DSP
The Technology Behind Multimedia

Presented to the
IEEE Central Texas Consultants Network
by
Sergio Liberman
2011-03-23

Abstract

- The multimedia products that we enjoy today share a common technology backbone: Digital Signal Processing (DSP).
- The digital revolution for consumers started with the introduction of the Compact Disc (CD) format for music distribution.
- With the increasing computing power and lower cost of digital processing resulting from Moore's Law, advanced signal processing algorithms have allowed consumer products such as HDTV, DVD and Blu-ray, MP3 players, multimedia capable smart phones, and audio/video streaming.
- This presentation will look into how analog signals such as sound and pictures are brought in and out of the digital domain for communication, storage, transmission, processing and playback, as well as some of the DSP algorithms.

Signals

- Mathematically, a signal is described as a function of one or more independent variables
- Signals in the physical world are any time varying or spatial varying quantity
- Signals in electronic systems are in the form of varying electrical voltage or current
- Signal variations carry information
- Examples of signals:

  Sound signals: f(t) => variation of air pressure in time

  Image signals: f(x,y) => variation of intensity in space

  Video signals: f(x,y,t) => variation of intensity in space & time

Signal Processing

- Modification of signals by processing elements.
- Linear processing: amplification, filtering, tone controls, equalization, etc.
- Non-linear processing: dynamic range control, noise gating, etc.
- Spectral analysis, synthesis, signal generation, etc.

Analog Signal Processing

- Analog signal processing is any signal processing conducted on analog signals by analog transducers and analog processing elements.
- Electrical analog processing elements include capacitors, resistors, inductors, transistors and operational amplifiers.
- Analog transducers include microphones, speakers, record player and guitar pickups, magnetic tape heads, etc.

Analog Signal Processing Disadvantages

- The primary disadvantage of analog signal processing is that any system has noise – i.e., random unwanted variation.
- As the signal is copied and re-copied, or transmitted over long distances, these apparently random variations become dominant.
- Analog processing components suffer from aging, drift with temperature, tuning issues, accuracy, repeatability, size, etc.

Digital Signal Processing Advantages

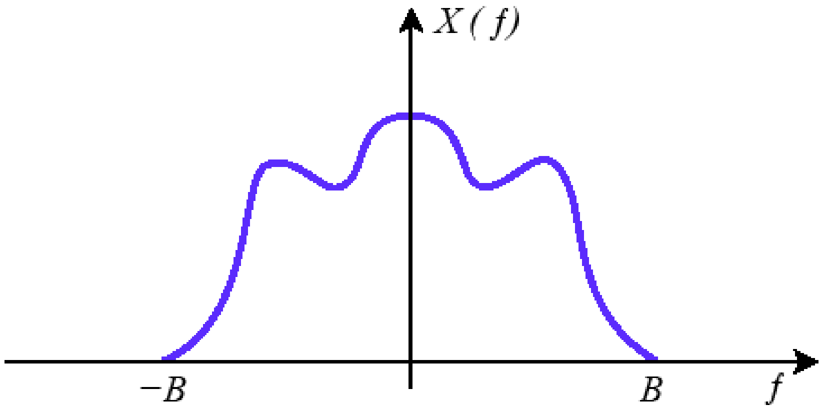
- Representation of signals by a sequence of numbers or symbols, which can be copied and transmitted with no noise or errors in a digital system.
- Digital processing components and their numerical results do not suffer from aging, drift with temperature, tuning issues, accuracy, repeatability, etc.
- There are signal processing functions only possible with DSP, such as linear phase filters.
- Moore's Law allows for huge size reduction of digital processing elements.
- Combination of advanced control and data processing with signal processing.

Digital Signal Processing

- Use digital computer processing elements for manipulation of signals.
- Digital processing elements cannot handle infinite values for the signal variables and parameters.
- Requires analog signals to be converted from continuous time/space to discrete time/space.
- Requires quantizing continuous analog signal amplitude values to a finite set of discrete signal amplitude values.

Nyquist–Shannon Sampling Theorem

- A band-limited analog signal, where B is its highest frequency, can be perfectly reconstructed from an infinite sequence of its samples if the sampling rate exceeds $2B$ samples per second.
- Allows sampling of analog signal, changing continuous time/space to a series of discrete signal value points.



