Available online at www.sciencedirect.com
SCIENCE DIRECT®
BioSystems 79 (2005) 199–206
BioSystems
www.elsevier.com/locate/biosystems

Pitch detection of dynamic iterated rippled noise by humans and a modified auditory model

Susan Denham

Plymouth Centre for Theoretical and Computational Neuroscience, University of Plymouth, Plymouth, UK

Abstract

Iterated ripple noise (IRN) is a broadband noise with temporal regularities, which can give rise to a perceptible pitch. Since the perceptual pitch to noise ratio of these stimuli can be altered without substantially altering their spectral content, they have been useful in exploring the role of temporal processing in pitch perception [Yost, W.A., 1996. Pitch strength of iterated rippled noise, J. Acoust. Soc. Am. 100 (5), 3329–3335; Patterson, R.D., Handel, S.,Yost, W.A., Datta, A.J., 1996. The relative strength of the tone and noise components in iterated rippled noise, J. Acoust. Soc. Am. 100 (5), 3286–3294]. A generalised IRN algorithm is presented, in which multiple time varying temporal correlations can be defined. The resulting time varying pitches are perceptually very salient. It is also possible to segregate and track multiple simultaneous time varying pitches in these stimuli. Temporal auditory models have previously been shown to account for the perception of IRNs with static delays [Patterson, R.D., Handel, S.,Yost, W.A., Datta, A.J., 1996. The relative strength of the tone and noise components in iterated rippled noise, J. Acoust. Soc. Am. 100 (5), 3286–3294]. Here we show that some simple modifications to one such model [Meddis R., Hewitt, M.J., 1991. Virtual pitch and phase sensitivity of a computer model of the auditory periphery I. Pitch identification, J. Acoust. Soc. Am. 89, 2866–2882] allow it to track moving correlations, and also improve its performance in response to static correlations.

Keywords: Dynamic iterated ripple noise; Pitch perception; Temporal processing

1. Introduction

Rippled noise, also known as delay and add noise, is created by delaying a copy of the original noise, multiplying it by a gain factor, and adding it to the original (Bilsen, 1966); Iterated rippled noise (IRN) is produced by repeating the operation. An IRN has a noise like waveform in which the degree of temporal regularity can be varied, and results in the compound perception of a ‘buzzy tone embedded in a noisy hiss’ (Patterson et al., 1996). IRN(d, g, n) is specified by three parameters; delay, d, gain, g, and number of iterations, n:

$$\begin{aligned} &y_i(t) = y_{i-1}(t) + g y_{i-1}(t-d); \quad \text{for } i = 1, \ldots, n \\ &y_0(t) = x(t), \quad \text{the input signal} \end{aligned} \tag{1}$$

The pitch of the tone corresponds to the inverse of the delay, and the relative strength of the tone to noise increases with increasing number of iterations. IRNs

E-mail address: sue@pion.ac.uk.

0303-2647/$ – see front matter
doi:10.1016/j.biosystems.2004.09.008

are useful for exploring temporal processing in the auditory system, since they vary considerably perceptually while having the same total energy and similar excitation patterns (Patterson et al., 1996). The perceived pitch is predicted by the lag of the first 'significant peak' in the autocorrelation function (Yost, 1996), or the summary autocorrelation function calculated in temporal models of auditory pitch perception (Meddis and Hewitt, 1991; Patterson et al., 1995). The perceptual tone/noise ratio was shown to be related to the height of the chosen peak in the autocorrelation function, but as Hartman points out this is not a very convincing model since the autocorrelation peak is extremely sharp for IRNs and very broad for pure tones, which is at odds with the relative saliency of the two types of pitch (Hartmann, 1998). The auditory models are better though with a sharp and very prominent peak in response to pure tones, and a far less prominent peak in response to IRNs; as shown in Fig. 1 below.

This paper describes preliminary investigations and modelling work using a generalised form of IRN, and is organised as follows. Firstly, the dynamic IRN (DIRN) stimulus is described and aspects of the pitch perception of such stimuli are discussed in relation to previous pitch studies. Extensions to an auditory model of pitch perception which enable it to the track time-varying pitches in DIRNs are then presented. Finally, the generalisation to multiple dynamic pitches is discussed. It is hoped that these stimuli will prove useful in investigations into the mechanisms of human pitch perception.

2. Dynamic iterated ripple noise

The IRN algorithm can be generalised to allow time dependent delays. Instead of adding a copy of the original noise back at some fixed delay, a mapping function is created so that each point can in principle be delayed by a different amount. The dynamic IRN is defined by:

$$\begin{aligned} y_i(t) &= y_{i-1}(t) + g y_{i-1}(t - d(t)); \quad \text{for } i = 1, \ldots, n \\ y_0(t) &= x(t), \quad \text{the input signal} \end{aligned} \tag{2}$$

In informal listening it was found that the resulting dynamic pitches were far more salient than the static ones, described above. Even with only one iteration, the upward sweep in the example below in Fig. 2 is easy to hear. It is important to note that the when the mapping function is applied to the original noise sample, the pitch sweep is not audible it only becomes audible

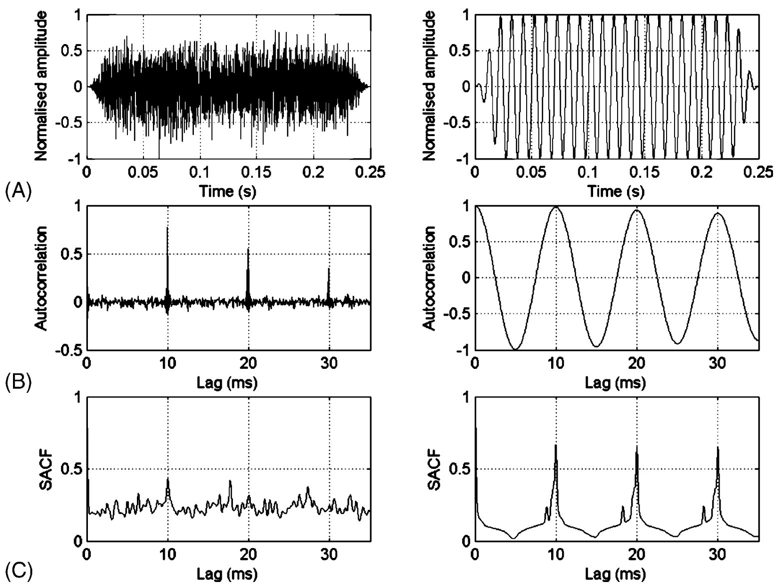

Fig. 1. Left column shows (A) an IRN stimulus (created using a 10 ms delay, gain 1, and 4 iterations), (B) the autocorrelation function for this stimulus, and (C) the auditory model (Meddis and Hewitt, 1991) summary autocorrelation (SACF) response. The right column shows the corresponding plots for a pure tone of the same pitch.