Design of Digital IIR Filter using Crow Search Optimization Algorithm for Cancer Cell Detection

[1]Ravjot Kaur , [2]Ranjit Kaur
[1]PG Student [ECE], Dept. of ECE, Punjabi University, Patiala, Punjab, India
[2]Professor, Dept. of ECE, Punjabi University, Patiala, Punjab, India

Abstract: In this paper a digital Infinite impulse response (IIR) digital filter is proposed which is designed using crow search optimization algorithm (CSA). CSA is population based metaheuristic algorithm which works on the intelligent activities of crows. CSA is based on the inspiration that crows hide their excess food in inaccessible places and get back it when needed. The intent of this work is to optimize coefficients of filter using CSA so that magnitude response error and passband and stopband ripples are minimized. The designed filter is further applied for recognition of cancer cells using Genomic signal processing (GSP) which is the study of genes. The ratio of mean amplitude and frequency is used to develop methodology for recognition of cancer cell from Homo sapiens DNA. Ratio of cancer cell differentiates from normal cell. This research will help in prediction of cancer in pre-stage and will further aid in the treatment.

Keywords: Infinite Impulse Response (IIR) Filter, Crow Search Algorithm (CSA), Magnitude Response, Optimization, Genomic Signal Processing (GSP), Cancer Cell, DNA.

I. INTRODUCTION

Digital filters are extensively used in field of signal processing. Filters are used to take out useful element from a signal or to eliminate unwanted part from the signal. Digital filters are classified as IIR and FIR(Finite impulse response) filters. IIRs are preferred over FIR because of the high selectivity of IIR at low filter order. IIR filters also have greater computational efficiency and reduced delay in system. IIR filters achieve desired magnitude response with a much lower filter order than FIR. IIR filters can be designed using two approaches: optimization approach and transformation approach. Transformation approach includes designing of analog filter and then conversion of analog filters to digital filters using transformation method. This technique is not very efficient in terms of phase response, coefficient quantization error and filter structure. On the other hand, filters designed with optimization approach easily stick in the local minima, this is because of the non-linear and multimodal nature of error surface. Optimization approach includes performance measurement of designed filter. Therefore, researchers have used design techniques which depends on modern heuristic optimization algorithms.[1].

The major hindrance for filter design is to obtain a low filter order so that specified magnitude and phase response can be attained. In most of the previous research work, conventional optimization methods have been used, which leads to local minima solution of the problem. Every design may not require local minima and may require global optimum. In that case, metaheuristic algorithms show promising results for solving multimodal and non-linear error surface problems. In most of the cases, optimal solution can be obtained with the help of metaheuristic algorithms. Nature inspired algorithms are metaheuristic algorithms. Conventionally, IIR filters are designed with the help of different nature inspired algorithms. In IIR filter design using real coded genetic algorithm(RCGA)[1] multiple conflicting objectives approach is used for designing. The main focus was to optimize the performance in stop band, transition band and pass band so that magnitude and phase response can be minimized. The weighted sum approach is used to find value of weights and to obtain filter coefficients various weight assignment to every population element gives numerous pareto-optimal solutions in single simulation iteration.[14]. It gives better performance than pre-existing techniques. In teacher-learning based optimization (TLBO) algorithm[2] the natural phenomenon of teacher and learning among teachers and learners is considered. This research showed that TLBO is suitable for designing of higher order filters. It introduced enhancement in the original algorithm to increase the exploration. The designed filter is compared to different algorithms to show its effectiveness in terms of performance and quality. Particle swarm optimization (PSO)[3,13] is remolded using concept of constriction factor and inertia weight approach(PSO-CFIWA)[3]. A multi-modal optimization problem is considered for design of 8^{th} order filter. The suggested technique finds the best coefficients to match the ideal phase response. The results showed that the recommended method is more efficient for sub-optimal problems. Adaptive IIR filter [4] to real world signal processing problems can be used along with PSO for designing of filter in noisy realistic time domain nature. Easy implementation and good convergence rate of PSO has given global minima for desired response.

