MIXED ANALOG-DIGITAL MULTIRATE SIGNAL PROCESSING

Sanjit K. Mitra
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106, U.S.A.

José E. Franca
Department of Electrical and Computer Engineering
Instituto Superior Técnico
Av. Rovisco Pais, 1, 1096 Lisboa Codex, Portugal

ABSTRACT
To achieve higher levels of integration there has been a growing interest in recent years in designing systems containing both analog and digital functions on a single integrated circuit. In most cases, these are inherently multirate systems because of the different sampling rates employed at various stages of the system. This paper reviews some recent developments in this area of integrated multirate analog-digital systems, with a special emphasis on their applications to communication systems.

1. INTRODUCTION
This paper discusses some recent developments in the area of multirate signal processing, including digital, analog and even mixed-signal analog-digital. We start, in Section 2, by reviewing fundamentals of multirate theory. Selected examples of mixed analog-digital multirate systems for signal processing functions are discussed in Section 3. A brief summary of the paper and its conclusions are finally presented in Section 4.

2. A REVIEW OF MULTIRATE THEORY
2.1 Multirate Fundamentals
To achieve different sampling rates at different stages, a linear multirate discrete-time system employs devices that either increases or decreases the sampling rates by an integer factor in addition to the three basic components of a linear single-rate discrete-time system, namely, the unit delay, the adder and the multiplier. The basic sampling rate alteration devices are the up-sampler and the down-sampler.

Likewise, the down-sampler with a down-sampling factor of M, shown symbolically in Figure 1(b), generates at its output a discrete-time sequence x[n] with a sampling rate that is M-times lower than its input discrete-time sequence x[n] where M is a positive integer. The down-sampling operation is implemented by keeping every M-th sample of the input at the output and removing M–1 in-between samples. The time-domain input-output relation of the down-sampler is given by

where X_d(e^{j\omega}) and X(e^{jL\omega}) denote the Fourier transforms of x_d[n] and x[n], respectively.

![Figure 1](a) Up-sampler and (b) Down-sampler.

Likewise, the down-sampler with a down-sampling factor of M, shown symbolically in Figure 1(b), generates at its output a discrete-time sequence x_d[n] with a sampling rate that is M-times lower than its input discrete-time sequence x[n] where M is a positive integer. The down-sampling operation is implemented by keeping every M-th sample of the input at the output and removing M–1 in-between samples. The time-domain input-output relation of the down-sampler is given by

As a result, if F_s is the sampling rate of x[n], the sampling rate of x_d[n] is F_s/M. In the frequency-domain the corresponding input-output relation is given by

where X_d(e^{j\omega}) and X(e^{j\omega}) denote the Fourier transforms of x_d[n] and x[n], respectively.

2.2 Interpolator and Decimator
As can be seen from Eq. (2), a factor-of-L sampling rate expansion results in an L-fold compression of X(e^{j\omega}) and L–1 repetitions, called images, of the compressed input spectrum. The images are filtered out by a lowpass filter H_u(z) which replaces the zero-valued samples in x_u[n] with interpolated sample values. The cascade of an up-sampler and the lowpass filter shown in Figure 2(a) is called an interpolator. The ideal interpolation is implemented by passing x_u[n] through an ideal lowpass filter with a frequency response [1]

Likewise, in the case of a down-sampler, as indicated by Eq. (2), X_d(e^{j\omega}) is the sum of M frequency scaled and
shifted images of \(X(e^{j\omega})\) with adjacent images of \(X_d(e^{j\omega})\) separated by \(2\pi\). Thus, unless \(X(e^{j\omega})\) is band-limited to the region \(0 \leq \omega \leq \pi/M\), there will be overlap of adjacent terms in Eq. (4) causing aliasing. As a result, to prevent any aliasing that may be caused by the down-sampling process, \(x[n]\) is usually first passed before down-sampling through a lowpass filter \(H_d(z)\) approximating the ideal frequency response

\[
H_d(e^{j\omega}) = \begin{cases} 1, & 0 \leq |\omega| \leq \pi / M, \\ 0, & \pi / M < |\omega| \leq \pi. \end{cases}
\]  

