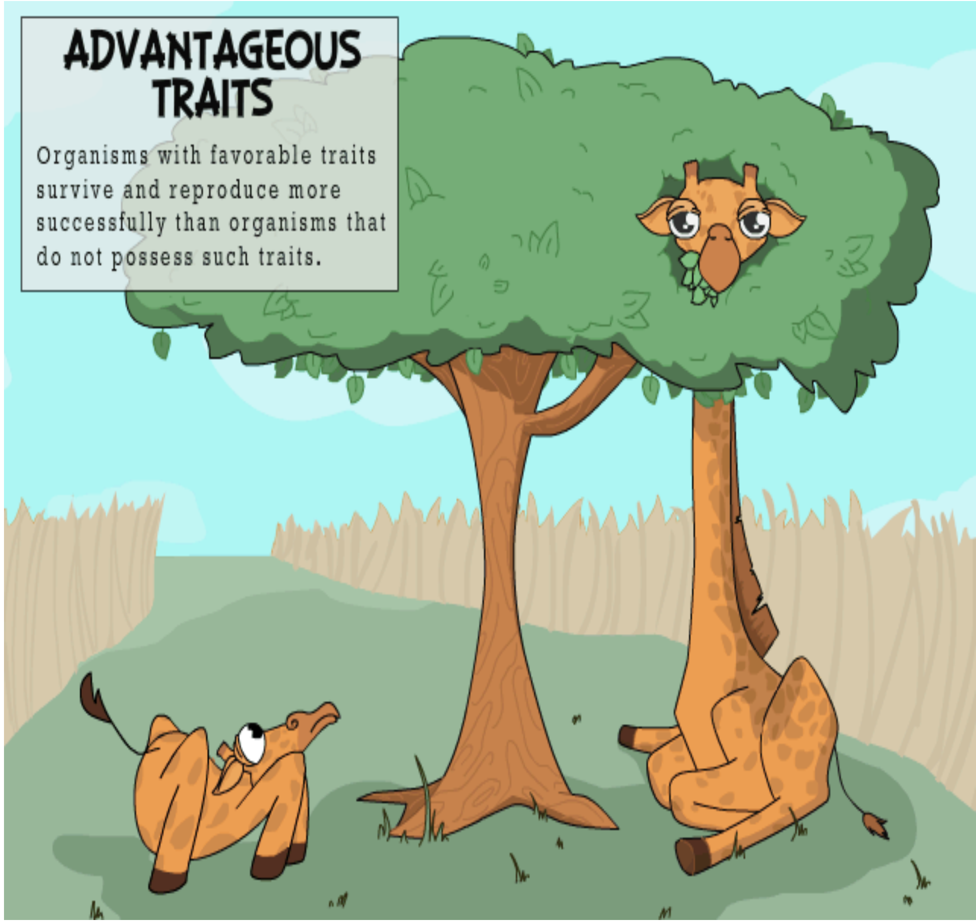
Cse634
Data Mining
Professor Anita Wasilewska

Genetic Algorithms Applications



Source : http://epicscience.epicsite.org/images/ChloeStewartgirafferevised.gif

Ok Google, what’s a genetic algorithm?

Genetic algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on bio-inspired operators such as mutation, crossover and selection.—Wikipedia

Areas where Genetic Algorithms are used...

Optimization
Economics
Neural Networks
Parallelization
Image Processing
Vehicle routing problems
Scheduling applications
Machine Learning
Robot Trajectory Generation
Parametric Design of Aircraft
DNA Analysis
Multimodal Optimization
Traveling salesman problem and its applications
…
And the saga continues

Let’s see something which you already know!

Consider a Neural Network model which has 4 parameters with five possible settings for every parameter. For a fully connected MLFF Neural Network, finding the right parameters or parameter tuning may take up some time.

Let’s say, it takes 5 minutes to train and evaluate a network on the given dataset. Total time considering all different combinations -

5 minutes * 5^4 = 3,125 minutes = 52 hours

Genetic Algorithms can optimize this and we can do this much faster.

How much faster?

Let's start with an initial population of 20 and we will keep top 25% (5 individuals) of best performing individuals and select 3 individuals at random from the others to ensure diversity and avoid getting stuck at local maxima. For 1st generation, we need (20 * 5 minutes = 100 minutes). Consider we go upto 10 generations. Then, as 12 are always replaced, the total time needed is:

$$100 \text{ minutes} + 12 * 9 \text{ generations} * 5 \text{ minutes} = 640 \text{ minutes} = 11 \text{ hours}$$

How to apply Genetic Algorithms to Neural Networks?

