(From *Proc. Second ARPA RASSP Conference,* (Washington, D.C.), pp. 257–261, July 1995.)

Approximate Processing and Incremental Refinement Concepts

J. Winograd*, J. Ludwig, H. Nawab*, A. Chandrakasan, A. Oppenheim
Massachusetts Institute of Technology, RLE, Cambridge, MA 02139
*Boston University, ECS Department, Boston, MA 02215

Abstract

Approximate processing and incremental refinement concepts are needed for applications where it is desirable to provide a systematic tradeoff between the quality of signal processing results and the availability of resources, such as time, bandwidth, memory, and power. We examine the impact of these concepts for three distinct application areas: (1) low-power frequency-selective FIR filtering, (2) real-time time-frequency analysis of signals and (3) DCT-based image encoding/decoding. Results from approximate processing of signal data illustrate the practical utility of these types of systems.

1 Introduction

Approximate processing of signals may be used by a signal processing system to potentially achieve a wide range of possible tradeoffs between utilization of system resources (cost) and system performance (output quality). We have formulated the concept of *incremental refinement* (IR) structures for signal processing transformations which in conjunction with appropriate *control strategies* may be used in designing approximate signal processing systems for various applications. These IR structures have the property that at intermediate stages of computation they produce approximations to the final output. In this paper, we report results from our exploration of three categories of practical applications for IR-based approximate processing systems.

2 A Low-Power Application

Recent investigations [1] [2] [3] related to low-power CMOS implementations of signal processing systems have been motivated by the growing demand for portable multi-media devices. A key issue in the design of such devices is to minimize the total power consumption in order to maximize the run time or minimize battery size. In this section we develop an algorithm designed for low-power frequency-selective FIR filters, an essential element in many communication devices. A key feature of this algorithm is that the number of filter taps used is dynamically varied to provide stopband attenuation in proportion to a simple estimate of the time-varying energy in the undesired components of the input signal.

To first order, the average power required to perform a signal processing task is: $P = \sum_i N_i C_i V_{dd}^2 f_s$, where C_i is the average capacitance switched per operation of type i (corresponding to addition, multiplication, storage, etc.), N_i is the number of operations of type i performed per sample, V_{dd} is the operating supply voltage, and f_s is the sampling frequency. One approach to power reduction is to lower N_i, which in the case of FIR filtering corresponds to the number of taps used to produce each output sample.

Consider a tapped delay line implementation (Fig. 1) of an FIR filter whose taps correspond to a rectangularly windowed version of the impulse response of an ideal frequency selective filter. This tapped delay line

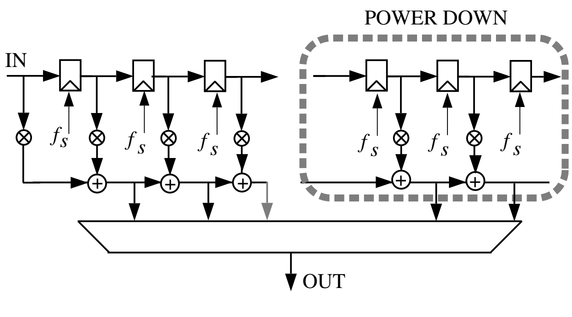

Figure 1: Illustration of dynamic tapped delay line

may be viewed as an incremental refinement structure for carrying out frequency selective filtering. If a subset of the taps are *powered down* (effectively disabling portions of the tapped delay line) in such a way that the length of the corresponding ideal impulse response is further truncated, then the net stopband attenua-

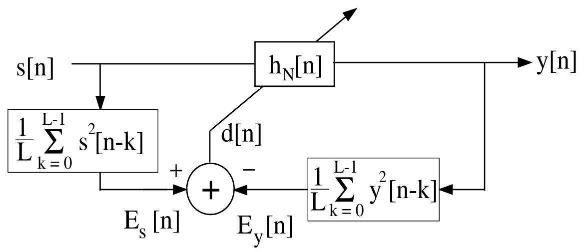

Figure 2: Overview of approximate filtering strategy

tion in the FIR filter output decreases proportionately. Conversely, *powering up* of additional taps can be used to increase the stopband attenuation in the FIR filter output. We have developed a *control strategy* to utilize this incremental refinement property of the tapped delay line for constructing approximate filtering systems whose stopband attenuation (and the corresponding power consumption) is adapted to the time-varying stopband energy of the input signal.