(6)

The combination of an anti-aliasing filter and a down-sampler as indicated in Figure 2(b) is usually called a **decimator** and the process of lowering the sampling rate by this arrangement is called **decimation**.

\[
x[n] \xrightarrow{L} H_u(z) \xrightarrow{y[n]}
\]

(a)

\[
x[n] \xrightarrow{H_d(z)} v[n] \xrightarrow{M} y[n]
\]

(b)

**Figure 2.** (a) Interpolator and (b) Decimator.

For a fractional-rate sampling rate interpolation (decimation), a cascade of an up-sampler, a lowpass filter and a down-sampler, as indicated in Figure 3, is usually employed where the interpolation (decimation) factor is given by \(L/M\).

\[
x[n] \xrightarrow{L} H(z) \xrightarrow{M} y[n]
\]

**Figure 3.** Sampling rate increase by a rational factor \(L/M\).

### 2.2 Polyphase Decomposition [2]

An arbitrary sequence \(\{x[n]\}\) with a z-transform \(X(z)\):

\[
X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}
\]

(7)

can be rewritten in the form

\[
X(z) = \sum_{k=0}^{M-1} z^{-k}X_k(z^M)
\]

(8)

where

\[
X_k(z) = \sum_{n=-\infty}^{\infty} x_k[n]z^{-n} = \sum_{n=-\infty}^{\infty} x[Mn+k]z^{-n}, \quad k = 0, 1, \ldots, M-1.
\]

(9)

The subsequences \(\{x_k[n]\}\) are called the **polyphase components** of the parent sequence \(x[n]\), and the functions \(X_k(z)\), given by the z-transform of \(\{x_k[n]\}\), are called the **polyphase components** of \(X(z)\). The relation between the sub-sequences \(\{x_k[n]\}\) and the original sequence \(\{x[n]\}\) is given by:

\[
x_k[n] = x[Mn + k], \quad k = 0, 1, 2, \ldots, M - 1.
\]

(10)

The polyphase decomposition of a causal FIR transfer function can be carried out by inspection. The polyphase decomposition of a causal IIR transfer function \(H(z) = N(z)/D(z)\), on the other hand, is not that straightforward. One way to arrive at an M-band polyphase decomposition of \(H(z)\) is to express it in the form \(N(z)/D(z^M)\) by multiplying the denominator \(D(z)\) and the numerator \(N(z)\) with an appropriately chosen polynomial and then apply an M-branched polyphase decomposition to \(P(z)\).

### 2.3 M-th Band Filters

A special case of the polyphase decomposition of a lowpass interpolation (decimation) filter for a factor-of-M sampling rate alteration is given by

\[
H_u(z) = \alpha + \sum_{i=1}^{M-1} z^{-j}G_i(z^M),
\]

(11)

where the constant \(\alpha\) is usually chosen as 1/M. A lowpass filter with a transfer function of the form of Eq. (11) is usually called an **M-th band or Nyquist filter** [2]. The impulse response \(h_d(n)\) of such a filter is given by

\[
h_d(n) = \begin{cases} \alpha, & n = 0, \\ 0, & \text{otherwise}. \end{cases}
\]

(12)

For a transfer function \(H_u(z)\) satisfying Eq. (11), it can be shown that

\[
\sum_{k=0}^{M-1} H_u(z W_M^k) = M \alpha = 1,
\]

(13)

where \(W_M = e^{\pm j2\pi/M}\). An M-th band filter for \(M = 2\) is called a **half-band filter**.

Under certain conditions, a causal and stable M-th band IIR transfer function \(H_u(z)\) can be decomposed as a sum of allpass transfer functions in the form

\[
H_u(z) = \frac{1}{M} \sum_{i=0}^{M-1} z^{\pm j} A_i(z^M),
\]

(14)

where \(A_i(z)\) are stable allpass transfer functions for a stable \(H_u(z)\). In the case of a 2-band (\(M = 2\)) decomposition, it has been shown that all Butterworth and elliptic odd-order half-band transfer functions can be analytically decomposed in the form of Eq. (19) [3]. For M-th band transfer functions with \(M > 2\), a computer-aided design approach to develop the decomposition has been proposed [4].

