Designing a Linear FIR filter

Leena Darsena, Himani Agrawal

Abstract- Digital filtering has a very important place in digital signal processing. There are various classes of digital filter like FIR filter and IIR filter. FIR filter is widely applicable due to stability, non recursive nature and linear phase. The Differential Evolution is a population based algorithm which has been proposed particularly for a number of numeric optimization problems. In this paper, various methods for design of digital FIR filter is studied.

Keywords: — FIR filter, Window Method, Differential Evolutionary Algorithm, Eclectic DE, Particle Swam Method, Genetic Algorithm, Soft Computing.

I. INTRODUCTION

FIR filters are filters which have a transfer function of a polynomial in $z^{-1}$ and are an all-zero filter because the zeroes in the $z$-plane determine the frequency response magnitude characteristic. The $z$-transform of an $N$-point FIR filter is given by

$$H(z) = \sum_{n=0}^{N-1} h(n) z^{-n}$$

FIR filters are particularly useful for applications in which exact linear phase response is required. FIR filter is generally implemented in a non-recursive way which guarantees a stable filter. FIR filter design basically consists of two parts

(i) approximation problem

(ii) realization problem

Four steps generation of a transfer function by approximation stage by taking the specifications are as follows:

(i) Choose a desired or ideal response in the frequency domain.

(ii) An allowed class of filters is chosen (e.g. the length $N$ for a FIR filters).

(iii) Choosing a measure of the quality of approximation.

(iv) Best filter transfer function is found by choosing an algorithm or method..

The realization part deals with choosing the structure to implement the transfer function which can be in the form of a program or form of a circuit diagram. For FIR filter design there are the following three well-known methods:

1. The window method
2. The frequency sampling technique
3. Optimal filter design methods

II. THE WINDOW METHOD

Most Common type filters include a low-pass filter, which pass through the frequencies below their cut-off frequencies, and attenuates frequencies above the cut-off frequency of a signal according to the requirement. There are many methods for designing FIR digital filters to meet arbitrary frequency and phase response specifications, such as frequency sampling techniques or window design method. The Window method is the most effective and popular method because this method is simple, convenient, fast and easy to understand. The main advantage of this design technique is that the impulse response coefficient can be obtained in closed form without solving complex optimization problems [4].

DESIGN STEPS FOR FIR FILTERS VIA WINDOW METHOD

1. Define the filter specifications.
2. Specify window functions according to the filter specifications.
3. Compute the filter order required for given specifications.
4. Compute the window function coefficients.
5. Compute the ideal filter coefficients according to the filter order.
6. Compute FIR filter coefficients according to the ideal filter coefficients and obtained window function.
7. If the resulting filter has too narrow or too wide transition region, it is necessary to change the filter order by decreasing or increasing it according to
needs, and after that steps 4, 5 and 6 are iterated as many times as needed.

Window functions can be divided into two groups; Fixed and Adjustable window functions. Fixed window functions which are used mostly are; Hanning window, Rectangular window, Hamming window and Blackman window. On the contrary Kaiser window is a kind of adjustable window function. Literature survey suggests that, these different widows are used for the spectral performance analysis and Digital FIR filter designing [5-7]. FIR filters design using a new window function, as given in [8]. The Performance Enhancement Study of FIR Filters Based on Adjustable Window Function is given in [9]. A new window function is described in [10] yielding suppressed main lobe width and minimum side lobe peak. In the study of Fourier transform of these different Fixed window functions, for the fixed length the Rectangular window yields smallest main lobe width and highest peak of side lobe among them. So Rectangular window is not widely used where applications of digital signal processing is involved. The Hanning and Hamming window provides good side lobe attenuation as compared to rectangular window, so these types of window are commonly used in different DSP applications. For higher side lobe attenuation Blackman window is used but the width of main lobe is wider in Blackman window as compared to Hamming and Hanning window. The Kaiser window is an adjustable window function which gives independent control of the ripple ratio and main lobe width. But the Kaiser window has the disadvantage of higher computational complexity because of the use of Bessel functions in the computation of window coefficients [8]. In some applications such as FFT, measurement and signal processing, higher side lobe attenuation is required as compared to a Hamming window [10]. The Blackman window function may be used for applications of this type, but the Blackman window has a wider main lobe width and if the main lobe width of any window functions increases the ability to distinguish two closely spaced frequency components decreases.