Goal - Find best parameters for Image Classification task

How we will do it? We will tune four parameters:

- Number of layers (or the network depth)
- Neurons per layer (or the network width)
- Dense layer activation function
- Network optimizer

Steps

- Initialize N random networks to create our population.
- Assign a rank to each network. As this is an image classification task, we'll use classification accuracy as our fitness function.
- Sort all networks in our population by their rank. Some percentage of the top-rankers will be promoted to next generation and only a few of the remaining will be chosen at random to ensure we don't get stuck at local maxima.
- Randomly mutate some parameters on some of the networks.
- Allow the chosen individuals to breed.

These steps were performed using CIFAR10 dataset.

Results

BRUTE-FORCE APPROACH FOR PARAMETER TUNING	
Layers	2
Neurons	768
Activation	elu
Optimizer	adamax
Time to Run	63 hours
ACCURACY	**56.03%**

VS

GENETIC ALGORITHMS APPROACH FOR PARAMETER TUNING	
Layers	2
Neurons	512
Activation	elu
Optimizer	adamax
Time to Run	7 hours
ACCURACY	**56.56%**

Genetic Algorithms in Computer Vision

Genetic Algorithms (GAs) are increasingly being explored in many areas of image analysis to solve complex optimization problems. [2]

It was proved that genetic algorithms are the most powerful unbiased optimization techniques for sampling a large solution space. [3]

Because of unbiased stochastic sampling, they were quickly adapted in image processing. They were applied for the image enhancement, segmentation, feature extraction and classification as well as the image generation. [3]

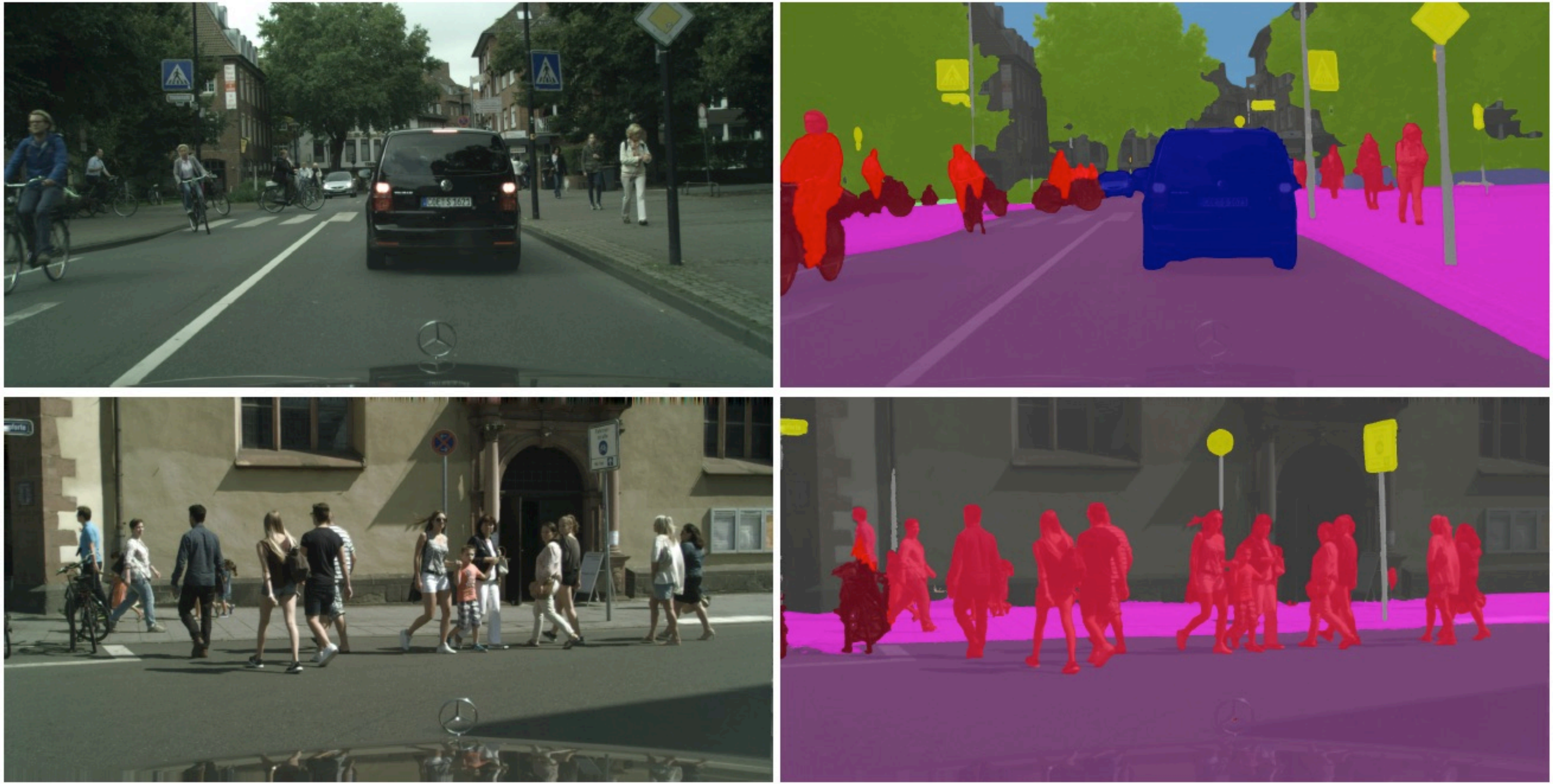
Genetic Algorithms in Computer Vision

- Image Segmentation
- Feature Selection for Image Classification

What is Image Segmentation?

