A CONTINUOUS-TIME SPEECH ENHANCEMENT FRONT-END FOR MICROPHONE INPUTS

Rich Ellis, Heejong Yoo, David W. Graham, Paul Hasler, and David V. Anderson

School of Electrical and Computer Engineering, Georgia Institute of Technology

ABSTRACT

In this paper, we propose a real-time noise suppression system implemented with analog VLSI. The algorithm implemented is designed to reduce stationary background noise while preserving the non-stationary signal component. Because the system relies on analog computation rather than digital, it has benefits such as extremely low power consumption and real-time computation. The algorithm is based on digital signal processing foundations that are slightly adjusted for use in the continuous-time domain. The analog components described as part of this system include a C^4 second-order section bandpass filter, peak and minimum detectors, a translinear division circuit, and a differential multiplier. Floating-gate circuits are used to set bias points and adjust offsets.

Audio signal enhancement by removing additive background noise from a corrupted noisy signal is not a new concept. However, with the prosperity of portable communication devices, it has recently received increased attention. Traditional methods of noise suppression include spectral subtraction, Wiener filtering, and a number of modifications on these methods that increase the intelligibility of the processed audio signal and/or reduce adverse artifacts.

While most noise suppression methods are focused on the processing of discrete-time sampled audio signals, we propose a technique for noise suppression in the continuous-time domain, before the eventual A/D conversion takes place. Our approach has roots in early noise suppression systems described elsewhere [1, 2]. Using advances in technology and new tools, including the computational potential of analog VLSI circuits, we are building on these and other concepts to design a system that operates in real time and uses extremely low amounts of power. The result is a system that performs a function normally reserved for digital computation, freeing those resources for other operations in the digital domain. This method of Cooperative Analog and Digital Signal Processing (CADSP) is the basic framework for this project.

1. STRUCTURE OF SUPPRESSION SYSTEM

A common model for a noisy signal, $x(t)$, is a signal, $s(t)$, plus additive noise, $n(t)$, that is uncorrelated with the signal

$$x(t) = s(t) + n(t). \quad (1)$$

The goal is to design a real-time system that generates some optimal estimate, $\hat{s}(t)$, of $s(t)$ from $x(t)$. We assume that the additive noise is stationary over a long time period relative to the short term non-stationary patterns of normal speech. The signal estimate, $\hat{s}$, may be found in the frequency domain using spectral subtraction or by applying a Wiener filter gain.

The authors thank Hewlett-Packard Corporation for partial funding of this project.

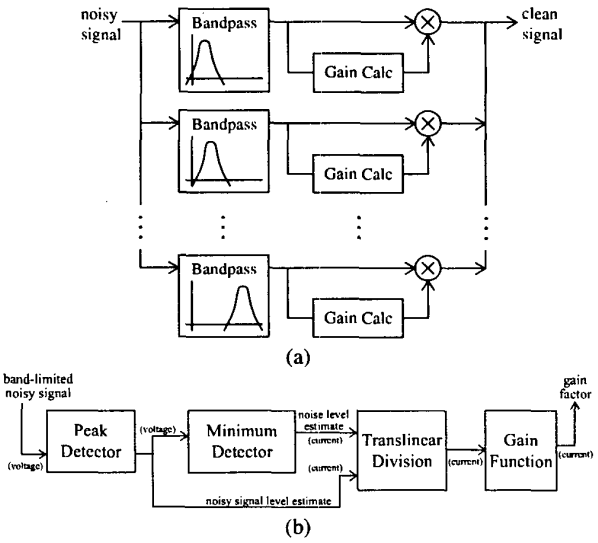

Fig. 1. The overall structure of the system and the details of the gain calculation block. The incoming noisy signal is divided into exponentially-spaced frequency bands. Within each frequency band, the noisy signal envelope is estimated using a peak detector. Based on the voltage output of the peak detector, the noise level is estimated using a minimum detector operating at a slower rate than the peak detector. The currents representing the noisy signal and noise levels are input to a translinear division circuit, which outputs a current representing the estimated signal-to-noise ratio. A nonlinear function is applied to the SNR current, and the resulting gain factor is multiplied with the band-limited noisy signal to produce a band-limited "clean" signal. Finally, the output of all of the bands are summed to reconstruct the signal with the noise components significantly reduced.

