A FAST REAL-TIME AUDITORY-NERVE MODEL

Satish G Iyengar[1], Jayant Datta[1] and Laurel H Carney[1,2]

[1]Dept. of Electrical Engineering and Computer Science, [2]Dept. of Biomedical and Chemical Engineering, Syracuse University, Syracuse, NY

ABSTRACT

We present the details of a real-time design and implementation of an auditory-nerve (AN) model on Motorola's DSP56367, a 150 MIPS audio digital signal processor (DSP). Sensory processing is highly nonlinear and time variant and so are the computational models that are derived to simulate these processes. Complex signal processing models have been derived to predict essential AN properties. However, their use is limited due to large processing time requirements. Our DSP solution makes fast real-time AN simulations possible and thus allows an experimentalist to change model parameters *on the fly*. This can speed up testing of several sound coding/decoding strategies hypothesized to occur in the human brain.

***Index Terms*—** DSP, auditory-nerve, real-time, MIPS, time-varying bandpass filters

1. INTRODUCTION

The goal of this study is to build a hardware prototype of an AN model to make fast real-time model simulations possible and thus expedite the study of neural sound-coding strategies. It is well known that sound processing by the auditory periphery (including AN) is highly nonlinear [3], [4]. These nonlinearities include compressive changes in gain and bandwidth as a function of stimulus level (cochlear compression), associated variations in the phase of the phase-locked responses, AN saturating nonlinearities and two-tone suppression [4]. In addition to the above nonlinear properties, frequency modulations *(df/dt)* or 'glides' have been reported in the impulse response estimates of AN fibers [2]. Further, these frequency modulations depend on the characteristic frequency (CF)[a] of the fiber in that a downward glide (*df/dt* < *0*) is seen for low CF fibers, no glide is seen for the mid CF fibers whereas an upward glide (*df/dt* > *0*) is seen for the high CF fibers (see Fig. 5). Such model nonlinearities and time varying properties pose numerous challenges in the real-time design and implementation of the system. In this paper, we describe in detail, a DSP-based approach to such highly nonlinear system implementation.

In the past, several approaches have been used to realize AN models in real-time such as VLSI-based models [6] and wave digital filter (WDF) implementations[1]. Though recent advances in these approaches have made it possible to include cochlear compression, the ability to model the instantaneous frequency glides is missing (e.g. [6]) or is erroneous for low CF fibers (e.g. [1]). To our knowledge, the Tan and Carney (2003) model is the only model at present that successfully incorporates level-independent frequency glides along with level-dependent compression and suppression. Also, the frequency analysis along the length of basilar membrane (BM) in the VLSI and the WDF models is simulated using a cascade of filters/resonators along a transmission line. The tuning properties at different places of the BM are thus achieved as a cumulative effect of all the resonators along the line. Although, this approach is closer to physiology, it is impossible to adjust the properties of only the required CF auditory-nerve fibers/filters. The entire cochlear model has to be implemented even if a small frequency range is of interest. The Tan and Carney (2003) AN model provides the flexibility to control and vary individual frequency band-specific filter parameters. This flexibility is attractive in hearing-aid applications where the DSP AN model is used as the front-end for speech processors.

2. THE AUDITORY-NERVE MODEL

In this section, an overview of the Tan and Carney (2003) AN model is given. The block diagram of the model is shown in Fig. 1. The model consists of four blocks: a fourth order linear time invariant (LTI) filter that simulates the middle ear [5], a time-varying band pass filter (BPF) as the signal path, a nonlinear control path whose output controls the gain and bandwidth of the signal path filter and the inner hair cell (IHC) and synapse block whose output is the instantaneous AN fiber discharge rate to the input acoustic stimuli. The focus of this paper is the real-time implementation of the signal path and the control path time-varying filters.

[a]Characteristic Frequency (CF) :
The sound frequency that produces the largest response in an AN fiber is called its characteristic frequency (CF).

