2017 16th IEEE International Conference on Machine Learning and Applications

Audio Signal Reconstruction Using Cartesian Genetic Programming Evolved Artificial Neural Network (CGPANN)

Nadia Masood Khan, Gul Muhammad Khan
Department of Electrical Engineering, University of Engineering and Technology Peshawar, Pakistan.
Email: [nadiakhan, gk502]@uetpeshawar.edu.pk

***Abstract*—We propose a novel audio signal reconstruction model that makes use of a non-linear estimation algorithm called Cartesian Genetic Programming evolved Artificial Neural Network (CGPANN). CGPANN estimates the non-linear graphs of audio signals with much better accuracy than its counterparts: the interpolation and extrapolation. We have compared them in terms of SNR improvement and ability to deal with disputed data. Unlike other conventional reconstruction algorithms, the proposed algorithm can restore the signal which is damaged up to 50% by noise. A state-of-the-art approach for reconstructing an audio signal utilizing machine learning is presented in this paper. The performance of algorithm is evaluated by measuring its Signal-to-Noise (SNR) improvement and difference between original and reconstructed signal in terms of Mean Absolute Percentage Error (MAPE). SNR improvement of up to 20 dB is recorded for single point estimation with 25% missing samples, 19 dB for multi-point (up to 5) estimation in which half of the data is missing and 16 dB for a signal with random variable noise.**

I. INTRODUCTION

In this work, a neuro-evolutionary algorithm for reconstruction of an audio signal through estimation of the missing samples is presented. Signal gets corrupted when passed through the communication medium, due to dynamic noise in various forms. There are two types of degradation [1] in audio sources: (1) localized degradation: it affects only few samples, and (2) Globalized degradation: it include noise that affect all samples of audio signals [2]. The current research focuses on restoring the signal degraded globally and propose a robust reconstruction model for regenerating the signal damaged by noise. The signal is restored by estimating lost and corrupted samples in time domain. Samples are completely missing/degraded, or its location is already known [3], so our emphasis is mainly on the reconstruction.

The problem of finding missing samples explored in literature mostly focus on short gaps in an audio signal. Methods proposed to date are based on interpolation and extrapolation depending on time limitations. In the first case, data segments surrounding the gap is used to find model parameters and generate the desired number of lost samples to fill the gap. While in case of extrapolation [4] previous samples are used to find the missing samples ahead. The problem becomes more difficult when there are long gaps in an audio signal. Filling the long gap is intangible due to static properties of an audio signal i.e. speech or music signal. In long gaps more information is lost which annoys the listener. To overcome this problem, there must be proper concealment technique. There are three recovery schemes when the correction is to be made at receiver side with no assistance from sender.

- **Insertion method**: A packet is inserted in the gap. As usually there is repetition of previous packets in audio signal, so previous packet is picked and inserted in the gap.
- **Interpolation method**: It generates a replacement packet which is expected to be similar with lost packet.
- **Regeneration method**: It develops decoder state from the data and generates an exact replacement of the missing samples [5].

Our work is based on an Interpolation method, namely, Cartesian Genetic Programming evolved Artificial Neural Network (CGPANN) is used to reconstruct the missing samples. The performance of proposed model is compared with the other signal reconstruction algorithms in Section 2 (Literature Survey) which are based on the statistical measurements, polynomial interpolation and extrapolation, and time series models. Noisy phase and its estimation can be a problem in speech recognition. For many years, researchers working on this believed that signal reconstruction should be independent of phase. In the proposed model, the signal reconstruction without phase information is made possible and longstanding hypothesis of the speech processing community is verified. Radu et al. successfully reconstructed signal without phase information by considering new classes of Parseval frames for an Hilbert space and used finite-dimensional frames [6]. The research proposed here involves efficient amalgamation of genetic programming and neural networks to train the system. The estimation algorithm is an Artificial Neural Network (ANN) with its parameters evolved using Cartesian Genetic Programming (CGP) and it proved to perform better than other neuro-evolutionary algorithms for estimating non-linear graphs, thus a suitable choice for exploring its capabilities in signal reconstruction domain. The best accuracy achieved by the CGPANN is because of its flexibility to optimize all the network parameters.