1.1. Frequency Decomposition

Figure 1a shows the structure of a continuous-time noise suppression system for real-time analog implementation. A filter bank separates the noisy signal into 32 bands that are exponentially spaced in frequency, similar to the human auditory system for frequency domain processing. Then a gain is calculated based on the envelopes of the observed subband signals and subband noise signals. This gain is applied to the subband signals and the signals are combined to form $\hat{s}(t)$.

The filters used in the filter bank are the Capacitively-Coupled Current Conveyor Second-Order Sections (C^4 SOS), shown in Figure 2a [5]. The C^4 SOS is composed of three C^4s (described in [4]) in which the feedback capacitor of the first stage filter is removed

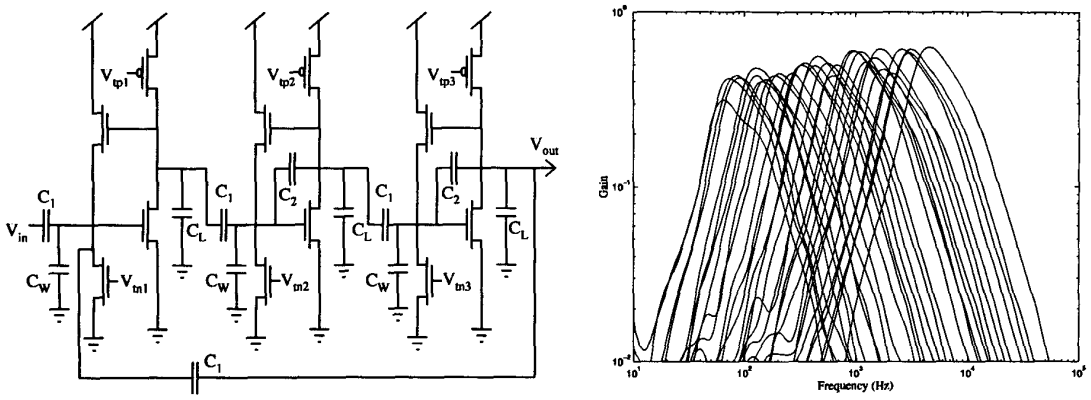

Fig. 2. **(a)** The schematic showing the C^4 SOS bandpass filter. The bias voltages are set using a resistive-divider network, which creates a logarithmically spaced filter bank. Future versions will implement floating gate elements instead of resistors to allow for programming and improved tuning [3]. **(b)** This plot shows the frequency response of each filter in a 32-band filter bank. In this configuration, the bandpass rolloff is 40 dB. The noted bias voltages refer to the voltages set at each end of the resistive network. This data is from a circuit fabricated in a 0.5 μm process available through MOSIS.

in order to make that stage a high gain amplifier. This filter can have a frequency response of any defined bandwidth, and outside that bandwidth, slopes of $\pm$40 dB/decade or greater occur. By adjusting the voltage biases, the response at either corner can be tuned to have a sharp transition or even a Q peak. A high Q peak is useful because it helps isolate the respective center frequency. We present detailed information on the C^4 SOS elsewhere [5].

After the incoming noisy signal has been band-limited by the filter bank, a gain factor is calculated based on the characteristics of each band-limited noisy signal. The resulting gain factor is then multiplied with the original band-limited signal. Finally, the band-limited signals are summed to reconstruct the full-band signal estimate, without the additive noise components. The algorithm for gain calculation and the elements that perform this functionality are explained in the following sections.

1.2. Gain Calculation Algorithm

The gain calculation circuits make up the processing block for the system. Each block operates independently of the others in the array, so effectively we have 32 parallel signal processors operating simultaneously on 32 band-limited signals. The block diagram shown in Figure 1b summarizes the gain calculation algorithm. The following sections will elaborate on the algorithm and relate the circuits that perform the computations. We present the details of the algorithm and the signal processing theory behind it elsewhere [6].

1.3. Peak and Minimum Detectors

The first step in the gain calculation algorithm is to estimate both the levels of the noisy signal and the noise. It is impossible to accurately estimate the actual signal component of the incoming signal, so the noisy signal is accepted as a reasonable estimate. The circuit in Figure 3a is a peak detector, which is used to give the noisy signal level estimate. The bias voltage V_{τ_d} sets the time constant at which the output will decay after a peak. When a speech signal is input, the peak detector will follow the envelope of the signal, rising rapidly with the increasing signal amplitude and decaying slowly enough to result in a smooth envelope. The peak detector will also follow the level of the additive noise, particularly in times where the signal is absent. The circuit outputs both a voltage and current that are representative of the noisy signal level.

