2-D FIR Filter Design

Lecture and notes of Prof. Brian L. Evans (UT Austin)

Scribe: Viswanath Sankaranarayanan (UT Austin)

Based on notes from Prof. Russell Mersereau (Georgia Tech)

1 Introduction

The problem with designing FIR filters based on windows is that it is not very efficient for filter design. In general, the filter length is twice the optimum length in each dimension. Optimal filter design seeks to minimize an error function $e(n_1, n_2)$ that is the difference between the designed filter impulse response $h(n_1, n_2)$ and the ideal impulse response $i(n_1, n_2)$.

We define an error function

$$\begin{aligned} E(\omega_1, \omega_2) &= H(\omega_1, \omega_2) - I(\omega_1, \omega_2) \\ &= -I(\omega_1, \omega_2) + \sum_{n_1} \sum_{n_2} h(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2} \end{aligned}$$

Choose the filter coefficients to minimize some measure of this error

$$\begin{aligned} E_2 &= \left[\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left| E(\omega_1, \omega_2) \right|^2 d\omega_1 d\omega_2 \right]^{\frac{1}{2}} \\ E_p &= \left[\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left| E(\omega_1, \omega_2) \right|^p d\omega_1 d\omega_2 \right]^{\frac{1}{p}} ; p \in \mathcal{I} \\ E_\infty &= \max_{(\omega_1, \omega_2) \in R_{\text{tile}}} \left| E(\omega_1, \omega_2) \right| \end{aligned}$$

where R_{tile} is the region of the fundamental frequency tile $-\pi \leq \omega_1 < \pi$ and

$-\pi \leq \omega_2 < \pi$. Note that

$$\lim_{p \to \infty} E_p = E_\infty$$

2 Optimum Least Squares Method

In this method, we truncate the ideal response (with $h(n_1, n_2)$ and $i(n_1, n_2)$ being real-valued):

$$\begin{aligned}
E_2^2 &= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left| E(\omega_1, \omega_2) \right|^2 d\omega_1 d\omega_2 \\
&= \sum_{n_1} \sum_{n_2} \left[h(n_1, n_2) - i(n_1, n_2) \right]^2 \quad (\text{Parseval's Relation}) \\
&= \sum_{(n_1, n_2)} \sum_{\in R} \left[h(n_1, n_2) - i(n_1, n_2) \right]^2 + \sum_{(n_1, n_2)} \sum_{\in R} \left[i(n_1, n_2) \right]^2
\end{aligned}$$

where R is the region of support of h. To minimize E_2^2, set

$$h(n_1, n_2) = \begin{cases} i(n_1, n_2) & \text{for } (n_1, n_2) \in R \\ 0 & \text{otherwise} \end{cases}$$

3 Optimum Design With Constraints

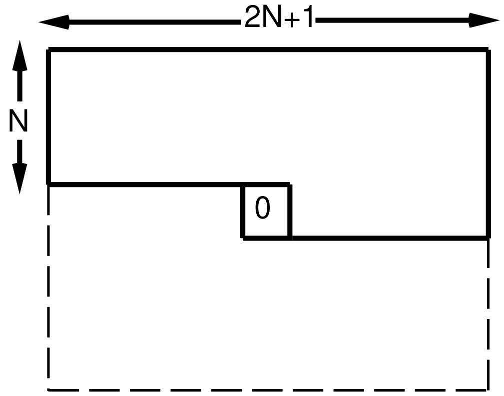
$$E(\omega_1, \omega_2) = \sum_{(n_1, n_2)} \sum_{\in R} h(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2} - I(\omega_1, \omega_2)$$

We have been treating the filter coefficients as if they are independent. They may be constrained, e.g. in the design of a zero-phase filter. If $h(n_1, n_2) = h^*(-n_1, -n_2)$, or since $h(n_1, n_2)$ is real-valued $h(n_1, n_2) = h(-n_1, -n_2)$ then

$$E(\omega_1, \omega_2) = h(0, 0) + \sum_{(n_1, n_2)} \sum_{\in R'} 2\, h(n_1, n_2) \cos(n_1, \omega_1 + n_2 \omega_2) - I(\omega_1, \omega_2)$$

$$= \sum_{i=1}^{F} a(i)\,\phi_i(\omega_1,\omega_2) - I(\omega_1,\omega_2)$$

where assuming that the filter is $(2N+1)\times(2N+1)$,



The number of degrees of freedom is $F = \dfrac{(2N+1)^2+1}{2} = 2N^2+2N+1$, $a(i)$ is the i^{th} free parameter. For a zero-phase filter,

$$a(i) = \begin{cases} h(0,0) & i = 1 \\ 2h(n_1,n_2) & i = (2N+1)n_2 + n_1 + 1 \end{cases}$$

$\phi_i(\omega_1,\omega_2)$ is the i^{th} basis function, which is

$$\phi_i(\omega_1,\omega_2) = \cos(n_1\omega_1 + n_2\omega_2)$$

for $i = (2N+1)n_2 + n_1 + 1$. The mapping from (n_1,n_2) to i follows a raster scan that begins at the origin and proceeds along the n_1 direction:

i	n_1	n_2
1	0	0
2	1	0
$\dots$	$\dots$	$\dots$
N	$N-1$	0
$N+1$	N	0
$N+2$	$-N$	1
$\dots$	$\dots$	$\dots$

For different choices of constraints, F, $a(i)$, and $\phi_i(\omega_1, \omega_2)$ will be different, but for any linear constraints, the error can always be written in this form.