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). [1]

The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze.



Source: http://vladlen.info/publications/feature-space-optimization-for-semantic-video-segmentation/

Applications of Image Segmentation

Medical Imaging

Object Detection

Face Recognition

Traffic Control

Content Based Image Retrieval

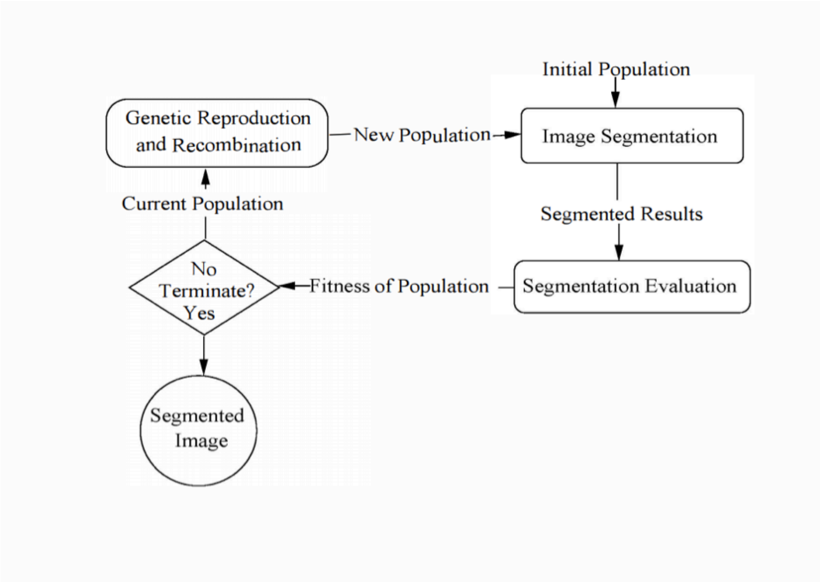
Genetic Algorithm based Image Segmentation

GA is proposed to explore the solution space

Each pixel is grouped into other pixels by distance function based upon both local and global already computed segments.[3]

Encoding: 32-bit binary string divided as 8 least significant bits identifies to which label they belong. Pixel position is encoded into the rest 24 bits.

The fitness function of the segmentation quality varies from image to image but generally is based upon the distance between each chromosome and the a label of which they are a part of. [3]



Source: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6320144

Comparison of Genetic Algorithm with other image segmentation algorithm

Method	Hand	House	Fish
K-means	0.66	0.61	0.63
Global Threshold Method (split-and-merge)	0.53	0.55	0.49
GA-Method	0.83	0.80	0.83

Agreement Indicators of the tested images. [1]

Feature Selection for Image Classification

Different techniques of getting Image Feature Vectors

- Histogram Mean
- Histogram of Gradients
- Scale Invariant Feature Transformation

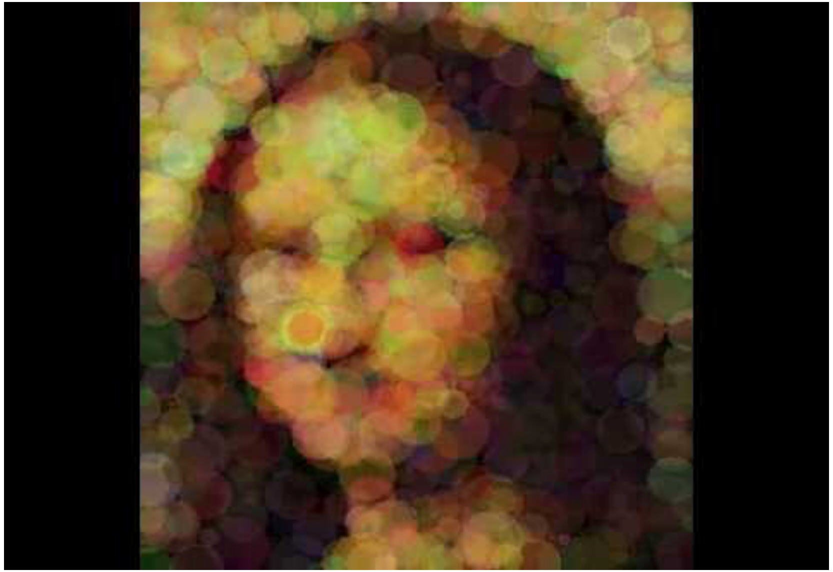
Genetic Algorithm Operators

Encoding : Bit string with 1 representing that feature is included and 0 for not.

