

Application of Genetic Algorithms for the Design of Digital Filters

Sabbir U. Ahmad and Andreas Antoniou



Dept. of Electrical and Computer Engineering University of Victoria, Canada

OVERVIEW



- Introduction
- Design based on genetic algorithms (GAs)
- Application I: Design of fractional-delay FIR filters
- Application II: Design of delay equalizers
- Other applications
- Conclusions





INTRODUCTION

- The design of digital filters by means of optimization involves multiple and often conflicting design criteria and specifications.
- The optimization problem is complex, highly nonlinear, and multimodal in nature.







CLASSICAL OPTIMIZATION ALGORITHMS

- Fast and efficient
- Very good in obtaining local solutions
- Unbeatable for the solution of convex (concave) problems
- In multimodal problems, they tend to zoom to a solution in the locale of the initialization point.
- Not equipped to discard inferior local solutions in favour of better solutions.



CLASSICAL OPTIMIZATION ALGORITHMS (Cont'd)



- Constraints can be imposed on the objective function but the mathematical complexity of the optimization problem is increased often by several orders of magnitude.
- Increased mathematical complexity usually introduces ill-conditioning and on occasion it renders the problem intractable.





- Are very flexible, non-problem specific, and robust.
- Can explore multiple regions of the parameter space for solutions simultaneously.
- Can discard suboptimal solutions in favour of more promising subsequent local solutions.
- They are more likely to obtain better solutions for multimodal problems.





- Owing to the heuristic nature of GAs, arbitrary constraints can be imposed on the objective function without increasing the mathematical complexity of the problem.
- They require a very large amount of computation.



DSP Group

8

GENETIC ALGORITHMS (Cont'd)

- GA milestones:
 - Influenced by Darwin's Origin of Species
 - Introduced by Holland in 1962
 - Investigated further by Rechenberg and Schwefel in 1965
 - First textbook on GAs with detailed analysis: Goldberg, 1989
 - First GA paper on filter design: Etter and Masukawa,1981







- In a nutshell, a GA entails four fundamental steps as follows:
 - Step 1: Create an initial population of random solutions (*chromosomes*) by some means.
 - Step 2: Assess the chromosomes for fitness using the criteria imposed on the required solution and create an elite set of chromosomes by selecting a number of chromosomes that best satisfy the requirements imposed on the solution.





- Step 3: If the top-ranking chromosome in the elite set satisfies fully the requirements imposed on the solution, output that chromosome as the required solution, and stop. Otherwise, continue to Step 4.
- Step 4: Apply crossover between pairs of chromosomes in the elite set to generate more chromosomes and subject certain chromosomes chosen at random to mutations, and repeat from Step 2.





- As in the natural evolution of living organisms,
 - Chromosomes with new traits are generated,
 - chromosomes with better traits tend to survive and transfer those traits to their descendant chromosomes, and
 - after a certain period of CPU time, a chromosome will emerge that best satisfies the requirements imposed on the solution.



CONCEPTUAL REPRESENTATION OF GAs





OBJECTIVE FUNCTION



- The objective function for GAs is formulated as in classical optimization algorithms.
- GAs do not need gradient information. Therefore, the mathematical structure of these algorithms is simple and flexible.
- Multiobjective variants of GAs can handle problems with multiple, often conflicting, optimization goals.



INITIALIZATION



- The initial population can be created in various ways as follows:
 - Through random selection
 - Through a deterministic uniformly distribution
 - By creating a *seed* such as a solution obtained by classical optimization and then applying random perturbations to it.
 - Using a combination of two or more of the above schemes.



CHROMOSOME CODING



- Chromosome coding is the way of representing the *design variables*.
- GAs use various coding schemes such as:
 - binary coding
 - integer coding
 - Gray coding
 - decimal coding





- *Binary coding* : Each variable is encoded into a bit string of predefined length.
- *Integer coding* : The elements of chromosome vectors are integers.
- *Gray coding* : A binary coding with minimum Hamming distance between adjacent numbers (adjacent numbers differ in one bit).
- *Decimal coding* : The elements of chromosome vectors are decimal numbers.





- The choice of coding scheme depends on the optimization problem at hand, e.g.,
 - binary coding is useful for discrete variables.
 - decimal coding might be necessary when highprecision is required.



GENETIC OPERATORS



- Crossover and mutation are used to produce new individuals from the parent chromosomes.
- There are many ways of performing crossover:
 - One-point, two-point, or uniform crossover is used with binary coding.
 - Simulated binary crossover or perturbation is used with decimal coding. (Simulated binary crossover is designed to imitate one-point binary crossover.)
- Mutation randomly changes an offspring after crossover.
 - Mutation is treated as supporting operator for the purpose of restoring lost genetic material.





