International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Genetic Optimization: A Survey

Kamini Viswakarma[1], Satish Dehariya[2]

[1, 2]Samrat Ashok Technological Institute, Vidisha, Madhya Pradesh., India-464001

Abstract: ***Genetic optimization for problems like classification, recognition, registration etc. related to different fields like military, astrophysics, medical science etc. is an emerging field. The same concept of genetic algorithm can be applied using different fitness function which will fit the given data model. Calculation of cost for each population completely depends upon the fitness function and it will change the way of learning and optimization of the problem. In this paper, we have gone through different types of methods those are implemented using Genetic Algorithm with various fitness function used for selecting the chromosome.***

Keywords: Genetic Algorithm, Image processing, Neural network, Fuzzy logic

1. Introduction

Genetic algorithms are inspired from the famous Darwin's theory about "Survival of Fittest". In biological aspect, crossover is done between parents to produce offspring. The offspring inherit the basic characters from the parents and nature automatically mutates some of the genes in child's chromosome to produce some new character apart from the parents. In natural biological reproduction process, from a generation i.e. the entire population, some parents are selected to reproduce the next generation. During, this process crossover and mutation occurs. [6] From the generated offspring, the fittest offspring survive to produce the next generation. The whole biological cycle is given in Figure no. 1.4.The genetic optimization process for optimizing any parameter for a certain mathematical model is derived from the above concept. To make a simulation in a computer system, we have to compare the natural process and the simulation to be carried out. [7]

In Image processing application, registering images is an important stage. Multiple images can be captured by different types of sensors, or same sensor can be used to take images at different span of time. After this, mathematical modelling techniques are implemented to correct the geometric errors namely translation, scaling and rotation of the input image to that of the reference image, so that these images can be used in various applications like change detection, image fusion etc.

The author used affine transformation method in order to create linear combinations. The equation used to create affine transformation is:

$$Z' = M.Z + X$$

The fitness function defined in order to get most suitable chromosome is:

Fitness, $$F = \frac{1}{S.L}\left(\sum_{i=1}^{S}\sum_{j=1}^{L}\frac{1}{m_{ij}}(z_i' - y_j')\right) \quad (1)$$

NOTE: The fitness function is very important stage in performing genetic based algorithms because the overall optimization using GA get affected due to the fitness equation used to perform further operations [17, 25]. It is author specific equation which works according to data set available with the researcher. If someone changes the fitness function equation then one will get different result than the existing one (either better result will be obtained or may be less efficient result). Focusing on this fitness function we are going to evaluate different genetic algorithm method implementation including this.

2. Genetic Algorithm

The author implemented GA to IR [14,15] using chromosome encoding, initializing, selecting, and reproduction principle of GA, finally terminating. We are focusing on fitness function used and chromosome selection method used throughout this survey paper. The author used 6 genes, reproducing translation (t1,t2), skewing(sk), squeezing(sq), rotation(r) and scaling(s) effects as shown below in figure 1

T1	T2	Sk	Sq	R	S

1. Chromosome encoding

Figure 1: Chromosome encoding

The objective of GA's is to find the optimal spatial transformation (chromosome) in a population of individuals.

2.1 Singhai and Singhai, 2012

Finding local maxima using direct search over matching the transformation is difficult due to which Rakhi et.al. applied Mutual Information(MI) to measure or calculate the statistical dependence of info redundancy between the image intensities of corresponding voxels in both floating image and reference image. They used Partial Volume Distribution Interpolation(PV) [7] in order to compute the MI criterion. Scope of the research is limited to pair of images, which are misaligned by rigid transformation (i.e. rotation and/or translation). Traditionally, simple hill climbing method is used to search for local maxima. But MI value function is monotonic, in that case, no local maxima. Hill climbing always works, due to presence of many local maxima, MI function values will mislead the process of fining global maxima. The author applied genetic algorithm to overcome this problem.

The representation of chromosome is drawn below:

Ci = [bi1,bi2,……..,bi32]

Where 12 bits represents the rotation angle, next 10 bits

represent X-translation and the last 10 bits represents Y-translation and the fitness value for the chromosome is given by

$$Fi=fi/sum\ fi \quad (2)$$

Where Fi is weighted fitness of a chromosome I and fi is the fitness of chromosome i.

In order to assign with probability to each chromosome I, using the Fi value, Roulette selection method is used.