Fitness : To evaluate how good a chromosome is we just have to evaluate the combination of features that the chromosome represents.

Selection : Roulette Wheel selection, Tournament Selection, Random Selection

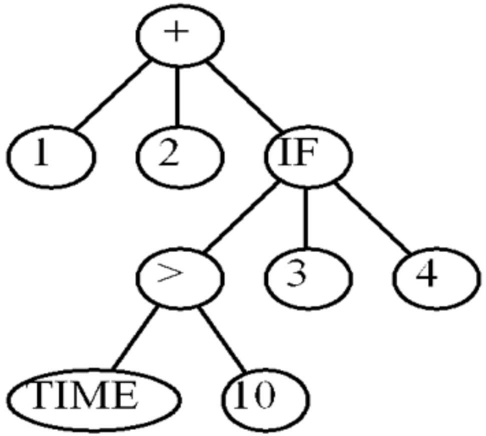
Approximating an Image using genetic algorithm



Genetic Algorithms in Data Science

Symbolic Regression using Genetic Programming:

The process of mechanically creating a computer program that fit certain numerical data.



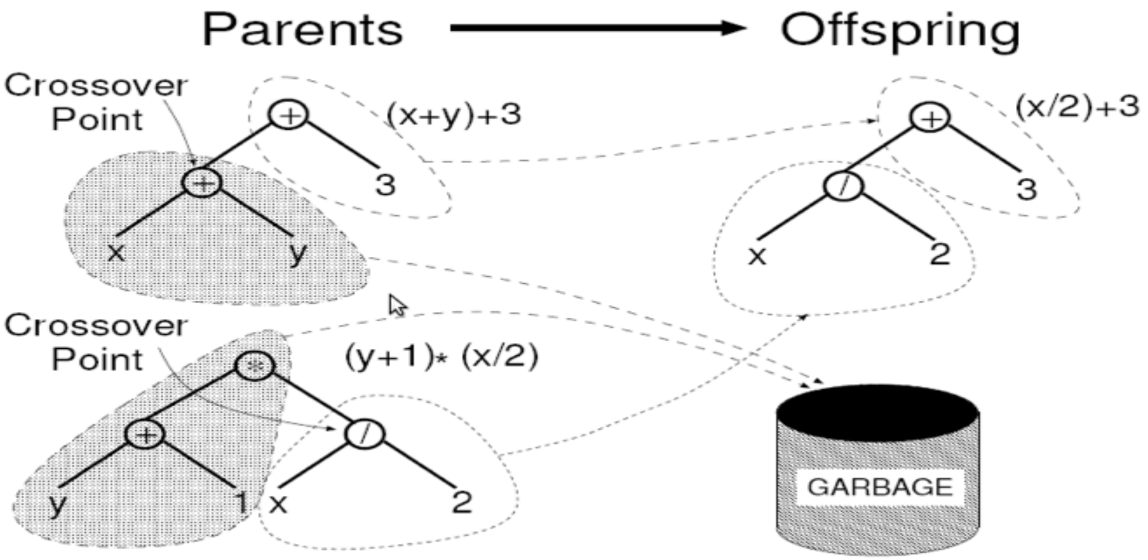
Source: “A Field Guide to Genetic Programming” by Poli, Langdon, McPhee. 2008.

Breeding Next Generation

Main operators applied on parents to generate offsprings are:

- Crossover
- Mutation
- Reproduction

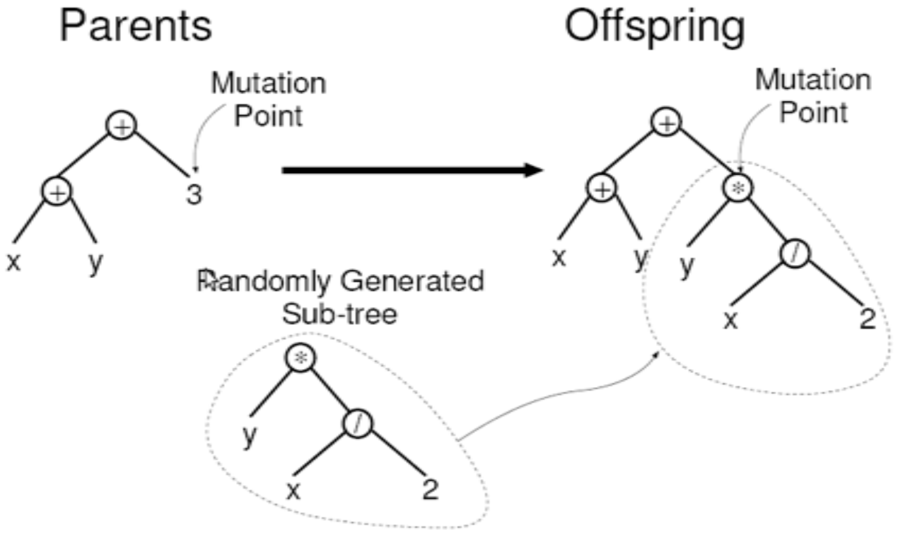
Subtree Crossover



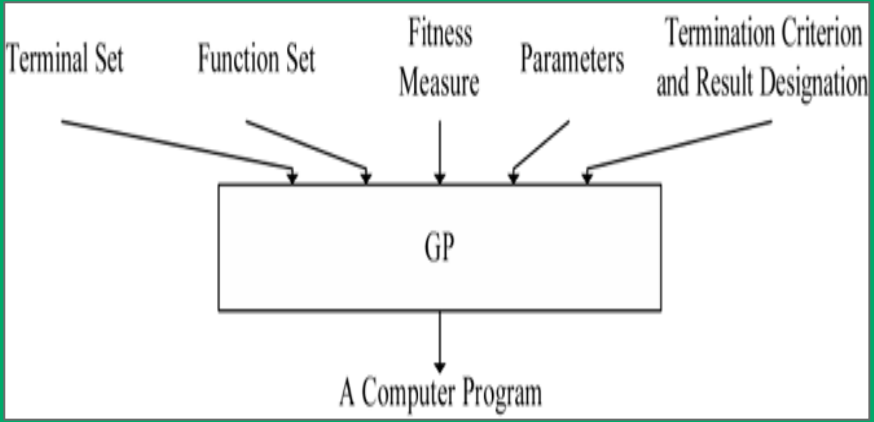
Source: “A Field Guide to Genetic Programming” by Poli, Langdon, McPhee. 2008.

Mutation

- **Subtree Mutation :** Replace the mutation point by randomly generated tree.
- **Point Mutation :** Randomly select a node and replace it with another primitive of same arity.



Source: "A Field Guide to Genetic Programming" by Poli, Langdon, McPhee. 2008.



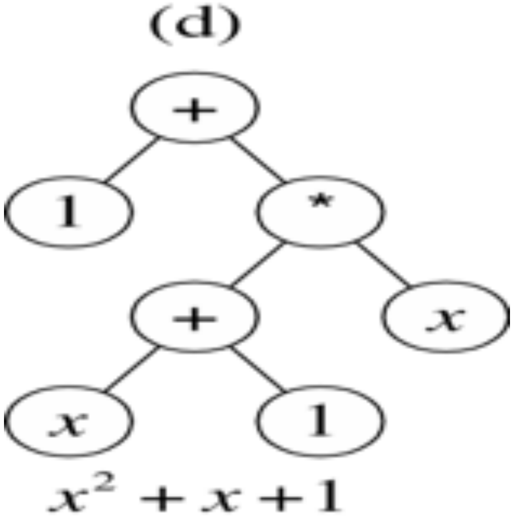
Five Preparatory steps to set up Genetic Programming

- Determining the set T of terminals.
- Determining the set F of functions.
- Determining the fitness measures.
- Determining the GP parameters.
- Determining the Termination criteria and result designation

Genetic Algorithms in Data Science

Go over the example

We want to evolve an expression whose values match those of the quadratic polynomial x2 + x + 1 in the range [−1,1].



Source: “A Field Guide to Genetic Programming” by Poli, Langdon, McPhee. 2008.

Areas where Genetic Programming will do well

- The interrelationships among the relevant variables is unknown or poorly understood.
- Finding the size and shape of the ultimate solution is a major part of the problem.
- Search domain is very large and solutions are sparsely distributed.
- There are good simulators to test the performance of tentative solutions to a problem, but poor methods to directly obtain good solutions.

Application of Symbolic Regression

GP was used in designing of an antenna for deployment on NASA's Space Technology 5 Mission.

Source: "An Evolved Antenna for Deployment on Nasa's ST5 Mission" by J. Lohn, G. Hornby, D. Linden, GTPT, May-2004

Intelligent Audio Watermarking using Genetic Algorithm

Watermark is a pattern of bits inserted into a digital image, audio or video file that identifies the file's copyright information (author, rights, etc.)

The basic idea of watermarking is to add a watermark signal into the host data to be watermarked such that the watermark signal is unobtrusive and secure in the signal mixture, but can partly or fully be recovered from the signal mixture later on if the correct cryptographically secure key is used.

An innovative watermarking scheme for audio signal is based on genetic algorithms (GA) which makes audio signals robust against watermarking attacks.

GA is employed for the optimal localization and intensity of watermark.

Overview

There are many techniques of audio watermarking algorithms can be grouped into third categories:

- **Patchwork** : The patchwork scheme embeds a special statistic into an original audio signal. The two major steps in the scheme are: (i) choose two patches pseudo-randomly and (ii) add the small constant value d to the samples of one patch A and subtract the same value d from the samples of patch B.
- **Echo hiding** : Echo hiding embeds data by introducing an echo according to the formula $y[n] = x[n] + \alpha[n - \delta]$

  where x[n] is the original audio signal, α is the echo's amplitude and δ is the audio signal delay

- **Spread Spectrum** : Spread-spectrum watermarking scheme is an example of the correlation method which embeds pseudo-random sequence and detects watermark by calculating correlation between pseudo-random noise sequence and watermarked audio signal.