### 2.4 Efficient Implementation of Interpolators and Decimators

It follows from above that an M-th band filter require fewer computations than an equivalent lowpass interpolation filter. Further savings in the computational efforts can be obtained by realizing the interpolation filter in a polyphase decomposition. In the implementation of the lowpass filter in the interpolator structure of Figure 2(a), the arithmetic
operations are carried out at the sampling rate \( LF_s \) of the output rather than at the lower sampling rate \( F_s \) of the input. A computationally efficient structure for the implementation of the decimator can be obtained by making use of an L-band polyphase representation of the lowpass filter:

\[
H_u(z) = \sum_{i=0}^{L-1} z^{-i} G_i(z^{L^{i}}),
\]

resulting in the structure of Figure 4 where the arithmetic operations are carried out at the lower sampling rate \( F_s \). Similar computationally efficient structures can be derived for the implementation of decimators based on a polyphase decomposition of the decimation filters.

**Figure 4:** Computationally efficient interpolator structure.

### 2.5 Filter Banks

The analysis filter bank is a set of digital bandpass filters \( H_i(z) \) with a common input and M outputs as shown in Figure 5(a). The passbands of the filters \( H_i(z) \) are usually contiguous and non-overlapping frequency bands of width \( 2\pi/M \). The input signal \( x[n] \) is thus decomposed into a set of M subband signals \( x_i[n] \) with each occupying a frequency band of width \( 2\pi/M \). As the subband signals have a narrow bandwidth, in many applications they are down-sampled by a factor of M to provide computational advantages.

Likewise, the synthesis filter bank is a set of digital bandpass filters \( F_i(z) \) with M inputs and a summed output as shown in Figure 5(b). The passbands of the filters \( F_i(z) \) are usually contiguous and non-overlapping frequency bands of width \( 2\pi/M \). The operation of the synthesis filter bank is exactly opposite to that of the analysis filter bank.

\[
\sum_{i=0}^{M-1} G_i(z^{L^{i}}) \hat{v}_i[n] = y[n]
\]

**Figure 5:** (a) Analysis filter bank, and (b) synthesis filter bank.

A novel multirate structure, developed primarily for efficient data compression, is the M-channel quadrature mirror filter (QMF) bank shown in Figure 6 [1].

In this structure, the input discrete-time signal \( x[n] \) is passed through an analysis filter bank consisting of a set of \( M \) bandpass filters \( H_i(z) \) decomposing \( x[n] \) into a set of \( M \) subband signals \( x_i[n] \) occupying contiguous nonoverlapping frequency bands of width \( 2\pi/M \). The subband signals are then down-sampled by a factor of \( M \). Next, these down-sampled signals are up-sampled by a factor of \( M \), and passed through a synthesis filter bank composed of a set of \( M \) bandpass filters \( F_i(z) \) to remove the images, resulting in an output signal \( y[n] \) whose sampling rate is the same as that of the input \( x[n] \). The input-output relation in the z-domain is given by

\[
Y(z) = \sum_{k=0}^{N-1} \sum_{i=0}^{M-1} H_k(z W_i^N) X(z W_i^N).
\]

**Figure 6** M-channel quadrature-mirror filter (QMF) bank.

The analysis and synthesis filters are usually designed to achieve a perfect reconstruction at the output implying that the output is a delayed and scaled replica of the input with complete aliasing cancellation [1]. In this case, Eq. (16) reduces to

\[
Y(z) = \frac{1}{N} \sum_{k=0}^{N-1} F_k(z) Y(z) = z^{-N} X(z).
\]

