A Mixed-Signal Correlator for Acoustic Transient Classification

R. Timothy Edwards[1], Gert Cauwenberghs[1], and Fernando J. Pineda[2]

[1]Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218
[2]Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723

Email: {tim,gert,fernando}@bach.ece.jhu.edu

***Abstract*— Correlation computations are widely used in template-based temporal pattern recognition, but are expensive to compute in real time on DSP systems. We have designed and fabricated a low-power chip which uses current-mode analog circuits to compute the correlation between an auditory input signal in the time-frequency domain and a stored binary template. We use a bucket-brigade device (BBD) to accumulate partial column sums over time at rates consistent with the auditory-band input. Experimental results of the correlator demonstrate its correct operation over a wide range of operating frequencies.**

I. Introduction

Many time-based classification systems compute the correlation of an incoming discrete-time signal or signals with a predetermined template [1]. While for speech and other complex long-term signals it is necessary to perform dynamic time warping (DTW) or similar weighting of the incoming signal [2], for transients (short-term acoustic events of less than approximately 1/10 of a second), a simple correlation in the time-frequency domain yields accurate classification results [3].

A general form of the simple correlation is

$$c_z[t] = \sum_{m=1}^{M} \sum_{n=1}^{N} x[t-n, m]\, p_z[n, m] \tag{1}$$

where M is the number of frequency channels of the input, N is the maximum number of time bins in the window, x is the array of input signals split into frequency bands, p_z is the matrix of template pattern values for pattern z, and t is the current time. This formula produces a running correlation $c_z[t]$ of the input array with the template z.

For large M and N, this algorithm can be expensive to execute on a DSP in terms of speed and power requirements. However, our approach lends itself elegantly to low-power parallel analog computation in the form of MOS transistor circuits operating primarily in the subthreshold mode.

In the following discussion, we present our algorithm as a set of incremental modifications to the baseline algorithm of Equation (1). First, we normalize the input, which is essential for the steps which follow. Next, we take time derivatives of the input and templates, In this form, we are able to make the template values binary without significantly increasing classification error rates. Finally, we move the time-differencing to the output, yielding the simplest possible form of the equation.

II. Simplifying The Correlation Equation

We assume the input x to the system is an acoustic signal, split into M frequency bands. We compute the (rectified) energy envelope for each band, denoted y, and then normalize these system inputs by the function

$$x[t, m] = \frac{y[t, m]}{\theta + \sum_{k=1}^{M} y[t, k]}, \tag{2}$$

where θ is a threshold value included to suppress noise during quiet invervals in the input. Pineda et al.[4] have shown that the normalized input representation is essential to significantly simplify the pattern classifier algorithm, greatly reducing the size and complexity of the hardware implementation but not degrading the classification result.

To make the computation less expensive, first we use the time difference between samples of the input and compare it to the difference between samples of the template. The effect of this is to subtract a constant from the result but otherwise does not affect the computation. We then replace the template values by the *sign* of the template time differences, indicating whether the energy is expected to be rising or falling at each time sample. Thus,

$$c_z[t] = \sum_{m=1}^{M} \sum_{n=1}^{N} (x[t-n, m] - x[t-n-1, m])\, {p'}_z[n, m] \tag{3}$$

where

$${p'}_z[n, m] = \operatorname{sign}(p_z[n, m] - p_z[n-1, m]). \tag{4}$$

Pineda et al.[4] have shown that binarization of the template has a negligible effect on classification performance. Note that this result does not hold if we do not compute the time difference, or if we binarize the input.

With this major simplification of the algorithm, the multiplication can be reduced to a multiplexing function. The required normalization of the input allows a further simplification that does not affect system performance, in which we make the template values binary $[0, 1]$ rather than binary $[-1, 1]$.