Quantization of Amplitude

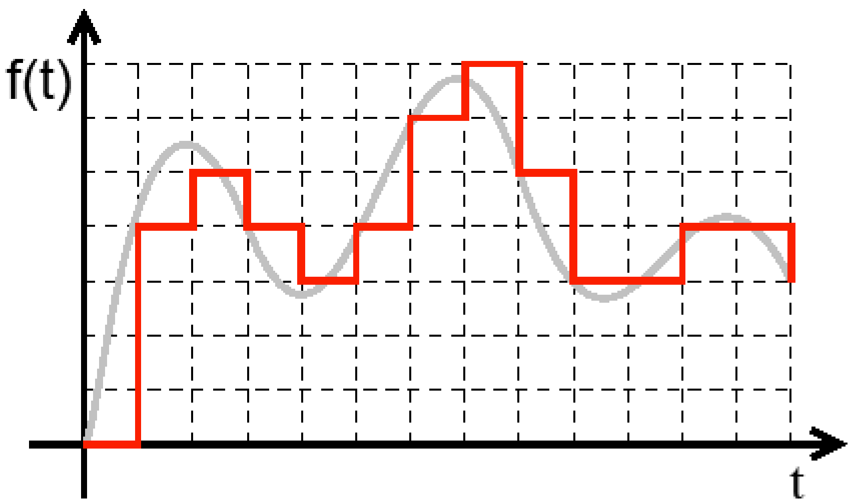
- After sampling of the analog signal, the amplitude values need to be quantized to a finite range.
- Quantization introduces amplitude error, resulting in quantization noise.
- Increasing resolution (the number of possible values for the amplitude) reduces quantization noise.
- Proper digital signal processing assumes enough bits of resolution in the A/D process so that quantization noise is negligible for the application.

How Many Bits Are Necessary?

- Every additional bit in resolution adds ~6dB to the SNR (signal to noise ratio) for uniform quantization.
- The dynamic range of human hearing is roughly 120 dB. (20 bits)
- The dynamic range of music as normally perceived in a concert hall doesn't exceed 80 dB (14 bits)
- Human speech is normally perceived over a range of about 40 dB. (7 bits)

ADC: Analog to Digital Conversion

- Converts the analog signal into a digital representation by sampling and amplitude quantization.



DAC: Digital to Analog Conversion

- DAC converts the digital signal abstract numbers into a concrete sequence of impulses that are then processed by a reconstruction filter using some form of interpolation to fill in data between the impulses.
- Delta-sigma modulation method generates a pulse-density modulated signal that can then be filtered to produce a smoothly varying analog signal. Most popular DAC for audio.
- PWM modulation method generates a pulse-width modulated signal that can then be filtered. Used in Class-D amplifers.

Example: Compact Disc digital audio First digital consumer format - 1980

- The CD disc contains audio signals converted from analog to digital using a linear encoding format.
- Audio bandwidth is assumed to be limited to 20 kHz (B from Nyquist-Shannon). The sampling rate for CD audio is 44,100 samples per second (> 2 x 20 kHz).
- The amplitude values are quantized to a fixed-point binary representation using 16 bits per audio sample providing 96dB SNR.
- The audio data bit rate for a CD disc is 16 bits x 44,100 samples/s x 2 channels = 1.4 Mbit/s
- .WAV file extension.