Windowing has following effects on the Fourier coefficients of the filter on the result of the frequency response of the filter are as follows:

(i) A major effect is that between values on either side of the discontinuity $H(w)$ become transition bands.

(ii) The width of the transition bands depends on the width of the main lobe of the frequency response of the window function, $w(n)$ i.e. $W(w)$.

(iii) Since the filter frequency response is obtained via a convolution relation, so the resulting filters are never optimal in any sense.

(iv) As $M$ (the length of the window function) increases, reducing the main lobe width of $W(w)$ which in turn reduces the width of the transition band, but more ripples are introduced in the frequency response.

(v) Elimination of the ringing effects at the band edge is observed due to windowing function, and it also result in lower side lobes at the expense of an increase in the width of the transition band of the filter.

There are following problems in filter design using window method:

(i) This method is applicable only if $H_d(w)$ is perfectly integrable, only if (2) can be evaluated. When $H_d(w)$ is complicated or cannot easily be put into a mathematical expression of closed form, evaluation of $h_d(n)$ becomes tedious.

(ii) The design flexibility is limited if windows are used e.g. in low pass filter design, the passband edge frequency generally cannot be specified precisely because the window smears the discontinuity in frequency. So, ideal LPF with cut-off frequency $f_c$ is smeared by the window to give a frequency response with stopband cut-off frequency $f_s$ and passband cutoff frequency $f_1$.

(iii) Window method is basically useful for design of prototype filters like lowpass, highpass, bandpass etc. Hence, its use in speech and image processing applications very limited.

III. THE FREQUENCY SAMPLING TECHNIQUE

The frequency sampling method allows us to design non-recursive and recursive FIR filters for both standard frequency selective and filters with arbitrary frequency response.

A. No Recursive Frequency Sampling Filters:

The problem of FIR filter design is to find a finite–length impulse response $h(n)$ that corresponds to desired frequency response. In this method $h(n)$ can be evaluated by uniformly sampling, the desired frequency response $HD(\omega)$ at the N points and finding its inverse DFT of the frequency samples as shown in the Figure 1.
\[ H (n) = \frac{1}{N} \sum_{k=0}^{N-1} H (k) e^{j \frac{2\pi}{N} nk} \]

where \( H (k), k = 0, 1, 2, \ldots, N-1 \), are samples of the \( H_D (\omega) \). To obtain a good approximation to the desired frequency response, we must take a good number of the frequency samples [9].

**Figure 1. Two type of the frequency sampling filter.**

**B. Types 1 and 2 frequency sampling filters:**

The frequency sampling filters are based on specification of a set of samples of the desired frequency response at \( N \) uniformly spaced points around the unit circle. The set of frequencies that are used till this point is determined by the relation.

\[ f_k = \frac{k}{N} F_s, \quad k = 0, 1, \ldots, N-1 \]

The frequencies in this equation are called frequencies of type 1.

Corresponding to the \( N \) frequencies at which an \( N \)-point DFT is evaluated. There is a second set of uniformly spaced frequencies for which a frequency sampling structure can conveniently be obtained. This set of frequencies determined by the relation [6].

\[ f_k = \frac{(k + \frac{1}{2})}{N} F_s, \quad k = 0, 1, \ldots, N-1 \]

This equation is called frequencies of type 2 for the case where \( N \) is either even or odd. The type 1 designs have the initial point at \( f = 0 \), but in the type 2 designs the initial point at \( f = 1/2N \). Type 2 frequency samples are important because they give additional flexibility to the design method to specify the desired frequency response at a second possible set of frequencies.