### 3. Mixed Analog-Digital Multirate Structures

#### 3.1 Oversampling A/D Converters

The oversampling A/D converters are inherently multirate because they relate an input sequence of samples obtained at high frequency to an output sequence of digital words produced at much lower frequency. Such converters achieve high resolution by means of oversampling and noise shaping, and have been shown to be a good alternative for the implementation of high resolution, low-speed A/D converters mainly because they can accommodate the limitations associated with the implementation of its analog processing in digital MOS technology [5]. As illustrated in the schematic diagram of Figure 7, an analog input signal is sampled at a frequency significantly higher than \( f_{\text{max}} \), the higher frequency present in the input signal. The correspondingly high frequency bit stream produced by a coarse quantizer is then decimated to a much lower frequency digital signal, until the required word length is formed. Such decimation is performed by means of a digital decimator consisting of a digital Nyquist lowpass filter together with a
down-sampler. The resulting A/D converter is called a delta-sigma A/D converter since it employs a delta-sigma modulator as the basic building block.

**Figure 7:** Typical architecture of a multirate analog-digital conversion system based on sigma-delta modulation.

A special case of Figure 7 is when \( R = 1 \). Here, the digital estimator is a simple voltage comparator and the D/A converter has one-bit resolution. The delta-sigma converter thus quantizes a very high sampling rate (e.g., 2 MHz) analog input sample sequence with very low resolution (e.g., 1 bit). With the aid of downsampling and digital filtering, the sampling rate is reduced (e.g., 8 kHz) and the resolution is increased (e.g., 10 bits). The basic idea for SNR enhancement in a delta sigma converter is to reduce the quantization noise by a high-pass filter leaving only a small fraction of the noise power inside the frequency band of interest. Further SNR improvement can be obtained by increasing the order of the filter [6]. Although its basic concept has been around for quite a few years, the integrated circuit technology needed to make the sigma-delta converters a viable solution has only recently become available. The delta-sigma A/D converter can be implemented on a single chip using MOS technology [5], with the analog part (delta sigma converter) usually implemented by SC circuits. The digital part consists of the decimator. Recently, five delta sigma modulators, combined in a parallel architecture, have been integrated onto a single 1.2 \( \mu \) CMOS chip [7]. A time-interleaved array of sigma delta modulators also has been reported [8, 9].

### 3.2 QMF-Based A/D Converters

The type of quadrature-mirror filter (QMF) bank described above can be used to implement a high-speed A/D converter, suitable for applications in the video-rate range (up to \( \approx 100 \) MHz) using a number of low-speed, low-cost A/D converters, as indicated in Figure 8 [10, 11]. Here, the analysis stage is an SC network composed of the analysis filter bank and the down-samplers, while the synthesis stage is a digital network composed of the synthesis bank filters and up-samplers.

**Figure 8:** High-speed analog-digital conversion system based on the quadrature-mirror filter.

Basically, in this approach the input signal is first decomposed into a number of contiguous frequency bands (subbands) so that a specific A/D converter (subconverter) can be assigned to each subband signal.

Although the structure of Figure 8 is a natural extension of the time-interleaved A/D converter [12, 13] (in fact, the latter can be seen as a particular case, where the analysis filters \( H_k(z) = z^{-k} \) and the synthesis filters \( F_k(z) = z^{-(M-1-k)} \)), the harmonic distortion due to mismatches among the subconverters is substantially reduced at a little additional cost [14]. The jitter problem due to uneven sample timing, which is another source of error in time-interleaved A/D converters, specially for high-frequency input signals, is reduced by the decimation stage in Figure 8. Similar idea has been used by Poulton et al. [15] and has been called a two-rank architecture. The QMF-based converters also incorporate the advantages of subband coding: By appropriately specifying the resolution of the subconverters throughout the respective subbands, the quantization noise can be separately controlled in each band, and the shape of the reconstruction error spectrum can be controlled as a function of the frequency. This strategy has been used in many speech and image coding applications.