In this study, image registration is done using genetic algorithm. Here, it is proposed that with help of direct search technique it is quite difficult to estimate the local maxima in the search space of matching transformation which can be automated with a suitable metric. In this paper, an effective search technique is proposed that is based on genetic algorithm which is used to register satellite imagery. The statistical dependence of registration parameter redundancy among corresponding voxel intensity in both master and slave images is measured by the application of mutual information in this method. For, chromosome selection procedure Roulette Wheel for chromosome selection method is adopted. In this proposed algorithm single point crossover method is used. In this testing process, satellite image of Bhopal city is taken and to create test data sets, these images are translated and rotated artificially. The image dimension taken in this study is 601x701 pixels. For GA optimization technique the translation parameters are varied from 0 to 100 pixels and the rotation angles are varied from 0 to 6^0. Image registration accuracy is computed and the timing performance is also analyzed. Two dataset are analyzed for GA. With the first dataset average RMSE found to be 0.043 with 200 generations which is an improved with comparison to hill climbing search method which is found to be 0.18. For second dataset, computed RMSE for hill climbing search is 5.93 which is improved to 1.65 with GA method. In case of timing performance, for hill climbing search it is 2.213 sec. whereas it is 1.9160 for GA approach. Hence, GA is found to be an efficient searching strategy in this experiment as proposed by the authors.

2.2 Valsecchi and Damas, 2012

The author introduced intensity-based image registration (IR) technique based on a modern, real-coded genetic algorithm. Instead of defining the translation component in terms of voxels, special units are considered. IR method is characterized into three main components: the transformation model, the similarity model and the optimization process.

The operators used in this method are common choices for real-coded GA: blend crossover (BLX-alpha)[] and random mutation [7]. The fitness value used for a solution f is simply the similarity between the two input images when registered using f, i.e.

$$f \rightarrow shi(Im,f(Is)) \quad (3)$$

where shi is similarity metric and considered both mutual information(MI) [7] and normalized correlation(NC). The parent selection is performed following the tournament approach. When images have different intensity distribution or multi-modal registration is required then MI is specially preferred.

In this study, an image registration technique using genetic algorithm is proposed. Real world MRI images are taken as test images in this experiment. Total of 16 scenarios are taken into consideration. In this method, translation parameters in x-axis and y-axis, rotation in x-axis, y-axis and z-axis and scaling factor are taken into consideration. These parameters are used to estimate the transformation model. Then, the registered image is formed. First, a chromosome is formed using these parameters. Here, a population is generated in a given range defined by the authors according to these parameters. Then, each chromosome is evaluated by a fitness function i.e. evaluating the image registration accuracy and most of the cases it is RMSE. Then the entire population is sorted and then the fittest chromosome is chosen among them i.e. the one with lowest RMSE. If the fittest one is not up to the mark for an optimal registration then the iteration continues to create next generation. For that some best chromosomes are chosen as parent and crossover is done among them to generate another population. Thus generation after generation continues until the fittest chromosome is chosen among them i.e. the one with lowest RMSE can create an optimal registration parameter set. Using this, optimal parameter set, a transformation model is generated to do the final image registration. Finally, the result is compared with the classical intensity based registration to check the difference. It is found that the registration using genetic algorithm based optimization is more accurate than that of the classical way of intensity based image registration.

2.3 Junli Lu *et al.*, 2011

Literature [9] explains the improvement of IFKNN and introduced the sample class in NFKNN. In this algorithm, sample class is pre-processed to get sample-edge-of-class. This led to better classification and improved execution time. This can be well understood by underlying flowchart:

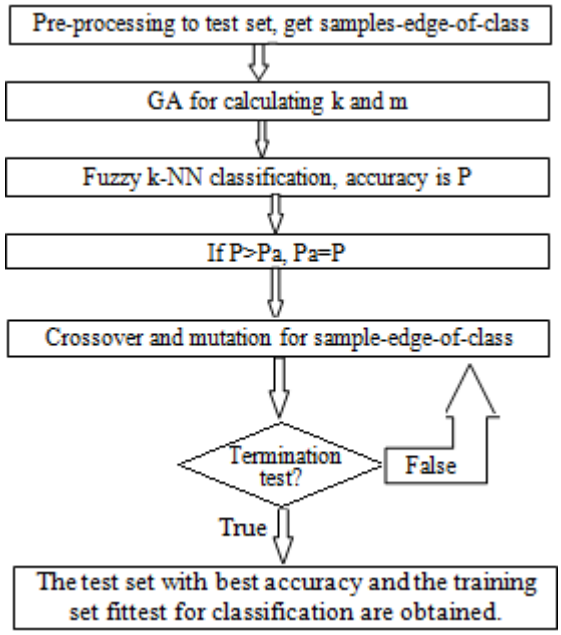

Figure 2: GA Flow Chart

The drawback still exists in pre-process stage as like the IFKNN in connection with execution speed. Pre-process is time consuming but once done rest of the NFKNN operation perform far better than IFKNN.