The use of digital filters for detection of cancer cell is very common. Researchers all over the world use different optimization methods for better recognition of effected cells. The designed filter using different optimization algorithms, as discussed in previous section, is used to detect the cancer cell along with other DSP techniques like Discrete Fourier transform (DFT) and Power spectral density (PSD). GSP deals with the study of genes for detection of abnormality in cells using DSP. The DNA samples are used to detect the variation in genes. Different DSP techniques and filters have been used over the past time for prediction of cancer cell from DNA samples. Morlet wavelets [5] are also used to expose the active sites in wavelet transforms. Oncogene proteins are investigated to identify their function and structure. Amplitude and phase factors of morlet wavelet are modified to attain optimization. The modified wavelet is then used to transform HIV, oncogenes and hemoglobin in blood cells. Probabilistic Boolean and Boolean networks [6] are studied for genetic regulatory. Dynamic nature of probabilistic Boolean networks is identified using markov chains. Boolean network's basin of attraction is analogous to probabilistic Boolean network in which corresponding class of communication is described. Exon and genes are explored using time domain algorithms [7]. Boss indicator sequences act as input to algorithms. These sequences are obtained from DNA sequences and then passed through resonant second filter. Average magnitude difference function is speech processing. Periodicity and pitch is detected using periodogram algorithm in time domain. Nano technology is also playing a leading role in cancer research [8]. In this GSP is considered to focus the hardware of nano tools. Spatial and temporal information is provided by dynamic cellular process which is applied using nanotechnology. GSP is appropriate for nano scale solutions. Link between proteins and genes within the cell and real world is provided by nanotechnology which helps in pre-mature detection of cancer cell. To achieve this different nano tools are used to measure fluid flow, energy metabolism and transportation of components. Tumor location is flooded with magnetic nano devices, to release drug load upon heating. This drug load destroys the cancer cell. Exon and gene are determined in eukaryotes using paired spectral content measure[9]. This method requires data to be driven from frequency domain. It requires training from DNA samples that detect axon and genes. DNA is converted into numeric sequence signal using numeric representation. After this, to and fro DSPs are coupled on same DNA sequences.

The intent of this paper is to design an IIR Low Pass (LP) filter using CSA and then to use the designed filter for detection of cancer cells. CSA is considered for designing of filter as it is very easy to implement with only two controlling parameters. This makes the convergence faster and also the designing much more easier. The filter coefficients are optimized using CSA to get the desired magnitude response. The designed filter is subject to different DNA samples to identify the infected cells. To distinguish between cancer and normal cells ratio of mean frequency and magnitude is calculated. The paper is structured as follows: Section 2 describes the design of LP filter using CSA. Technique used to detect cancer cell is explained in section 3. section 4 describes the detailed methodology and its analysis. In section 5,results are discussed and compared with butterworth filter. Finally, conclusions and future scope are described in section 6.

II. DESIGN OF IIR FILTER

In this section design of low pass IIR filter is explained. Transfer function of IIR filter is given as:

$$H(z) = \frac{\sum_{k=0}^{M} a_k z^{-k}}{1+\sum_{k=1}^{N} b_k z^{-k}} \qquad (1)$$

The first and second order filter sections are cascaded to obtain a third order filter. The equation used for designing of third order filter is:

$$H(Z) = A\left(\prod_{u=1}^{K} \frac{1 + a_{2u}Z^{-1}}{1 + a_{2u+1}Z^{-1}}\right)\left(\prod_{v=1}^{L} \frac{1 + a_{4v+2K-2}Z^{-1} + a_{4v+2K-1}Z^{-2}}{1 + h_{4v+2K}Z^{-1} + h_{4v+2K+1}Z^{-2}}\right) \qquad (2)$$

$X = [A, a_2,..., a_{2K+4L+1}]\ N_{TX1}$ is a vector decision variable of dimension $T \times 1$, with $T = 2K + 4L + 1$. A represents the gain and $[a_2, a_3,... a_{2K+4L+1}]$ denotes the filter coefficients of the first- and second-order sections.

