A MICROPOWER ANALOG VLSI PROCESSING CHANNEL FOR BIONIC EARS AND SPEECH-RECOGNITION FRONT ENDS

Timothy K.-T. Lu, Michael Baker, Christopher D. Salthouse, Ji-Jon Sit, Serhii Zhak, and Rahul Sarpeshkar

timlu@mit.edu and rahuls@mit.edu

Analog VLSI and Biological Systems Group, Research Laboratory of Electronics

Massachusetts Institute of Technology

77 Mass. Ave.

Cambridge, MA 02139

ABSTRACT

Next-generation bionic ears or cochlear implants will be fully implanted inside the body of the patient and consequently have very stringent requirements on the power consumption used for signal processing. We describe a low-power programmable analog VLSI processing channel that implements bandpass filtering, envelope detection, logarithmic mapping and analog-to-digital conversion. A bionic ear processor may be implemented through the use of several such parallel channels. In a proof-of-concept 1.5 μm AMI MOSIS implementation, the most power-hungry channel of our system (7.5kHz center frequency) consumed 7.8 μW of power, had an internal dynamic range (IDR) of 51dB, and provided 64 discriminable levels of loudness per channel. Such numbers already satisfy the requirements of today's commercial bionic ear processors and can lower the power consumption of even advanced DSP processing schemes of the future by an order of magnitude. Our processing channel is also well suited for use in low power speech recognition front ends, which commonly require the same sequence of operations in cepstrum-like front ends. Future improvements in the interfaces between the various stages of our processing channel, which were not optimized in this implementation, promise a potential internal dynamic range of more than 60dB with little or no increase in power.

1. INTRODUCTION

In the past 25 years, the development of bionic ears (BEs) has been successful in restoring hearing to the profoundly deaf by stimulating the auditory nerve with implanted electrodes to mimic the natural response of the ear to sounds.

Figure 1 shows a common approach to signal processing in BEs. The microphone transduces audio signals to electrical signals which are then fed into the audio front end (AFE). The AFE senses and pre-amplifies the microphone signal, pre-emphasizes important frequencies in speech through filtering, and uses an automatic gain control (AGC) system to compress the 80dB input dynamic range of sounds (30dB SPL-110dB SPL) into a narrower internal dynamic range (IDR) for subsequent processing. The IDR is typically 50dB. The subsequent processing is composed of several parallel channels each of which extracts the envelope energy in a defined frequency band and maps it via a logarithmic function into a narrower dynamic range of perceivable electrode stimulation. Today's processors typically have 16 channels that together span the entire audio frequency spectrum from 250Hz—10kHz. Figure 1 only shows two such channels. In our implementation, the output of the logarithmic map is digitized and sets the programmable DAC bits of electrode-stimulation circuits. The stimulation currents from the electrodes excite the auditory nerve to evoke the sensation of hearing [1].

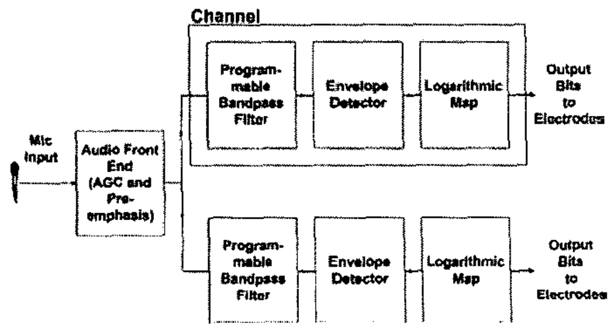

Figure 1. A Common Approach to Signal Processing in Bionic Ears

In this paper, we will not report on implementations of the AFE and AGC circuits, since they are not part of the parallel channels of processing but common to all of them. Our own implementations of these circuits, however, yield a combined power consumption for the AFE and AGC which is about 125 μW.

Analog implementations afford power savings over the combination of an A-to-D converter and a DSP processor: Even with Moore's law scaling, the latter schemes are unlikely to lower the power consumption of 32-channels of processing below a few milliwatts or 10 mW. The total power consumption of a system that uses 32 of our processing channels amounts to less than 0.4 mW. Thus, our analog implementation promises an order-of-magnitude improvement in power consumption over that of even advanced DSP designs. Furthermore, like digital implementations, our analog processing channel is programmable.

