraction using MFCC*

Feature extraction is the process of converting the raw voice signal to feature vector which can be used for classification. There are various types of feature extraction techniques, among which are cepstral coefficient feature based and prosodic-based feature such as fundamental frequency and formant frequency. MFCC is one of the cepstral coefficient based extraction techniques, which is used to extract the feature of the animal sound.

The MFCC process shown in Figure 2 involves signal preprocessing, Hamming windowing of each data block so as to redce the spectral distortion due to abrupt data truncation. A fast Fourier transform (FFT) algorithm is then applied to the resultant preprocessed signal. Mel-filtering, logarithmic processing and inverse FFT are applied to derive the MFCC. The scale of Mel is given as [21]. The MFCC are then calculated. The scale Mel is given by

$$Mel(f) = 2595 \log_{10}\left(1+\frac{f}{700}\right) = 1127 \log_e\left(1+\frac{f}{700}\right) \quad (1)$$

where f is the frequency in Hz.

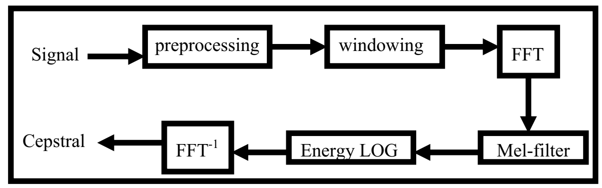

Figure 2. MFCC Process

C. *Classification Algorithm*

SVM has been used as the classification algorithm in this preliminary work. Since SVM involve mapping non-linearly training data into higher-dimensional feature space through the kernel function.

In the linear separable SVM case, there exists a separating hyperplane whose function is given as

$$\mathbf{w}\cdot\mathbf{x} + b = 0 \qquad w \in R^N,\ b \in R, \quad (2)$$

which implies $y_i(\mathbf{w}\cdot\mathbf{x} + b = 0) \geq 1,\ i = 1,\ldots,N.$ (3) This leads to optimized linear division that is the construction of hyper plane that separates the data into two classes. By minimizing $\|\mathbf{w}\|$ subject to this constrain, this approach tries to find a unique separating hyperplane. Here $\|\mathbf{w}\|$ is the Euclidean norm of $\mathbf{w}$, and the distance between the hyperplane and the nearest data points of each class is $1/\|\mathbf{w}\|$. By introducing Lagrange multiplier α_i, the SVM training procedure amounts to solving a convex quadratic problem (QP). The solution is a unique globally optimized result, which has following properties

$$W = \sum_i^N \alpha_i y_i x_i \quad (4)$$

If α_i is not equal to zero, then x_i is called the support vectors (SV). When SVM is trained, the decision can be obtained by comparing each new example x with *only* the support vector $\{x_i\}$, $i \in SV$;

$$y = \text{sign}\left(\sum_{i \in SV} \alpha_i y_i (x_i^{\mathbf{T}} \cdot x) + b\right) \quad (5)$$

However, in a non-linear separable case, SVM performs a non-linear mapping of the input vector $\mathbf{x}$ from the input space into a higher dimensional feature space, where the mapping data is determined by kernel function (see Table II). The choice of the kernel function depends on the characteristic of the data. Different kernel function can be selected to obtain optimal classification results. The choice of the degree in polynomial kernel and choice of weight in Gaussian RBF kernel also depend on the characteristic of data.

IV. EXPERIMENT RESULTS

This research involves data collection of the different kind of bird sounds. The database provided by animal bioacoustics contains many species of bird [22-25]. The proposed ASAD is shown in Fig. 1. The animal sound is firstly processed in ASAD block, which separate sound and silence frame. The short-time energy and zero crossing rates are calculated which will be used to differentiate the frames. The animal sound was sampled at the rate of 16 kHz, and then the samples are segmented into frames with 30 ms overlap. Each frame is multiplied with suitable window function w(n) $\{n = 0,1,2,3,..., N-1\}$, so as to compute the short time energy E_m according to

$$E_m = \sum_m [s(n)w(m-n)]^2 \quad (6)$$

where $w(m)$ Hamming window which is define as

$$w_N(n) = 0.54 - 0.46\cos(2n\pi/(N-1)) \quad (7)$$

Zero Crossing Rate (ZCR) was then applied for frame classification . The ZCR is defined as

$$ZCC_i = \sum_{k=1}^{N-1} 0.5\left|sign(s[k]) - sign(s[k-1])\right| \quad (8)$$

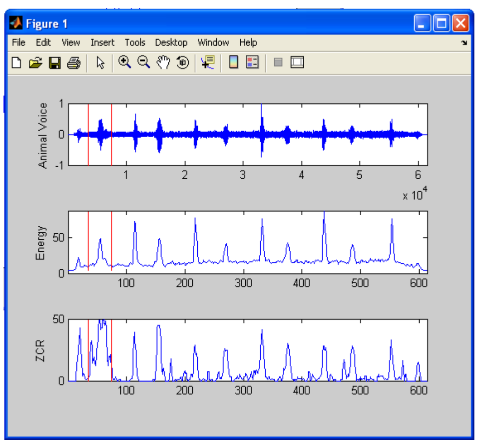

Figure 3. Voice from the animal, Enery and ZCR

Figure 3 shows a typical results obtained from the use of ZCR in this work. All the stages explained previously are then applied on the signals. Another important process worthy of mentioning is the enrollment of the training data in SVM. The SVM-based multi-class classification is applied to perform animal sounds activity detection. Gaussian kernel with degree of 7 and a penalty term C of 10^{15} are used in this experiment. This proposed system has been used to classify 7 different kind of bird the training data consists of 12 ceptral coefficient each bird.