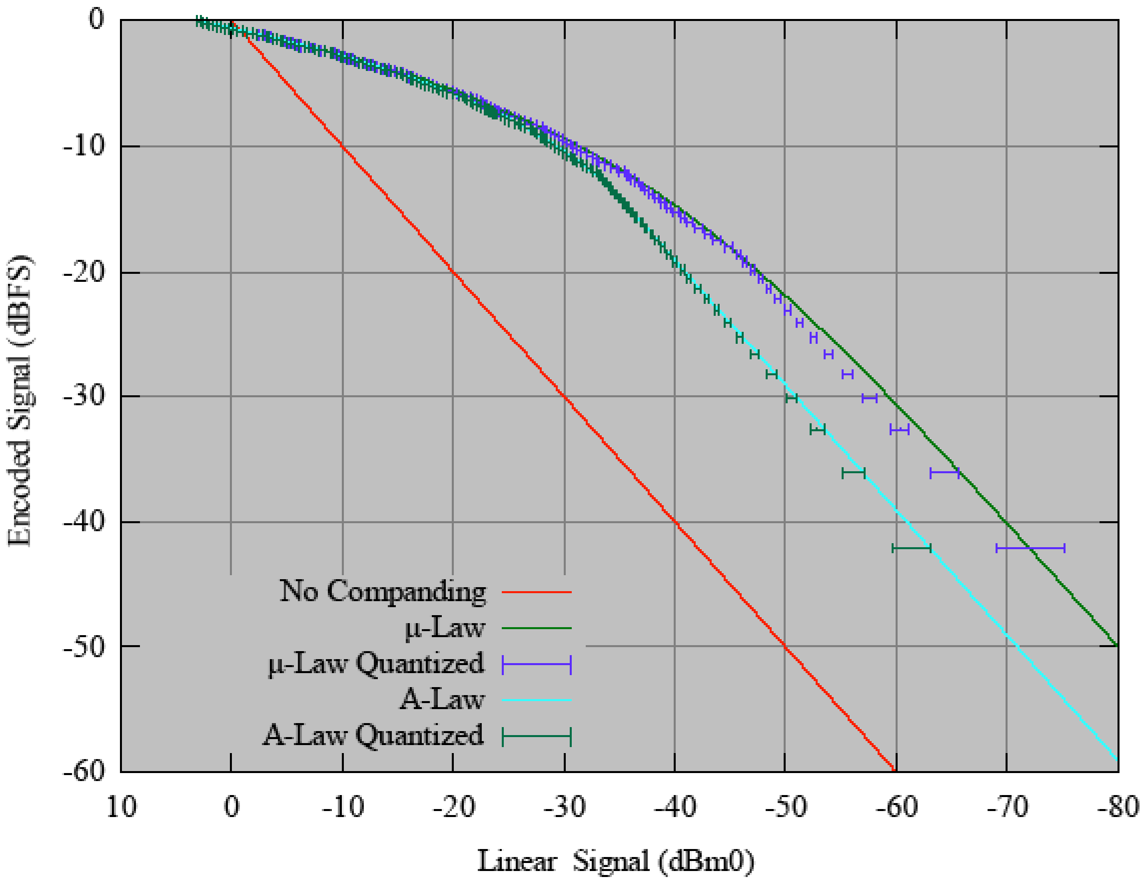
Example: T-Carrier Digital Telephony First extensive use of DSP - 1961

- Telephony audio bandwidth is 3.5 kHz (B).
- Voice in digital telephony is sampled at 8,000 samples per second (> 2 x 3.5kHz).
- Amplitude of voice signals are quantized to ~ 13.5 bits per sample (equivalent linear encoding).
- The amplitude is compressed using A-law/u-law non-linear companding algorithm to reduce the number of bits per sample to 8 bits.
- Bit rate for one voice circuit: 8 bits x 8,000 samples/s = 64 kbit/s

Signal Compression - 1961

- A-law/u-law non-linear companding reduces the dynamic range of the signal, thereby increasing the coding efficiency and resulting in a signal-to-distortion ratio that is superior to that obtained by linear encoding for a given number of bits.
- Bit rate for one voice circuit: 8 bits x 8,000 samples/s = 64 kbit/s
- Bit rate before companding: 14 bits x 8,000 samples/s = 112 kbit/s

Companding Transfer Function



Coding

- Coding, in the sense used here, is the process of reducing the bit rate of a digital signal.
- Reduces storage requirements and transmission bandwidth for the digital signal.
- The coder input is a digital signal.
- The coder output is a lower bit rate digital signal.
- The decoder reverses the process and provides (an approximation to) the original digital signal.

Types of Coding

- Lossless Coding commonly refers to coding methods that are completely reversible, where the original signal can be reconstructed bit for bit. Equivalent to data compression in computers such as Zip, RAR, etc.
- Lossy coding commonly refers to coders that create an approximate reproduction of their input signal. The nature of the loss depends entirely on the kind of lossy coding used. MP3 is an example of lossy coding.