A more efficient structure for implementation can be obtained by realizing the analysis and the synthesis filters in polyphase form, and then moving the down-samplers to the left of the polyphase sub-filters in the analysis bank and moving the up-samplers to the right of the polyphase sub-filters in the synthesis bank, as indicated in Figure 9 for a two-band decomposition. Using a tree structure, QMF banks with more than two bands can be developed from the two-band structure of Figure 9 [16].

**Figure 9:** Example of a more efficient structure for implementation of a high-speed A/D conversion system based on the QMF bank.

### 3.3 Front-End Filtering Sub-Systems

Multirate signal processing filter structures are often used in front-end sub-systems, specially for high-frequency baseband
signals as is the case in video processing applications. The practical advantages to be gained with this approach are best illustrated with the example given in Figure 10(a) where a continuous-time filter and a sampled-data decimator are optimally combined to meet the requirements of a single-chip front-end sub-system for digitizing video signals according to the CCIR 601 recommendation [17]. This predominantly multirate analog system solution can lead to significant savings in power and silicon consumption when compared with the functionally equivalent but predominantly multirate digital solution illustrated in Figure 10(b) [18, 19].

Two solutions have been studied for the SC realization of the decimation and filtering function, one based on an active-delayed block polyphase structure that realizes a finite impulse response transfer function with 19 coefficients and a sampling rate reduction factor of $M = 5$ [20], and the other based on the optimum implementation of a 5th order elliptic bilinear z-transfer function with a down-sampling factor of $M = 3$ [21]. In both circuit solutions, it has been shown that all amplifiers have to settle within a time interval of approximately 25 nS, which can be comfortably achieved using state-of-the-art CMOS amplifiers [22].

An area of potential interest for the N-path approach is in mobile radio system applications, where low power is a major concern. For this purpose, the design of a mixed analog/digital N-path architecture is currently under investigation [28]. A predominantly digital form of implementation of such system is illustrated in Figure 11(a), consisting of an analog front end serial/parallel converter and an A/D converter followed by a digital filter in each path. It should be noted that all of the attractive features of the all-analog N-path filters are also present in such mixed-signal implementation. For example, relatively simple digital filters in each path can be used to generate an overall filter with a highly selective bandpass filter. Moreover, as in the pseudo N-path approach, a high-speed digital filter can be time-multiplexed N times to make it look as N separate filters, thus ensuring that each path of the N-path structure is identical. In the alternative, predominantly analog form of implementation shown in Figure 11(b), an analog sampled-data decimator is placed in front of the channel ADCs working at a lower sampling frequency and thus allowing for possible savings in power dissipation in the conversion blocks. In this case, however, the analog decimation filter blocks may not be as well matched as their digital counterparts in Figure 11(a) and hence yielding a probable increase of inband spurious signals in the overall bandpass filter responses of the system.

### 3.4 Mixed Analog-Digital Implementation of N-path Filters

The N-path filter, originally introduced for the processing of analog continuous-time signals [23], consists of N identical time-invariant filters, to each of which a common input signal path and a common output signal path are periodically connected synchronously, through appropriate modulators. For certain bandlimited modulating functions, or certain bandlimited input and output signals, the overall time-varying structure can be made to appear as a time-invariant network, which leads to its possible practical advantages pointed out earlier. However, the difficulty in ensuring the analog filter paths to be identical has made it difficult to implement the overall structure in a monolithic integrated circuit to reap its benefits. One way to overcome the problem of mismatches between the N paths is by using the pseudo N-path principle [24]. The SC technology has provided a practical method for the fully integrated realization of high-quality narrowband filters based on the N-path concept, as, for instance, in [25, 26]. The N-path concept has also been extended to an all-digital structure by Mitra et al. [27].

