RESEARCH ARTICLE

Spectral Dynamics of Non-Stationary Audio Waves, VLF& ULF Signals: Complexity of BIS Processes in the Environment

Tuhin Dutta[1], Nandkishore[3], Deepak Kumar[2] and Jagjivan Ram[4]
1. Department of Physics, ARSD College, University of Delhi, Delhi, India, tuhindutta38@gmail.com.
2. Department of Physics, Ramjas College, University of Delhi, Delhi, India.
3. Department of Physics, Singhania University, Rajasthan, India, deepakm256@gmail.com
4. Department of Physics, Ramjas College, University of Delhi, Delhi, India.

Manuscript Info

Abstract

BIS stands for Breakdown of Integrated System. We have considered our universe as an integrated system and it becomes chaotic while dealing with these signals with these variant non-stationary signals. We propose the rate of spectral dynamics as a possible scheme to categorize sounds in the environment. Sound in the natural environment form an important class of biologically relevant non-stationary audio signals. We propose a dynamic spectral measure to characterize the spectral dynamics of such non-stationary sound signals (Audio Waves, VLF & ULF signals) and classify them based on rate of change of spectral dynamics. We categorize sounds with slowly varying spectral dynamics as simple and those with rapidly changing spectral dynamics as complex.

1. Introduction

The human auditory system is capable of discriminating a large variety of complex sounds in the natural environment. Interestingly, anatomical studies of the adult human brain indicate that specialized regions of the brain analyse different types of sounds [1]. Music, speech and environment noise are processed in areas that are anatomically distinct. [2] However, the reasons for this kind of functional organization are not clearly identified. We study the spectral dynamics of different environmental sounds and develop indices to quantify rate of change of spectral dynamics. We propose rate of change of spectral dynamics to explain sound categorization.