The final modification of the algorithm involves noting that the time difference commutes with the summation, so that we can write Equation (1)

$$\begin{aligned} c_z[t] &= \sum_{m=1}^{M} \sum_{n=1}^{N} x[t-n, m]\, {p'}_z[n, m] \\ &- \sum_{m=1}^{M} \sum_{n=1}^{N} x[(t-1)-n, m]\, {p'}_z[n, m]. \end{aligned} \tag{5}$$

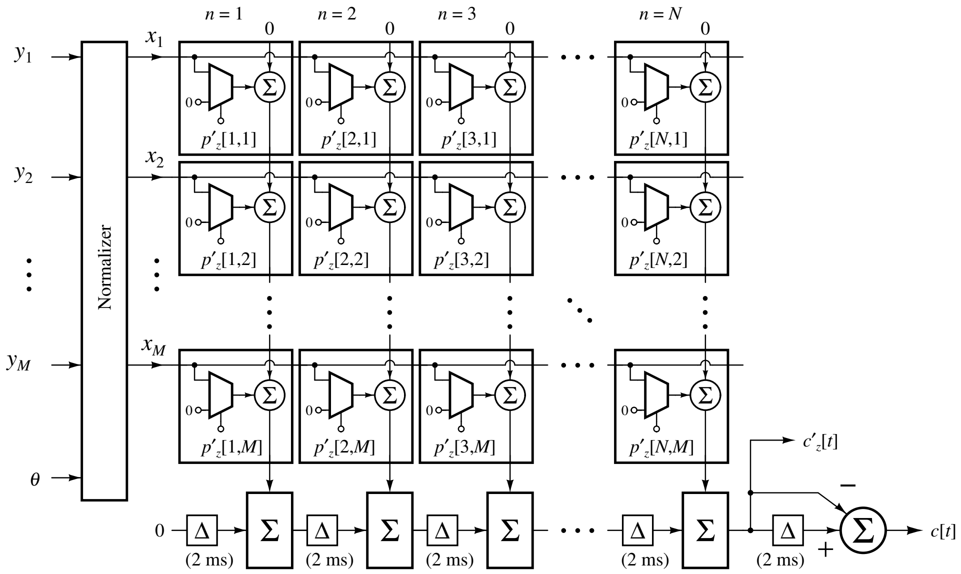

Fig. 1. Block diagram of the temporal current correlator.

If we let

$$c'_z[t] = \sum_{m=1}^{M} \sum_{n=1}^{N} x[t-n,m]\, p'_z[n,m], \tag{6}$$

then

$$c_z[t] = c'_z[t] - c'_z[t-1]. \tag{7}$$

In this form, we can compute the time difference at the output of the correlator rather than at the inputs, yielding three advantages:

1. We need to compute only one difference rather than M differences.
2. Architecturally, the algorithm is less affected by device mismatch when computing the output based on the difference of successive outputs rather than the absolute value of the output.
3. Most importantly, since the inputs x are rectified, the product xp' (when p' is binary $[0,1]$ is always positive and equals either x or zero, which allows us to conveniently implement the entire convolution as an array of simple on/off current switches carrying current in one direction only.

Figure 1 shows the system as presented, where each template value is a single bit controlling a switch (multiplexer) which adds either zero or the unidirectional current input to the sum. The summed current is accumulated over time, with the difference taken at the output. Our choice of an analog implementation stems from the ability to make such a system compact and extremely power-efficient. With a separate correlator circuit for each template, the system is fully parallel, and only about as complex as a RAM array.

Figure 3 shows a schematic of the correlator array for one template, minus the refresh and write circuitry for the dynamic memory. The correlator accepts inputs in the form of currents, one for each channel m. The diode-connected transistor at the input mirrors this current simultaneously to all cells across the array.

The other subcomponents of the correlator are explained in the following sections.

III. Current-switching Memory Array

To implement the summation over m in Equation (7), we utilize the simplest form of summation available to analog circuits, that of summing currents onto a single line. Each cell in the array contains a single pMOS transistor switch with the template bit value stored on its gate; the switch allows or disallows the input current for channel m to be added to the total current for column n.

Figure 3 shows an extremely compact dynamic circuit which can be used for the memory and switch. A value of V_{dd} on the gate of transistor M_2 will cause the transistor to be completely off and pass no current. A value of V_{dd} less approximately a volt or more will cause the transistor to be on, and pass the input current which is mirrored in transistor M_1. Transistor M_3 is used to write the bit value onto the storage capacitor (which is also a pMOS transitor), and can store a high bit value of V_{dd} or a low bit value of approximately one transistor threshold. The capacitor stored value tends to leak toward V_{dd}, ensuring that an "off" state leaks zero current. An "on" state must be periodically refreshed. All N columns in a given row m are written or refreshed together.