**C. Recursive frequency sampling filter:**

In recursive frequency sampling method the DFT samples \( H (k) \) for an FIR sequence can be regarded as samples of the filters \( z \)-transform, evaluated at \( N \) points equally spaced around the unit circle. [8]

\[ H (k) = H (z)|_{z = e^{j \frac{2\pi}{N} k}} \]

The main idea of the frequency sampling design method is that a desired frequency response can be approximated by sampling it of \( N \) evenly spaced points and after that obtaining an interpolated frequency response that passes through the frequency samples. For filters with reasonably smooth frequency responses, usually the interpolation error is small. In the case of band select filters, in which a radical change occurs across the band in the desired frequency response, the frequency samples which occur in transition bands are made unspecified variables whose values are chosen by an optimization algorithm which reduces some function of the approximation error of the filter, so it was shown that depending on where the initial frequency sample occurred [6], there were two types of frequency sampling filters.

**Merits of frequency sampling technique**

(i) Unlike the window method, this technique can be used for any given magnitude response.

(ii) This method is helpful in the design of non-prototype filters where the desired magnitude response can take any irregular shape.

Some of the disadvantages with this method are that the frequency response obtained by interpolation is equal to the desired frequency response only at the sampled points. At the other points, there will be a finite error present.

**III OPTIMAL FILTER DESIGN METHODS**

Many methods are present under this category. The basic idea in each method is to design the filter coefficients repeatedly until a particular error is minimized. The various methods are as follows:

2. Particle Swarm Optimization
1. Genetic Algorithm :

The designing of IIR filter by bilinear transformation require sufficient prior data to build up filter with desired parameter and show bad performance in maximum cases. To overcome these unsolved problems we use optimization approach to solve most of the problem and provide more accuracy, high efficiency, achievement, less prior knowledge to design any digital IIR filter. On the other hand genetic algorithm work on the population of candidate’s solution and various other constraints under the strategy set before. It also requires fitness function after which selection, crossover mutation and many process is performed [1, 3]. New powerful approaches are met with Genetic optimization method to solve the more difficult optimization problem.GA uses stochastic process and provide differently non-random solution or better solution. Genetic algorithm can be used in various application areas. Its flow chart is shown in fig 2.

Basic steps of GA are as follows:

A. Population

Encode each bit as a gene and a string of genes are called as chromosome and set of chromosomes is called population. First step to start with is to initialize the population i.e., is called as initial population.

B. Evaluation

Each chromosome has to be assessed and to be assigned a value called fitness value, larger the fitness value, (that says i.e., a good gene) probability will be more to select it for reproduction. Fitness is the measure of goodness of a chromosome.

B. Selection

The individual chromosomes which have best fitness values are selected and proceeded for next step called reproduction where blending of the both guardians would be carried to process new offspring’s [8]. Two regularly utilized methods are „roulette wheel” & „tournament” selection Over roulette wheel, every individual will be allocated a sector (slice) size proportional to their fitness evaluated The wheel is then spun and the individual inverse to the marker turns into a standout amongst those as parents[2].In this paper roulette wheel selection is used.

D. Reproduction

Two chromosomes which are selected based on the fitness value [9] from the population undergo a process called reproduction to produce offspring’s. Parents who have better fitness values have superior possibilities to be selected for production of finer offsprings.

E. Crossover

Choosing a random point and Splitting the parents at this crossover point and Creating children by trading their tails is called crossover process. Crossover probability is typically in the range (0.6, 0.9) [10]. Fig .1 describes the crossover process. It can be classified as

a. Simple Crossover: The Process we discussed above is a simple crossover process [2].

b. Arithmetic Crossover: It is a process where two complimentary linear combinations of the parents were produced [2].

c. Heuristic Crossover: It gets profit from fitness information with furthermore extrapolation of the two individuals to enhance the offspring [2].