The watermarking problem can be viewed as an optimization problem. Therefore, it can be solved by genetic algorithms (GA). GA improve the performance of watermarking schemes

The Embedding Algorithm

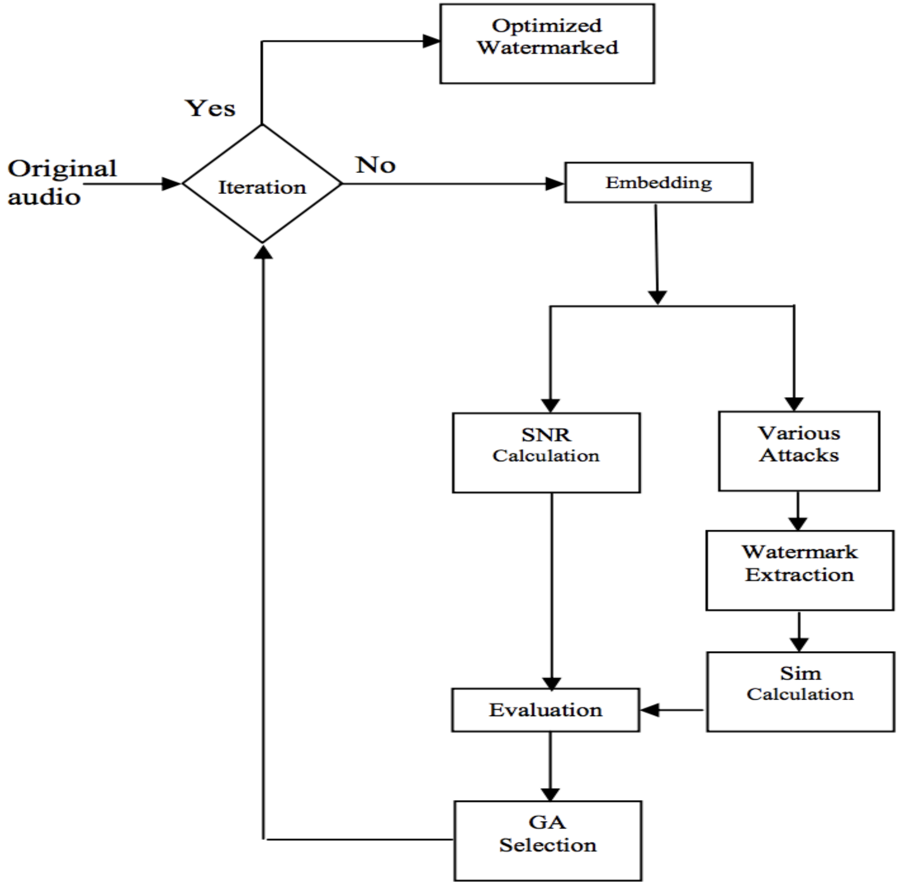
The embedding algorithm is performed to the wavelet coefficients of the audio segment.

1. The watermark data is transformed into a uni-dimentional antipodal sequence. Then, a random sequence is generated which is used to encrypt watermark to ensure security.

2. Wavelet coefficients of the input audio signal are obtained which are divided in k segments.

3. Then, each bit of watermark data is embedded into each segment.

Block Diagram for watermark embedding with GA



Source:http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.2037&rep=rep1&type=pdf

We applied the signal-to-noise-ratio (SNR) for evaluating the quality of the watermarked signal, and similarity (Sim) to evaluate the robustness. The definitions of SNR and Sim are:

$$SNR = 10\log_{10}\left(\frac{\sum_i (Y_i^2)}{\sum (Y_i - y_i)^2}\right)$$

Where Y and y are audio signal original and audio signal embedding respectively.

$$Sim = \frac{\sum_{i=1}^{M1} \sum_{j=1}^{M2} w(i,j) w^*(i,j)}{\sqrt{\sum_{i=1}^{M1} \sum_{j=1}^{M2} W(i,j)^2} \sqrt{\sum_{i=1}^{M1} \sum_{j=1}^{M2} W^*(i,j)^2}}$$

Where W and W* are original and extracted watermarks, respectively, i and j are indexes of the binary watermark image.