GENETIC OPERATORS (Cont'd)







- Both crossover and mutation are probabilistic operations and their frequencies of occurrence are controlled by predefined probabilites.
- As crossover plays the key role in improving the solution, it is assigned a high frequency of occurrence (typically 80-90%).
- The frequency of occurrence of mutation is kept fairly low (typically 5-10%) to prevent the GA from producing a large number of random solutions.



SELECTION METHODS



- Chromosomes are selected from the population based on the requirements imposed on the solutions in order to create a new population on the principle of the "survival of the fittest".
- The common selection methods used are
 - *roulette-wheel* selection
 - tournament selection
 - rank selection
 - Elitist selection





- *Roulette-wheel selection* : Each individual's probability of being selected is proportional to its fitness value.
- *Rank selection* : The individuals are ranked from 'best' to 'worst' on the basis of their measured fitness values and new fitness values are then assigned to the individuals that are inversely related to their ranking.
- *Tournament selection* : A group of individuals are chosen at random from the population and the one with the best fitness value is selected.
- *Elitist selection* : A number of individuals deemed to be the best are always passed on to the next generation unchanged.



APPLICATION 1: FRACTIONAL-DELAY FILTER DESIGN



- Fractional-delay FIR filters are needed for many DSP applications that require a tunable fractional delay (FD), e.g.,
 - speech coding and synthesis,
 - sampling-rate conversion,
 - time-delay estimation,
 - analog-to-digital conversion.



FRACTIONAL-DELAY FILTER (Cont'd)



- Three approaches are available for the design:
 - Recompute the coefficients.
 - Use lookup tables.
 - Design an FD filter based on the Farrow structure.
- The FS was introduced by Farrow in1988.
 - An FD is tunable on line without redesigning the filter.



FRACTIONAL DELAY APPROXIMATION



• Fractional delay (FD) filter:

$$x(n) \rightarrow z^{-(d+\mu)} \rightarrow y(n) = x(n-d-\mu)$$

d is the integer delay and μ is the fractional delay

 The impulse response of an ideal fractional delay filter can be represented by a *sinc* function shifted by the amount of FD.



FD FILTERS BASED ON FARROW STUCTURE (FDFS)



 An FS consists of (P+1) parallel FIR subfilters, each of length N.



where

$$B_k(z) = \sum_{n=0}^{N-1} b_{kn} z^{-n}$$





THE GA APPROACH



- GA structure:
 - A population of potential solutions is created from an initial least-squares solution.
 - Adaptive crossovers and mutations are applied.
 - The objective function used to evaluate the fitness of the individual solutions is based on both the amplitude response and delay errors.
 - A two-stage termination criterion is used.





- Chromosome (candidate solution) :
 - Matrix B consists of decimal-valued coefficients.
 - Is called the phenotype representation.
 - The phenotype representation along with the objective function are used for fitness evaluation.

$$\mathbf{B} = \begin{pmatrix} b_{00} & b_{10} & \mathbf{L} & b_{P0} \\ b_{01} & b_{11} & \mathbf{L} & b_{P1} \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ b_{0(N-1)} & b_{1(N-1)} & \mathbf{L} & b_{P(N-1)} \end{pmatrix}$$



ENCODING SCHEME

- Binary encoding:
 - Uses a fixed number of bits.
 - Is called the genotype representation.
 - The genotype representation is used for genetic operations, i.e., crossover and mutation.
- Example :
 - Subfilter (SF) length = 5
 - No. of subfilters = 3
 - 4-bit binary coefficients







INITIALIZATION



- Initialization of the GA:
 - An LS solution is used as seed for the initial population.
 - Half of the population is generated from points in the neighborhood of the seed.
 - The remaining is generated randomly (to maintain diversity).
- Subsequent generations:
 - Two-thirds of the population is selected from the previous generation.
 - One-third is generated randomly.



CROSSOVER



- Two complementary offspring chromosomes are generated from two randomly selected parent chromosomes.
 - The frequency of crossover is controlled by the crossover probability, P_x .
 - A randomly created mask is used to select genes from the two parents.





MUTATION



- Uses random bit inversion.
- The frequency of mutation is controlled by the mutation probability, P_m .
- $-P_m$ is much smaller than P_x .
- Serves as supporting operator for restoring lost genetic materials.
- Less effective than crossover in reducing the objective function.