IV. Bucket Brigade

We perform the summation over time dynamically using a bucket brigade device [5]. This device is similar to a CCD line, but is more appropriate for this application, in which the system is clocked at a rate of 1 or 2 ms: while the charge-transfer efficiency in a bucket brigade is less than that of a CCD [6], the CCD is adversely affected by dark currents in the quiescent state and cannot operate at slow (auditory) rates. Large poly1-poly2 capacitors, which are significantly less affected by leakage currents than CCD capacitors, store the charge at each bucket brigade node.

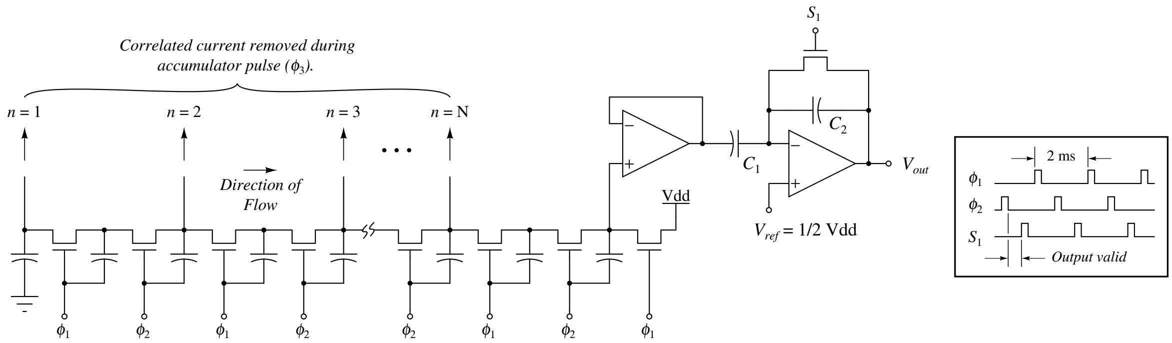

Fig. 2. The bucket brigade device (BBD).

Figure 2 shows the bucket brigade line. It is driven by a nonoverlapping two-phase clock. Each BBD transistor's gate is coupled to its drain through a capacitor (preferably several hundred fF or greater). At the start of a cycle, the analog value to be transferred (a voltage measured negatively from V_{dd}) is stored on the transistor source, and a value of $V_{dd} - V_{th}$ is stored on the transistor drain, where V_{th} is the transistor threshold voltage. The clock raises the gate-to-source voltage V_{gs} of the transistor while ensuring a positive drain-to-source voltage V_{ds}, initiating current flow from drain to source. Provided that the capacitors at drain and source of the transistor are the same size, the voltage drop at the drain will equal the voltage rise at the source. Current flow becomes negligible when $V_{gs} = V_{th}$. The final voltage at the source is therefore $V_{dd} - V_{th}$, and the final voltage at the drain, when the clock voltage returns to ground, is the value originally at the source.

In our bucket brigade accumulator, values are accumulated by integrating the summed current from each column of the correlator array off of the bucket brigade capacitor. The integration occurs over a constant amount of time determined by the pulse ϕ_3. Since the input currents vary slowly with respect to the pulse width of ϕ_3, the integration of current closely approximates a current-to-voltage conversion, resulting in a voltage drop at the bucket brigade stage proportional to the input current.

A simple single-transistor switch controlled by the signal ϕ_3 (active low) enables the integration. This circuit is shown in Figure 3. In order to prevent charge injection on the sensitive bucket brigade node, we switch the transistor via its source rather than its gate. When the source is grounded, the transistor mirrors the current from the correlator column.

One requirement of this scheme is that the current drawn from the bucket brigade stage must reach the value of the desired integrating current in a time which is short compared to the pulse width of ϕ_3, and must drop to zero in equally short time after the ϕ_3 pulse ends. The parasitic capacitance between the source and gate of M_1 will tend to pull down the gate voltage when the source is lowered. If the gate is part of a simple mirror, then the parasitic capacitance is charged directly by the input current, which is too small to meet the requirement. One solution to the problem is a MOS version of the Wilson current mirror, shown in Figure 3. Transistor M_5 is 10 times wider than the other transistors, and can charge the parasitic capacitance quickly.