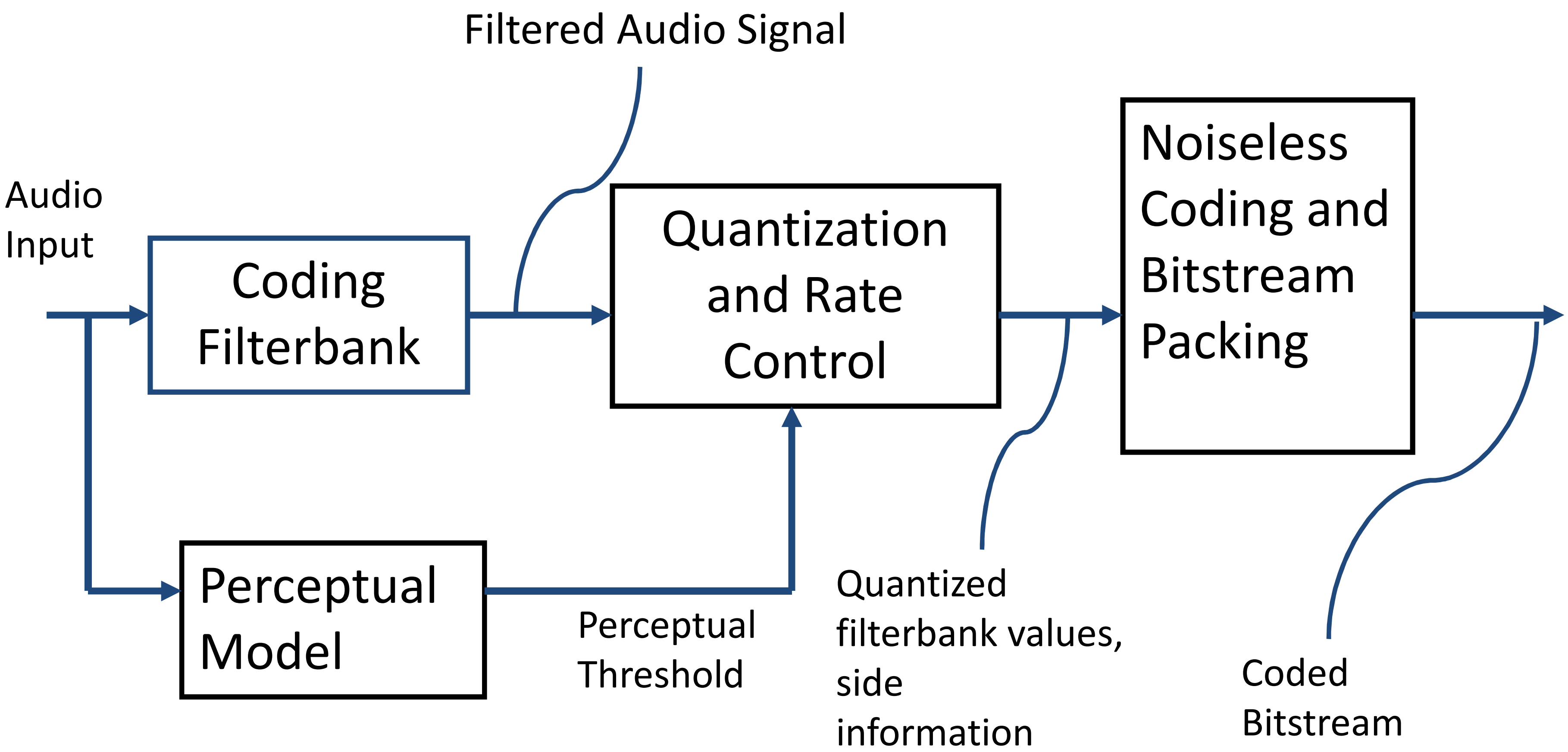
Perceptual Coding

- Perceptual coding uses a model of the destination, i.e. the human being who will be using the data.
- Perceptual coding attempts to remove parts of the signal that the human cannot perceive.
- Perceptual coding is lossy.
- Perceptual coders will, in general, have a *lower* SNR than a source coder, and a higher perceived quality than a source coder of equivalent bit rate.