The left panel of figure 1 shows examples of sound-pressure waveforms from the natural environment. A striking feature of these different waveforms is that the successive disturbances are not equally spaced in time and are not of constant shape. In fact, a characteristic feature of these waveforms is the variation of spectral content as a function of time. Such non-stationarity in spectral content, which is a common feature of biological signals (electroencephalography, for example) makes it difficult to study such signals using standard analysis techniques. New methods of analysis, which use joint time-frequency representation (TFR) have emerged as convenient methods to describe such non-stationary dynamics. A TFR is obtained by mapping a one-dimensional signal (continuous or discrete) is the time domain into a two-dimensional time-frequency representation. It allows a simultaneous analysis in the time and frequency domains. TFRs provide localization both in time and frequency, within limits of resolution allowed by the uncertainly principle [3]. We study one such class of TFRs called spectrograms.
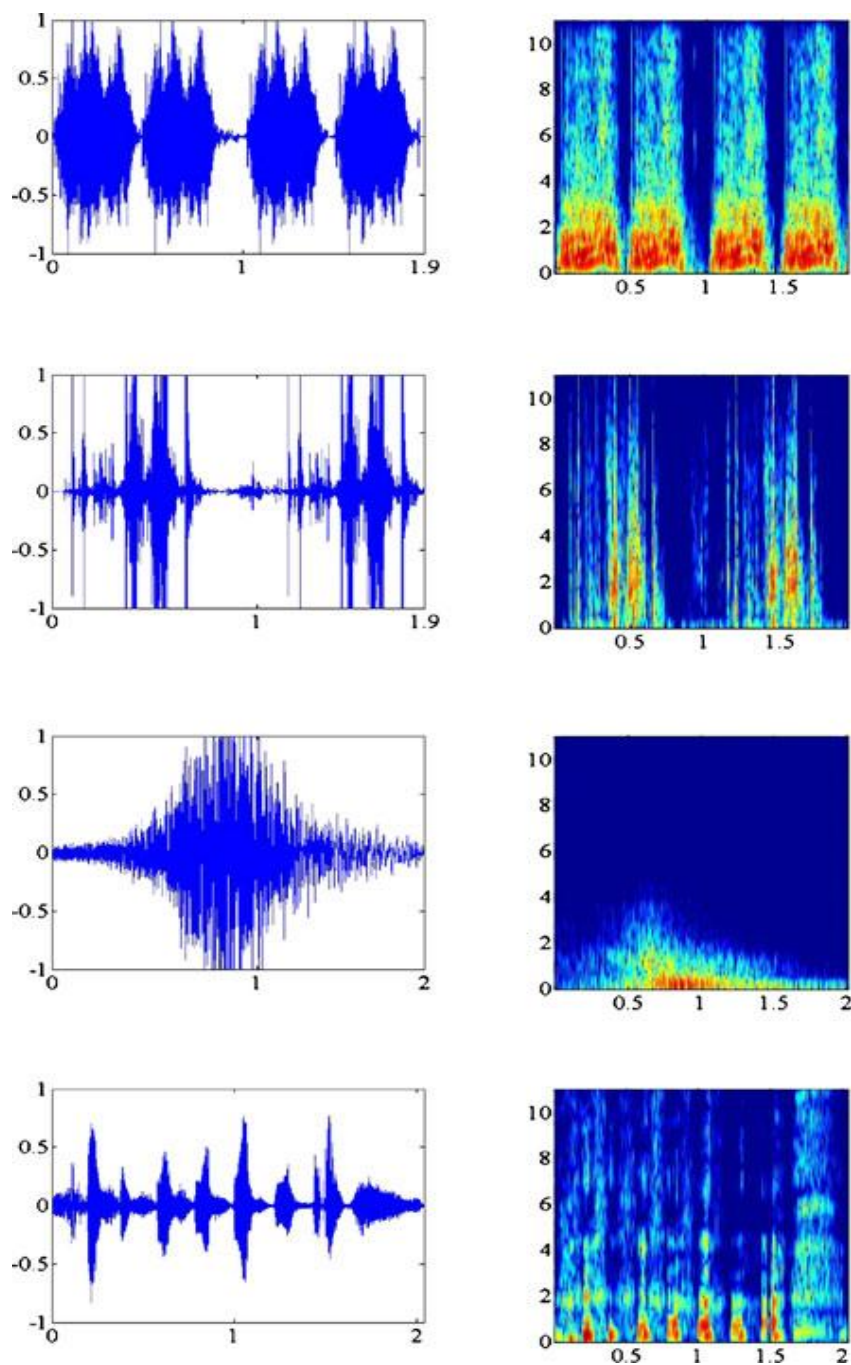
Figure 1. Left panel shows time–amplitude waveforms for some environmental sounds. Tool (saw), page turn, aeroplane and laughter show time-varying spectral structure which is shown in the right panels in the spectrographic representation using a 45 Hz Hamming window. Frequency (in Hz) is plotted on the y-axis while time (in s) is plotted on the x-axis with intensities (in dB) represented in colour. Red indicates maximum power while blue indicates minimum power. The colour index is relative to the highest and lowest intensities for each signal.

In the following sections, we identify a data set of sounds in the environment and describe them using the spectrographic representation. We find that the spectral structures, one that has a periodic or harmonic spectral distribution and the other that has noisy spectral distribution. We identify a measure to characterize such spectral structures and propose that the spectral dynamics of any sound in the environment can be described in terms of these spectral structures. We define an index to characterize 'sound complexity' in terms of the number of distinct spectral structures and estimate the

complexity of different environmental sounds. We suggest that spectral features of sounds in the natural environment could be a basis for the evolution of specialized auditory processing areas in the human brain.

2. Data Analysis

Sounds are collected from online databases and were drawn from several different classes-animal cries (e.g. cow moo), environmental sounds (telephone ring, airplane noise), and human non-verbal vocalizations (e.g. laughter). The sampling frequency of all sounds was 22,050 Hz. The sounds were pre-processed using Goldwave (version 5.10) software for noise reduction. Noise reduction is the elimination of unwanted noise, such as a background hiss or a power hum within a sound. Goldwave was also used to ensure that all sounds were matched for 2-s length.