In this research work, using genetic algorithm, a new fuzzy k-NN classification techniques is developed and used for classification of random data and real data as well. It need less time for calculation and have a higher accuracy rate in comparison to standard classification process. It is based on two ideas. The first idea explains on optimizing selection rate 'k' and mutation rate 'm'. Before optimizing k and m, the training set is categorized into two sets as one named as "sample inside class" and the other as "sample edge of class". The first set will not participate in the GA algorithm flow, the second set will. This saves time for the entire process as we don't have to process the data which is already in the class. In the second idea, it has been proposed that the classification rule already been set by the test dataset may not fit for the test dataset. Hence, for that confusion, the test data set will undergo the preprocessing of the algorithm. By this operation, the original classification rule will not change as they will be already inside the sample edge of class category. The real dataset comprises of 3902 record or objects and each having 8 attributes. There are 14 classes in the taken dataset. The accuracy of the proposed method is 84.27% accurate in the real dataset and when there is only two classes, the accuracy of classification is 91.2%.

2.4 Selma Ayse Ozel, 2011

The fitness of chromosome in this literature is determined by using predictive accuracy [11] which is defined as:

$$Predictive\ Accuracy = TP\_rate \times TN\_rat \quad (4)$$

where TP_rate is the true positive rate and TN_rate is the true negative rate. The main objective is to obtain maximum predictive accuracy. Following this method the fitness of the chromosome is calculated using the cosine similarities [11] of each Web page's chromosome in the training dataset. The cosine similarity for chromosome vector c and a document vector d is computed by:

$$sim(c,d) = \frac{\sum_{i=1}^{m}\sum_{=1}^{N_i} c_{ij}\ x\ d_{ij}}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{N_i} {c_{ij}}^2}\ x\ \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{N_i} {d_{ij}}^2}} \quad (5)$$

where m equals 6 since we use the terms in the title, header, anchor, emphasize, paragraph, and list item tags in the features set, and Ni is the number of terms belongs to tag i in the features set. Since cij and dij are in the range [0,1], sim(c,d) takes values in the range [0,1] and shows the degree of similarity between a chromosome c and a document d.

For reproduction purpose roulette wheel is divided into pop_size slots, each proportional to individual's fitness, is spun, and an individual is selected. Followed by this method furthe GA concept is implemented focusing mainly over feature extraction. In this experiment, an optimal feature selector is designed for the webpages using genetic algorithm techniques. Webpage classification is a problem where the dataset is the webpage itself. The HTML tags of the webpages are objects. The main problem is to select the HTML tags and push them into different classes. In this study, A GA based automatic feature selector is designed which will select the best possible feature for an unknown webpages. The whole research work consists of 3 steps. The first step is to extract the features. Then in the next step, a GA feature selector comes to play. Its work is to give the features best possible weights. After that, in the third step, Ga based classification is done to set the classification rule for the corresponding web page features and the weights assigned to them. In this research work a total of 5 classes is taken into consideration. The classification accuracy of this method is 96%. As there are limited no. of HTML tags are taken into consideration, the authors have recommended that, the proposed methodology can be applicable for other HTML tags and URLs.

2.5 Zhang Lianmei & Jiang Xingjun, 2010

They proposed a combination classification method in which the output of multiple decision trees are obtained by related probability then the result is parallel combined [14]. After obtaining the desired output genetic algorithm is being applied for further optimization of connected weight matrix in combination algorithm. Using more than one decision tree has led to more accurate classification and sustaining result in comparison to using single decision tree result. Precisely we can say that, it is a hybrid genetic algorithm method which includes policy making before applying genetic algorithm.

In data management policy making tree (decision tree) is one of the classified method comprising pitch point expressing an attribute test and test out is represented by branches. But the leaf node represents the kind of distribution. Further the policy making tree is transformed into classified rules. Policy making rule is well explained in[]. With the creation of policy making tree, now comes the hybrid algorithm in which we focus mainly on classified function, success match, sufficiency function and cumulative probability computation method.

Classification function is a formula that uses the sample for classification. After classification, success match equation is used for successfully matching the sample with the given value. The sufficiency function is used to perform roulette trial-and-error method in order to get individual sufficiency. The sufficiency function implementation is followed by cumulative probability computation method. Following this the concept of Genetic Algorithm is used for optimization using simulation analysis.