Using Perceptual Coding

- Perceptual audio coding is applicable where the signal will NOT be reprocessed, equalized, or otherwise modified before the final delivery to the consumer.
- NEVER use multiple passes through the perceptual encoders.
- Popular audio perceptual coding: MP3, AAC, Dolby Digital, DTS

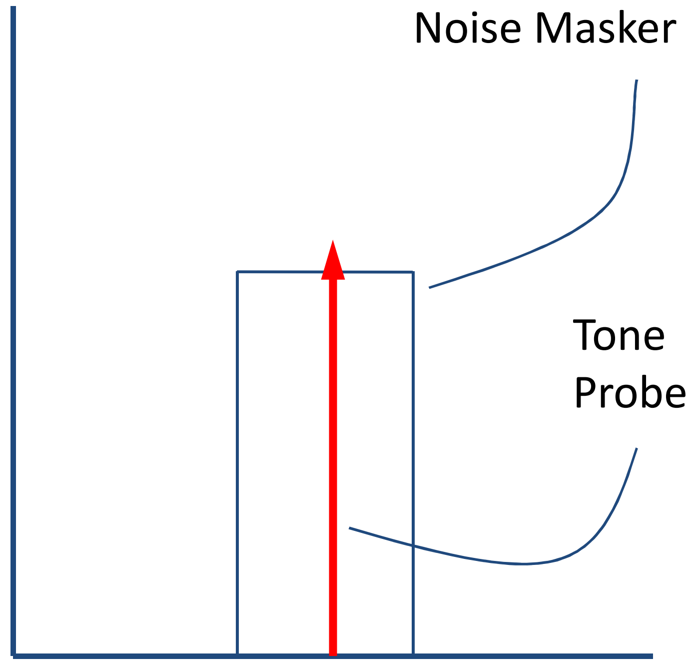
Perceptual Audio Coder Block Diagram



What is Auditory Masking?

- The Human Auditory System (HAS) has a limited detection ability when a stronger signal occurs near (in frequency and time) to a weaker signal.
- In many situations, the weaker signal is imperceptible even under ideal listening conditions.

Auditory Masking



The Tone is NOT Audible

Auditory Filterbank

- The mechanism in the human cochlea constitute a mechanical filterbank.
- The shape of the filter at any one position on the cochlea is called the ***cochlear filter*** for that point on the cochlea.
- A ***critical band*** is very close to the passband bandwidth of that filter.

Critical Band

- Critical band is a range of frequencies over which the masking SNR remains more or less constant.

- For example, any noise signal within +- .5 critical band of the tone will produce nearly the same masking behavior as any other, as long as their energies are the same.

Time Considerations in Masking

- Simultaneous Masking

- *Forward Masking* – Masking of a signal by a masker that precedes the masked signal.

- *Backward Masking* – Masking of a signal by a masker that comes after the signal.

Forward Masking

- The length of forward masking is >20ms, and is sometimes stated to be as long as several hundred milliseconds.
- In practice, the decay for post masker masking has two parts, a short *hangover* part and then a longer *decaying* part.
- Different coders take advantage of this in different ways.

Backward Masking

- Backward masking appears to be due to the length of the impulse response of the cochlear filter.
- At high frequencies, backward masking is less than 1ms for a trained subject who is sensitive to monaural time-domain masking effects.
- Subjects vary significantly in their ability to detect backwardly masked probes.

