Zhak, S., S. Mandal, and R. Sarpeshkar, "A proposal for an RF Cochlea", invited paper, Proceedings of the Asia Pacific Microwave Conference, New Delhi, India, December 15-18, 2004.

Author’s Information

Corresponding Author

Professor Rahul Sarpeshkar
Room 38-294
Massachusetts Institute of Technology
77 Massachusetts Av
Cambridge, MA 02139-4307
USA

Phone: 1-617-253-4872
Fax: 1-617-253-5210
E-mail: rahuls@mit.edu

Technical Contents

The following topics from the list of conference topics best describe the contents of this paper.

1. Solid State Devices and Circuits
2. Monolithic Integrated Circuits
3. Passive Devices and Circuits
4. Medical and Biological Applications

Abstract

The human ear is one of the greatest marvels of nature: the inner ear or cochlea performs at least 1GFLOPS of real-time sensing, filtering, amplification, gain control, and data-compression computations in a tiny volume. The ear consumes about 14 µW of power while running on a 150mV battery; it could run on a pair of AA batteries for 15 years. The ear can sense 0.05 angstroms of eardrum motion at its best frequency and has an input dynamic range that spans 12 orders of magnitude in sound intensity. The ear operates over a frequency span of about 3 decades (10 octaves). Our ears report information with enough fidelity such that the auditory system can make a sound-location discrimination that corresponds to an inter aural time difference of a few microseconds even though the component parts of the system have 1-10 millisecond time constants. These impressive specifications were produced by at least 220 million years of evolution.

The overall architecture of the ear is much like that of a wideband radio receiver with many parallel outputs systematically spanning a range of corresponding input carrier frequencies. Alternatively, the ear functions as a sensitive wavelet-based spectrum analyzer. Of course, the ear’s structures are designed to sense signals at audio frequencies rather than at radio frequencies. The outer ear is composed of a pinna and eardrum that receive sound much like a complex directional antenna. The middle ear is composed of three bones that implement an impedance matching network to match the low-pressure, high-velocity reception at the eardrum to the higher-pressure, lower-velocity capabilities of the inner ear or cochlea. The cochlea is a distributed transmission line with inductors implemented by its fluid mass and capacitors implemented by its exponentially-tapered-in-stiffness basilar membrane. The tapered transmission line successively low-pass filters the incoming signal such that late sections of the transmission line primarily amplify and sense low frequencies, and early sections of the transmission line primarily amplify and sense high frequencies. In other words, the cochlea transforms frequency to spatial position. The signal attenuation in the transmission line due to fluid viscosities and other absorptive processes is partially compensated for by electromechanical nonlinear amplifiers implemented by the cochlea’s outer hair cells. The signals at various points on the transmission line are demodulated by peak-detecting inner hair cells and reported to the brain to yield a wideband wavelet-like spectral output.

Nonlinearity and gain control in the distributed outer-hair-cell amplifiers implement compression of the 120dB input dynamic range to a narrower 40dB output dynamic range in the auditory nerve fibers. The gain control is local allowing for strong amplification of soft sounds at one frequency with reduced amplification of loud sounds at another frequency ensuring good audibility. The nonlinear tone-to-tone suppression and gain-control properties of the ear allow good preservation of spectral peaks in the signal, naturally enhancing channel outputs with high signal-to-noise with respect to neighboring channels that have a reduced signal-to-noise ratio. The ear’s algorithm for enhancing signals in noise allows it to serve as the state-of-the-art front end for speech recognition in noise, an area where humans still outperform machines by two orders of magnitude in spite of decades of research.

Both biology and electronics suggest that cochlea-like structures represent a very promising direction for designing ultra-low-power wideband RF amplifiers and spectral analyzers. This paper describes an architecture for realizing such a wideband amplification, spectral analysis and enhancement system at RF frequencies that approximates the behavior of the biological cochlea. Preliminary simulation data is presented.

A PROPOSAL FOR AN RF COCHLEA

S. M. Zhak, S. Mandal and R. Sarpeshkar
Research Laboratory of Electronics, Massachusetts Institute of Technology
Cambridge, MA 02139, USA
E-mail: rahuls@mit.edu

This paper describes an architecture for realizing a wideband amplification, spectral analysis and enhancement system at RF frequencies that approximates the behavior of the biological cochlea. Preliminary simulation data is presented.