Minimization of magnitude response error is the main requirement of the design. L_1 norm approximation error for magnitude response is to be minimized. $E_P(X)$ denotes the L_P norm approximation error of magnitude response, which is denoted as :

$$E_p(X) = \{ \sum_{i=0}^{K} \left| H_d(w_i) - |H(w_i, x)| \right|^p \}^{1/p} \qquad (3)$$

For L_1 norm approximation error p is taken as 1 and ideal response equals to 1, the $E_1(X)$ is:

$$E_1(X) = 1 - H(Z) \qquad (4)$$

According to the requirement of design criteria, the objective function to be achieved is given by:

$$Minimize\ O(X) = E_1(X) \qquad (5)$$

The best position in terms of the objective function value gives the solution to the optimization problem i.e. best filter coefficients to attain minimum magnitude response error.

III. DSP TECHNIQUE

Every signal cannot be discovered in time domain. Some signals belongs to frequency domain which make use of DFT. Genomic sequences in DNA are converted into numerical sequences using several mapping techniques.[11]. One of the mapping technique depends upon the hydrogen bonds among nucleotide bases. For each nucleotide one binary sequence is obtained, the generated sequence depends upon the strong and weak bonds within nucleotides bases.

A,C,T and G are four bases of nucleotides. $x[n]$ defines a N length DNA sequence with alphabets and $x_i[n]$ defines a single binary sequence for DNA .

If x[n]=A T T A G C T C G A, then xi[n]= 0 0 0 0 1 1 0 1 1 0

Xi[k] gives DFT of binary sequence by equation:

$$x_i[k] = \sum_{n=0}^{N-1} x_i[n] e^{-\frac{j2\pi nk}{N}} \quad (6)$$

Where $k=0,1,2,3,4........,N\text{-}2, N\text{-}1$ and $n=0,1,2,3,........,N\text{-}1$.

PSD of sequence is :

$$P_D[k] = \sum_{k=0}^{\infty} |Xi[k]|^2 \quad (7)$$

For indication of cancer cell or normal cell $P_D[k]$ is plotted in coding region. The plot describes the spectral characteristics of the cells. The designed filter using CSA is then used to further improve the spectral characteristics of the cells. The variations in spectral peaks distinguish between cancer and non cancer cells.

IV. METHODOLOGY USING CSA

The main aspect of the work is to design an IIR filter using CSA and to obtain filtered PSD of DNA sequence. CSA works on the principle that crow flocks have several similarities with optimization problem. Crows hide their excess food in secret places and regain it when required. Other crows in the flock try to steal food from one another. They chase each other to do pilfering. But doing thievery from a hidden place of a crow is not easy task as if the crow finds out that another crow is chasing it, the crow will try to trick the other one by going to some other place in environmental space. If we consider this scenario as an optimization problem, crows act as searchers which corresponds to population in LP IIR filter, the environmental space is search space, each location in environment space is equivalent to a possible solution, the quality of food is fitness function and the best food source of the environmental space is the global solution of the problem which gives best filter coefficients.[10].