MP3

- MP3 = MPEG-1 or MPEG-2 Audio Layer III
- Standard released in 1993
- De facto standard for digital audio compression in portable players
- Lossy perceptual coding compression
- MP3 compression to 128 kbit/s data rate achieves ~11x data rate reduction as compared to CD audio
- Highest bit rate is 320 kbit/s. Decent audio quality for popular music achieved at about 192 kbit/s

Limitations of the MP3 Format

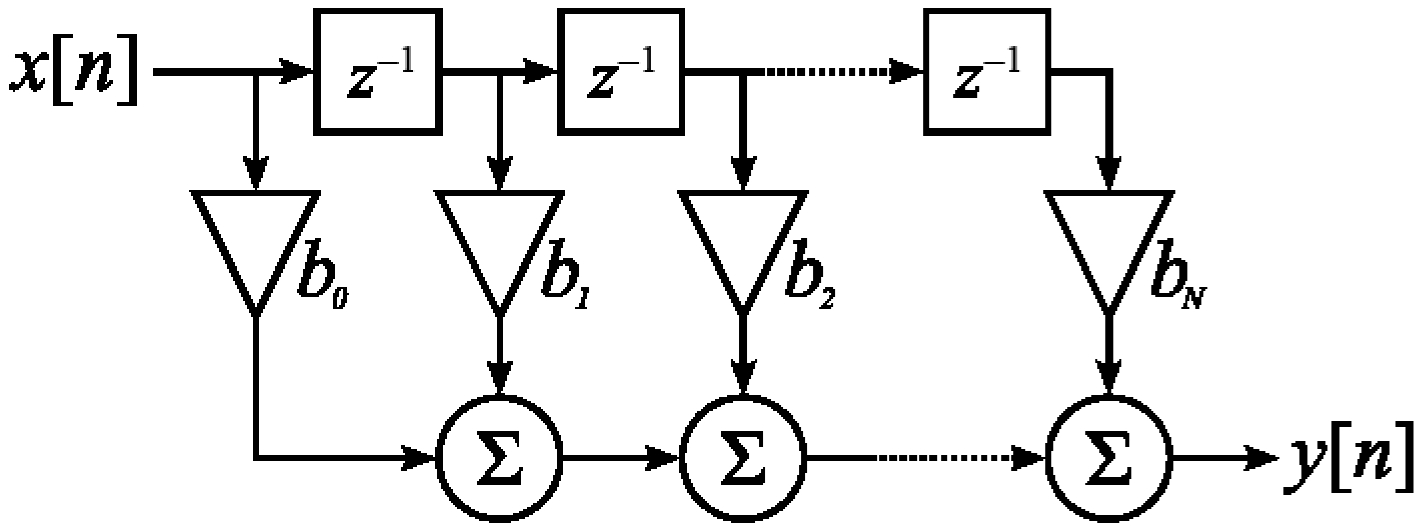
- Time resolution can be too low for highly transient signals and may cause smearing of percussive sounds.
- Due to the tree structure of the filter bank, pre-echo problems are made worse, as the combined impulse response of the two filter banks does not, and cannot, provide an optimum solution in time/frequency resolution.
- The combining of the two filter banks' outputs creates aliasing problems that must be handled partially by the "aliasing compensation" stage; however, that creates excess energy to be coded in the frequency domain, thereby decreasing coding efficiency.
- Frequency resolution is limited by the small long block window size, which decreases coding efficiency.
- There is no scale factor band for frequencies above 15.5/15.8 kHz.
- Joint stereo is done only on a frame-to-frame basis.
- Internal handling of the bit reservoir increases encoding delay.
- Encoder/decoder overall delay is not defined, which means there is no official provision for gapless playback. However, some encoders such as LAME can attach additional metadata that will allow players that can handle it to deliver seamless playback.
- The data stream can contain an optional checksum, but the checksum *only* protects the header data, not the audio data.

AAC

- Advanced Audio Coding, designed to replace MP3, used in the Apple iPod as the standard algorithm.
- Lossy perceptual coding compression.
- Achieves better sound quality for the same bit rate as compared to MP3.

FIR Filter

- A discrete-time FIR filter of order *N*. The top part is an *N*-stage delay line with *N*+1 taps.