The output of the bucket brigade must be time-differenced according to Equation (7). We accomplish this with the switched-

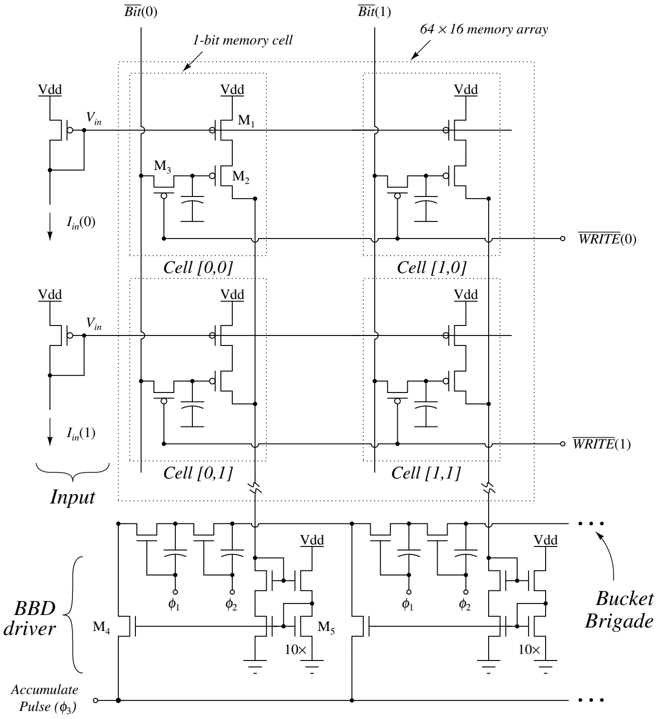

Fig. 3. The correlator array.

capacitor circuit shown in Figure 2. We measure the output relative to the voltage V_{ref}, and scale it by the ratio C_1/C_2. The circuit must be reset by S_1 at the same frequency as the bucket brigade clocks. The output is valid between the ϕ_2 and S_1 clocks. The bucket brigade is fully pipelined, yielding one full correlation at every time step.

V. Experimental Results

We have designed and fabricated this chip in a standard 1.2µm CMOS technology.

We have tested this chip by computing a correlation between simple input signals and a stored template on each frequency channel. These tests confirm the correct operation of the current correlation and the shifting of the bucket brigade over a range of clock frequencies much wider than that at which the chip is expected to operate normally.

Figures 4, 5, and 6 show the bucket brigade output of $c'_z[t]$ from Equation (6) when a constant input is integrated on a cluster of bucket brigade cells and then shifted to the output. In all the figures, the midline is $V_{dd} - V_{th}$; the output pulses above that baseline during clock phase ϕ_2, while the output pulses below the baseline represent the value in the bucket brigade cell, which should be measured from the baseline.

In Figure 4, the first 12 buckets accumulate current during the ϕ_3 phase. These appear at the output on the first twelve clock cycles. Differences in the bucket values are due to transistor mismatch in the bucket brigade driver mirror (M_4 and M_5 in Figure 3); from the standpoint of the correlation equation these can be viewed as small weights distributed around unity gain that are multiplied by each column sum. Figure 5 shows that the output is exactly the same when the clocking frequency is 200 times slower (about 0.5 ms per cycle). In Figure 6, the stored pattern in memory is located further back in the bucket brigade and the resulting pulse must shift seventeen stages to reach the output. Some diffusion of the signal occurs due to charge transfer inefficiency in the bucket brigade shifts.

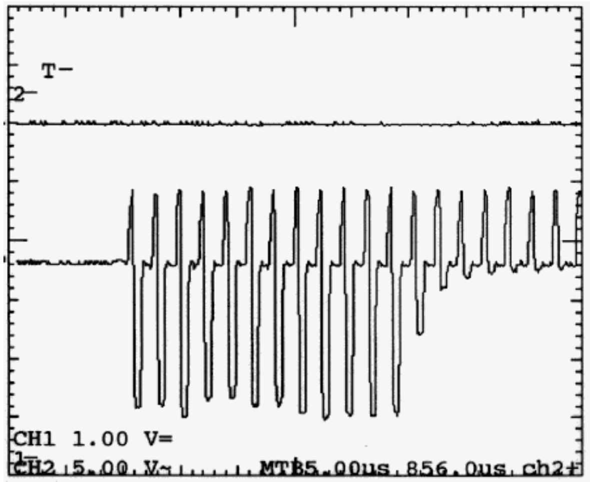

Fig. 4. Oscilloscope trace of pulse on the bucket brigade.