Our control strategy for approximate filtering is depicted in Figure 2. Note that only two additions and two multiplications are required to obtain $E_s[n+1]$ from $E_s[n]$. To analyze the control strategy, we assume that the input signal $s[n] = x[n] + w[n]$, where $x[n]$ and $w[n]$ correspond to wide sense stationary, independent random processes. The output $y[n]$ is produced by filtering $x[n]$ at each time, n, with an $N[n]$-tap filter having frequency response $H_{N[n]}(\omega)$. Given that the output power spectral density is

$$P_y(\omega) = P_s(\omega)|H_{N[n]}(\omega)|^2 + P_w(\omega)|H_{N[n]}(\omega)|^2, \quad (1)$$

our objective is to choose $N[n]$ to be the smallest odd integer which assures that

$$\int_{SB} P_y(\omega)|H_{N[n]}(\omega)|^2 d\omega < \gamma, \quad (2)$$

where the parameter γ is the maximum tolerable stopband energy in the output, and SB denotes the stopband region in frequency. Defining $A_{SB}(N[n]) = \int_{SB} |H_{N[n]}(\omega)|^2 d\omega$ and $Q = d[n]A_{SB}[n-1]$, the decision rule for choosing $N[n]$ is:

$$\begin{array}{rcl} Q > \gamma & \longrightarrow & \text{increase } N \text{ by } N_0 \\ \gamma - \delta < Q < \gamma & \longrightarrow & N \text{ unchanged} \\ Q < \gamma - \delta & \longrightarrow & \text{decrease } N \text{ by } N_0 \end{array} \quad (3)$$

The parameters δ and N_0 control the sensitivity of the dynamically-varying filter length. The values of $A_{SB}(N[n])$ can be precalculated and stored in a lookup table. In the case of stationary interferences, the filtering algorithm will converge to the length which satisfies eqn. (2) and remain fixed thereafter. For nonstationary interferences, the filter length will dynamically vary to approach the appropriate length at each time.

We have successfully applied our approximate filtering system to the problem of separating a low-frequency FDM voice channel from a high-frequency one and demonstrated that significant power savings over conventional methods may be obtained. The incorporation of adaptive approximate filters into a binary-tree structured filterbank for low-power source coding applications has also been successfully explored.

3 A Real-Time Application

The discrete Fourier transform (DFT) is an important component of many real-time signal processing systems. For those in which real-time deadlines vary or computational resources are dynamically allocated, the incremental refinement approach to DFT computation provides a framework in which the availability of an approximate result can be guaranteed across a wide range of resource allocations.

In our recent work, we have developed a class of incremental refinement structures for DFT computation. These structures allow many different quality/cost refinement paths to be achieved. Such structures may be derived by considering the real-valued N-point signal under analysis to be represented in radix complement form using a fixed-point mixed-radix encoding[1] [4]. When a D-digit signal representation based upon the radices $(m_{D-1}, m_{D-2}, \ldots, m_1, m_0)$ is employed and m_{D-1} is even, the value of each signal point $x(n)$ can be related to the value of the digits with which it is encoded via:

$$x(n) = \sum_{d=0}^{D-1} \alpha_d(x_d(n))\beta_d \quad (4)$$

where $x_d(n)$ is the dth digit from the least-significant digit of the word representing $x(n)$,

$$\alpha_d(x) = \begin{cases} x, & (d \neq D-1)\vee \\ & (0 \leq x \leq (m_{D-1}/2) - 1), \\ x - m_{D-1}, & (d = D-1)\wedge \\ & (m_{D-1}/2 \leq x \leq m_{D-1} - 1), \end{cases} \quad (5)$$

and

$$\beta_d = \begin{cases} 1, & d = 0, \\ \prod_{j=0}^{d-1} m_j, & 1 \leq d \leq D-1, \end{cases} \quad (6)$$

[1] We note that fixed-point binary and, in some instances, floating-point representations can be considered as special cases of mixed-radix.