Structure and correlation in the detection of melodic sequences

Jason Lee*, Frances Leung, Janakan Satkunasingham*, Winnie Wong** and Willy Wong*,****

*Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto **Institute of Biomaterials and Biomedical Engineering, University of Toronto

Introduction

It is known that higher animals perform well in the detection and recognition of familiar patterns under noisy conditions. Why is this so? We postulate that it is the inherent structure or redundancy in patterns that allows the sensory system to reconstruct familiar patterns from a noisy scene. In an earlier study by Wong and Barlow[1], it was found that for a well-known melody, the amount of signal and tolerable noise follows a power relationship, with the exponent approximately equaling two. I.e., doubling the amount of melody notes results in a four-fold increase of tolerable noise. This is to be contrasted with the results obtained from detecting signals without structure where tolerable noise level goes as a *linear* function of signal level (exponent equals one).

This experiment continues this study by measuring the detection performance of randomly generated signals of varying degrees of structural content. Fractal melodies were used and the degree of self-correlation (structural content) was changed in a continuous manner. If our postulate is correct, we would expect the exponent relating the tolerable noise and signal level to increase monotonically when the structural content of the signal increases.

This paper provides an overview of the experiment and discusses the hypothesis used in the experiment. Results of the experiment will be shown during the oral presentation.

Generating fractal melodies

It is known that signals with $1/f^{\beta}$ power spectrum, where β is constant and non-negative, have fractal properties [2]. Increasing β increases the self-correlation within the signal and vice versa. In particular, studies have shown that most music possess a 1/f power spectrum [3], [4]. Attempts have been made at generating music from $1/f^{\beta}$ spectrum. According to Voss [3], melodies generated from $1/f^2$ spectra lacks diversity (too correlated) and $1/f^0$ is too random (uncorrelated), while 1/f melody sounds more pleasant to most people. Signals with $1/f^0$, 1/f and $1/f^2$ spectrums are most commonly associated with white, pink and brown noise processes respectively.

In this experiment, we assume that the self-correlation of the signal represents its structural content, and by varying the parameter β, we can generate random melody with varying degree of structural content. In generating the melody, the discrete Fourier transform of the melody is first constructed with the following relationship to ensure $1/f^{\beta}$ power spectrum.

$$|X(k\,\Delta f)|^2 \propto 1/k^{\beta}, \quad k = 0, 1, 2, 3\ldots\ldots 511$$

The phase of the Fourier transform is randomly chosen between 0 to 2 PI for each value of k. The Fourier transform of negative frequency is set as the complex conjugate of that of positive frequency to ensure a real time series. FFT is then performed on the Fourier transform to generate the signal in time series with 1024 samples. The samples are then quantized into 21 levels, each corresponding to a musical note in the C major scale over a 3-octave range. Random rhythm is also added to the melody to make the melody sound more musical.

Experiment

In the experiment, the subject is presented with either a series of noise notes generated uniformly over the 3 octave range (a noise-only trial) or a fractal melody masked by noise notes (a signal-plus-noise trial). Both the noise-only and the signal-plus-noise trials have the same total number of notes. The melody is generated dynamically and differs from trial to trial. The subject chooses a response of either *noise-only* or *signal-plus-noise* after listening to an audio segment of approximately 10 secs in duration. Experiments are carried out over different values of β and over different lengths of the melodic sequence (both parameters are varied separately). The noise level or total number of noise notes changes dynamically, trial-by-trial, according to a Bayesian adaptive algorithm called QUEST [5]. The answer of the subject is recorded, and the percentage of correct answer is used to construct the psychometric function. We then noted the value of tolerable noise permitting 75% correct answers and tabulated such values for different values of beta and total signal or melodic length.

Discussion

The study by Wong and Barlow [1] showed that for a well-known melody, doubling the amount of melody notes results in a four-fold increase of tolerable noise. Hence, a log-log plot of tolerable noise level and signal length results in a line of slope two, meaning a power relationship with exponent equal to two. This is in contrast with the linear relationship between tolerable noise and signal for an unstructured signal, where a line of slope one was found on a log-log plot. Similar results have also been found in other recognition tasks in other modalities [6]. We postulate that it is the inherent structure of the signal that gives rise to the superior signal detection performance (a higher slope on a log-log plot). However, their experiment used the same tune throughout the experiment, and one may suggest that the *a priori* knowledge about the melody alone can contribute to the superior performance achieved by the subjects.

The major difference between this experiment and the above experiment is that the subject does not know ahead of time what the melody will be. Hence memory does not play a part in the detection of the melody, and the major cue is the structure, or the correlation in the melody. If the inherent structure of the melody can achieve superior performance in detection, the slope of the log-log plot of noise vs signal will increase when β, hence the correlation within the signal increases (see figure 1).

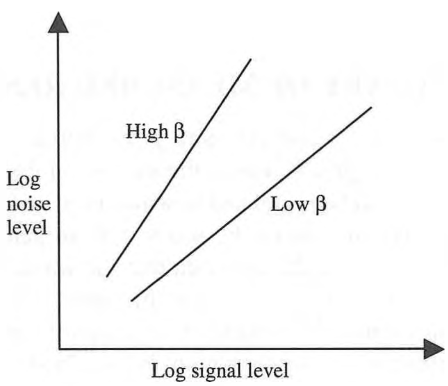

Figure 1: Expected Result from different β

References

[1] H. Barlow and W. Wong, "Tunes and templates," *Nature* vol. 404, pp. 952 – 953, Apr. 2000.

[2] P. Addison, "The colour and power of noise," in *Fractals and Chaos: An Illustrated Course*, London: Institute of Physics Publishing, 1997, pp. 80 – 81.

[3] R. F. Voss and J. Clarke, "'1/f noise' in music: Music from 1/f noise," *J. Acoust. Soc. Am.*, vol 63(1), pp. 258 – 263, Jan. 1978.

[4] R. F. Voss and J. Clarke, "'1/f noise' in music and speech," *Nature*, vol. 258, pp. 317 – 318, Nov. 1974.

[5] A. B. Watson and D .G. Pelli, "QUEST: A Bayesian adaptive psychometric method," *Perception and Psychophysics*, vol. 33(2), pp. 113 – 120, 1983.

[6] P. Neri, M. C. Morrone and D. C. Burr, "Seeing biological motion," *Nature*, vol. 395, pp. 894 – 896, Oct. 1998.