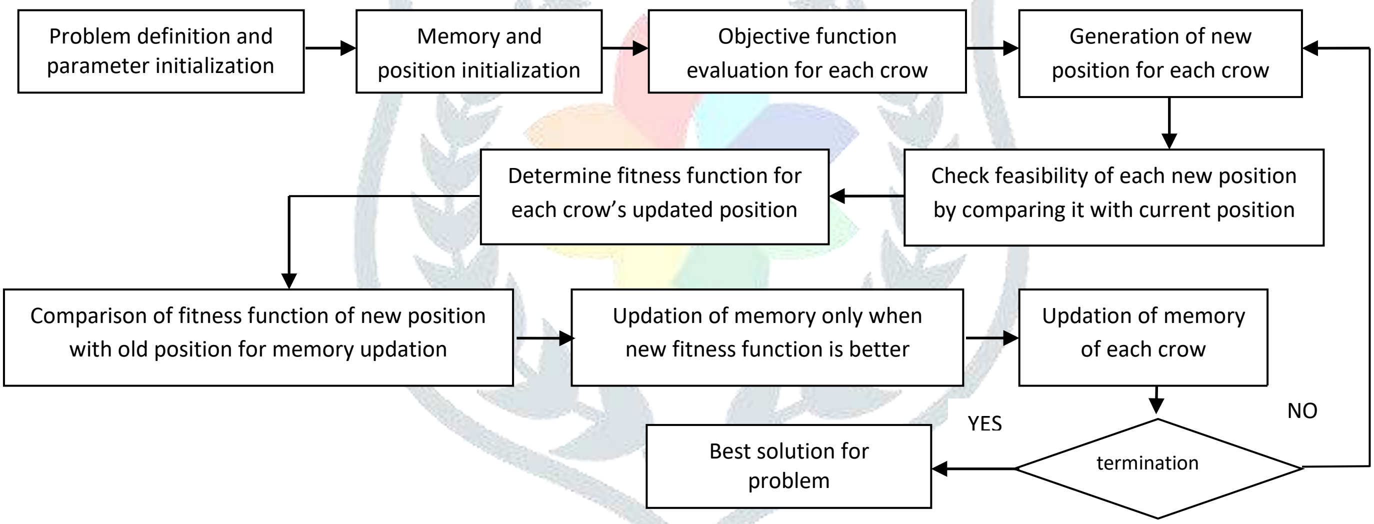

Fig. 1: *Flow chart of Crow Search Algorithm*[10]

The steps involved in the proposed work are discussed as follows:

Step 1: Design a third order LP IIR filter using CSA.

Step 2: Input the different base pair length DNA coding samples and assign single binary sequence for each alphabet base on the basis of weak and strong bonds.

Step 3: DFT of obtained sequence is calculated using eq.(6).

Step 4: PSD is calculated using eq.(7).

Step 5: Designed LPF is used to obtain the filtered PSD of sequence.

Step 6: Mean amplitude and mean frequency of filtered PSD is obtained.

Step 7: Ratio of mean amplitude and mean frequency is calculated.

Step 8: Ratio of normal cell and cancer cell is compared.

V. RESULTS AND DISCUSSION

Optimization algorithm along with DFT power spectrum gives effective and precise results. Different DNA samples are taken for test. Figure 1 shows the response graph and eq.(8) shows the obtained magnitude response of designed LP IIR filter using CSA. The designed filter have transfer function:

$$H_{LP}(z) = 0.2032 \times \frac{(z + 0.2184)(z^2 - 0.9519z + 0.9593)}{(z - 0.4082)(z^2 - 1.2261z + 0.6491)} \quad (8)$$

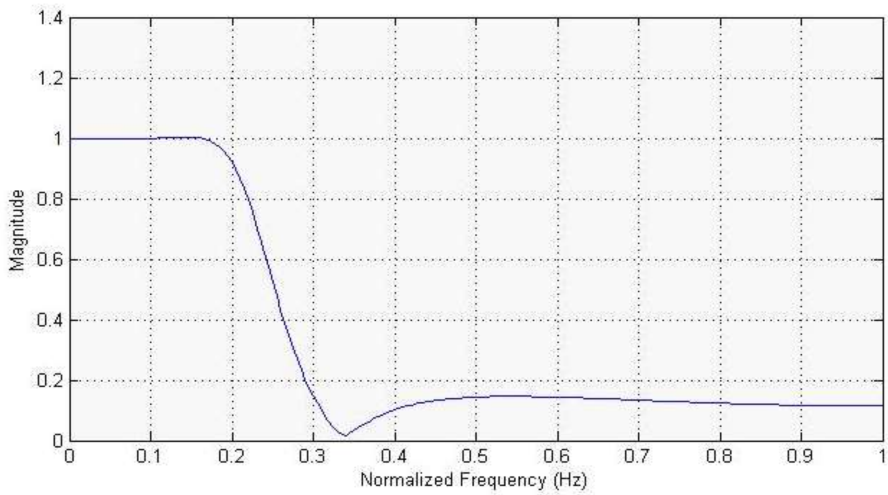

Fig 2: *Magnitude Response of 3rd order Low pass IIR digital filter using CSA*

The designed filter has minimum passband magnitude response, *H(Z)*= 0.920174.

Ratio of mean amplitude and frequency distinguish between cancer cell and normal cells. It is seen that samples with ratio greater than 1 belongs to normal cell whereas others belong to cancer cells. High accuracy is achieved in low order filter as compared to higher order filter.CSA provides better performance in designing of filters with less controllable parameters and fast convergence rate. It is noticed that normal cells are less spiky as compared to cancer cells. Figure 3,4,5,6,7 and 8 clearly depicts the filtered PSD to distinguish among cells. Different DNA samples over which methodology is applied are listed as follows:

Table 1. *DNA samples for cancer and normal cells*

Homosapien cell	Accession number	DNA Length
Cancer cell	NM_178439	4326
Cancer cell	NM_033085	1071
Cancer cell	NM_001318101	2076
Normal cell	NM_032464	2199
Normal cell	NM_001048265	731
Normal cell	NM_144654	685

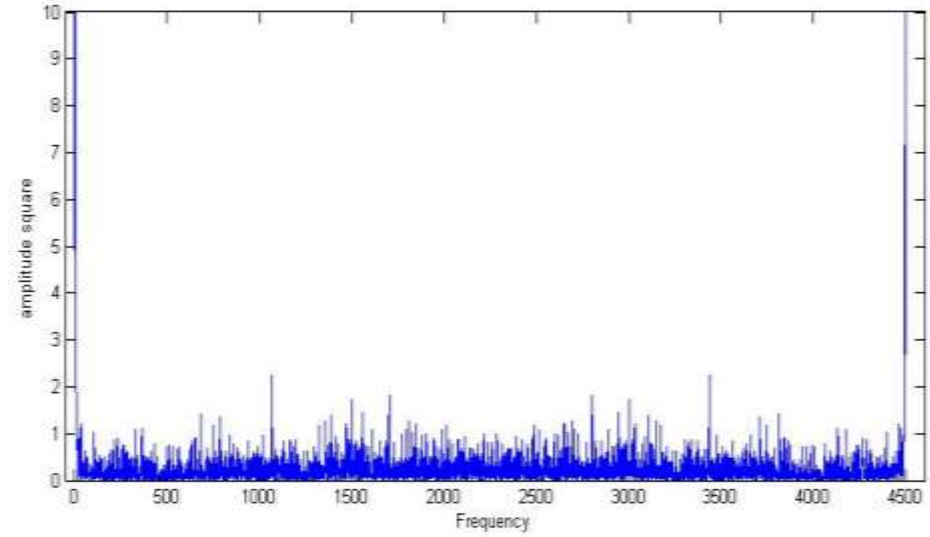

Fig 3.*PSD of cancer cell (NM_178439) and length of Exon 4326 bp with the use of CSA algorithm*

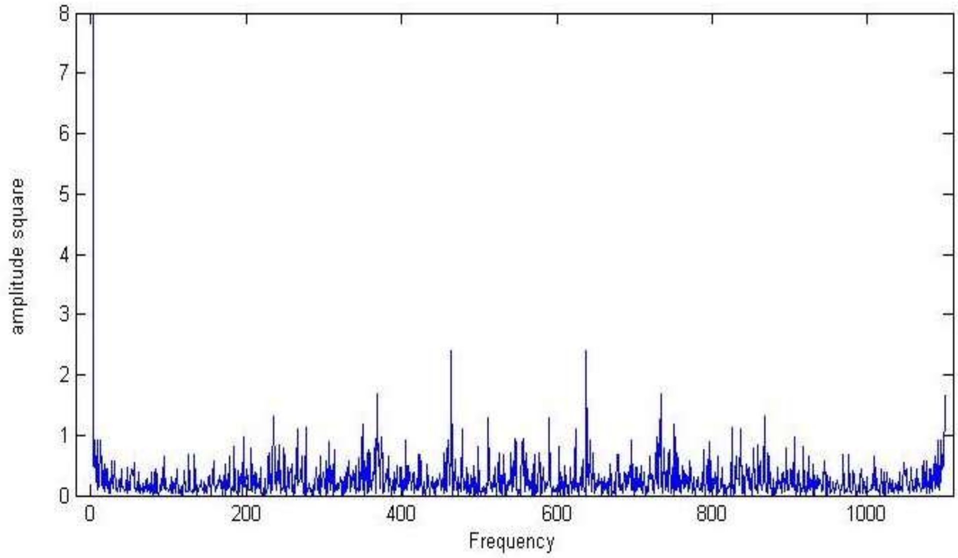

Fig 3 *.PSD of cancer cell (NM_033085) and length of Exon 1071 bp with the use of CSA algorithm*

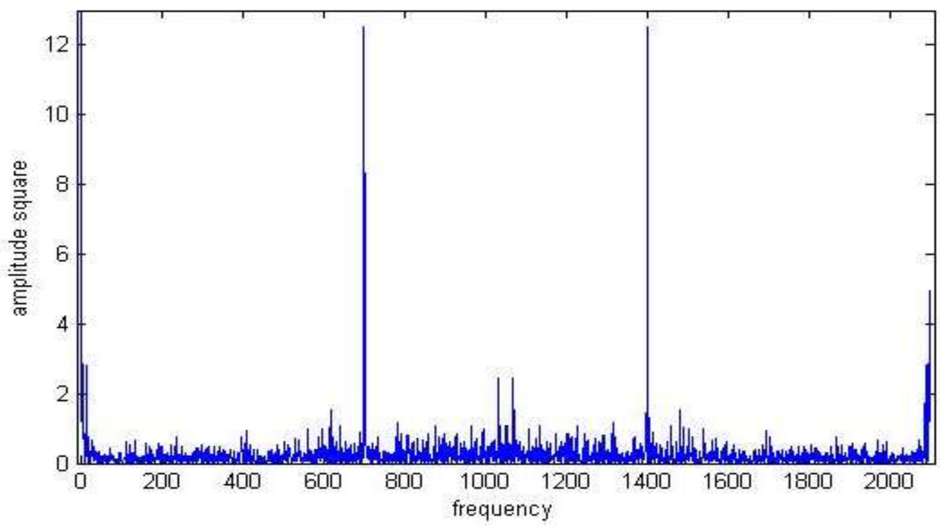

Fig 5 *.PSD of cancer cell (NM_001318101) and length of Exon 2076 bp with the use of CSA algorith*

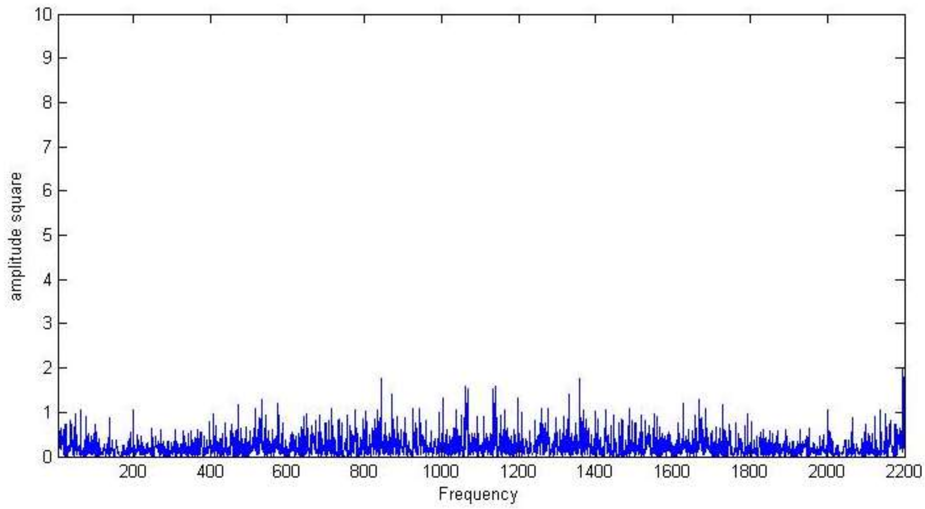

Fig 6 *.PSD of Normal cell (NM_032464) and length of Exon 2199 bp with the use of CSA algorithm*