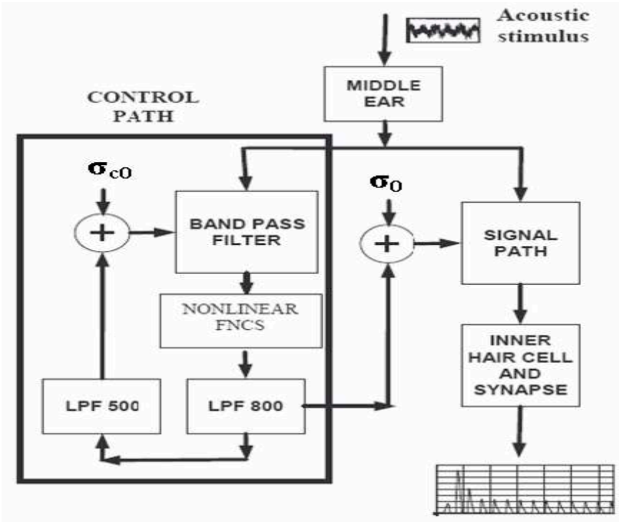

Fig.1. Schematic of the Tan and Carney (2003) AN model.

2.1. The Signal Path

The Signal path is a twentieth order BPF centered at CF of an AN fiber. The filter gain and bandwidth are level dependent and vary on a sample-by-sample basis to capture the compressive nonlinearity of the healthy human ear. The gain-bandwidth dynamics are achieved by changing the real part of the poles in the Laplace domain. The poles and zeros are initialized for the 'quiet' condition, i.e., for low input sound levels. As the input level becomes more positive, the control path generates a 'control signal' (output of LPF 800 in Fig. 1) that shifts the poles away from the imaginary axis, thereby decreasing the filter gain and broadening the filter bandwidth. Conversely, as the input level becomes more negative, the poles are shifted towards the imaginary axis, thereby increasing the filter gain and decreasing the filter bandwidth. These time-varying dynamics achieve the cochlear compression phenomenon.

2.2. The Control Path

The control path determines the signal path pole locations at any instant based on the input signal level. It includes a sixth order BPF that limits the range of frequencies affecting the signal path dynamics. The filter is followed by nonlinear functions and low pass filters that limit the dynamics of the compressive nonlinearity. The nonlinear functions are derived to match several properties from AN physiological studies [4]. Also, as in the signal path, the control path BPF gain and bandwidth are input-dependent and vary at each time step.

3. SYSTEM DESIGN

3.1. System Hardware

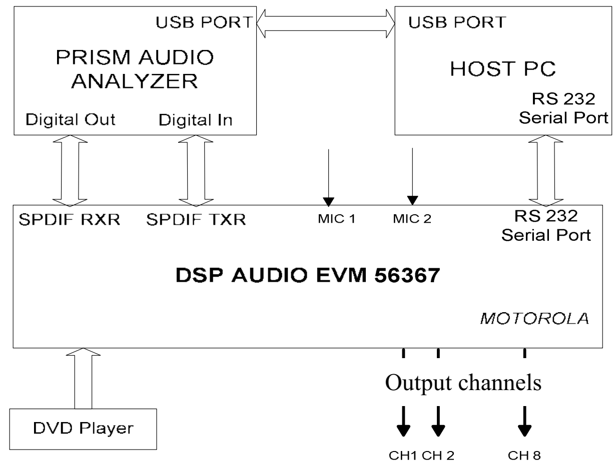

Fig.2. Hardware set up and the DSP interfaces

Motorola's DSP56367 is central to the system layout and is interfaced to an audio analyzer (*dScope III Prism Sound),* a personal computer (PC) and a DVD player. The audio analyzer can provide several test inputs to the DSP-model and is capable of performing several run-time analyses (e.g. FFT, step responses, etc.). The DVD player is used to provide real-world speech/sound input signals. The DSP is run in the Software Architecture (SA) mode [7], a framework which enables highly modular software development and rapid prototyping. The SA mode also allows running multiple AN CFs in parallel and/or change CFs on the fly. The PC includes several software tools like MATLAB for offline signal analyses, DSP simulator for debugging the DSP assembly code and the DSP assembler to generate the assembly *(.asm)* files.