The circuit shown in Figure 3b is a minimum detector and is used to estimate the noise level in the signal. The input to the minimum detector is the voltage output from the peak detector. In this way, we estimate the noise level by following the minimum values of the noisy signal envelope. The bias voltage V_{τ_a} sets the attack time constant; it is set to run much slower than the peak detector in order to follow the slow changes found in the amplitude of relatively stationary noise. When the signal is present, the output will maintain a slowly rising level; when the signal is not present, the minimum detector will track the noise level.

In both of these circuits, the floating-gate pFET device shown at the top of the schematics is used for offset cancellation [3]. With no input signal, that is for its bias voltage, an offset (bias) current will be produced at the output node. The floating gate will be programmed in a such as way as to cancel this offset current, causing zero current at the output node until a signal is present. The current outputs of the peak and minimum detectors are the inputs to the next component in the system.

1.4. Translinear Division

In terms of (1), an estimation of the *a posteriori* signal-to-noise ratio (SNR) is defined by a ratio of the envelopes of the signals,

$$\widehat{SNR}(t) = \frac{e_{\hat{s}}(t)}{e_{\hat{n}}(t)} \approx \frac{e_{\hat{x}}(t) - e_{\hat{n}}(t)}{e_{\hat{n}}(t)} = \frac{e_{\hat{x}}(t)}{e_{\hat{n}}(t)} - 1 \qquad (2)$$

where $e_{\hat{x}}(t)$ is the noisy signal envelope estimate, which can be represented as the sum of the actual signal envelope $e_s(t)$ and the noise envelope estimate $e_{\hat{n}}(t)$.

Multiplication and division operations can be performed using the translinear principle [7]. The circuit shown in Figure 4

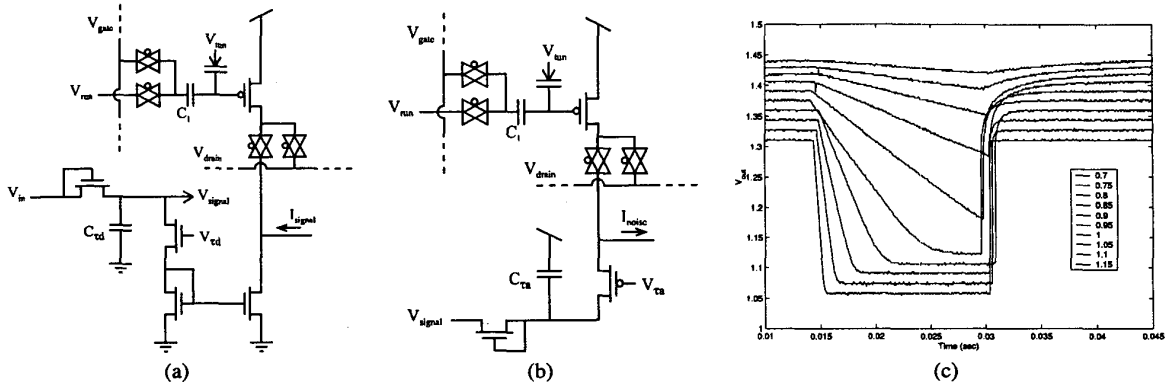

Fig. 3. (a) The peak detector circuit shown in this schematic tracks the incoming band-limited noisy signal, giving the noisy signal level estimate. This circuit has two outputs: the current I_{signal} goes into the division circuit, while the voltage V_{signal} is the input to the minimum detector. The bias voltage V_{τ_d} sets the time constant at which the output will decay after a peak. (b) The minimum detector is biased with V_{τ_a} to have a much slower time constant than the peak detector, so as to follow the slowly changing curve that results at the bottom of the noisy signal envelope. The output I_{noise} is an estimate of the noise level. (c) This plot shows the behavior of the peak detector at varying time constant biases with a stepped input signal.