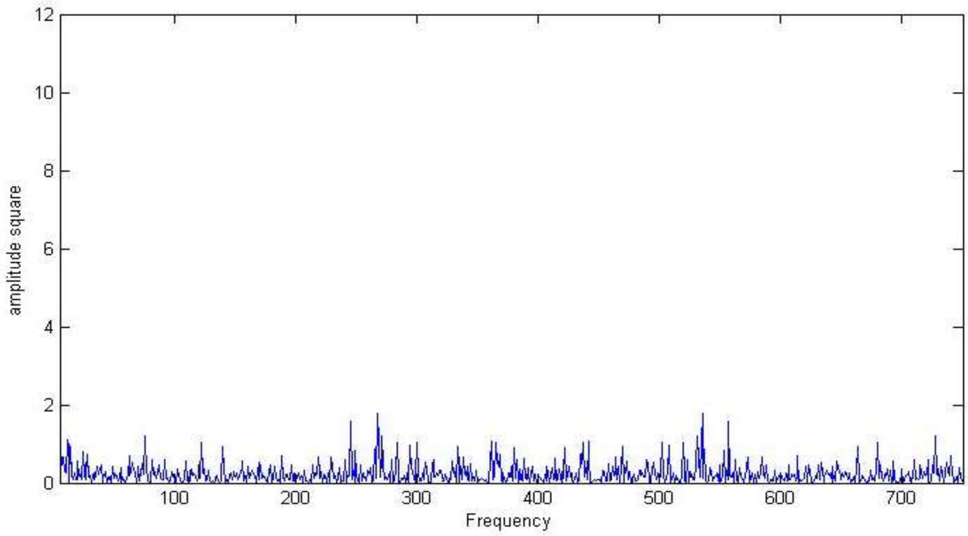

Fig 7 *.PSD of Normal cell (NM_001048265) and length of Exon 731 bp with the use of CSA algorithm*

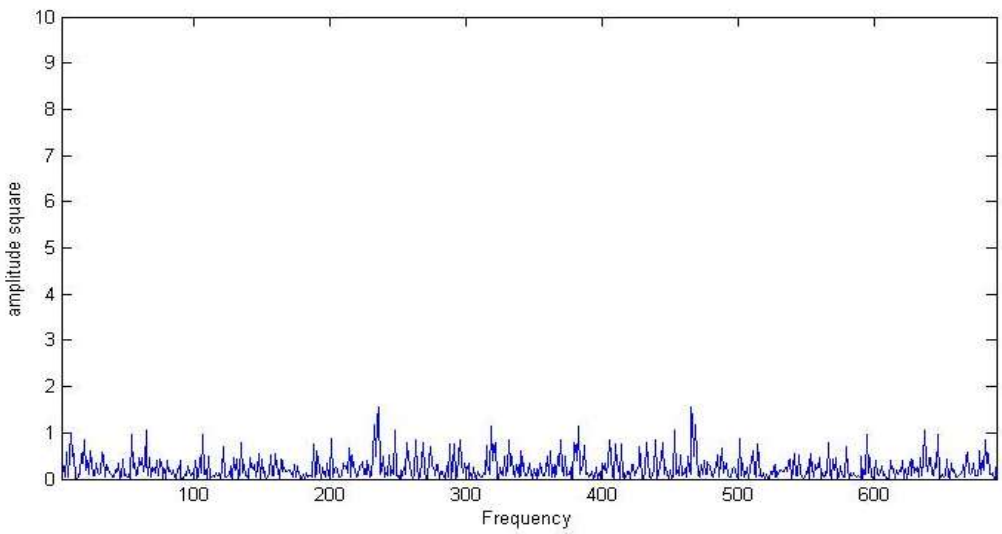

Fig 8 *.PSD of Normal cell (NM_144654) and length of Exon 685 bp with the use of CSA algorithm*

Table 2.*Comparison between Butterworth 12th order and CSA 3rd order LP Filter*.

Homosapien cells	Accession number of DNA sample	Ratio of mean amplitude and mean frequency using IIR Butterworth 12th order filter	Ratio of mean amplitude and mean frequency using CSA with 3rd order filter
Cancer Cells	NM_178439	0.5059	0.5193
	NM_033085	0.7521	0.7541
	NM_001318101	0.8217	0.8267
Normal Cells	NM_032464	1.0986	1.0992
	NM_001048265	1.1800	1.2157
	NM_144654	1.1181	1.1302

VI. CONCLUSION

In this paper a metaheuristic algorithm CSA is applied for designing of LP IIR filter. The designed filter is further used for cancer cell detection. The designed filter gives better performance in terms of reduced L_1 approximation error and provides greater convergence. The proposed technique is tested for various samples of DNA. The results revealed that the proposed algorithm along with GSP is an easy implementation for cancer cell detection. Filter designed with CSA improved the accuracy of cell detection by improving mean amplitude and frequency ratio. The ratio comparison reveal that low order filters are better in decreasing cost and power consumption. The proposed method can further be used for designing of FIR filters to improve efficiency. For further advancement in proposed method, it can be extended to other oncogenes and can be applied for detection of other diseases.

REFERENCES

[1]. Ranjit Kaur, Manjeet Singh Patterh and J.S. Dhillon, "Design of Optimal L1 Stable IIR Digital Filter using Real Coded Genetic Algorithm," *IAENG International Journal of Computer Science,* vol. 39 Issue 4, pp.329-338,Nov,2012.

[2]. Damanpreet Singh and J.S. Dhillon, "Fuzzy based design of digital IIR filter using ETLBO," *Turk J Elec Eng & Comp Sci*,,vol.24 , pp.4042-4062,2016.

[3]. Suman Kr. Saha, Rajib Kar, Durbadal Mandal and S. P. Ghoshal "Digital Stable IIR Low Pass Filter Optimization using PSO-CFIWA," *IEEE1st Int'l Conf. on Recent Advances in Information Technology,*2012 .

[4]. Sheng Chen and Bing L. Luk ,"Digital IIR filter design using particle swarm optimization," *International journal of Modelling, Identification and Control,* vol. 9, 2010.

[5] E.Pirogova,Q.Fang,M.Akay,andI.Cosic,"Investigationofthestructural and functional relationships of oncogene proteins," Proc. IEEE, vol. 90, no. 12, pp. 1859–1867, Dec. 2002.

[6] I. Shmulevich, E.R. Dougherty, and W. Zhang, "From boolean to probabilistic boolean networks as models of genetic regulatory networks," *Proc. IEEE,* vol. 90, no. 11, pp. 1778–1792, 2002.

[7] E. Ambikairajah, J. Epps, and M. Akhtar, "Gene and exon prediction using time-domain algorithms," *in Proc. IEEE 8th Int. Symp. Signal Processing and its Applications*, pp. 199–202,2005.

[8] ChenJie and Wong Stephen T.C.: "Nanotechnology for Genomic signal Processing in cancer Research". *IEEE Signal Processing Magazine,* January, 2007.

[9] M.Akhtar,J.Epps,andE.Ambikairajah,"Paired spectral content measure for gene and exon prediction in eukaryotes," *in Proc. IEEE Int. Conf. Information and Emerging Technologies,* pp. 127–130, 2007.

[10]. Alireza Askarzadeh ,"A novel metaheuristic method for solving constrained engineering optimization problems: Crow Search Algorithm," *ELSEVIER,* vol.169,pp.1-12,Jun 2016.

[11].M.Roy, S.Biswas and S.Barman (Mandal): " Identification andanalysis of coding and non-regions of a DNA sequence by Positional Frequency Distribution of Nucleotides (PFDN)algorithm*". International Conference on Computers and Devices for Communication (CODEC-09)*, December 14-15, 2009.

[12]. Lee scargall "EEE305", "EEE801 Part A": Digital Signal Processing.

[13]. Suman Kumar Sahib, Rajib Kar, Durbadal Mandal and S. P. Ghoshal, " IIR Filter design with Craziness based Particle Swarm Optimization Technique," *World Academy of Science, Engineering and Technology*,vol.5,pp.1628-1635,Nov,2011.

[14]. Ranjit Kaur , Manjeet Singh Patterh and J. S. Dhillon, "Real Coded Genetic Algorithm for Design of IIR Digital Filter with Conflicting Objectives," *Applied Mathematics and Information Sciences,* vol. *8*,pp. 2635-2644 ,2014.

[15] National Centre for Biotechnology Information (NCBI).[Online]. Available:http://www.ncbi.nlm.nih.gov/