DIGITAL FILTERING

1. Applications
2. Digital and analog filters: Pros and cons
3. IIR digital filters
4. FIR digital filters
5. FIR and IIR filters: Pros and cons
6. Ideal filters
7. The approximation problem
8. The realization problem

APPLICATIONS

1. Noise suppression
   (a) imaging devices (medical, etc)
   (b) biosignals (heart, brain)
   (c) signals stored on analog media (tapes)
2. Enhancement of selected frequency ranges
   (a) equalizers for audio systems (increasing the bass)
   (b) edge enhancement in images
3. Removal or attenuation of selected frequencies
   (a) removing the DC component of a signal
   (b) removing interferences at a specific frequency, for example those caused by power supplies
4. Bandwidth limiting
   (a) anti-aliasing filters for sampling
   (b) ensuring that a transmitted signal occupies only its alloted frequency band.
5. Special operations
   (a) differentiation
   (b) integration
   (c) Hilbert transform
6. Simulation/Modeling
   (a) simulating communication channels
   (b) modeling human auditory system

DIGITAL AND ANALOG FILTERS: PROS AND CONS

It was said of early digital filters that they

1. Cost too much
2. Were too large
3. Used too much power.

But these considerations have become less important with advances in hardware. Digital filters have the following advantages

1. Programmable (filter characteristics easily changed)
2. Reliable and repeatable
3. Free from component drift
4. No tuning required
5. Superior performance in some cases

IIR DIGITAL FILTERS

IIR filter: A filter with an Infinite Impulse Response.

The transfer function $H(z)$ of a realizable IIR filter must be a rational transfer function in z^{-1}:

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots a_N z^{-N}}.$$

IIR digital filters are implemented using ARMA (*autoregressive moving average*) difference equation:

$$\begin{aligned} y(n) = b_0\, x(n) + \cdots + b_M\, x(n-M) \\ - a_1\, y(n-1) - \cdots - a_N\, y(n-N) \end{aligned}$$

FIR DIGITAL FILTERS

FIR filter: A filter with a Finite Impulse Response.

The transfer function $H(z)$ of a causal FIR filter is a polynomial in z^{-1}:

$$H(z) = b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}.$$

FIR digital filters are usually implemented using MA (*moving average*) difference equation:

$$y(n) = b_0\, x(n) + b_1\, x(n-1) + \cdots + b_M\, x(n-M).$$

FIR AND IIR FILTERS: PROS AND CONS

FIR digital filters have several desirable properties in relation to IIR filters.

1. FIR filters can have exactly linear phase.
2. FIR filters are automatically stable.
3. There are several very flexible methods for designing FIR digital filters.
4. FIR filters are convenient to implement.

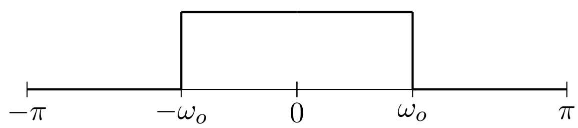
On the other hand,

1. Linear-phase FIR filters can have long delay between input and output.
2. If the phase need not be linear, then IIR filters can be much more efficient to implement.

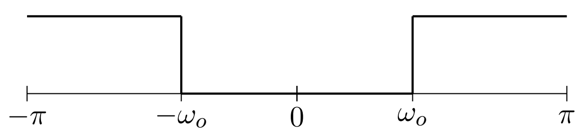
IDEAL FILTERS

The frequency responses of the common *ideal* filters:

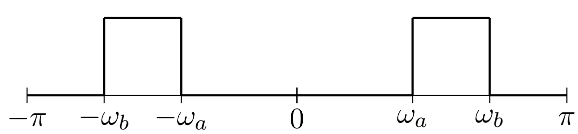
1. Low-pass



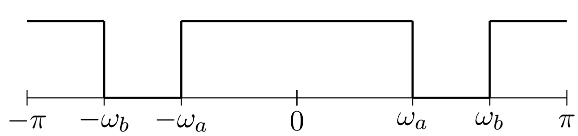
2. Hi-pass



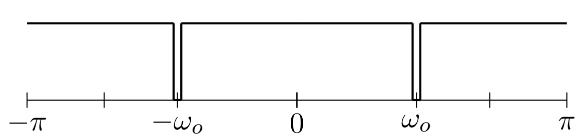
3. Band-pass



4. Band-stop



5. Notch



DIGITAL FILTER DESIGN

The filter design process consists of two parts: the approximation problem and the realization problem. The approximation problem deals with the choice of parameters or coefficients in the filter's transfer function. The realization part of the design problem deals with choosing a structure to implement the transfer function.
The approximation stage can be divided into 4 steps:

1. A desired or ideal response is chosen (usually in the frequency domain).
2. A class of filters is chosen (for example, FIR vs IIR).
3. A design criteria is chosen (least square or minimax).
4. An algorithm is selected to design the transfer function.

The realization stage can also be divided into 4 steps:

1. A set of structures is chosen.
2. A criteria for comparing different implementations is chosen.
3. The best structure is chosen, and its parameters are calculated from the transfer function.
4. The structure is implemented in hardware or software.

The impulse response of ideal filter

$$d(n) = \mathsf{IDTFT}\left\{D^f(\omega)\right\},$$

where $D^f(\omega)$ is an ideal low-pass response, for example, is unrealizable because it is noncausal and of infinite duration.

The actual frequency response of the filter will be denoted by $H^f(\omega)$

$$H^f(\omega) = \mathsf{DTFT}\left\{h(n)\right\} = \sum_{n=-\infty}^{\infty} h(n)\, e^{-jn\omega}.$$

The DTFT is the Z-transform evaluated on the unit circle $z = e^{j\omega}$.

$$H(z) = \mathsf{Z}\left\{h(n)\right\} = \sum_{n=-\infty}^{\infty} h(n)\, z^{-n}$$

$$H^f(\omega) = H(e^{j\omega})$$

The desired (or target) response of the filter will be denoted $D(\omega)$.

Common approximation criteria:

1. The *square error* criterion is written as

$$E_2 = \int_0^{2\pi} W(\omega) \cdot |H^f(\omega) - D(\omega)|^2 \, d\omega$$

2. The *Minimax (or Chebyshev)* error criterion is written as

$$E_\infty = \max_{\omega} |W(\omega) \cdot (H^f(\omega) - D(\omega))|$$

3. The L_p error criterion is written as

$$E_p = \int_0^{2\pi} W(\omega) \cdot |H^f(\omega) - D(\omega)|^p \, d\omega$$

APPROXIMATION WITH CONSTRAINTS

In addition to these error functions to evaluate a filter's frequency response, one can place *constraints* on the frequency response.

1. It may be required that the filter response meet certain tolerances. For example, for a low-pass filter, it may be required of the filter that it satisfy

   $$1 - \delta_p \le |H^f(\omega)| \le 1 + \delta_p$$

   for all frequencies ω in the passband, and

   $$|H^f(\omega)| \le \delta_s$$

   for all frequencies ω in the stopband.

2. Specified null constraint:

   $$H^f(\omega_o) = 0$$

   for a specified frequency ω_o.

3. One can ask that the filter response have a specified degree of tangency with the desired response. For example, one may whish to specify several derivatives $H^f(\omega)$ at certain frequencies.

THE REALIZATION PROBLEM

Once the transfer function is determined, it can be realized using different structures. For example, one can implement an IIR filter using the direct form, or a cascade of second order sections, or one of several other structures. While they are all equivalent when infinite precision is used, different structures behave differently when the coefficients and the arithmetic operations are quantized.

SOME NOTATION

A_s and A_p represnt the attenuation in the stopband and the pass-band of a lowpass filter.
The meaning of A_p and A_s is given by

$$\delta_p = 1 - 10^{-A_p/20} \tag{1}$$

$$\delta_s = 10^{-A_s/20} \tag{2}$$

Equivalently,

$$A_p = -20 \log_{10}(1 - \delta_p) \tag{3}$$

$$A_s = -20 \log_{10}(\delta_s) \tag{4}$$

The constants δ_p represents the size of the ripple in the pass-band, and δ_s represents the size of the ripple in the stop-band. A_p, and A_s are just δ_p and δ_s in decibels. The pass-band ripple size δ_p is the maximum deviation of the actual frequency response from 1 in the pass-band. The stop-band ripple size δ_s is the maximum deviation of the actual frequency response from 0 in the stop-band.