F. Mutation

Mutations are random changes done to a chromosome in order to get a good fitness value as shown in Fig. 2. Mutations are random changes which reintroduces the genetic diversity into the population [9].
2. Particle Swarm Optimization (PSO):

The particle swarm optimization (PSO) is a evolutionary parallel computation technique motivated by the simulation of social behaviour. The PSO algorithm was first introduced by Eberhart and Kennedy [13]. The PSO algorithm is initialized with a population of random candidate solutions, conceptualized as particles. Every particle is assigned a randomized velocity and is iteratively moved through the problem space, which is, updated by applying some kinds of operators according to the fitness information which can be generated from given problem statement to be simulated so that the individuals of the population can be expected to reach better solution areas. It is attracted towards best fitness position obtained thus far by the whole population (social part) and by the individual itself (cognitive part). The position and velocity associated with each particle are arranged in a form of matrix. For the ith particle it can be given as: The length of the vector D represents the dimensions of the specified problem with (i=1, 2, ..., N) specifying the population set of the particles. An evaluation of the fitness function is done for all swarm particles and optimization of this fitness function is the aim, like a swarm of birds stop flocking when the food is found. For the fitness function optimization the position and the velocity of the particle is updated according to the social best and individual best results found so far.

Step I. Specify the fixed parameters used for the design of low pass linear phase FIR filter. This includes passband edge frequency=0.25; passband ripple=0.1; stopband edge frequency=0.3; stopband ripple=0.01; order of the filter n=20, thus filter has M=21 which is the length.

Step II. Initialize the population of particles with, N=30; number of dimensions (filter coefficients), D=11 (n/2+1); maximum and minimum value for the population of filter coefficients is equal to -2; velocity range is equal to 0.01; sampling frequency is 1Hz; number of samples taken are 128; maximum no of iterations, itmax is 250.

Step III. Generate the initial particle vectors using above defined parameters and calculate initial error fitness value for the whole population using Eq. (5) and (3).

Step IV. Minimum fitness value from error fitness vector is calculated and group best (hibest) values and individual best (hibest) from whole population is computed.

Step V. Update the velocity and position (filter coefficient value) according to Eq. (9) and (8) while checking against the limits of filter coefficients, now taking them as the initial particle vectors; also calculate the error fitness value from these updated parameters and hibest and hgbest accordingly.

Step VI. If values of vector hgbest and hibest calculated in Step V are better than those calculated in Step IV, then replace the vectors and no change if otherwise.

Step VII. Iterate continuously from Step IV to Step VI until convergence criteria (reaching itmax or error fitness value equals minimum error fitness) is met.

IV CONCLUSION

Filter design technique is now switching from traditional formula based design to intelligent design. We have seen lots of paper on the filter design using the Genetic algorithm and Particle swarm Optimization. These techniques as described and proved in discussed papers that they gives optimal result as compared to the traditional filter design techniques. The evolutionary algorithm required a fitness function which is to be minimized and the authors who uses genetic algorithm tried to develop a fitness function which gives better result. This is not guaranteed that using this type of algorithm will gives better result than traditional method with every fitness function until we use proper fitness function. In this paper we discussed about the design of the IIR and FIR filters and how to design the fitness function for specific problems. The mathematical equation of the system with proper assumptions can increase the accuracy of the algorithm. We also discussed about the multi objective problem and find that the simplest way to use multi objective problem is to convert the multi objective problem into single objective problem by providing the weight to every objective according to its importance. These algorithms are highly recommended for robotics with neural network called Hybrid intelligence. Using techniques for optimizing the neuron weight can gives better result for particular network. Algorithm for neural networks can be replaced by this algorithm. The most interesting fact of these algorithms are that they are inspired from the nature and we knows that nature is most optimum system we ever seen. Also the performance can be improved by modifying algorithm parameter.
REFERENCES


First Author: Leena Darsena, Electronics and telecommunication deptt, SSGI, Bhilai, India, 08109141386

Second Author: Mrs Himani Agarwal Associate Professor, Electronics and Telecommunication Deptt, SSGI Bhilai (C.G.)