A pseudo-code for GA Algorithm

```
Pmut = 0.3;
 Pcross = 0.9;
 Popsize = 500;
Generation = 100;
Initialization (pop1 ,..., pop1000);
 Evaluation (pop1 ,..., pop1000)  /* Embedding ();
                                   Detection ();
                 Various attacks ();
                           */
For loop = 1 to generation do
          Selection (tournament selection size 2);
          Mutation ();
         Evaluation (pop1 ,..., pop1000)
End For
```

Source : http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.2037&rep=rep1&type=pdf

Objective Function Evaluation The objective function of GA is composed of the signal-to noise-ratio (SNR) of the watermarked signal versus the host signal used as a quality measure and the similarity (Sim) between extracted watermark and original watermark. The objective value fi can be computed by the following equation:

$$f_i = SNR_i + \frac{1}{p}\sum_{n=1}^{p}(Sim_{n,i})$$

where p is the number of attacking schemes.

Genetic Algorithm Optimization Applied to Electromagnetics: A Review

Authors :-
Daniel S. Weile
Eric Michielssen
Member, IEEE

Published :-
IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 45, NO. 3, MARCH 1997

Abstract

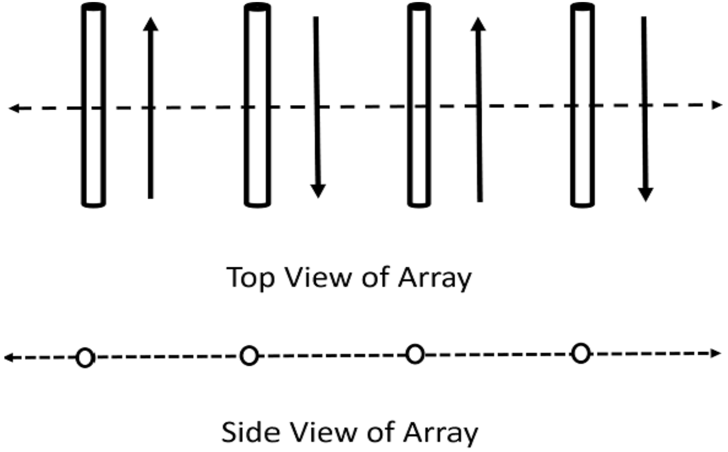
Genetic algorithms are on the rise in electromagnetics as design tools and problem solvers because of their versatility and ability to optimize in complex multimodal search spaces. This paper describes the basic genetic algorithm and recounts its history in the electromagnetics literature. Also, the application of advanced genetic operators to the field of electromagnetics is described, and design results are presented for a number of different applications.

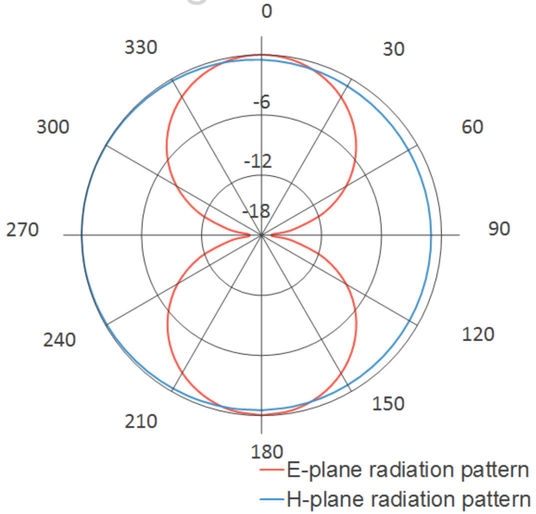
GENETIC ALGORITHMS IN ELECTROMAGNETICS:Applications

Genetic Algorithm in Antenna Design

- A plethora of studies have investigated GA-based methods of reducing the sidelobes of an array by thinning, amplitude or phase tapering, or element position perturbations.
- Closely related to this problem is reducing the scattering of strip arrays by thinning or perturbation. Most of these studies involve digital parameters and, thus, are textbook cases for binary or other finite alphabet GA’s.
- For instance, an array built with ten 3-b phase shifters can be encoded into a chromosome of 30-b, composed of a concatenation of ten 3-b phase shifter settings.

- Single antenna elements have been designed as well. For example, techniques exist for the fast GA optimization of wire antennas loaded with passive resonant circuits (RLC) to broaden their frequency response.
- The most intriguing application of GA's to antenna design is the so-called "genetic antenna". In this study, a wire antenna was designed to radiate circularly polarized (CP) waves by choosing endpoints for straight wire segments comprising the antenna and then "connecting the dots."





Source:- TutorialsPoint and ResearchGate.net

Design of Layered Electromagnetic Devices using Genetic Algorithm

- Among the earliest GA studies in electromagnetics is the design of multilayered optical filters by minimizing the difference between the observed and desired filter characteristic.
- Certain filters can be realized by layering two preselected materials, the only design parameters considered are layer thicknesses which are real valued and arrayed into a real-coded chromosome.
- Very similar to the design of optical filters is the design of absorbers consisting of a perfect electric conductor (PEC) coated with lossy materials to minimize reflection of impinging waves.