The results obtained shows an average training time of 1.7 seconds. In the training and testing phase using a small database, the proposed SVM-based ASAD system has been able to achieve bird classification rate of 100% accuracy. Further experiment is currently being carried out using bird data from other sources so as to evaluate the effectiveness of the proposed system.

REFERENCES

[1]**http://www.commondreams.org/headlines04/1229-07.htm** , accessed online April 2010**.**

[2] Bhargava, N., Katiyar, V., K., Sharma, M.,L., Pradhan, P., (2009). Earhquake Prediction through Animal Behavior: A Review. Indian Journal of Biomechanics: Special Issue.

[3]Thanassoulas, C., Klentos, V., (2010). How"Short" a "Short-term prediction" can be? A review of the case of Skyros Island, Greece, EQ (26/7/2001, Ms=6.1). Geophysics.1002.2162.

[4]Thanassoulas, C., Klentos, V., Verveniotis, G. Zymaris, N. (2009). Seismic electric precursors observer prior to the 6.3R Eq of July 1st 2009, Greek and their use in short-term earthquake prediction. Geophysics.1003.2772.

[5]Hayakawa, M., T. Itoh, K. Hattori, and K. Yumoto, (2000): ULF electromagnetic precursors for an earthquake at Biak, Indonesia on February 17, 1996. *Geophys. Res. Lett.*, **27**, 1531-1534.

[6]Ikeya, M. 2004. *Earthquake and Animal: from Folk Legend to Science*. World Scientific.

[7]Thanassoulas, C., and Tsatsaragos, J., 2000. The earthquakes of Izmit , Turkey (Ms=7.5R, 17/8/1999) and Athens, Greece (Ms=5.9R, 07/09/1999) as being detected by precursory electrical signals (T=24h period oscillation). Open File report E3906, Inst. Geol. Min. Expl. (IGME), Athens, Greece, pp. 1-18.

[8]Thanassoulas, C., Klentos, V., Vervebiotis, G., Zymaris, N. 2009. Seismic electric precursor observed prior to the 6.3R EQ of July 1th 2009, Greece and their use in short-term earthquake prediction..

[9]Uyeda, S., Nagao, T., Orihara, Y., Yamaguchi, T., Takahashi, I. 2000. Geoelectric potential change: Possible precursor to earthquake in Japan….

[10]Buskirk,E.,R, Frohlich, C., Latham,G.,V,1981, 'Unusual Animal Behavior Before Earthquakes: A Review of Possible Sensory Mechanisms, Review of Geophysics and Space Physics, Vol.19,No.2,Pages 247-270.

[11] Harma,A,(2003). Automatic Recognition of bird species based on sinusoidal modeling of syllables. In: Proc. Of IEEE Internat. Conf. on Acoustic, Speech and Signal Processing. CICASSP 03. Vol. 5, pp. 545-548.

[12] Chou, C,. and Liu, P,. (2009). Bird Species Recogniiton by Wafelet Transformation of a Section of Birdsong. In: Proc. Of symposia and workshop on ubiquitous, Autonomis and Trusted Computing. Pp 189-193.

[13] Miguel A. A., Carlos J. B., Hector C., Luis J.V., and Aide. T.M. (2009). Automated Classification of Bird and Amphibian calls using Machine learning: A comparison of Methods. Elsevier J. Ecological Informatics. Volume 4,

Issue 4, pp 206-224.
[14] Huang, C., Yang, Y., Yang, D., Chen, Y. (2009). Frog Classification Using Machine learning techniques. Elsivier J. Expert System with Application. Volume 36, Issue 2 Part 2, pp 3737-3743.
[15]Nikerson, C.M., Bloom Field, L.,L., Dawson, M.,R.,W., Sturdy, C.,B. (2006). Artificial Neural Network Discrimination of Black-Capped Chickadee (poecile atricapillus) call notes. Appl. Acoust. 67(11-12), 1111-1117.
[16] Fagerlund, S. (2007). Bird Species Recognition using Support Vector Machines. Eurasip J. Appl. Signal Process. (EURASIP JASP).8p.
[17]Bardeli, R., Walff, D., Kurth, F., Koch, M., Tauchert, K.,H., Frommolt, K.,H. (2009). Detecting bird sounds in a complex acoustic environment and application to biocastic monitor. Elsivier J. Pattern Recognition Letters.
[18] Cristianini,N., Shawe,T.,J. (2000) *An introduction to Support Vector Machine and other kernel-based learning methods.* Cambridge University Press,
[19] L.R.Rabiner, R.W. Schafer, *Digital processing of speech signals.* Prentice Hall, 1978.
[20] Bannet.P.Kristin. C.Campbell, *Support Vector Machines: Hype or Hallelujah?*. SIGKDD exploration. Vol. 2, 2000.
[21] Jose Boris Sanchez, *Speaker Identification based on an Integrated System Combining Cepstral Feature Extraction and Vector Quantization,* (Master Thesis), The Florida State University College of Engineering, United State of America, 2005.
[22] Bi-Acoustic: http://www.avisoft.com/sounds.htm
[23] http://www.bird.cornell.edu/
[24].http://www.math.sunysb.edu/~tony/gulls.html
[25].http://thryomanes.tripod.com/B