Subthreshold-MOS, silicon-cochleas, and analog circuits for cochlear-implant processing have been previously proposed [2-7] as means for implementing complex signal processing with very low power. This work proves the promise of such prior work by achieving numbers that make an analog processing system commercially feasible in the near term.

0-7803-7761-3/03/$17.00 ©2003 IEEE

We implemented our system with the MOSIS AMI 1.5μm BiCMOS process with 2.8V power supplies. The chip contains an analog front end, two complete channels, and other circuitry for diagnostic purposes. We built two complete channels to address issues of matching that are natural to ask in analog implementations. The layout of the system is shown in Figure 2. The two channels and the audio front end can entirely fit in half of the 4.6 mm x 4.7 mm chip used, although this was not done in this implementation.

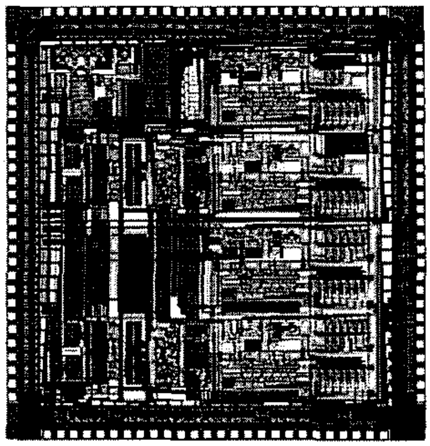

Figure 2. VLSI Layout of Two Channels of Our System. This Chip is 4.6 mm x 4.7 mm on a 1.5 μm Process.

The organization of this paper is as follows. Section 2 describes the programmable bandpass filters. Section 3 presents the operation of the envelope detectors. Section 4 discusses the implementation of the logarithmic map circuit used to provide dynamic range compression and A/D conversion. Section 5 presents experimental data demonstrating the operation of the entire system. Section 6 describes how our system could also be used as a front-end for speech recognition system. We summarize and conclude our findings in Section 7.

2. THE PROGRAMMABLE BANDPASS FILTER

Bandpass filters are a crucial component in the signal-processing chain in BEs. The array of bandpass filters mimics the frequency-to-place transformation of the biological cochlea: High frequency sounds stimulate the basal region of the auditory nerve while low frequency sounds stimulate the apical region of the auditory nerve. In a BE, the electrodes corresponding to high-frequency channels primarily stimulate basal regions of the auditory nerve, while the electrodes corresponding to low-frequency channels primarily stimulate apical regions of the auditory nerve.

Subthreshold Gm-C filters are well suited for use in BEs because they offer low power consumption and can be tuned over a wide frequency range to cover the spectrum of hearing [8]. The capacitive-attenuation filter used in the channel described here had first-order rolloffs and was single-ended like that described in [8]. Higher-order and fully differential versions of such filters are described in [9].

The center frequency of the filter is programmed via 5 DAC input bits which set the bias current of the transconductor and yield a total of 32 possible different configurations to span the frequency range of hearing. Alteration of the DAC reference current provides a further degree of freedom if needed. The programmability of the filter is important in accommodating variations in patients.

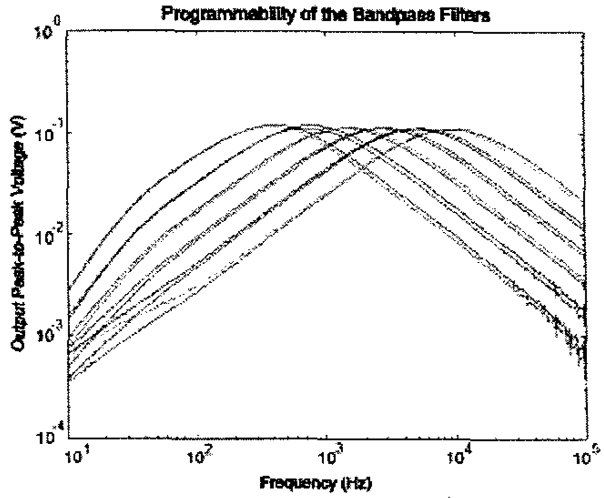

Figure 3. Programmability of the Bandpass Filter with DAC Settings.

Experimental results from two separate channels are shown in Figure 3, which presents the programmability of the capacitive-attenuation bandpass filters. As the DAC bits are adjusted, the center frequencies of the filters move from about 250 Hz to 10 kHz. Furthermore, Figure 3 exhibits the close matching between filters from different channels on the same chip. Measurements from this chip demonstrate that the filters can swing 848 mVrms with 5% THD and have noise floors of 200 μVrms, yielding a minimum dynamic range of 72dB for this part of the channel.

The power required for high-frequency filters is larger than that required for low-frequency filters. For a 5kHz – 10kHz filter, our measured power dissipation was 2.1 μW per channel; power consumption dropped to 0.14 μW per channel for a 100 Hz – 200 Hz filter. Fully differential filters with second-order rolloff slopes have power consumptions of 0.23 μW and 6.36 μW for the same frequency ranges [9].

3. THE ENVELOPE DETECTOR

Envelope detectors are important in the design of BEs since they transform the energy of bandpassed audio signals to information for patients to process. The envelope detection strategy used in this paper is similar to that described in [3] and consists of a rectifying stage and a peak-detector stage. Circuit innovations in the rectifying stage allow us to achieve superior dynamic range at the same power consumption and are described in some detail in a companion paper at this conference [10]. We shall only briefly describe the operation of the envelope detector in this paper.

The envelope detector uses a wide-linear-range transconductor [11] to drive a class B mirror and thus perform rectification and voltage-to-current conversion. The rectified currents from the mirror are summed to produce a full-wave output and are fed into a current-mode peak detector with asymmetric attack and release times. The peak detector is identical to that used in [3]. In our implementation, the release time is adjustable to suit the preferences of the patient. A slow feedback loop performs offset correction and ensures that offsets in the wide-linear-range transconductor and in the two halves of the class-B mirror do not greatly degrade the minimum detectable signal of the envelope detector.

Figure 4 shows experimental results obtained from the output current of the envelope detector circuit in response to varying input amplitudes at frequencies of 100Hz, 1 kHz, and 10 kHz. The linear dynamic range of the envelope detector circuit at 100 Hz is demonstrably 60 dB if we only allow ±10% deviations from linearity; at 1 kHz, the linear dynamic range is 59 dB. At 10 kHz, as described in [10], high-frequency operation of the envelope detector results in residual dead-zone effects from the rectifier in the circuit, and the linear dynamic range of the envelope detector is degraded to 49dB. However, Figure 4 shows that a monotonic response with good signal-to-noise is still obtained over the entire 60dB range of operation. We measured a power consumption of 2.8 μW. Further unpublished optimizations in the envelope detector or an increase in power consumption to 5μW [10] can yield almost 60dB of linear dynamic range at all frequencies. In this paper, we did not explore these issues further since the overall dynamic range for the channel was lowered by interstage coupling anyway.

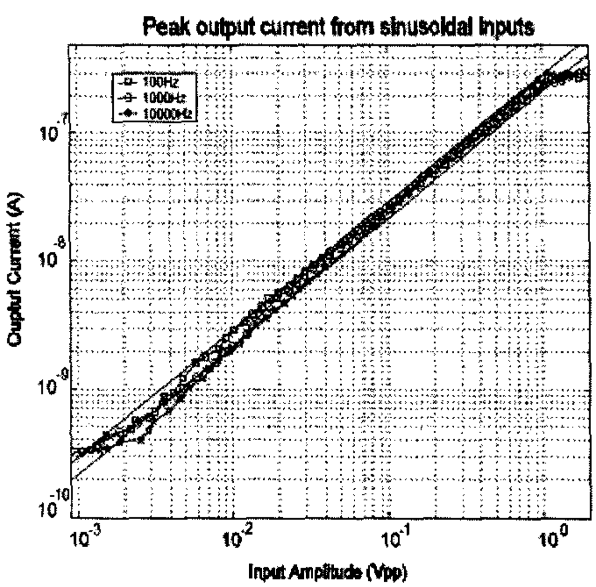

Figure 4. Envelope Detector Output Current as a Function of Input Amplitude for Different Frequencies. Deviations from Linearity of ±10% are also shown [10].

4. LOGARITHMIC MAP

The electrical dynamic range that is psychophysically observed in deaf patients is usually between 3 dB and 30 dB, with a typical value being about 10 dB [12]. The logarithmic-map stage maps the 40dB-60dB internal dynamic range in envelope-energies range into this electrical dynamic range. It does so by having the electrode stimulation currents be a linear function of the logarithm of the envelope energy. Independent of their electrical dynamic range, deaf patients appear to perceive changes in sound intensity of about 1 dB. Thus, a good patient with 30dB of electrical dynamic range may be able to resolve about 30 discriminable levels; to ensure that such perception is possible, the logarithmic map must be precise to at least 5 output bits. We may achieve all of these specifications by building a low-power current-input logarithmic A/D converter that is at least 5-bit precise.

It is generally accepted that the envelope in each band of speech does not vary significantly faster than 100 Hz. Thus, the logarithmic A/D converter need only sample at a rate greater than 200 Hz, the Nyquist frequency.

The low power and relatively slow bandwidth requirements were met with a 6-bit diode-based logarithmic dual-slope A/D converter. Several circuit innovations to cancel offset and temperature dependence in the logarithmic map circuit were employed but are beyond the scope of this paper.

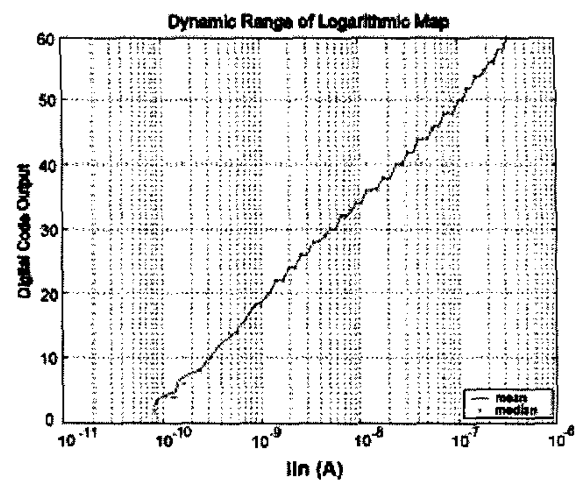

Figure 5. Performance of the Logarithmic Map Circuit.

Figure 5 shows the overall performance of the logarithmic map circuit. It is able to output a linear range in its digital code as the measured input current from the envelope detector varies in constant ratio increments over a 60dB dynamic range from 200 pA to 200 nA. Figure 5 was measured for a 7.5kHz input. We measured 1.68 μW in analog power and 1.26 μW in digital power for this stage yielding a total of about 3 μW.

5. OVERALL SYSTEM PERFORMANCE

The overall dynamic range of a single channel in this system was measured for a 7.5kHz input to a bandpass filter centered at 7.5Khz. Figure 6 demonstrates linear operation from a 4 mVpp input to a 1.5 Vpp input, which is a dynamic range of 51 dB. Our overall dynamic range is less than that of each of our stages primarily because our proof-of-concept design did not ensure that the minimum output noise floor stage of each stage of in the cascade matched the minimum input noise floor of the next stage of processing in the cascade. If these optimizations are made in future designs, we expect to achieve an overall dynamic range of operation of 60dB with little or no increase in power.

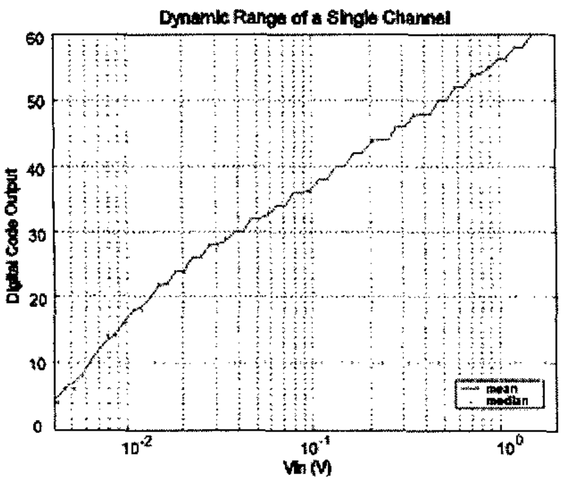

Figure 6. Overall System Performance of the Channel

The total analog power measured for each channel is comprised of 2.1 μW in the bandpass filter, 2.8 μW in the envelope detector, and 3 μW in the logarithmic map. Thus, each channel requires a total of 7.9 μW. For a 32-channel implant, the expected power consumption is thus about 256 μW. If we add the 125 μW power consumption of the AFE, a fully functional 32-channel bionic ear processor would require less than 400 μW of power. Using filters with second-order rolloffs [9] do not alter this conclusion if the power scaling with center frequency in the filters is included. Thus, our analog processing channel offers an order-of-magnitude improvement over even advanced A-D-and-DSP implementations of the future.

6. USE FOR SPEECH-RECOGNITION FRONT ENDS

The processing channel that we have described may be easily programmed to create an array of filters that form a Mel filter bank [13]. The logarithm of the envelope energies of these filters then yields a very cepstral-like representation except that the final step of the computation, the discrete cosine transform is omitted [14]. The latter transform, or alternate transforms, may easily be performed on the digital numbers output by our channel in a relatively cheap fashion by a subsequent DSP. The relatively expensive filtering and log operations are performed by our analog channel, saving power.

7. CONCLUSIONS

Experimental data from our chip proves that analog VLSI implementations of processing channels for bionic ears or low-power speech-recognition front ends can yield order-of-magnitude power savings over even advanced DSP implementations of the future. Such implementations are therefore, likely to be very useful in fully implanted bionic ear systems or in portable speech recognition systems of the future, especially if they are programmable like our implementation.

REFERENCES

[1] P. C. Loizou, "Mimicking the human ear," *IEEE Signal Processing Mag.*, Vol. 5, pp. 101-130, Sept. 1998.

[2] R. J. W. Wang, R. Sarpeshkar, M. Jabri, and C. Mead, "A low-power analog front-end module for cochlear implants," presented at the XVI World Congress on Otorhinolaryngology, Sydney March 1997.

[3] 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, pp. 245-274, 1998.

[4] W. Germanovix and C. Tourmazou, "Design of a micropower current-mode log-domain analog cochlear implant," *IEEE Trans. on Circuits and Systems II: Analog and DSP*, Vol. 47, pp. 1023-1046, Oct. 2000.

[5] T. S. Lande, J. T. Marienborg, and Y. Berg, "Neuromorphic cochlea implants," In *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS 2000)*, Vol. 4, pp. 401-404, 2000.

[6] R. F. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 36, pp. 1119-1134, Jul. 1988.

[7] W. Liu, A.G. Andreou, and M.H. Goldstein, "Voiced-speech representation by an analog silicon model of the auditory periphery," *IEEE Trans. on Neural Networks*, Vol. 3, No. 3, pp. 477-487, May 1992.

[8] C. D. Salthouse and R. Sarpeshkar, "A micropower band-pass filter for use in bionic ears," In *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS 2002)*, Vol. 5, pp. 189-192, 2002.

[9] C. D. Salthouse and R. Sarpeshkar, "A practical micropower programmable bandpass filter for use in bionic ears," *IEEE Journal on Solid State Circuits*, in press, 2003.

[10] M. Baker and R. Sarpeshkar, "A micropower envelope detector for audio applications", submitted for review to *the IEEE International Symposium on Circuits and Systems*, 2003.

[11] R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A low-power wide-linear-range transconductance amplifier," *Analog Integrated Circuits and Signal Processing*, Vol. 13, No. 1/2, May/June 1997.

[12] J.-J. Sit, "A low-power analog logarithmic map circuit with offset and temperature compensation for use in bionic ears," M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2002.

[13] B. C. J. Moore, *Introduction to the Psychology of Hearing*, Academic Press, London, 1997.

[14] L. Rabiner and B.-H. Juang, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ, Prentice Hall, 1993.