- Adaptive crossover and mutation rates:
 - Adaptivity enables the GA to explore new areas of the parameter space when progress toward a solution is slow.
 - Initially P_x and P_m are set to relatively low values.
 - If no improvement is achieved in the solution after a number of generations, P_x and P_m are increased by letting

$$P_x = 1.05P_x$$
 and $P_{m=} 1.1P_m$



OBJECTIVE FUNCTION



- Objective function: $\Delta = W_a \delta_a + W_d \delta_d$
 - Peak amplitude-response error:

$$\delta_a = \max_{0 < \omega < \omega_p, \, |\mu| \le \mu} \left\| 1 - \left| H(e^{j\omega}, \mu) \right\| \right\|$$

- Peak phase-delay error:

$$\delta_{d} = \max_{0 < \omega < \omega_{p}, |\mu| \le \hat{\mu}} \left| \frac{\omega \mu - \arg \left\{ H(e^{j\omega}, \mu) \right\}}{\omega} \right|$$

- W_a and W_d are positive weighting factors.



SELECTION



- Ranking process
 - Chromosomes are ranked on the basis of fitness.
 - A small number of top-ranked chromosomes are recorded as elite chromosomes.
 - A fixed number of best-fit chromosomes are selected for the next generation.

The best fitness value in a generation is defined as

$$\psi = \min(\Delta_i)$$
 for $i = 1, ..., N_P$

where N_p is the population size.



TERMINATION CRITERION

- The GA terminates
 - when a predefined maximum number of generations is exceeded, or
 - if it does not improve the solution after a prespecified number of 'unproductive' generations.
- Two-stage termination
 - In early stages, a larger number of unproductive generations is allowed before termination.
 - In later stages, a smaller number of unproductive generations is allowed before termination.





EXAMPLE OF FDFS FILTER DESIGN



• Passband ω_p = 0.5, Farrow structure designed with 3 subfilter, each of length 9.





COMPARISON WITH LS METHOD



Maximum error vs fractional delay plots







EVOLUTION OF OBJECTIVE FUNCTION



Reduction in objective function through the evolution of the GA





- The GA yields quantization-error-free FD filters.
- An objective function based on the amplituderesponse and delay errors criteria offers flexibility.
- The GA yields an improved design with respect to the initial LS design.
- GA requires a large amount of computation.



- Linear-phase filters are usually designed as FIR or IIR filters.
- For highly selective filters, equalized IIR filters are often preferred.
 - An IIR filter is first designed to meet the amplitude response specifications.
 - A delay equalizer is constructed to equalize the group delay of the IIR filter.
 - Allpass filters are used as delay equalizers.



DELAY EQUALIZER DESIGN (Cont'd)



- Delay equalizers are usually designed by using gradient-based optimization methods (quasi-Newton methods work very well).
- The stability of the equalizer obtained cannot be guaranteed.
- To assure stability constrained optimization is often used which causes the objective function to become highly nonlinear.





A delay equalizer can be characterized by a transfer function of the form

$$H_{E}(z) = \prod_{j=1}^{L} \frac{1 + c_{1j}z + c_{oj}z^{2}}{c_{oj} + c_{1j}z + z^{2}}$$

where L = N/2 is the number of equalizer sections.

The equalizer coefficient vector can be written as

$$\mathbf{x} = \begin{bmatrix} c_{01} c_{11} c_{02} c_{12} \mathbf{L} & c_{0L} c_{1L} \end{bmatrix}$$

The stability condition for an equalizer is

$$c_{0j} < 1, c_{1j} - c_{0j} < 1, c_{1j} + c_{0j} > -1$$
 $(j = 1, 2, ..., L)$



OBJECTIVE FUNCTION



The group-delay flatness of the equalized filter can be measured in terms of a parameter Q which is given

$$Q = \frac{100(\tau_{FE} - \tau_{FE})}{(\tau_{FE} + \tau_{FE})}$$

where

by

 $\tau_{FE} = \text{Max group delay}$ $\tau_{FE} = \text{Min group delay}$

and

$$\tau_{FE} = \tau_F + \tau_E$$

• Q is used as the objective function



THE GA APPROACH



- GA structure:
 - The initialization, crossovers, and mutations are done as in the design of FD filters.
 - The objective function used to evaluate the fitness of the individual solutions is based on the flatness of the group delay of the filter-equalizer combination.
 - A sequential optimization is used.





- Chromosome (*candidate solution*) :
 - The coefficient vector *x* expressed in matrix form is used as the candidate solution:

$$c_{01}$$
 c_{02} c_{03} ... c_{0L} c_{11} c_{12} c_{13} ... c_{1L}

- Each column represents an equalizer section.
- To avoid very long binary strings, a floating-point representation is used in encoding the chromosomes.