4 Example

Suppose we want to design a circularly symmetric lowpass filter

$$\begin{aligned} I(\omega_1, \omega_2) &= \begin{cases} 1 & \text{for } \omega_1^2 + \omega_2^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases} \\ i(n_1, n_2) &= f(\sqrt{n_1^2 + n_2^2}) \\ h(n_1, n_2) &= \hat{f}(\sqrt{n_1^2 + n_2^2}) \end{aligned}$$

Consider a 5×5 filter:

$$
\begin{array}{ccccc}
h(2,2) & h(2,1) & h(2,0) & h(2,1) & h(2,2) \\
h(2,1) & h(1,1) & h(1,0) & h(1,1) & h(2,1) \\
h(2,0) & h(1,0) & h(0,0) & h(1,0) & h(2,0) \\
h(2,1) & h(1,1) & h(1,0) & h(1,1) & h(2,1) \\
h(2,2) & h(2,1) & h(2,0) & h(2,1) & h(2,2)
\end{array}
$$

Only 6 of the 25 coefficients are free because all the coefficients that lie at the same distance from the center (radius) must have the same value.

$$
\begin{aligned}
H(\omega_1, \omega_2) \;=\; & h(0,0) \\
& + h(1,0)\,[2\cos(\omega_1) + 2\cos(\omega_2)] \\
& + h(2,0)\,[2\cos(2\omega_1) + 2\cos(2\omega_2)] \\
& + h(1,1)\,[4\cos(\omega_1)\cos(\omega_2)] \\
& + h(2,2)\,[4\cos(2\omega_1)\cos(2\omega_2)] \\
& + h(2,1)\,[2\cos(\omega_1 + 2\omega_2) + 2\cos(\omega_1 - 2\omega_2) + 2\cos(\omega_2 + 2\omega_1) + 2\cos(\omega_2 - 2\omega_1)]
\end{aligned}
$$

$F = 6$:

$$
\begin{array}{ll}
a(1) = h(0,0) & \phi_1(\omega_1, \omega_2) = 1 \\
a(2) = h(1,0) & \phi_2(\omega_1, \omega_2) = 2\cos(\omega_1) + 2\cos(\omega_2) \\
\vdots & \vdots \\
a(6) = h(1,2) & \phi_6(\omega_1, \omega_2) = \begin{array}{l} 2\cos(\omega_1 + 2\omega_2) + 2\cos(\omega_1 - 2\omega_2) \\ +2\cos(\omega_2 + 2\omega_1) + 2\cos(\omega_2 - 2\omega_1) \end{array}
\end{array}
$$

5 Least Squares Design

$$E_2^2 = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left(\sum_{i=i}^{F} a(i)\phi_i(\omega_1,\omega_2) - I(\omega_1,\omega_2)\right)^2 d\omega_1 d\omega_2$$

By taking the partial derivative of E_2^2 with respect to each of the $a(i)$ terms and setting the result to zero, we obtain the following system of linear equations:

$$\begin{aligned}
\sum_{i=1}^{F} a(i)\phi_{ik} &= I_k \\
\phi_{ik} &= \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi_i(\omega_1,\omega_2)\phi_k^*(\omega_1,\omega_2)d\omega_1 d\omega_2 \\
I_k &= \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} I(\omega_1,\omega_2)\phi_k^*(\omega_1,\omega_2)d\omega_1 d\omega_2
\end{aligned}$$

If $\phi_{ik} = 0$ for $i \neq k$, i.e., the basis functions are orthogonal, then these equations become particularly easy to solve. We can have a symbolic math environment such as Maple or Mathematica solve these $F^2 + F$ integrals for us.

6 Discrete Solution

The plain Least Squares solution has the following disadvantages:

- Need to evaluate integrals (let a computer program do that)
- No possibility of frequency weighting

A solution that solves both problems is to sample the points in the 2-D frequency plane and use frequency weighting, e.g.

$$\hat{E}_2^2 \triangleq \sum_m W_m \left[H(\omega_{1m}, \omega_{2m}) - I(\omega_{1m}, \omega_{2m})\right]^2$$

- $W_m \geq 0$
- Error can be controlled by the values of the weights and locations of the samples.
- The coefficients that minimize $\hat{E}_2^2$ are given by

$$\begin{aligned} \sum_{i=1}^{F} a(i)\phi_{ik} &= I_k \\ \phi_{ik} &= \sum_m W_m \phi_i(\omega_{1m}, \omega_{2m}) \phi_k^*(\omega_{1m}, \omega_{2m}) \\ I_{ik} &= \sum_m W_m I_i(\omega_{1m}, \omega_{2m}) \phi_k^*(\omega_{1m}, \omega_{2m}) \end{aligned}$$

  where $k = 1, 2, \ldots, F$.