3.2. System Firmware

Motorola's DSP56367 is a 24-bit fixed-point processor with a 56-bit accumulator. The permissible data representation range is thus limited to the closed-open interval [-1.0, 1.0), and follows the two's complement arithmetic rules. Thus each block in the model has to be scaled to ensure there is no clipping at any stage of the model. Our system is designed to treat a full-scale DSP signal as the equivalent of a 90 dB SPL (sound pressure level) signal.

The Tan and Carney (2003) model implementation presents a number of interesting challenges due to several of its properties. First, all filters designed to match physiological characteristics are allowed to have gains greater than unity (positive gains). This further emphasizes the need to scale their transfer functions (frequency responses) so that they are *'well-behaved'*. Second, the filter coefficients vary on a sample by sample basis and are largely dependent on the input signal. Thus, the DSP is faced with the task of computing these coefficients '*on the fly*' in addition to filtering. Third, the factor with which the newly derived transfer function (corresponding to the new filter coefficients at each time step) is required to be scaled is a function of time. Thus, it is not possible to compute and scale the coefficients by the required amount beforehand (offline). These methods have to be a part of the DSP firmware running in real-time; thus ensuring normalized transfer functions at all time. Fourth, the normalization routine is further complicated by the presence of several nonlinear functions. We now describe the model implementation details and discuss methods used to circumvent the issues mentioned above.

3.2.1. Realizing time-varying filters in real-time

All filters are implemented as a cascade of second order (biquad) sections in the z-domain using a sampling rate of 48 kHz. For example, the twentieth order signal path filter is implemented as a cascade of ten biquad sections. Fig. 3 shows a schematic of the signal path filter pole-zero pattern on the Laplace plane. It also illustrates the movement of the poles at each time step as a function of input signal. With no signal at its input, the set of poles closest to the imaginary axis is at 'Xquiet'. We call this the 'quiet condition'. As the signal amplitude becomes more positive, the poles are pushed away from the imaginary axis, thus decreasing the overall gain of the filter. Poles are pulled in the opposite direction (towards the imaginary axis) as the signal amplitude becomes more negative. Let 'Xmax' denote the pole-location farthest from the imaginary axis. We call this the 'max' condition. 'Xmin' denotes the pole-location closest to the imaginary axis and we term this as the 'min' condition. The filter has the largest gain in the 'min' condition.

The signal path filter, if normalized for this *'min'* condition, avoids the possibility of clipping. However the problem with this approach lies in estimating the maximum possible gain since it is largely input signal dependent. We know that the LPF 800 Hz output in the control path determines the signal path filter pole-location at each time step. Thus, we combine all (positive) gains from the ME through the control path blocks and estimate the signal path filter pole-locations closest to the imaginary axis as follows:

$$\hat{X}_{\min} = \sigma_{quiet} - \hat{\sigma}_{\min}$$

where $\hat{\sigma}_{\min}$ = (Middle Ear gain) * (Control path BPF gain) * (min(output of control path nonlinearity)) * (LPF 800 Hz gain)