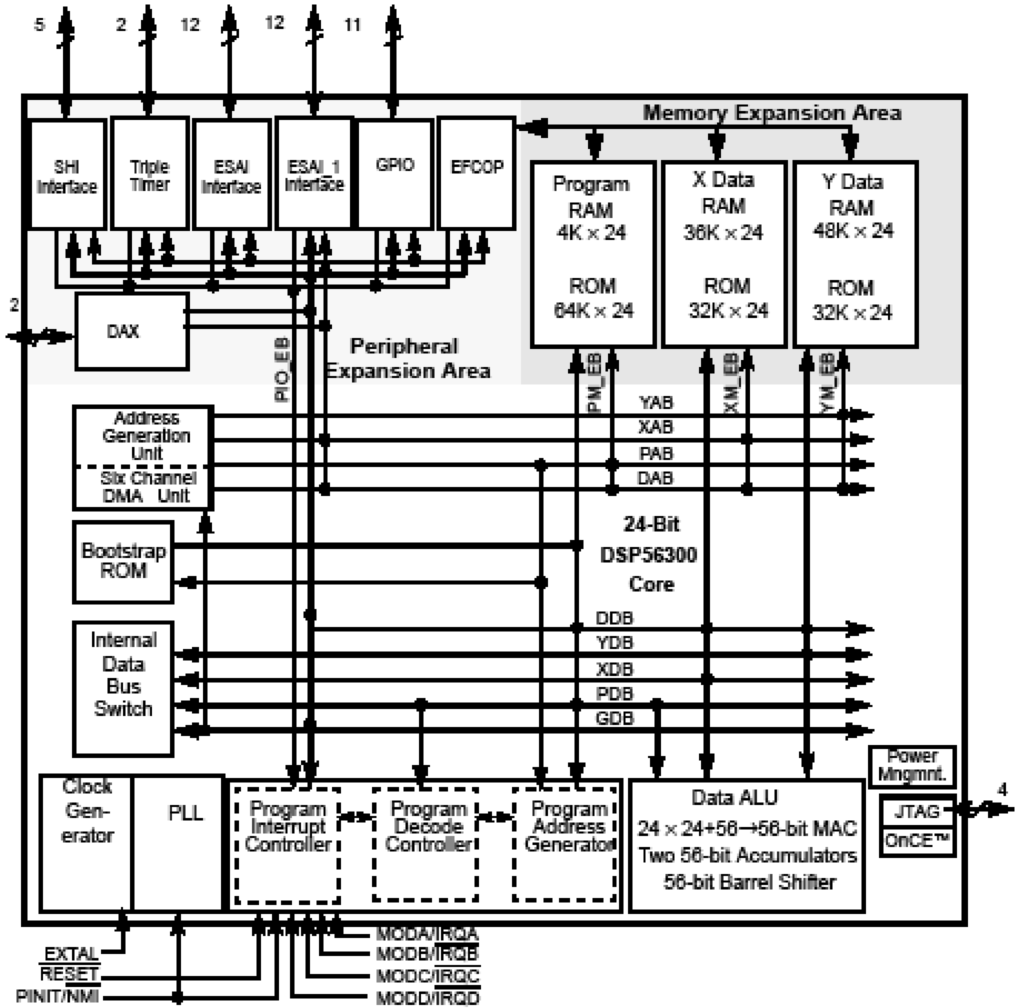
FIR Filter (cont.)

- A **finite impulse response (FIR)** filter is a type of a signal processing filter whose impulse response is of *finite* duration.
- Only filter capable of linear phase.
- No feedback, so it is inherently stable, and there is no recirculation of rounding errors.
- The output is calculated by multiplying each data sample in the delay line by the correspondent filter tap coefficient, and accumulating all multiplies.

DSP Processor Features

- Separate program and data memories, concurrent access on multiple data busses.
- Special SIMD (single instruction, multiple data) operations. Some processors use VLIW techniques so each instruction drives multiple arithmetic units in parallel.
- Fast multiply-accumulates (MACs). Many fundamental DSP algorithms, such as FIR Filter or the Fast Fourier Transform (FFT) depend heavily on multiply-accumulate performance.
- Hardware modulo addressing, allowing circular buffers to be implemented without having to constantly test for wrap around.
- Bit-reversed addressing, a special data memory addressing mode, useful for calculating FFTs.
- Special loop controls, such as architectural support for executing a few instruction words in a very tight loop without overhead for instruction fetches or exit testing.
- Fast interrupt response with low overhead context switching.

Freescale DSP56371 Block Diagram



Thank You!

- Sergio Liberman
- Sergio.Liberman@gmail.com