To see the effects of frequency sampling in 1-D, run `matlab` and then type filtdemo and compare a Remez (Parks - McClellan) design with an FIR-LS (FIR Least Squares) design. The FIR Least Square filters are up to 4 times longer than Remez filters in each dimensions. About half of the FIR-LS length is necessary to meet the stopband specification. Also the filter meets its specifications only at the sampling points, and between these points the response may shoot up or fall to a large extent.

7 Design Using Transformations

7.1 Background

Consider a 1-D zero-phase filter of length $2N+1$

$$
\begin{aligned}
h(n) &= h(-n) \\
H(e^{j\omega}) &= \sum_{n=-N}^{N} h(n)e^{-j\omega n} \\
&= h(0) + \sum_{n=-N}^{-1} h(n)e^{-j\omega n} + \sum_{n=1}^{N} h(n)e^{-j\omega n} \\
&= h(0) + \sum_{n=1}^{N} h(n)(e^{-j\omega n} + e^{j\omega n}) \\
&= \sum_{n=0}^{N} a(n)\cos(\omega n)
\end{aligned}
$$

where $a(n) = \begin{cases} h(0) & , n = 0 \\ 2h(n) & , 1 \leq n \leq N \end{cases}$

A common trigonometric identify expresses $\cos(\omega n)$ as a polynomial of degree n in the variable $\cos(\omega)$:

$$\cos(n\omega) = T_n\left(\cos(\omega)\right)$$

where $T_n(x)$ is the n^{th} Chebyshev Polynomial

$$
\begin{aligned}
T_0(x) &= 1 \\
T_1(x) &= x \\
T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x)
\end{aligned}
$$

$$\Rightarrow H(\omega) = \sum_{n=0}^{N} a(n)T_n(\cos(\omega))$$

7.2 Derivation of the Chebyshev Recursion

$$\cos(A)\cos(B) = \frac{1}{2}\cos(A+B) + \frac{1}{2}\cos(A-B)$$

Let $A = \omega$ and $B = (n-1)\omega$. Then,

$$\begin{aligned}
2\cos(\omega)\cos(n-1)\omega &= \cos(n\omega) + \cos(n-2)\omega \\
\cos(n\omega) &= 2\cos(\omega)\cos(n-1)\omega - \cos(n-2)\omega \\
T_n(\cos(\omega)) &= 2\cos(\omega)T_{n-1}(\cos(n-1)\omega) - T_{n-2}(\cos(\omega)) \\
T_x &= 2xT_{n-1}(x) - T_{n-2}(x)
\end{aligned}$$

7.3 The Transformation

McClellan suggested that we can obtain a 2-D zero-phase FIR filter if we make the substitution

$$F(\omega_1, \omega_2) \rightarrow \cos(\omega)$$

Then

$$H(\omega_1, \omega_2) = \sum_{n=0}^{N} a(n)T_n\left(F(\omega_1, \omega_2)\right)$$

F should be chosen to be the frequency response of a zero-phase FIR filter

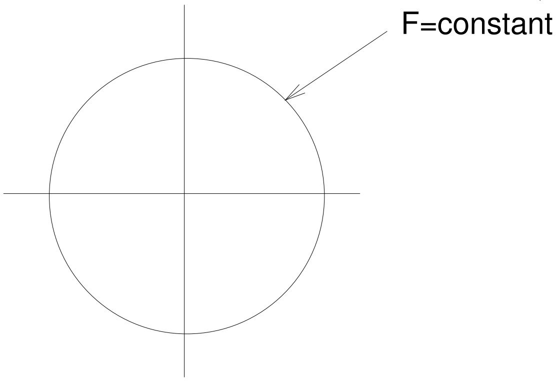
- F is real $\Rightarrow$ H is real $\Rightarrow$ zero phase.
- F is $(2Q+1) \times (2Q+1) \Rightarrow H$ is $(2NQ+1) \times (2NQ+1)$ because H is a polynomial in F.

The discrete-time Fourier transform of a 3×3 zero-phase filter reduces to

$$F(\omega_1, \omega_2) = A + B\cos(\omega_1) + C\cos(\omega_2) + D\cos(\omega_1 - \omega_2) + E\cos(\omega_1 + \omega_2)$$

What does the frequency response look like?

Consider the set of points $\{ \ (\omega_1, \omega_2) : F(\omega_1, \omega_2) = \text{const} \ \}$. These points define a closed surface or series of closed surfaces in the (ω_1, ω_2) - plane.



- For all the points on the contour H is constant.
- Contour shape depends upon A, B, C, D, and E.
- Value of H also depends upon $\{a(n)\}$ for $n = 0, 1, \ldots N$.

7.4 Procedure

- Design transformation
- Design prototype