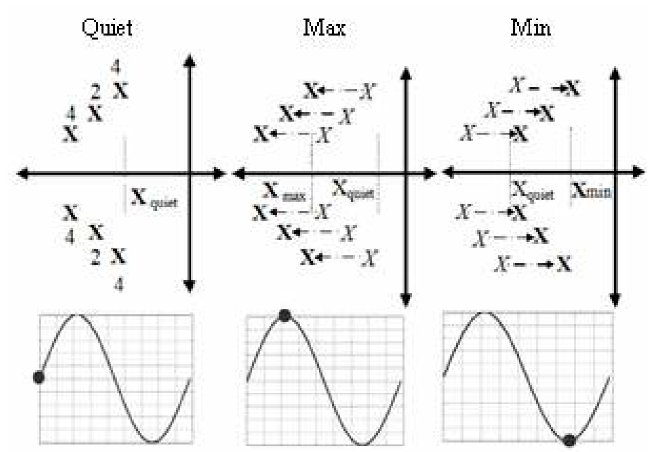

Fig.3. Signal Path Filter Dynamics

A simlar scaling method is used for the control path BPF. Now, given the pole-zero locations on the Laplace plane (at each time step), filter coefficient expressions are derived using bilinear transform for discrete-time filter implementation in the z-domain. The discrete-time filter coefficients for all the estimated control signal points were found to be 'nearly-linear' and monotonic (as a function of control signal) for all CFs. Hence, they are calculated offline and stored as a look up table (LUT) in the DSP memory. The processor is programmed to linearly interpolate between these 'seed' values in the LUT to compute the actual biquad coefficients in real-time at every sampling period.

4. RESULTS

In this section we present a comparison between the real-time simulations and the offline AN model implemented in C.

4.1. Click Responses

Fig. 4 is a plot comparing the real-time and offline model click responses at different levels. The plot shows responses for fibers with CF 500 Hz (low CF), 1100 Hz (mid CF) and 2000 Hz (high CF). All the output signals are normalized by their respective maximum values. This is to enable comparisons between the DSP and the C models. Also, it can be seen that the responses to a 60 dB SPL click show quantization (finite word length) effects. Performance loss due to quantization is pronounced due, in part, to the low energy content of the input click and also to model nonlinear effects.

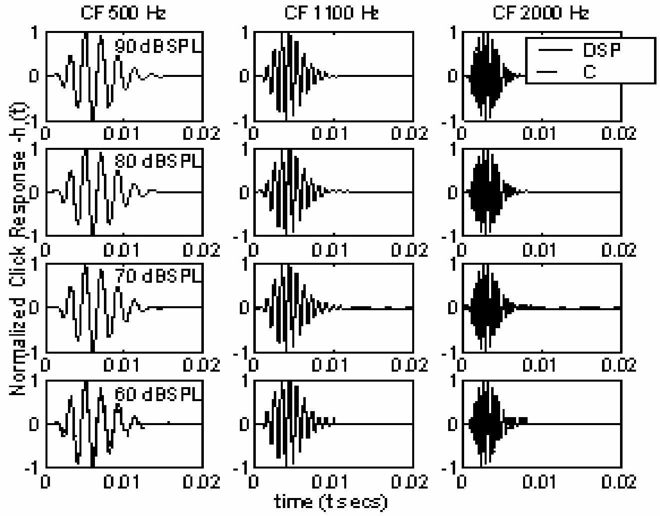

Fig.4. Click responses from the DSP model compared to its offline implementation (C model)

4.2. Frequency glides

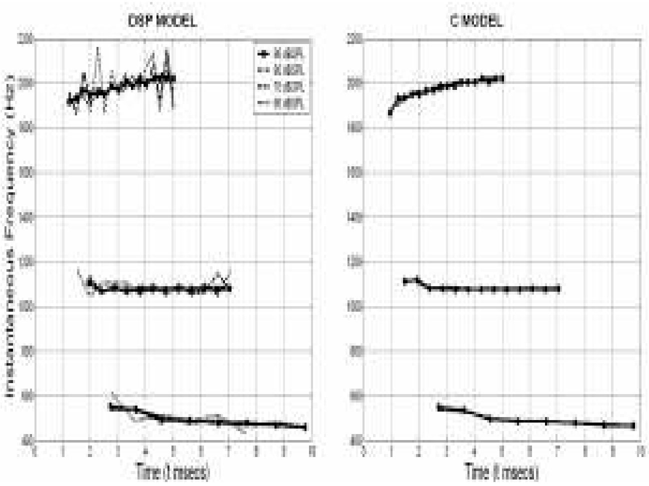

Fig.5. Instanteneous Frequency plots for 500 Hz, 1100 Hz and 2000 Hz CF AN fibers.