Applications of Genetic Algorithms in Statics

- Many applications of GA's involve shaping magnetic pole pieces or insulators to produce a desired magnetic or electric field distribution in a given region of space.
- One notable study of this sort is, is one in which a one can optimize a pot-core transformer not only to produce a specific field pattern, but to minimize the device dimensions as well.
- As the stated goals are physically incommensurable, they are combined in an ad hoc manner to achieve a palatable result. This technique may lead to extensive tinkering to find a suitable algebraic combination of objectives.

DESIGN EXAMPLES USING ADVANCED GENETIC ALGORITHMS OPERATORS

Genetic Algorithm Design of Broadband Loaded Wire Antennas in a Complex Environment

- Simple GA to the design of a wire antenna situated in an arbitrary, possibly complex, environment for operation over a broad range of frequencies.
- The wire is to be augmented with a given number of parallel RLC loads whose locations and element values are to be determined by the GA to produce maximum gain over the frequency band of interest.
- In addition, the GA simultaneously designs a matching network to achieve an acceptable standing wave ratio (SWR) over the band.
- GA’s typically require many objective function evaluations, which permits the fast analysis of loaded structures residing on a complex platform.

- A simple way to optimize network topology is to encode all matching network component values with two extra control bits.
- The first of these bits could indicate whether the element in question is a capacitor or an inductor, and the second whether the element is series or shunt connected.
- The matching network could then be built element by element from the antenna back to the source. This would lead to a chromosome structured as follows. The first part of the chromosome would contain descriptions of the antenna loads and locations spanning the first N_{ant} bits. If the antenna has 2^{Nt} possible load locations, and N_{RLC} RLC loads with components that can be described by N_c bits, then:-

$$N_{mn} = 4N_S \times (N_c + 3N_d)$$

The specific GA implemented to accomplish this used stochastic binary tournament selection with $N_{RLC} = 3, N_t = N_c = 7$, $N_s = 2$ and $N_d = 1$ resulting in a 148b - chromosome. The optimized antenna gain and SWR can be seen in the figure on the right.

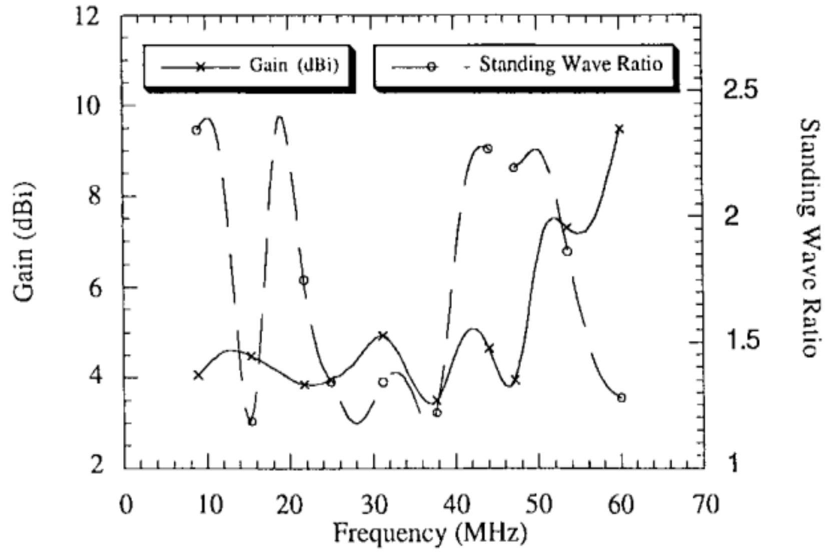


Fig. 3. The gain and SWR of the GA optimized antenna.

References

1. “Genetic Algorithms: A Tool for Image Segmentation” by Alaa Sheta, Malik S. Braik and Sultan Aljahdali, ISBN: 978-1-4673-1520-3
2. “Object Detection for Computer Vision Using a Robust Genetic Algorithm” by Tania Mezzadri Centeno, Heitor Silvério Lopes, Marcelo Kleber Felisberto, and Lúcia Valéria Ramos de Arruda
3. “A SURVEY OF GENETIC ALGORITHMS APPLICATIONS FOR IMAGE ENHANCEMENT AND SEGMENTATION” by Mantas Paulinas, Andrius Ušinsk, ISSN 1392 – 124X, 2007
4. “Feature Selection for Image Classification Based on a New Ranking Criterion” by Xuan Zhou, Jiajun Wang School of Electronic and Information Engineering, Soochow University, Suzhou, China
5. www.genetic-programming.com
6. https://en.wikipedia.org/wiki/Genetic_programming
7. “An Evolved Antenna for Deployment on Nasa’s ST5 Mission” by J. Lohn, G. Hornby, D. Linden, GTPT, May-2004
8. http://www.naun.org/main/UPress/cc/cc-12.pdf
9. https://blog.coast.ai/lets-evolve-a-neural-network-with-a-genetic-algorithm-code-included-8809bece164
10. https://lethain.com/genetic-algorithms-cool-name-damn-simple/