INITIALIZATION



- Initialization of the GA:
 - The initial population is created randomly.
- Subsequent generations:
 - Two-thirds of the population is selected from the previous generation.
 - One-third is generated randomly (to maintain diversity).



ADAPTIVE PERTURBATION



- Crossovers are replaced by an adaptive perturbation technique.
- Initially, a relatively large perturbation is applied.
- As time advances, the level of perturbation is reduced exponentially using the control factor

$$\alpha = 0.4e^{-K/15}$$

- When no improvement is achieved after a specified number of generations, *K* is increased by one.
- Mutations are replaced with fixed but occasional perturbations.



SELECTION



- Ranking process
 - Only stable solutions are involved in the ranking.
 - A fixed number of best-fit chromosomes are selected for the next generation.

The best fitness value in a generation is defined as

 $\psi = \min(Q_i)$ for $i = 1, ..., N_P$

where N_{ρ} is the population size.



SEQUENTIAL DESIGN



- The required equalizer order cannot be predicted.
- New equalizer sections are added sequentially.
- The best solution for the *k*-section design is used to construct an initial (*k*+1)-section design that can be used as seed for the next design. Random values are used for the coefficients of the new section.



DESIGN EXAMPLE: BANDPASS FILTER



Filter specs.:

DSP Group

 $\delta_p = 1, \quad \delta_a = 40 \text{ dB}$ $\omega_{a1} = 0.2, \quad \omega_{p1} = 0.3$ $\omega_{a2} = 0.7, \quad \omega_{p2} = 0.5$ $\omega_s = 2 \text{ rad/s}$ Maximum delay error: 2%



DESIGN EXAMPLE (Cont'd)



Results:

- Number of equalizer sections required: 5.
- •The value of Q was reduced from 52.46% to 1.95%.



evolution of the GA.



REMARKS ON APPLICATION #2



- The GA can minimize an objective function based on the passband filter-equalizer group delay deviation.
- It discards unstable solutions.
- Equalizers can be designed that would satisfy arbitrary prescribed specifications.



OTHER APPLICATIONS



- Design of cascade-form multiplierless FIR filters
- Design of asymmetric FIR filters
- Hybrid GA-LS approach for IIR filters



CASCADE-FORM MULTIPLIERLESS FIR FILTER DESIGN



- The approach uses a recently introduced GA called *orthogonal GA* (OGA).
- OGA is based on the so-called *experimental design technique*.
- A fixed-point design of a linear-phase FIR filter is obtained.
- The effects of a finite word length are minimized by considering the filter as a cascade of two subfilters.



DESIGN EXAMPLE



• Lowpass filter specs.:

 ω_p = 0.35, ω_a = 0.5 rad/s, Cascade sections: *N* = 15+13

The initial cascade and direct-form designs were both obtained with the *Remez exchange algorithm* except that in the first design the transfer function was factorized before coefficient quantization.

	Initial cascade design	Design by OGA	Direct- form (Remez)
$\delta_{_{p}}\left(dB ight)$	0.229	0.105	0.174
$\delta_{\scriptscriptstyle S} ({\sf dB})$	34.30	37.70	30.15



Amplitude Response



DESIGN OF ASYMMETRIC FIR FILTERS

University of Victoria

- Symmetric Coefficients
 - Efficient design methods, e.g., window method, Remez algorithm.
 - Large group delay.
- Asymmetric Coefficients
 - Approximately linear phase response in passband
 - Arbitrary amplitude response in the baseband
 - Relatively small group delay
 - Can be designed by using classical optimization methods with a multiobjective formulation.
 - Can also be designed by using a variant of multiobjective GA known as *elitist non-dominated sorted GA* (ENSGA).



MULTIOBJECTIVE OPTIMIZATION



- The design problem requires simultaneous optimization of several objective functions.
- The approach yields a set of compromise solutions known as *Pareto optimal* solutions.



DESIGN EXAMPLE



- **Design specs.:** $\omega_p = 0.25$, $\omega_a = 0.4$ rad/s, N = 23
- A *weighted LS* (WLS) solution was used as seed for the initial set of solutions.





DESIGN EXAMPLE (Cont'd)



Results:

	$\delta_{_p}$ (dB)	δ_{s} (dB)	Q (%)
WLS	1.06	36.74	2.31
NSGA	0.58	38.07	0.77



3-D plot of Pareto optimal solutions





- Two types of algorithms, a GA and a least-squares quasi-Newton algorithm, are used in a hybrid algorithm that combines the advantages of the two algorithms and avoids their limitations. The objective is to
 - achieve a robust optimization method for IIR filters,
 - reduce the computational effort associated with the GA.
- A GA is used for global search.
- A quasi-Newton algorithm is used for the local search.