As shown in Fig.5, the low-frequency CF fiber shows a downward glide; the mid-frequency CF fiber shows a constant glide and the high-frequency CF fiber shows an upward glide. Also, the glides have to be independent of the input sound level. Though the instantaneous frequency curves across input levels (constant CF) show the same trends for the real-time model, they do not exactly match each other. This error can be attributed to higher quantization errors with decreasing click amplitudes. Low energy input clicks cause notable loss of bits as the stimulus is processed through different stages of the model. Such a quantized response causes errors in instantaneous frequency estimation.

4.3. Cochlear Compression/Gain Control mechanism

Cochlear compression is one of the main property of a healthy ear [3]. Fig. 6 plots the root mean square (RMS) value of the signal path output as a function of the input sound pressure level for several CFs.

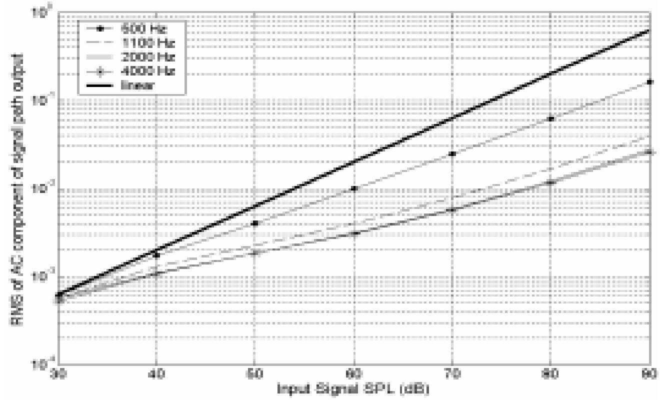

Fig.6. Compression in the signal path output

The input is a pure tone at CF. It can be seen from the plot that as the input level increases the signal path output gets compressed. A linear curve is also included for reference. The plot shows higher amount of compression for higher CF in agreement with physiological results [3].

5. REFERENCES

[1] C. Giguere, and P.C. Woodland, “A computational model of the auditory periphery for speech and hearing research. I . Ascending Path”, *J. Acoust. Soc. Am., Vol. 95 (1)*, pp. 331-342, 1994.

[2] L.H. Carney, M.J. McDuffy, and I. Shekhter, “Frequency glides in the impulse responses of auditory-nerve fibers”, *J. Acoust. Soc. Am., Vol.105*, pp. 2384-2391, 1999.

[3] M.A. Ruggero, N.C. Rich, A. Reico, S.S. Narayan and L. Robles, “Basilar-membrane responses to tones at the base of the chinchilla cochlea”, *J. Acoust. Soc. Am., Vol. 101*, pp. 2151-2163, 1997.

[4] Q. Tan, and L.H. Carney, “A phenomenological model for the responses of auditory-nerve fibers: II. Nonliner tuning with frequency glide”, *J. Acoust. Soc. Am., Vol. 114 (4)*, pp. 2007-2020, 2003.

[5] Rosowski, J.J., Models of External- and Middle-Ear Function, *in Auditory Computation,* Springer-Verlag, New York, 1996.

[6] R. Sarpeshkar, R.F. Lyon and C.F. Mead, “An analog VLSI cochlea with new transconductance amplifiers and nonlinear gain control”, *Circuits and Systmes, ISCAS International Symposium on, Vol. (3)*, pp. 12-15, 1996.

[7] T.C. Chen, J. Datta, and B.A. Karley, “An open DSP software architecture for consumer audio”, *IEEE Trans. on Consumer Electronics, Vol.45,* pp. 1253-1258, 1999.