In this research work, using hybrid genetic algorithm, a new datamining method is proposed. This method is a hybrid form of multiple decision trees and genetic algorithm. Instead of using single decision tree for the classification problem in a standard genetic algorithm, the researchers used a multiple decision tree. Optimization of the connecting weights in the network is done by using genetic algorithm. A parallel combination of multiple decision trees is done and each decision tree adopts probability measurement level output method. The research work uses two groups of dataset given

by . The first dataset has been classified to the accuracy of 76.00% and the second group is classified to the accuracy of 63.86%. Moreover, it is concluded that, the proposed method optimizes the classification rules and helps in improving classification results.

2.6 Ingole *et al.*, 2009

In this study, image registrations of medical images are done using genetic algorithm. Using genetic algorithm based optimization technique; a mathematical model is created to optimize geometric errors present in the image like translation, rotation and scaling [15]. So that, these corrected and registered images can be used in change detection and image fusion like applications. The objective of this paper is to create a mapping function which can map to set of matched points. The entire method is divided into three basic steps like selection, reproduction and termination of the optimization process. Here, a chromosome vector is defined by six genes that is translation in x-axis and y-axis, skewing, squeezing, rotation and scaling effect. In the selection process, while each successive generation is generated, each population goes through a fitness function as to be evaluated. For the fitness function in this proposed method, the point matching is done by nearest neighbor method and the transformation model is applied on the basis of affine transformation. Then, a portion of the entire population is selected to breed the next generation. For, this procedure always the fittest chromosomes are selected as parent. In reproduction section, the selected parents breed a whole new generation and here mutation is applied to some portion of the entire population to change their properties that is inherited from the chosen parents. For termination process, there are two major conditions for this method. First one is, if the solution has a fittest chromosome which can satisfy the minimum criteria defined by the user and can give highest registration accuracy. Another stopping criterion is as if the maximum no. of generation is created that is the no. of iteration defined by the user. The best fitness value is found to be 1.0742 and mean is found to be 1.8143. The test images are brain MRI images.

3. Result Analysis

Table 1: GA Methodologies

S.No	GA method used	Methodology
1.	GA-IR	MI
2.	GA-PV	MI & PV
3.	GA-IR	MI, MI-Re, NC and NC-Re

According to the above table, first method's implementation is good but the fitness function can be used more efficiently in order to get better result. Second method is an extension to the first method which along with MI method PV method is also implemented. Besides that, second method also implements roulette wheel method to find the best chromosome for implementing final GA method. The third method is the best among all others. It implements GA using IR method and concept of MI and NC is used along with restart mechanism. Restart mechanism helps in lowering the probability of one "bad" run. This helps in reducing the rate of error in the obtained after implementing the concept. Generally, one pass GA gives better result in comparison to traditional image registration techniques but restart mechanism lowers the chance of error and also improves the efficiency.

4. Conclusion

After analyzing the results obtained by the above mentioned algorithm in GA implementation the last GA i.e. GA-IR using MI, MI-Re, NC, and NC-Re method is the best method. The result varies by magnitude of 1 over time, in other words the rate of error is decreased to minimum and result produced is optimum at any given time. Genetic algorithm can optimize the classification, rule discovery, effective registration parameters, multiple decision trees or as a feature selector for the image or webpage classification problem. As almost every author has recommended a new idea of classification for new type of data.

References

[1] Dayanand Savakar, "Identification and Classification of Bulk Fruits Images using Artificial Neural Networks", International Journal of Engineering and Innovative Technology (IJEIT) Volume 1, Issue 3, March 2012
[2] Le Hoang Thai, Tran Son Hai & Nguyen Thanh Thuy, "Image Classification using Support Vector Machine and Artificial Neural Network", I.J. Information Technology and Computer Science, 2012
[3] M. Umadevi & S. Srinivasulu, "Parameters for Tsunami Classification Model using BPN Based Artificial Neural Network", European Journal of Scientific Research, ISSN 1450-216X Vol.76 No.4, pp.633-647, 2012
[4] S. Daniel Madan Raja & A. Shanmugam, "ANN and SVM Based War Scene Classification Using Invariant Moments and GLCM Features: A Comparative Study", International Journal of Machine Learning and Computing, Vol. 2, No. 6, December 2012
[5] S. R. Patil & S. R. Suralkar, "Fingerprint Classification using Artificial Neural Network", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 2, Issue 10, September 2012
[6] Singhai Rakesh & Singhai Jyoti, "Registration of Satellite Imagery Using Genetic Algorithm", Proceedings of the World Congress on Engineering 2012 Vol II WCE, July 4 - 6, 2012
[7] Valsecchi Andrea & Damas Sergio, "An Image Registration Approach using GeneticAlgorithms", WCCI 2012 IEEE World Congress on Computational Intelligence, 10-15, 2012
[8] Haoyu Ma, Bin Hu, Mike Jackson, Jingzhi Yan & Wen Zhao, "A Hybrid Classification Method using Artificial Neural Network Based Decision Tree for Automatic Sleep Scoring", World Academy of Science, Engineering and Technology, 2011
[9] Junli Lu, Guang Zhao, Cheng Yang & Junjia Lu, "New Fuzzy k-NN Classification by Using Genetic Algorithm", Seventh International Conference on Natural Computation, 2011