DESIGN EXAMPLE



• IIR bandpass filter specs.:

 ω_{a1} = 120, ω_{p1} = 175, ω_{p2} = 220, ω_{a2} = 320, ω_s = 1000 rad/s





DESIGN EXAMPLE (Cont'd)





Passband Error



Stopband Error



CONCLUSIONS



- The design of digital filters and equalizers through the use of GAs has been explored.
- Five different types of classical design problems have been investigated.
- In all projects, improved designs have been achieved relative to designs produced by well-known state-of-the-art techniques.
- Evolution is a very slow process. Consequently GAs require a large amount of computation. However, this is not a critical demerit nowadays unless the filter design needs to be carried out in real or quasi-real time.



REFERENCES



- [1] A. Antoniou, *Digital Signal Processing: Signals, Systems, and Filters*, McGraw-Hill, 2005.
- [2] D. E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*, Addison-Wesley, San Francisco, 1989.
- [3] M. Gen and R. Cheng, *Genetic algorithms and engineering design*, Wiley, New York, 1997.
- [4] D. Suckley, "Genetic algorithm in the design of FIR filters," *IEE Proc.,* pt. G, vol. 138, pp. 234-238, Apr. 1991.
- [5] C. W. Farrow, "A continuously variable digital delay element," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 3, pp. 2641-2645, Espoo, Finland, Jun. 7-9, 1988.
- [6] T. I. Laakso, V. Valimki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay," *IEEE Signal Processing Mag.*, pp. 30-60, Jan. 1996.
- [7] S. U. Ahmad and A. Antoniou, "A Genetic Algorithm Approach for Fractional Delay FIR Filters," *Proc. IEEE Int. Symp. on Circuits and Syst.*, Kos, Greece, pp. 2517-2520, May 21-24, 2006.
- [8] C. Charalambous and A. Antoniou, "Equalization of recursive digital filters," IEE Proc., pt. G, vol. 127, pp. 219-225, Oct. 1980.
- [9] S. U. Ahmad and Andreas Antoniou, "A Genetic Algorithm for the Design of Tunable Fractional-Delay Allpass IIR Filter Structures," to appear at *Proc. IEEE Canadian Conf. on Elec. and Comp. Eng.*, Vancouver, Canada, Apr. 22-26, 2007.
- [10] S. U. Ahmad and A. Antoniou, "A Genetic-Algorithm Based Approach for the Design of Delay Equalizers," *Proc. IEEE Canadian Conf. Elec. and Comp. Eng.*, pp. 566-569, Ottawa, Canada, May 2006.



REFERENCES (Cont'd)



- [11] Y. C. Lim, "Design of discrete-coefficient-value linear phase FIR filters with optimum normalized peak ripple magnitude," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1480-1486, Dec. 1990.
- [12] Y.W. Leung and Y. Wang, "An orthogonal genetic algorithm with quantization for global numerical optimization," *IEEE Trans. Evolutionary Comp.*, vol. 5, Issue 1, pp. 41-53, Feb. 2001.
- [13] S. U. Ahmad and A. Antoniou, "Cascade-Form Multiplierless FIR Filter Design Using Orthogonal Genetic Algorithm," *Proc. IEEE Int. Symp. on Signal Proc. and Information Technology*, pp. 932-937, Vancouver, Canada, Aug. 27-30, 2006.
- [14] N. Srinivas and K. Deb, "Multi-objective function optimization using non-dominated sorting genetic algorithms," *Evolutionary Computation*, vol. 2, no. 3, pp. 221-248, 1994.
- [15] J. W. Adams, "FIR digital filters with least-squares stopbands subject to peak-gain constraints," *IEEE Trans. Circuits Syst.*, vol. 39, pp. 376-388, Apr. 1991.
- [16] S. K. Kidambi and R. P. Ramachandran, "Design of nonrecursive filters satisfying arbitrary magnitude and phase specifications using a least-squares approach," *IEEE Tarns. Circuits Syst. II*, vol. 2, no 11, pp. 711-716, Nov. 1995.
- [17] S. U. Ahmad and Andreas Antoniou, "A Hybrid Genetic Algorithm for the Design of IIR Digital Filters," to appear at *Proc. IEEE Canadian Conf. on Elec. and Comp. Eng.*, Vancouver, Canada, Apr. 22-26, 2007.







DSP Group

Thank you