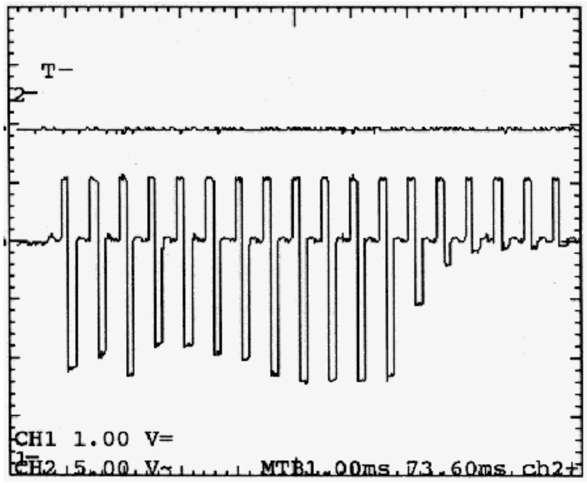

Fig. 5. Oscilloscope trace of pulse on the bucket brigade.

Results of system-level operation will be available when a chip with a revised circuit largely eliminating diffusion is fabricated by May 1997. We have also designed a current-mode frontend filterbank to interface to the correlator.

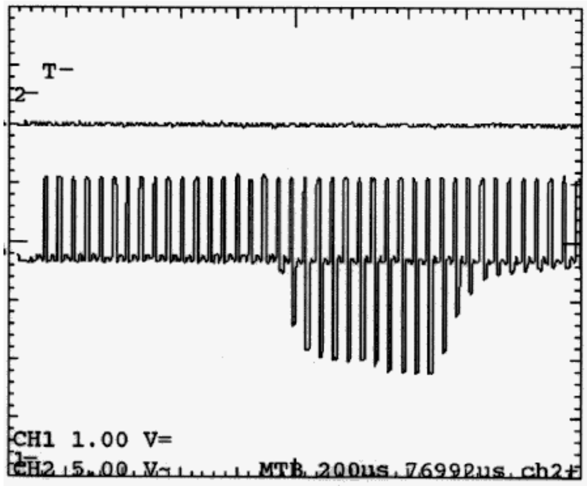

Fig. 6. Oscilloscope trace of pulse on the bucket brigade.

VI. Conclusions

We have designed and fabricated a chip intended for use as a classifier of transient (short-term) acoustic signals. The chip uses analog current-mode circuits to compute a correlation between an acoustic signal and a stored template of binary values, in which the input, presented in real time, is decomposed into an array of energy envelopes across a set of frequency channels, and normalized across all channels. A bucket-brigade device accumulates partial column sums over time. The pattern-classification algorithm which the chip implements has been proven through simulation to be robust to binarization of the template values when the template values encode the expected sign of the time derivative of each of the normalized input channels.

We have tested the correlator chip. Results show that the correlation output is valid over a large range of operating rates encompassing the entire auditory range and beyond. Upon fabrication of a frontend filterbank chip, we will perform tests to determine how effective the correlator is when used as part of a complete temporal pattern recognition system.

References

[1] K. P. Unnikrishnan, J. J. Hopfield, and D. W. Tank, "Connected-Digit Speaker-Dependent Speech Recognition Using a Neural Network with Time-Delayed Connections," *IEEE Transactions on Signal Processing*, **39**, pp. 698–713, 1991.

[2] F. Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-23, pp. 66–72, 1975.

[3] F. J. Pineda, K. Ryals, D. Steigerwald, and P. Furth, "Acoustic Transient Processing using the Hopkins Electronic Ear," World Conference on Neural Networks, Washington, D.C., 1995.

[4] F. J. Pineda, G. Cauwenberghs, R. T. Edwards, "Bangs, Clicks, Snaps, Thuds, and Whacks: an Architectue for Acoustic Transient Processing," (to appear in) *Neural Information Processing Systems*, 1996.

[5] R. Coggins, M. Jabri, B. Flower, and S. Pickard, "A Hybrid Analog and Digital VLSI Neural Network for Intracardiac Morphology Classification," *IEEE Journal of Solid-State Circuits*, **30**, pp. 542–550, 1995.

[6] E. R. Fossum, "Wire Transfer of Charge Packets Using a CCD-BBD Structure for Charge-Domain Signal Processing," *IEEE Transactions on Electron Devices*, **38**, Feb. 1991.