![Figure 10: Multirate architectures for CCIR 601 digitization of video signals. (a) Predominantly analog multirate and (b) predominantly digital multirate.](image)
3.5 Block Implementation of Digital Filters

Consider a linear, time-invariant causal FIR digital filter characterized by an impulse response $h(n)$ of length $N$. The input-output relation of this filter is given by the convolution sum

$$y(n) = \sum_{r=0}^{N-1} h(i)x(n-r), \quad n \geq 0,$$

where $x(n)$ and $y(n)$ are, respectively, the input and output sequences. The above equation permits a sequential computation of the output, one sample at a time, beginning with $y(0)$. The output sequence of an FIR filter can also be computed in blocks of length $L$ with $L \leq N$ by rewriting Eq. (17) in a block-convolution form given by

$$Y_k = H_k X_k$$

where $X_k$ and $Y_k$ are, respectively, the $k$-th input and output blocks of length $L$:

$$X_k = \begin{bmatrix} x(kL) & x(kL+1) & \ldots & x(kL+L-1) \end{bmatrix},$$

$$Y_k = \begin{bmatrix} y(kL) & y(kL+1) & \ldots & y(kL+L-1) \end{bmatrix},$$

and $H_k, k = 0, 1, 2, \ldots, L-1$, are $L \times L$ matrices composed of the impulse response coefficients. A block implementation of an FIR filter for $L = 3$ is sketched in Figure 12 [29].

Likewise, in the case of an M-th order IIR filter characterized by a transfer function given by

$$H(z) = \sum_{i=0}^{M} a_i z^{-i},$$

one possible block representation is of the form

$$Y_{k+1} = -B_0^{-1}B_1 Y_k + B_0^{-1}A_0 X_{k+1} + B_0^{-1}A_1 X_k,$$

Figure 13: Illustration of the block implementation of an IIR filter.

A multirate interpretation of block processing is indicated in Figure 14, also for $L = 3$.

3.6 Mixed Analog-Digital Implementation of Block Filtering

Monolithic implementations of digital signal processors with on-chip A/D and D/A converters have become feasible due to the developments of mixed analog-digital CMOS-based technologies, including BiCMOS technology which combines the high integration density and low power consumption of CMOS with the high-speed and driving capability of bipolar transistors. With such technology it is also possible to fabricate other types of analog and digital circuitry on the same chip, in which signal processing can be performed partially in analog form and partially in digital form. On the other hand, the difficulties of implementing single chip processors for high-frequency applications include the large die areas occupied by high-speed A/D converters, and the realization of very fast digital filters. For video applications, the flash conversion approach is still the most popular method for A/D conversion. However, in addition to...
the large die area associated with this technique, the inevitable input capacitance resulting from paralleling many comparators may seriously limit the achievable conversion speed.

In the alternative mixed analog-digital architectures discussed herein, we exploit the savings in the die area achieved with the use of an array of low-speed successive approximation A/D converters, and the high-speed implementation of FIR and IIR filters by using block filtering techniques. The potentially high-speed processing achieved with this approach comes from the parallelism of both the A/D conversion and the digital filtering stages.

The block processing approach lends itself easily for the digital processing of analog signals. A monolithic signal processor can be implemented if an array of low-speed A/D converters is placed between the serial/parallel register (which can be implemented by CCD devices or SC networks) and the block filter. Alternatively, the block convolution can be realized using DFT methods which, in turn, can be implemented by CCD or SC networks. In this case the DFT operation can be removed from the block digital filtering stage and placed between the decimators and the A/D array.

![Figure 15: Illustration of a monolithic signal processor implemented by placing an array of low-speed A/D converters between a serial/parallel register and a block filter.](image)

4. CONCLUDING REMARKS

This paper discussed some recent developments in the area of multirate signal processing, including digital, analog and even mixed-signal analog-digital. After reviewing fundamentals of multirate theory a few selected examples of mixed analog-digital multirate systems for signal processing and conversion functions were discussed. These included the popular oversampling converters as well some new architectures based on quadrature-mirror filter banks, analog-digital interfacing for video processing, and mixed analog-digital implementations of N-path and block filters. In such architectures the benefits of multirate analog-digital integrated systems can be fully exploited for integrated circuit implementation using modern technologies.

Acknowledgment: This work was supported in part by a grant from the Rockwell International Corporation.

References:


