A Time Correlated Approach to Adaptable Digital Filtering

Hy Grossman
Senior Design Engineer

Steve Pellarin
VP of Engineering

Teletronics Technology Corporation
Newtown, PA USA

ABSTRACT

Signal conditioning is a critical element in all data telemetry systems. Data from all sensors must be band limited prior to digitization and transmission to prevent the potentially disastrous effects of aliasing. While the 6^{th} order analog low-pass Butterworth filter has long been the de facto standard for data channel filtering, advances in digital signal processing (DSP) techniques now provide a potentially better alternative.

This paper describes the challenges in developing a flexible approach to adaptable data channel filtering using DSP techniques. Factors such as anti-alias filter requirements, time correlated sampling, decimation and filter delays will be discussed. Also discussed will be the implementation and relative merits and drawbacks of various symmetrical Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) filters. The discussion will be presented from an intuitive and practical perspective as much as possible.

KEY WORDS

Digital Filter, Adaptive, DSP, FIR, IIR, Time Correlation, Aliasing

INTRODUCTION

All modern telemetry systems that collect continuous analog data from accelerometers, strain gauges and a host of other sensors, convert each analog signal to a digitally sampled data stream prior to transmission and/or storage. This is such a basic aspect of telemetry that we rarely consider that a continuous analog signal is really a very different animal from a digital data stream. Properly sampled digital data can be considered properly sampled if and only if a high quality replica of the pre-sampled analog signal can be reconstructed from the digital data. Two things are required for proper sampling:

1. Sufficient Analog-to-Digital (A/D) converter resolution. If the A/D converter step size is on the order of the analog noise in the system, no low-level data fluctuations present in the analog data will be lost.
2. The frequency spectrum of the sampled signal must be band-limited to no more than ½ the sampling frequency. This requirement is known as the Nyquist criterion and although there is some confusion about the term's definition, we will refer to the Nyquist frequency or Nyquist rate as meaning half the sampling frequency. Violate

the Nyquist criterion and you get aliasing, which can make a high frequency electrical noise signal indistinguishable from a low frequency vibration and render your data useless.

Requirement 2 is the primary reason a precision low pass filter is part of the entry price of every analog data channel in a telemetry system.

ANALOG VS. DIGITAL

Low-pass filters used to band-limit channel data can be analog or digital, which is to say, they can employ resistors, capacitors and op amps and operate on analog signals or they can mathematically process the digital data stream. Analog filters are simple to build and understand and have low power consumption. On the other hand, they are relatively inflexible and their accuracy is limited by available component tolerances. Digital filters require a DSP or other high speed numeric processor and associated support hardware and ironically, also require a reasonably good analog filter for anti-aliasing (which will be discussed in greater detail in the next section). Once implemented, they provide very high accuracy and flexibility way beyond what can be implemented with analog circuitry.

TIME CORRELATION

At this point, it would be useful to contrast how low pass filtering occurs in an analog filtered telemetry channel vs. a digitally filtered one. In telemetry channels that employ programmable analog filters such as those on TTC's SCD-108S signal conditioning card, each channel is periodically sampled in the data format at perhaps four times highest data frequency of interest and the analog filter has been set to cut off at a filter setting slightly above the highest data frequency of interest. Assuming the channel is configured for sequential sampling, each time the channel comes up in the format, an A/D convert command is generated and the output signal of the channel filter at that moment is digitally sampled. At the same time, the previous A/D sample is placed in the format data stream. The important point is that the analog filter continuously processes channel data but the data is sampled only when the channel appears in the format. The result is that every digital sample in the format is time-synchronous with its place in the format. This is time-correlated sampling.

In digitally filtered systems such as those employed by TTC's SCD-116D card, the channel A/D converter samples the data at a substantially higher rate than the channel appears in the format. These samples are then mathematically processed to provide each filtered data sample. So how do you achieve time-correlation in digitally filtered systems?

There are two ways. The most straightforward is to sample the channel data at a fixed rate that is very high compared to the format sample rate. Let's assume we have a channel that appears periodically in the format at a 4KHz rate and we want our channel cut off frequency to be 1KHz. If we sampled the channel data at a 100KHz rate, we would need to have a digital filter with a –3dB frequency of $0.01f_S$, where f_S is the data sample frequency. Since sampling is asynchronous to the format, the 100KHz sampled data used to calculate each filtered data point will have an age-uncertainty of one sample period or 10us. If we assume the data signal is a full scale, 1KHz sine wave, the maximum possible error in counts at zero-crossing resulting from time uncertainty is given by the formula:

$$\mathbf{E_C = (2^{BITS}/2) * SIN (\Delta t*360^\circ/P_D)} \qquad (1)$$

Where E_C is the maximum possible error in counts, BITS is the bit resolution of the ADC, Δt is the time-uncertainty (10µsec) and P_D is the maximum data frequency period (1000µsec).

For 12-bit sampled data, the maximum error can be as large as 128 counts or better than 3% of full scale. Not good. If we raise the filter sample frequency to 1MHz, we get about 13 counts of error or about 0.3% of full scale. Better but not outstanding. In addition, we incur the huge DSP processing burden needed to handle the higher sample rate.

The second approach is somewhat more complicated to implement but eliminates the problem described above by sampling the A/D converter at a frequency that is an exact multiple of the channel format rate and is phase-locked with it. Digitally filtered data collected this way will be time-correlated with the format, as are analog filtered samples taken each time the channel appears in the format (such as those collected with the SCD-108S card).

In the SCD-116D and the SCD-108D cards, digital phase locked loops (PLL) are used to multiply the format sample rate by 2^N where N is an integer selected to provide a sampling rate in the octave from 28KHz to 56KHz (PLLs in the SCD-608D, MSCD-104D and MSCD-604D operate from 56KHZ to 112KHz). Phase locking the DSP sample rate to the channel format rate eliminates sample latency error. Using the example cited earlier, the channel PLL will multiply the 4KHz format sample rate x8 (2^3) to provide an A/D sample rate of 32KHz. If the digital filter's -3dB frequency is set to 0.031 (1/32) of f_S, our channel -3dB frequency will be 1KHz as required.

By varying f_S to select the filter's -3dB frequency within a given octave, we greatly reduce the number of possible filter characteristics that need to be stored or calculated. On the other hand, we have made the job of the analog anti-aliasing filter more difficult. In a perfect world, we would like our anti-aliasing filter to have no effect on the highest data frequency of interest yet attenuate frequencies above the Nyquist rate ($f_S/2$) into oblivion. In reality a reasonable goal is <-0.1dB attenuation (about 1%) at the highest data frequency of interest and >-40dB attenuation at the lowest possible value of $f_S/2$. In TTC's SCD-116D, the maximum data frequency is specified as 2.8KHz (higher frequencies may be selected provided the attenuation of the anti-aliasing filter is allowed for) while the lowest possible value of $f_S/2$ is 14KHz. Achieving the required performance requires an anti-aliasing filter with a minimum of a 4^{th} order Butterworth characteristic.

The end result is a fully time correlated system, not just channel to channel on a given card, or card to card within a given chassis, but channel to channel throughout an entire Distributed Data Acquisition System. This applies whether the channels are sampled sequentially or simultaneously.

DECIMATION

Earlier references were made to digital filters having -3dB frequencies of $0.031f_S$ and lower. In most instances, such FIR and to a lesser extent IIR filters are not directly realizable. Consider a 40-tap symmetrical FIR filter. The form of an FIR filter is:

$$\mathbf{y_0 = a_0{*}x_0 + a_1{*}x_1 + a_2{*}x_2 + \ldots\ldots + a_{N-1}{*}x_{N-1}} \qquad (2)$$

where N = 40 for a 40-tap filter, x_0 to x_{N-1} are progressively older input sample values, a_0 to a_{N-1} are the filter coefficients, and y_0 is the current output value.

Basically, a FIR filter is nothing more than a weighted moving average filter where the weight of the k^{th} sample is determined by the value and sign of coefficient a_k. If we applied a sinusoidal input to the above filter whose period was greater than 80 samples, it would be impossible to fully block this signal from appearing in the output no matter what filter coefficients we selected. Table 1 shows a typical set of symmetrical FIR filter coefficients.

Table 1: Typical Set of 40-Tap Symmetrical FIR Filter Coefficients, Normalized to 1

a_0= -0.00002	a_5= 0.00011	a_{10}= 0.00937	a_{15}= -0.03896	a_{20}= 0.26842	a_{25}= -0.03941	a_{30}= 0.00179	a_{35}= 0.00052
a_1= -0.00003	a_6= -0.00133	a_{11}= 0.01344	a_{16}= 0.00662	a_{21}= 0.19940	a_{26}= -0.01635	a_{31}= -0.00275	a_{36}= 0.00032
a_2= 0.00007	a_7= -0.00305	a_{12}= 0.00547	a_{17}= 0.09640	a_{22}= 0.09640	a_{27}= 0.00547	a_{32}= -0.00305	a_{37}= 0.00007
a_3= 0.00032	a_8= -0.00275	a_{13}= -0.01635	a_{18}= 0.19940	a_{23}= 0.00662	a_{28}= 0.01344	a_{33}= -0.00133	a_{38}= -0.00003
a_4= 0.00052	a_9= 0.00179	a_{14}= -0.03941	a_{19}= 0.26842	a_{24}= -0.03896	a_{29}= 0.00937	a_{34}= 0.00011	a_{39}= -0.00002

If we carefully examine these coefficient values, we observe two things:

1. The first half of the coefficients is a mirror image of the second half. That is $a_0 = a_{39}$, $a_1 = a_{38}$, …. , $a_{19} = a_{20}$. This is what makes the filter symmetrical and also what makes its phase linear and its delay constant for all frequencies.
2. The mid-range coefficients are much larger than the end coefficients. This means samples 15 thru 24 are weighted much heavier than samples 0 thru 4 and 35 thru 39. It also means that the lowest frequency sinusoid we could effectively attenuate with a 40-tap FIR filter has a period no greater than about 20 samples rather than 80 samples. This makes the minimum practical $f_C \sim 0.05f_S$.

Enter decimation. In the earlier example, data with no significant frequency content above 14KHz was sampled at 32KHz, comfortably above the Nyquist rate. If we pass this data through a 40-tap FIR filter having a -3dB frequency of $0.095f_S$ (a decimation filter), we will attenuate all frequencies above $0.125f_S$ by greater than -40dB. In effect we have now band-limited our data to 0.125*32KHz or 4.0KHz. Next we decimate by 4. Theoretically, this means discarding three filtered data samples out of every four. In reality, we only execute the decimation filter algorithm once for every four input samples we add to the filter input buffer. The filtered output data is now applied to another 40-tap FIR filter with $f_C = 0.125f_S$. The effective sample rate of this filter is 32KHz/4 or 8KHz which again satisfies the Nyquist criterion. The filter -3dB frequency of this second filter is then 0.125*8KHz or 1.0KHz as required. To achieve lower -3dB frequencies, decimation can be repeated as many times as necessary.

STRENGTHS AND WEAKNESSES

The form of a FIR filter is given by **equation 2** which is repeated below for clarity. By examining the equation, it becomes clear that every output sample of an n-tap FIR filter is a function of only the filter coefficients, the current input sample and n - 1 previous input samples. Any change in the signal that occurred greater than n samples ago will have no effect on the filter output, hence the name ***F***inite ***I***mpulse ***R***esponse filter. This type of filter is also referred to as non-recursive since there is no feedback path from output to input. The FIR filter's non-recursive nature is also why it is inherently stable.

$$\mathbf{y_0 = a_0{*}x_0 + a_1{*}x_1 + a_2{*}x_2 + \ldots\ldots + a_{N-1}{*}x_{N-1}} \qquad (2)$$

A 6-pole IIR filter is composed of three cascaded 2-pole stages and can be mathematically represented by equations 3, 4 and 5. The output of the first cascaded stage is y_2, the second is y_1 and the third is y_0. Note the output of each stage is not just a function of the current input and two previous inputs but is also a function of two previous outputs as well. The output terms are what makes this a recursive filter and are also the source of any tendency toward oscillation.

$$\mathbf{y_0 = a_0{*}x_0 + a_1{*}x_1 + a_2{*}x_2 + b_1{*}y_1 + b_2{*}y_2} \qquad (3)$$
$$\mathbf{y_1 = a_3{*}x_1 + a_4{*}x_2 + a_5{*}x_3 + b_3{*}y_2 + b_4{*}y_3} \qquad (4)$$
$$\mathbf{y_2 = a_6{*}x_2 + a_7{*}x_3 + a_8{*}x_4 + b_5{*}y_3 + b_6{*}y_4} \qquad (5)$$

Referring to the figures below:

Figure 1 displays step responses for 120-tap and 40-tap FIR filters as well as those for analog and IIR digital versions of a 6-pole Butterworth filter. The –3dB frequency of all four filters is 14 Hz and the channel sample rate is 5x this value or 70 Hz.

Figure 2 displays the frequency domain performance for 120-tap and 40-tap FIR filters in the pass and transition bands. The –3dB frequency of both filters is $0.125f_S$.

Figure 3 displays the frequency roll-off and stop band characteristics of the same 120-tap and 40-tap FIR filters as in Figure 2.

Figure 4 displays the frequency domain performance for 6-pole Butterworth, 8-pole Butterworth, 6-pole Bessel and 6-pole Chebyshev IIR filters in the pass and transition bands. The –3dB frequency of all four filters is $0.050f_S$.

Figure 5 displays the frequency roll-off and stop band characteristics of the same four IIR filters as in Figure 4.

Figure 1: Analog/Ditital Filter Step Response and Delays

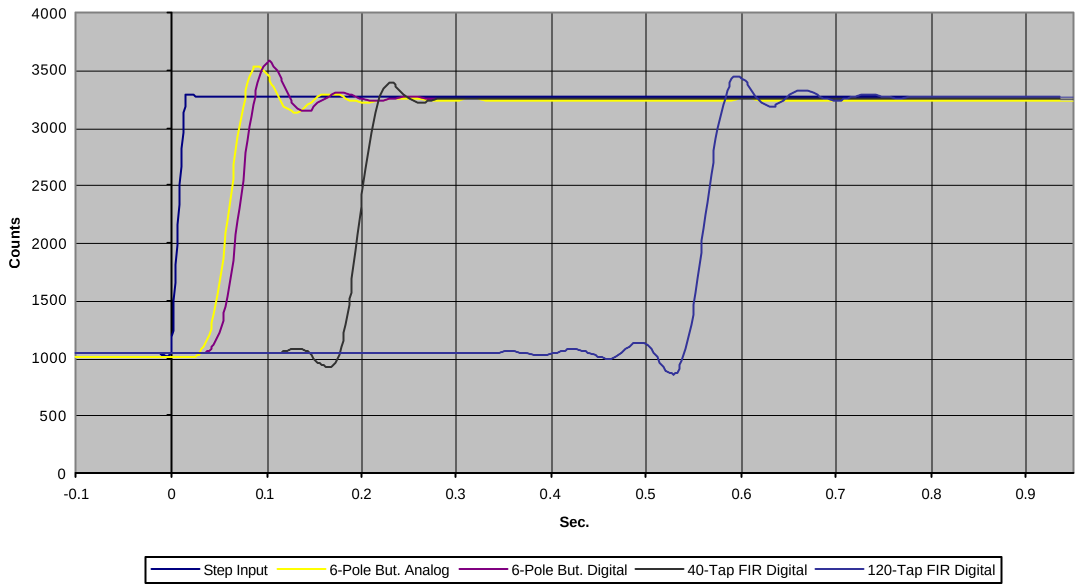


Figure 2: FIR Filters Pass and Transition Bands

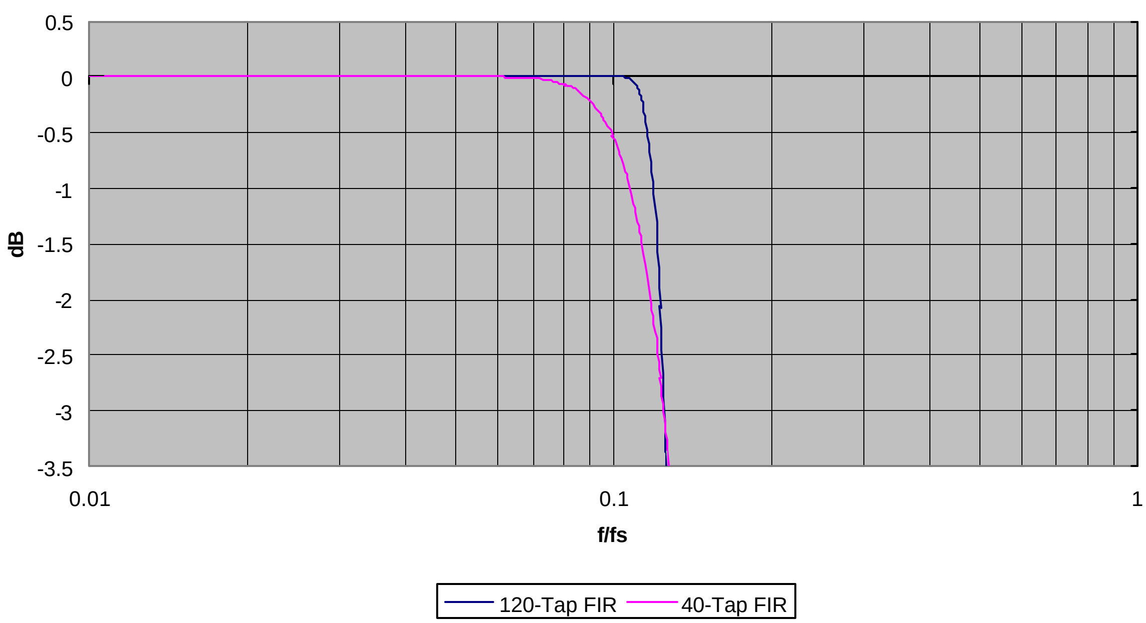


Figure 3: FIR Filters Frequency Roll-Off

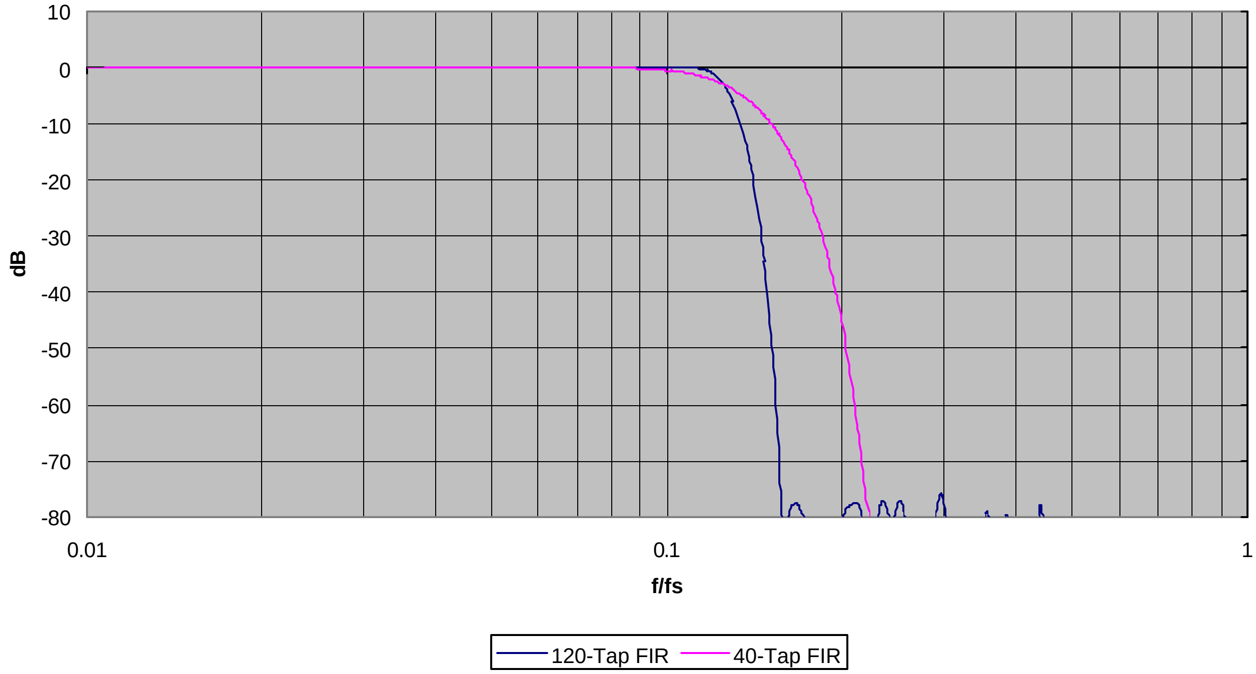


Figure 4:

IIR Filters Pass and Transition Bands

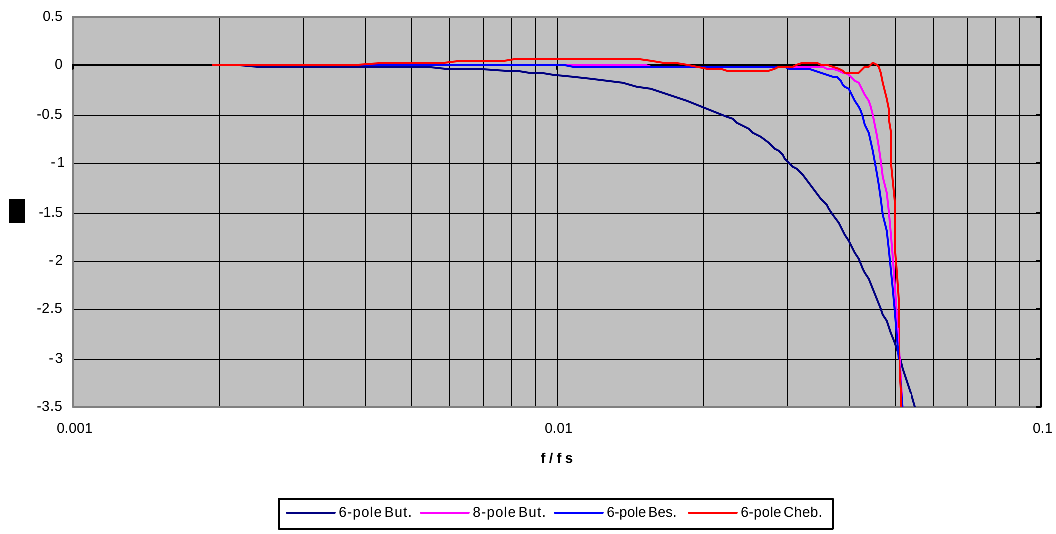
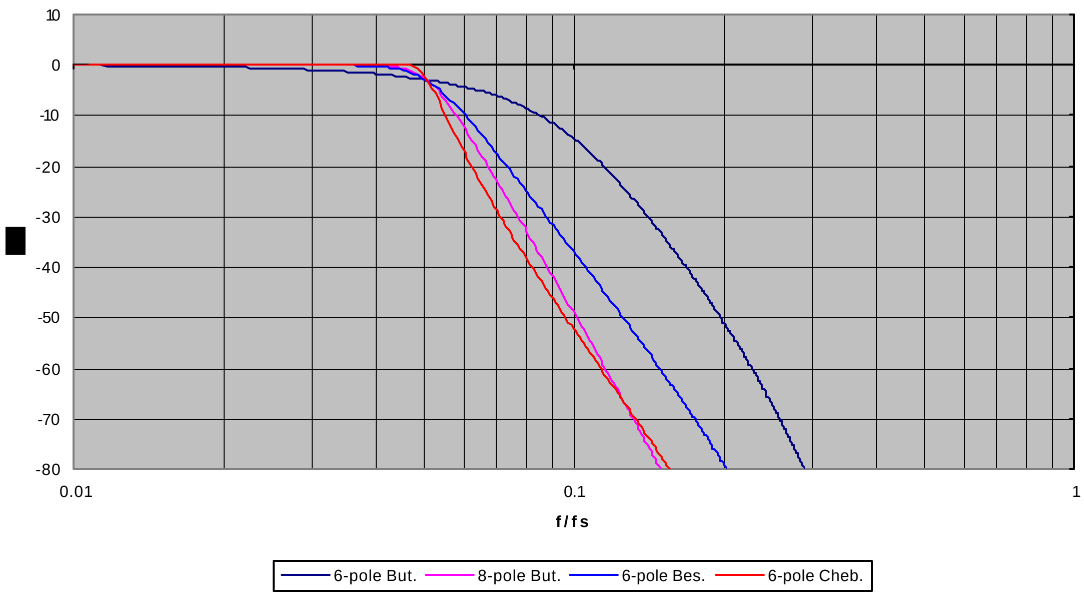


Figure 5:

IIR Filters Frequency Roll-Off



Comparing the frequency roll-off characteristics of the filters depicted in **Figures 3 and 5**, it is clear that the 120-tap FIR has by far, the sharpest roll-off. The 120-tap FIR will outperform a 12-pole Butterworth or 10-pole Chebyshev filter in this respect while remaining completely stable and non-oscillatory. From **Figures 2 and 4**, we see that the 120-tap FIR also has the flattest pass band of the group. The –0.1dB point of this filter is at 86% of the –3dB frequency which beats all but the 6-pole Chebyshev filter. However, unlike the Chebyshev filter, there is no oscillation in attenuation throughout the pass band. If you are looking for the proverbial brick-wall filter, it would be hard to beat the 120-tap FIR filter.

If the FIR filter has an Achilles heel, it is depicted in **Figure 1**. A step function was applied to the input of the four filters shown at time, t = 0. The relatively slow apparent rise time of the input step is an artifact of the plotting program. Note that the propagation delay of the 120-tap FIR filter is 3 times that of the 40-tap FIR filter and almost 10 times that of the 6-pole Butterworth IIR filter. Theoretically, this should not be a problem since the filter has linear phase or constant delay. Having said that, some people may simply not be comfortable with such a long latency.

The 40-tap FIR filter offers frequency roll-off and pass band flatness that is fairly comparable to that of a 6-pole Butterworth filter. Like other symmetrical FIR filters, it has linear phase. Its step response shows only a slight overshoot (about 5%) with very little ringing that appears symmetrically at the leading and trailing edges of the step. Only the Bessel IIR filter (not shown), which has almost no overshoot and ringing, offers a better step response. Propagation delay of the 40-tap FIR filter is about 3 times that of one of the 6-pole IIR filters.

The 6-pole Butterworth has long been the standard by which other low-pass filters are measured and is why the step response of an analog version of this filter was included in **Figure 1**. This filter features maximally flat pass band and –36dB/octave frequency roll-off beyond the –3dB frequency. Note the nearly identical rise time, overshoot (about 14%) and ringing in the analog and IIR digital versions of this filter. The slight additional delay in the digital version comes from a combination of decimation filter delay and slightly longer card latencies.

The 8-pole Butterworth filter has the flattest pass-band of all the IIR filters shown with less than -0.1dB attenuation to beyond $0.6f_S$. Its frequency roll-off is –48dB/octave beyond the –3dB frequency. Its step response has slightly more overshoot and slightly longer delay than a 6-pole Butterworth filter.

On the basis of its frequency domain performance illustrated in **Figures 4 and 5**, the 6-pole Bessel filter looks like a miserable performer. However, this linear phase filter alone provides a symmetrical step response with essentially no overshoot or ringing on both leading and trailing edges of the step. If the information you are interested is primarily in the time domain, that is if it is carried in the shape of the filtered waveform, the Bessel filter is the one to choose. An electrocardiogram (ECG) signal is an example of a waveform whose information is primarily in the time domain. More to the topic of telemetry, pulse coded modulation (PCM) data is usually passed through a Bessel filter to band-limit the pulse train without adding ringing before being sent to an RF transmitter.

Of the IIR filters discussed, the 6-pole Chebyshev filter offers the sharpest transition from pass band to stop band. Its performance in this area is matched only by that of the 120-tap FIR filter. As shown in **Figure 5**, the frequency roll-off 6-pole Chebyshev filter exceeds that of the 8-pole Butterworth to beyond –60dB. Part of the price paid for this sharp transition is oscillation of gain in the pass band (see **Figure 4**). The Chebyshev characteristic chosen for use in TTC's products limits this oscillation to $\pm$0.1dB or about $\pm$1%. Other potential issues are the amount of overshoot and ringing present in the filter's step response and its non-linear phase response.

CONCLUSION

Successful signal conditioning of analog data requires attention to detail in several important areas. The analog signal must be band-limited by a well-chosen analog filter and properly sampled at a rate that will not cause significant aliasing. The resolution of the analog-to-digital converter needs to be high enough to provide negligible quantization noise.

When the data is being collected for inclusion in a PCM or similar system, a strategy must be in place to ensure that the collected data is, and remains, time-correlated with the data stream. Finally, the characteristic of the channel filter chosen to band-limit the data (other than the anti-aliasing filter) must be selected in accordance with the kind of information that the data is collected to provide.

REFERENCES

1. Smith, Steven W. The Scientist and Engineer's Guide to Digital Signal Processing, Second Edition. California Technical Publishing, USA, 1999.
2. Lyons, Richard G. Understanding Digital Signal Processing. Addison Wesley Longman, Inc. USA, 1997.