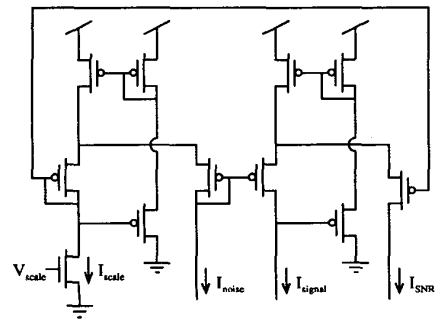

Fig. 4. The translinear circuit shown in this schematic implements the division of I_{noise} into I_{signal}, yielding a current representing the signal-to-noise ratio, I_{SNR}. We subtract an I_{scale} from this result to get the function described in Equation 3.

performs a division operation where the output, I_{SNR}, can be represented by

$$I_{SNR} = I_{scale}\frac{I_{signal}}{I_{noise}} - I_{scale} \tag{3}$$

and represents the estimated SNR. The current I_{scale} is set by the bias voltage V_{scale} and is used to put the output current into the proper range for the gain function.

1.5. Gain Function and Multiplication

The final element of the gain calculation algorithm is the gain function and multiplier. In this element a gain is calculated and applied to the band-limited signal to realize the noise suppression. The gain may be expressed as a function of the SNR. Several different gain functions may be used but all of them have the general characteristics of low gain for low SNR and high gain (at or near unity) for high SNR, with varying smoothing between these two regions. The transition between low gain and high gain must be smooth to avoid adverse auditory effects, yet it must quickly follow the SNR changes to ensure an acceptable amount of noise suppression. We present further details elsewhere [6].

In the circuit implementation, the gain function and the multiplier comprise one circuit, shown in Figure 5a. The transistors at the top of the schematic are the differential pair where the actual multiplication takes place; the multiplier will be described shortly. The transistors below the differential pair create the behavior of the gain function. The output of the gain function circuit is I_{gain}, which can be approximated by

$$I_{gain} = \frac{(I_{SNR}^2/I_o + I_{min1})I_{max}}{(I_{SNR}^2/I_o + I_{min1}) + I_{max}},$$

where I_{max} and I_{min1} are set by the voltage biases V_{max} and V_{min1} and effectively create the upper and lower bounds of the gain factor output; I_o is a device dependent parameter of a subthreshold MOSFET. The voltage bias V_{min2} can be used to further adjust the range of input current that the circuit will accept. The two diode-connected transistors on the input branch cause the current that is mirrored onto the gain branch to be squared (the increase may be slightly larger due to back-gate effects); this function is important to ensure a quick transition from low to high gain as I_{SNR} increases.

The multiplication portion of the circuit results from the interaction between the inputs to the differential pair V_{in1} and V_{in2} and the gain factor current I_{gain} from the gain function circuit below. The circuit is biased to operate in the linear range of the following equation:

$$I_{out1} - I_{out2} = I_{gain}\tanh\frac{\kappa(V_{in1} - V_{in2})}{2U_T}, \tag{4}$$

where $U_T = kT/q$. The input to the multiplier is the band-limited noisy signal, i.e. the output of the C^4 SOS bandpass filter. Because the initial implementation is single-ended rather than differential, the "negative" input voltage, V_{in2}, is held constant. The data shown in Figure 5b,c show the functionality of the multiplier and the form of the gain function.

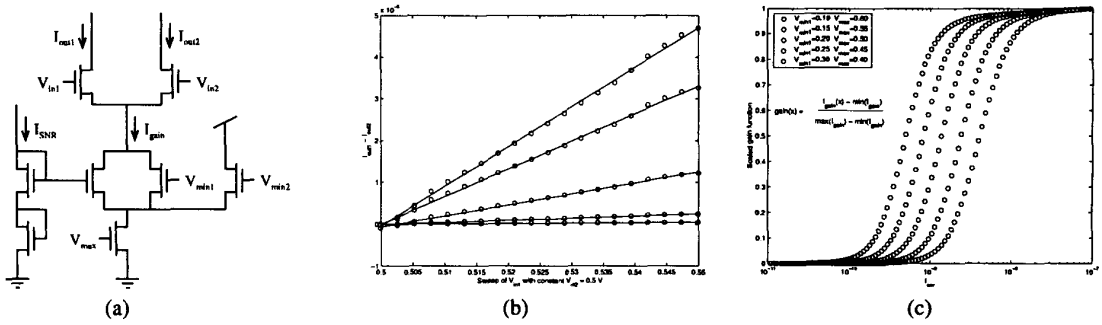

Fig. 5. **(a)** The schematic of the multiplication and gain function circuit. The output current, $I_{out1} - I_{out2}$, is the product of the band-limited input signal, $V_{in1} - V_{in2}$, and the gain factor, I_{gain}. Each voltage bias sets a current that forms the overall behavior of the gain function circuit. **(b)** The output current, $I_{out1} - I_{out2}$, varies with its two factors: the gain factor and the differential input voltage. This plot shows sweeps of the input voltage for several values of I_{SNR}. **(c)** The gain function of the circuit depicted in (a) is plotted versus I_{SNR}. At the extremes of the signal-to-noise ratio either a low gain factor or a unity gain factor is the result. Multiple curves show the bias voltages being adjusted.

2. SYSTEM RESULTS

The experimental results presented in this paper are from tests on individual components that have not yet been integrated into a larger system. Figure 6 shows a noisy speech signal that has been processed by the components in our system. The system is effective at adaptively reducing the amplitude of noise-only portions of the signal while leaving the desired portions relatively intact.

2.1. Circuit Noise and Distortion

Any noise or distortion created by the gain calculation circuits minimally affects the output signal because these circuits are not directly in the signal path. While the bandpass filters and the multipliers will inject a certain amount of noise into each frequency band, this noise will be averaged out by the summation of the signals at the output of the system. Distortion in the signal path will arise from the bandpass filters and the multiplier. In the bandpass filter array, the C^4 SOS structure is not cascaded as in cochlea models, therefore eliminating the typical distortion or noise accumulation. In speech, particularly in noisy environments, the signal power is more evenly distributed across a broad frequency range than a simple tone, and therefore allowing for large input amplitudes with minimal output distortion (higher system signal-to-noise ratio). As a result, we typically have signal amplitudes through each filter that are 10mV to 30mV or less, resulting in harmonic distortion through the system less than -30dB; future versions using differential signals will further reduce these effects.

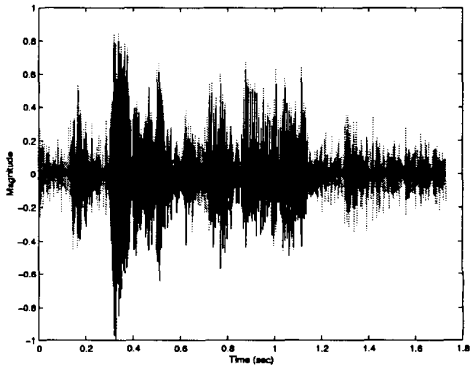

Fig. 6. Noise suppression in one frequency band. The light gray data is the subband noisy speech input signal; the black waveform is the corresponding subband output, after the gain function has been applied.

3. REFERENCES

[1] E. J. Diethorn, "A subband noise-reduction method for enhancing speech in telephony and teleconferencing," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, New Paltz, New York, October 1997.

[2] S. L. Gay and J. Benesty, *Acoustic Signal Processing for Telecommunication*, Kluwer Academic Publishers, 2000.

[3] M. Kucic, P. Hasler, J. Dugger, and D. Anderson, "Programmable and Adaptive Analog Filters using Arrays of Floating-Gate Circuits," *IEEE Advanced Research in VLSI*, Salt Lake City, UT, March 2001.

[4] P. Hasler, M. Kucic, and B. A. Minch, "A transistor-only circuit model of the autozeroing floating-gate amplifier," in *IEEE Midwest Symposium on Circuits and Systems*, Las Cruces, August 1999.

[5] D. Graham and P. Hasler, "Capacitively-coupled current conveyer second-order section for continuous-time bandpass filtering and cochlea modeling," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Phoenix, Arizona, 2002.

[6] H. Yoo, D. V. Anderson, and P. Hasler, "Continuous-time audio noise suppression and real-time implementation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, Florida, 2002.

[7] T. Serrano-Gotarredona, B. Linares-Barranco, and A. G. Andreou, "A general translinear principle for subthreshold mos transistors," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 46, no. 5, pp. 607–616, May 1999.

4. ACKNOWLEDGEMENTS

The authors wish to thank Dr. Bradley A. Minch for several useful circuit discussions.