1. Introduction

The human ear is one of the greatest marvels of nature: the inner ear or cochlea performs at least 1GFLOPS of real-time sensing, filtering, amplification, gain control, and data-compression computations in a tiny volume [1]. The ear consumes about 14 μW of power while running on a 150mV battery; it could run on a pair of AA batteries for 15 years. The ear can sense 0.05 angstroms of eardrum motion at its best frequency and has an input dynamic range that spans 12 orders of magnitude in sound intensity. The human ear operates over a frequency span of about 3 decades (10 octaves). Our ears report information with enough fidelity such that the auditory system can make a sound-location discrimination that corresponds to an inter aural time difference of a few microseconds even though the component parts of the system have 1-10 millisecond time constants. These impressive specifications were produced by at least 220 million years of evolution.

The overall architecture of the ear is much like that of a wideband radio receiver with many parallel outputs systematically spanning a range of corresponding input carrier frequencies. Alternatively, the ear functions as a sensitive wavelet-based spectrum analyzer. Of course, the ear's structures are designed to sense signals at audio frequencies rather than at radio frequencies. The outer ear is composed of a pinna and eardrum that receive sound much like a complex directional antenna. The middle ear is composed of three bones that implement an impedance matching network to match the low-pressure, high-velocity reception at the eardrum to the higher-pressure, lower-velocity capabilities of the inner ear or cochlea. The cochlea is a distributed transmission line with inductors implemented by its fluid mass and capacitors implemented by its exponentially-tapered-in-stiffness basilar membrane. The tapered transmission line successively low-pass filters the incoming signal such that late sections of the transmission line primarily amplify and sense low frequencies, and early sections of the transmission line primarily amplify and sense high frequencies. In other words, the cochlea transforms frequency to spatial position. The signal attenuation in the transmission line due to fluid viscosities and other absorptive processes is partially compensated for by electromechanical nonlinear amplifiers implemented by the cochlea's outer hair cells. The signals at various points on the transmission line are demodulated by peak-detecting inner hair cells and reported to the brain to yield a wideband wavelet-like spectral output.

Nonlinearity and gain control in the distributed outer-hair-cell amplifiers implement compression of the 120dB input dynamic range to a narrower 40dB output dynamic range in the auditory nerve fibers. The gain control is local allowing for strong amplification of soft sounds at one frequency with reduced amplification of loud sounds at another frequency ensuring good audibility. The nonlinear tone-to-tone suppression and gain-control properties of the ear allow good preservation of spectral peaks in the signal, naturally enhancing channel outputs with high signal-to-noise with respect to neighboring channels that have a reduced signal-to-noise ratio. A signal processing architecture for companding and spectral enhancement inspired by the tone-to-tone suppression, gain control, and masking properties of the biological cochlea was described in [2]. The ear's algorithm for enhancing signals in noise allows it to serve as the state-of-the-art front end for speech recognition in noise, an area where humans still outperform machines by two orders of magnitude in spite of decades of research.

The energy efficiency of the ear arises largely because of four reasons: 1) It uses a custom analog electromechanical physical architecture to implement its complex dynamical system rather than using millions and millions of switches as current digital computers do. 2) The architecture uses feedback and nonlinearity to adapt to signals over different intensity levels enabling wide-dynamic-range operation with low power. 3) Distributed traveling-wave-amplifier architectures such as those in the cochlea represent the state-of-the-art in amplification in terms of power efficiency for a given gain-bandwidth product; the transmission lines in the cochlea are tapered in contrast with most electronic implementations. 4) The ear's exponentially tapered transmission-line-architecture allows it to convert a spectral-analysis problem that scales quadratically in the order of its filters to one that scales linearly in the order of its filters [3]. The filters implemented by the ear are of extremely high order with slopes of at least 240dB/octave having been measured. A bank of such filters, even those considerably more modest than those implemented in the ear, would be extremely expensive to implement in a conventional digital paradigm and extremely power hungry. Thus, this paper will focus primarily on analog microelectronic implementations of cochlea-like structures for real-time and low-power spectral analysis at RF frequencies.

2. Architecture

We want to exploit the clever computational ideas of the cochlea to build an “RF traveling-wave cochlea” in silicon. Such a signal processor would be tremendously useful for building an ultra-low-power wideband wide-dynamic-range adaptive spectral analysis system. MATLAB simulations of a novel traveling-wave architecture [4] suggest that such processors can indeed be built at audio frequencies and that it would work much better than the state-of-the-art cochlea in [1] whose input-output curves were too compressive, and whose compression characteristics were not easily programmable. We will now focus on the related problem of how such traveling-wave processors (with and without sophisticated gain control) can be built at RF frequencies.

Figure 1 shows a traveling-wave cochlea architecture using an exponentially tapered transmission line [5, 6]. The inductors that model the fluid in the cochlea are now real inductors and the capacitors and resistors that model the springs, damping and active outer hair cell (OHC) amplifiers in the cochlea are now implemented by transconductances *G* and current or voltage-mode filters. Such a transmission line model is capable of emulating some of the sophisticated signal processing of the cochlea at RF frequencies. This model has been successfully simulated in the GHz range (see Section 3) suggesting that our ideas can be made to work in a real circuit. In fact, in electronics simple distributed traveling-wave architectures have long been used to build very energy-efficient wideband amplifiers [7]. However, distributed amplifiers in electronics are not as sophisticated as the nonlinear-and-active transmission-line topologies in the cochlea. Additionally, none of the ideas, algorithms, or circuits presented in this paper are confined to any particular frequency range. Thus, in addition to the RF and audio frequency ranges, the architectures we describe could be useful in several fields of sensing including NMR, bio-detection, sonar, radar, and implants.

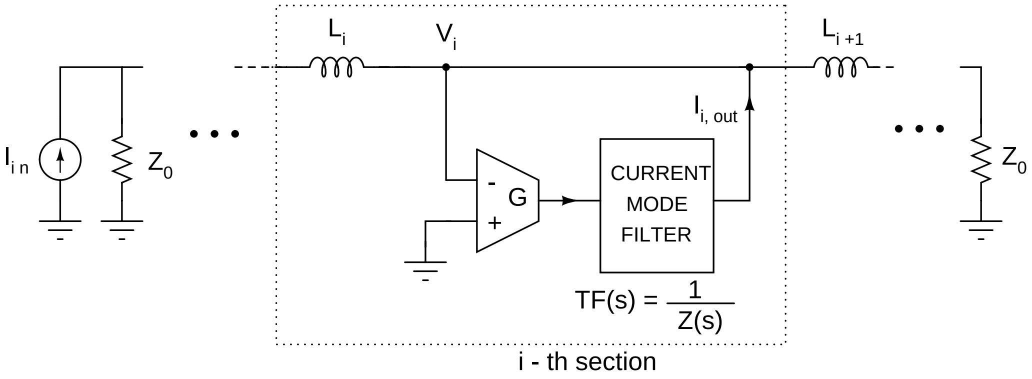

Figure 1. One incremental segment of an RF-cochlea model

3. Implementation

The cochlear model shown in Figure 1 consists of a one-dimensional array of shunt admittance elements $\frac{G}{Z(s)}$ coupled together by inductors. The current mode filter in Figure 1 has a transfer function $TF(s) = \frac{1}{Z(s)}$. It can be replaced by a voltage-mode filter if the transconductance *G* is placed after the filter rather than before it. The precise value of the terminating resistance Z_0 on the right hand side of Figure 1 does not matter, because the signal essentially decays completely before it reaches the end of the cochlea. The biological cochlea is a continuous structure and must be modeled using an infinite number of circuit elements. However, in practical electronic implementations only a finite number of active elements can be used. This limits us to implementing a discrete approximation of the continuous biological cochlea at RF frequencies. The cochlea is discretized by using a finite number of filter sections per frequency octave. Each section has properties that scale exponentially with position along the cochlea, with the exponential taper being gentle enough to prevent unwanted signal reflections. Thus the input impedance of the cochlea should be real over the entire range of operating frequencies. In practice, discretization of the continuous cochlea structure and component mismatches along sections can lead to inter-section reflections that degrade performance.

The signal processing properties of the architecture shown in Figure 1 depend critically on *Z(s)*. Types of *Z(s)* that lead to cochlea-like behavior in this architecture can be obtained analytically from a Wentzel-Kramer-Brillouin (WKB) approximation to the wave equation in inhomogeneous media. The WKB method essentially assumes that the medium varies slowly compared to the wavelength of the propagating wave. This approximation is found to be good both for the biological cochlea and also for most electronic cochlea-like structures that we design. In other words, the exponential taper of the cochlea is gentle enough that its properties appear almost constant on the scale of the wavelength. Using WKB and insights from [5], a suitable transfer function was found to be

$$sZ(s) = \frac{\left(s^2 + 2ds + 1\right)^2}{s^2 + s\frac{\mu}{Q} + \mu^2} \qquad \textbf{(1)}$$

where $s \equiv \frac{j\omega}{\omega_c(i)}$ is a scaled, dimensionless frequency variable and d, μ and Q are system constants. The center frequency $\omega_c(i)$ varies exponentially with position (filter section number) i and is defined as

$$\omega_c(i) = \omega_c(0)2^{-\frac{i}{m}} \qquad \textbf{(2)}$$

where $\omega_c(0)$ is the maximum frequency at which the cochlea operates, i is the position index (i.e., section number; $0 < i < i_{\max}$) and the cochlea has been discretized such that there are m sections per octave of frequency. The inductance L_i increases exponentially along the cochlea and is given by

$$L_i = L_0 2^{\frac{i}{m}} \qquad \textbf{(3)}$$

where L_0, the inductance of the first stage, may be chosen to set the transconductance G and the source and terminating resistance Z_0 are as follows

$$G = \frac{1}{\omega_c(0)L_0}\left(\frac{N \ln 2}{m}\right)^2 \qquad \textbf{(4)}$$

$$Z_0 = \omega_c(0)L_0 \frac{m}{\mu N \ln 2} \qquad \textbf{(5)}$$

Here N is a system parameter that is proportional to the steepness of the exponential taper of the cochlea. As one might expect, the behavior of this cochlear model approaches that of a real cochlea as m increases and the discrete approximation to continuous cochlear mechanics improves. This is primarily because the number of active elements participating in the wave amplification increases and the amplitude of the inter-section reflections decreases as the number of sections per octave increases. In simulation, it was found that cochlea-like transfer functions were typically obtained for $m > 40$. Better ways to obtain discrete approximations of the continuous differential equations that model the actual cochlea are currently being studied. The goal is to reduce the number of sections per octave needed for obtaining acceptable performance, thereby reducing power consumption, complexity and layout area.

Figure 2 shows simulated transfer functions $I_{i,out} / I_{in}$ (magnitude and phase) obtained along the length of the structure for different input frequencies. The curves are similar to those found in biological cochleas, and demonstrate the conversion from frequency to position which is among their most important characteristics. The position i along the cochlea at which the maximum gain is obtained (the ‘best position’) is a logarithmic function of the input frequency. This property makes the architecture useful for wideband, real-time signal detection and spectral analysis. The height of the peak in the amplitude transfer function increases as N increases. However, the value of m required for acceptable performance also increases with N, i.e., to obtain large gains with this architecture a large number of stages are required per octave. Alternatively, the peak gain can be increased without increasing m if d is decreased. Filters with high quality factors however suffer from stability problems, offer reduced dynamic range and consume more power. Thus there is a trade-off between gain and power consumption, complexity and layout area in this architecture.

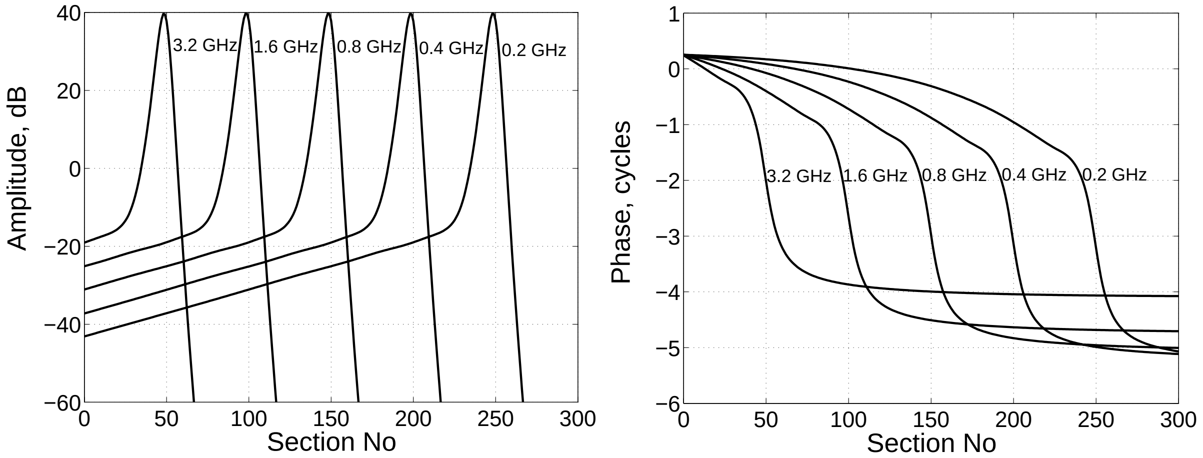

Figure 2. Simulated amplitude and phase transfer functions of the RF cochlea for different input frequencies.

Typical parameter values used for these simulations were $N = 20.96$, $d = 0.1$, $\mu = 0.76$ and $Q = 3.8$. These values were chosen to loosely fit transfer functions obtained from biological cochleas [5]. With values of $Z_0 = 50\ \Omega$, $\omega_c(0) = 2\pi \times 6.4$ GHz, $L_0 = 0.275$ nH, $G = 7.6$ mA/V and $m = 50$ sections per octave, the cochlea operated over six octaves of frequency (i.e., there were a total of $i_{\max} = 300$ sections). The maximum inductance value was 17.6 nH and the total inductance (summed over all 300 sections) was 1.2 µH. Transfer function characteristics were found to be reasonably robust to parametric fluctuations. This should considerably simplify VLSI implementation of the architecture. In addition, the relatively low values of filter quality factors required means that it should be relatively straightforward to construct and stabilize each filter section using typical integrated circuit components. Integrated spiral inductors can be used to couple the sections together, leading to a completely integrated implementation. Alternatively, the integrated filter sections can be combined with off-chip inductors, resulting in a hybrid system that saves on silicon area.

4. Conclusions

In this paper, we have proposed a cochlea-like distributed traveling wave architecture for wideband amplification, spectral analysis and enhancement at RF frequencies. Preliminary simulations of this architecture appear promising.

Acknowledgements

The authors acknowledge useful discussions with T. H. Lee, A. Hajimiri, E. Vittoz, C. Enz and C. Toumazou regarding the ideas in this paper.

References

[1] R. Sarpeshkar, R. F. Lyon and C. A. Mead, "A Low-Power Wide-Dynamic-Range Analog VLSI Cochlea," *Analog Integrated Circuits and Signal Processing*, vol. 16, no. 3, pp. 245-274, Aug. 1998.

[2] L. Turicchia and R. Sarpeshkar, "A Bio-Inspired Companding Strategy for Spectral Enhancement", Accepted for publication in *IEEE Transactions on Speech and Audio Processing*, 2004.

[3] R. Sarpeshkar, "Traveling Waves Versus Bandpass Filters: The Silicon and Biological Cochlea," *Proceedings of the International Symposium on Recent Developments in Auditory Mechanics*, Editors H. Wada et al, pp. 216-222, World Scientific, 2000.

[4] L. Turicchia and R. Sarpeshkar, "The Silicon Cochlea: From Biology to Bionics," *Proceedings of the Biophysics of the Cochlea: From Molecules to Models*, pp. 417-424, Titisee, Black Forest, Germany, July 2002.

[5] G. Zweig, "Finding the Impedance of the Organ of Corti", *Journal of the Acoustical Society of America*, vol. 89, no. 3, pp. 1229-1254, March 1991.

[6] A. Hubbard, "A Traveling-wave Amplifier Model of the Cochlea", *Science*, vol. 259, pp. 68-71, Jan. 1993.

[7] E. L. Ginzton, W. R. Hewlett, J. H. Jasberg and J. D. Noe, "Distributed Amplification", *Proceedings of the IRE*, pp. 956-959, August 1948.