3. Materials and Methods

As described earlier, new analysis techniques, which use joint time-frequency representation (TFR) within the limits of resolution allowed by the uncertainly principle [3] have emerged as convenient methods to describe non-stationary dynamics. For signals, where the dynamics can be considered to be stationary in short time windows, the short time Fourier transform (STFT) [3], has been found to be extremely useful. A display of the sound signal using the STFT in the time-frequency representation is called the spectrogram. A spectrogram is obtained by first partitioning the signal into small overlapping equal segments of time t and then carrying out a STFT, for each segment [3]. The STFT of a function is defined as

$$S(t,f) = \int_{-\infty}^{\infty} e^{-i2\pi f \tau} s(\tau) h(\tau - t) d\tau,$$

where s(t) is the signal, f is the frequency and h(t) is the window function. For signals where temporal resolution is required, h(t) is narrow and spectral resolution is poor. On the other hand, for good frequency resolution, h(t) is broad and provides poor temporal resolution [4]. The energy-density spectrum of STFT is defined as a spectrogram (right panel of figure 1). The spectro-temporal structure of complex sounds viewed in the spectrographic representation exhibits essentially two kinds of spectral structures: (1) harmonic and (2) noisy. The spectral structure in some regions is highly patterned (see the vertical stripes in the top right panel) suggesting periodic or harmonic structure whereas in other regions the

underlying spectral distribution is noisy (see the right panel, third from top).

A standard method to measure the amount of spectral structure in a stationary signal is the spectral flatness measure (SFM) [5]. The SFM estimates the number of peaks in the power spectrum as opposed to a flat spectrum and is defined as the ratio of the geometric mean to the arithmetic mean of the power spectrum. A distribution of the power spectrum is expressed as

$$SFM = \log \frac{\left[\prod_{f=1}^{N} S(f)\right]^{1/N}}{(1/N)\sum_{f=1}^{N} S(f)}$$

where S(f) is the magnitude of each frequency component in Hz and N is the number of FFT points used to estimate the power spectral density of s(t). For a pure tone, which has a single peak in the power spectrum and has the simplest spectral structure, SFM is 0, whereas for white noise, which has infinite peaks, SFM is 1. To expand the dynamic range it is expressed on a logarithmic scale and thus, for a pure tone, SFM is minus infinity whereas for a white noise signal, SFM is 0. Low SFM sounds are, therefore, tonal while high SFM sounds are noisy.

For non-stationary sounds, we define a time-dependent SFM(t), which estimates the spectral structure in each temporal segment. SFM(t) defined in terms of S(t,f) is obtained from the spectrographic representation as

$$SFM(t) = \log \frac{\left[\prod_{f=1}^{N} S(t,f)\right]^{1/N}}{(1/N)\sum_{i=1}^{N} S(t,f)}$$

where S(t, f) is the power associated with each frequency component is that particular temporal segment. To describe environmental sounds which have varying spectral dynamics, we propose an index of spectral variability, namely spectral structured index (SSI) in terms of the variance of SFM(t) as

$$SSI - \sum \left[\frac{SFM(t) - \frac{\sum SFM(t)}{N}}{N}\right]^2$$

where N is the number of time frames and SSI is the average spectral variance for a given signal.

We calculate SSI for different environmental sounds and propose a categorization of environment sounds

in terms of SSI. For sounds with spectral distributions fluctuating rapidly across time frames, SSI is large and we classify them as complex sounds. On the other hand, when variation in the spectral distribution across time frames is small we classify them as simple sounds. We suggest that the SSI defines degree of special complexity and can be used to categorize sounds into varying levels of complexity.