[10] K. Usha Rani, "Analysis Of Heart Diseases Dataset Using Neural Network Approach", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.1, No.5, September 2011
[11]Selma Ayşe Özel, "A Genetic Algorithm Based Optimal Feature Selection for Web Page Classification", IEEE, 2011
[12]Steve Diersena, En-Jui Leeb, Diana Spearsc, Po Chenb & Liqiang Wanga, "Classification of Seismic Windows Using Artificial Neural Networks", Science Direct, 2011
[13]RC Chakraborty, "Fundamentals of Genetic Algorithms", AI Course Lecture 39-40, 2010
[14]Zhang Lianmei & Jiang Xingjun, "Research on New Data Mining Method Based on Hybrid Genetic Algorithm", International Symposium on Computer, Communication, Control and Automation, 2010
[15]Ingole V. T., Deshmukh C. N., Joshi Anjali & Shete Deepak, "Medical Image Registration Using Genetic Algorithm", Second International Conference on Emerging Trends in Engineering and Technology, ICETET, 2009
[16]Ji Peirong, Hu Xinyu & Zhao Qing, "A New Genetic Algorithm for Optimization", IEEE, 2008
[17]Xian-Jun Shi & Hong Lei, "A Genetic algorithm-Based Approach for Classification Rule Discovery", International Conference on Information Management, Innovation Management and Industrial Engineering, 2008
[18]Insung Jung & Gi-Nam Wang, "Pattern Classification of Back-Propagation Algorithm Using Exclusive Connecting Network", World Academy of Science, Engineering and Technology, 2007
[19]Abdulhamit Subasia & Ergun Ercelebi, "Classification of EEG signals using neural network and logistic regression", Computer Methods and Programs in Biomedicine, 2005
[20]E. Hosseini Aria, J. Amini & M.R.Saradjian, "Back Propagation Neural Network for Classification of IRS-1D Satellite Images", 2003
[21]N Kannathal, U Rajendra Acharya, Choo Min Lim, PK Sadasivan & SM Krishnan, "Classification of cardiac patient states using artificial neural networks", Exp Clin Cardiol Vol 8 No 4 2003
[22]Guoqiang Peter Zhang, "Neural Networks for Classification: A Survey", IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 30, No. 4, November 2000
[23]Howard E. Michel & A. A. S. Awwal, "Enhanced Artificial Neural Networks Using Complex Numbers", 2000
[24]Gisbert Schneider & Paul Wredeb, "Artifcial neural networks for computer-based molecular design", Progress in Biophysics & Molecular Biology, 70,175-222, 1998
[25]Philipp Koehn, "Combining Genetic Algorithms and Neural Networks: The Encoding Problem", December 1994
[26]A. Pomerleau, Jay Gowdy & Charles E. Thorpe, "Combining Artificial Neural Networks and Symbolic Processing for Autonomous Robot Guidance", Engng App/ic. ArliJ. Inrell. Vol. 4. No. 4, pp, 279-285, 1991
[27]W. McCulloch and W. Pitts, "A Logical Calculus of the IdeasImmanent in Nervous Activity," Bulletin of Mathematical Biophysics,vol.5, pp. 115-133, 1943.

Author Profile

Kamini Vishwakarma received the B.E. degree in Computer Science and Engineering from Technocrats Institute of Technology, Bhopal in 2015 and pursuing M.Tech. degree in Computer Science and Engineering from Samrat Ashok Technological Institute, Vidisha. Implementation of E-business pattern using object oriented approach has been her major project topic during her bachelor degree. Optimization using genetic algorithm has been the broad research area in her master's.

Satish Dehariya received the B.E. degree in Information Technology from Samrat Ashok Technological Institute, Vidisha in 2007 and M.Tech. degree in Computer Science and Engineering from Barkatullah University Institute of Technology, Bhopal. He has been working in Samrat Ashok Technological Institute, Vidisha as teacher and Cloud computing has been the research topic since his master's. He holds ten years of experience in the various fields of data mining techniques.