4. Results

A total of 15 sounds were analysed. To deal with silences in sounds, we extracted epochs in the sound signal where power is <1 dB and assigned them an SFM value of 0. Narrowband spectrograms were obtained using a 45Hz Hamming window for all the sounds. Figure 2 shows computed values of SFM (t) plotted on a logarithmic scale for some of the sounds. As seen in figure 2, SFM (t) does not change much across time windows for airplane noises (for example), a feature which is also reflected in the spectrographic representation (figure 1). On the other hand, for laughter, SFM (t) shows fluctuations across time windows. Thus SFM(t) follows the spectral dynamics in successive time frames.
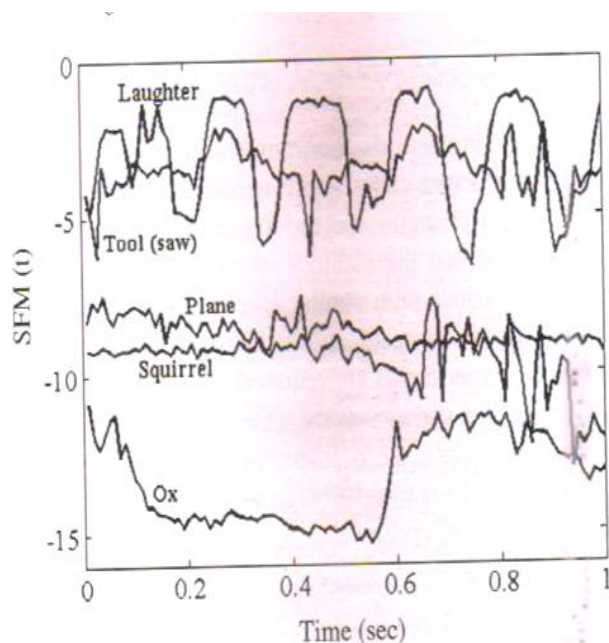

Figure 2. Plot of SFM(t) vs, time for different environmental sounds

The variation in spectral structure across time windows for different environmental sounds, as estimated by SSI, is shown in table 1. For signals with similar spectral dynamics across time windows SI<1 (airplane noise, for example), while for signals with varying spectral dynamics across time windows

SSI > 1 (laughter). We therefore suggest that, based on spectral dynamics, sounds in the natural environment may at least be classified into two categories, namely simple and complex. Signals with SSI < 1, can be classified as simple sounds, whereas sound signals with SSI > 1 can be classified as complex sounds.

Table 1: SSI for various environmental sounds (BIS Processes, BIS-communication)

Complex sounds		Simple sounds	
Cow	*1.0532*	*Tool (Saw)*	*0.219*
Doorbell	*1.1103*	*Breaking glass*	*0.3525*
Coin drop	*1.2509*	*Phone ring*	*0.423*
Crow	*1.4835*	*Ox*	*0.5219*
Laughter	*1.8827*	*Bagpipes*	*0.5747*
Chickens	*2.0167*	*Aeroplane*	*0.7471*
Crying	*2.3601*	*Horn*	*0.8361*
Squirrel	*6.9204*	*Page turn*	*0.899*

5. Conclusions

We propose a classification of sounds in environment in terms of spectral dynamics. Sounds for which the spectral structure varies slowly across time windows are categorized as simple and sounds with rapidly changing spectral dynamics are categorized as complex. Based on our results we suggest that the auditory system may adopt processing strategies that might be similar for sounds with similar spectral dynamics, which could be a crude explanation for their anatomical organization in different regions of the human brain [1]. Functional neuroimaging experiments are required to validate our proposal and are currently in progress. Our analysis shows that the spectrographic representation presents a convenient representation to describe the rich spectral dynamics of non-stationary signals. The spectral structure index (SSI) could emerge as a novel measure to study spectral complexity in physical and biological systems.

Acknowledgements

References
[1]   **O Chiry, E Tardif, P J Magistretti and S Clarke**, *Eur.J. Neurosci.* **17,** 397 (2003).

[2]   **R J Zatorre, P Belin and V B Penhune**, *Trends in Cog. Sci.* **6,** 37 (2002).

[3]   **L Cohen,** *Time frequency analysis* (Prentice-Hall, New Jersey, 1995).

[4]   **R Reddy, V Ramachandra, N Kumar and Nandini C Singh**, *Biol. Cybern.* **100(4),** 299 (2009).

[5]   **N S Jayant and P Noll**, *Digital coding of waveforms* (Prentice-Hall, 1984).