0-7695-6321-X/17/31.00 ©2017 IEEE
DOI 10.1109/ICMLA.2017.0-100

II. LITERATURE SURVEY

A. Reconstruction In Time Domain

Wolfe et al. [7] explored Gabor regression model for finding lost data in audio time series with one third of data missing. They have reported SNR improvement of 10 dB when 36.5% of signal is missing for the solo piano signal and 5.94 dB improvement for the jazz trumpet recording when 37.5% values are missing. The audio signals like voiced speech or music extracts are distorted when reconstructing with autoregressive modeling [8]. In [7] Gabor regression model is used to overcome autoregressive inadequacies for reconstruction of signals. The Gabor regression model [9] represents time series as a superposition of time-frequency shifted versions of simple window functions. Scott et al. [10] presented Multiresolution Fourier Transform (MFT) for restoration of noisy audio signals and improved the SNR from 10 dB to 11 dB. Matched Sign Pursuit (MSP) algorithm presented in [11] for sparse signal reconstruction has achieved the Mean Squared Error (MSE) gain of 12 dB over Compressive Sampling Matching Pursuit (CoSaMP) and 20dB over Consistent Reconstruction. Restoring the lost samples is done by interpolation when impulsive noise such as clicks, bursts or scratches damages one or two samples. Oudre and Laurent [12] designed an automated system which takes a degraded audio signal as input and detects the location of lost samples, used interpolation to replace them with appropriate values and hence restore the original signal. Stallmann et al. [13] used Time Delay Artificial Neural Networks (TDANN) for audio signal noise detection in gramophone. A detailed comparison of TDANNs and other reconstruction algorithms is given in [14]. Algorithms presented in literature are evaluated using three aspects: estimator efficiency, sampling rate and computational complexity. Our proposed algorithm gives better estimator efficiency by using standard sampling rate without disturbing the properties of original audio signal and is computationally efficient as well. When average sampling rate exceeds or meets Nyquist criteria only then a band limited signal can be uniquely recovered. Interpolation [15] is used for reconstruction of a band limited continuous time signal when the signal is uniformly sampled and Lagrange interpolation [16] is used when sampling is non-uniform. However, in both cases there is constraint of sampling on or above Nyquist rate. The proposed research is focused on the reconstruction of signals which are sampled below Nyquist rate and removes the limitation associated with signal reconstruction.

B. Reconstruction In Frequency Domain

Compressive sensing or sparse sampling is one such method in which signal is reconstructed from far fewer samples usually less than required by Shannon-Nyquist sampling theorem. But for the reconstruction of signal through compressive sensing, there are two conditions. The signal should be sparse in some domain but this is not the case usually. Second is incoherence between sparse signals. These two constraints make recovery of signal difficult using compressive sensing. Compressed Sensing (CS) used in [17] is applied to the harmonic part of sinusoidally-modeled audio signals. Their investigation showed that CS can be used to encode signals at low bit rates instead of modeling sinusoidal parameters because the harmonic part is sparse in frequency domain. Meignen et al. [18] used time frequency domains to correctly estimate the signal when the Signal-to-Noise Ratio (SNR) is 10 dB. But the signals having SNR below 10 dB cannot be restored using algorithm proposed in [18]. To enhance speech quality and recognition accuracy, the harmonic structure of speech in high frequency resolution has been examined [19] in which signal is reconstructed from estimated posteriors and phase from original noisy signal. They have attained a SNR gain of 8.38 dB for enhancement of speech at 0 dB.

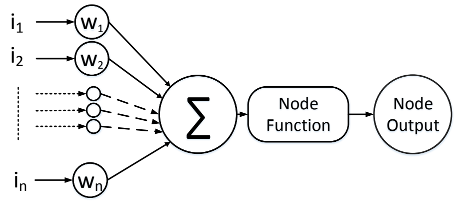
C. Signal Reconstruction Using Neural Network

Use of neural network for signal reconstruction is in practice for quite long. Dipartimento et al. [20] reviewed neural network architectures suitable for nonlinear real-time audio signal processing. Neural Networks can be used for nonlinear audio signal processing applications e.g. audio signal recovery, speech quality enhancement and nonlinear transducer linearizion. Neural Networks like Multi-Layer Perceptron (MLP), Time Delay Neural Network (TDNN) and Recurrent Neural Network (RNN) represents an adaptive circuit which can extend and generalize the simple adaptive linear filter in nonlinear domain. The reconstruction techniques proposed in literature involve complex signal processing and hence have low computational efficiency but this paper introduces a method in which signal is reconstructed through estimation of lost samples and is computationally efficient, since it evolves all the parameters of ANN to obtain an optimum architecture.

III. RECONSTRUCTION MODEL CARTESIAN GENETIC PROGRAMMING ARTIFICIAL NEURAL NETWORK (CGPANN)

Cartesian Genetic Programming (CGP) was introduced for the evolution of feed forward digital circuits by Miller and Thompson [21]. CGP is composed of genotype which consists of genes or nodes and phenotype shows the system connectivity of inputs via the nodes to the outputs. CGPANN is explored in many applications in past producing competitive results. The aim of this paper is to find the best fit model of CGPANN which has the ability to predict the lost samples with extraordinary accuracy. The exploration of neuro-evolutionary based machine learning algorithm for signal reconstruction is a novel approach. ANN Parameters of the estimation model are extracted from original audio signal which is continuously evolved using CGP for best performance. The evolution of CGPANN continues until reasonably accurate solution is reached. The idea of CGPANN is proposed by Khan et al. [22] as it replaces each computational node in CGP with an artificial neuron, thus producing an artificial neural network. The CGPANN [23] inherits key characteristics of Artificial Neural Networks (ANN) i.e. system inputs, nodes/neurons, node's inputs, connection weights, activation functions within the node, node's outputs and system outputs.

Fig. 1. CGPANN Node - Basic Structure

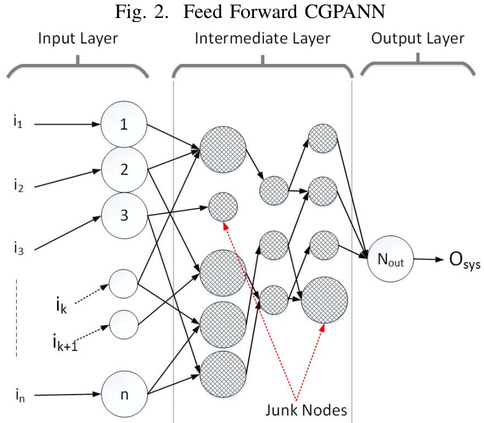


A node, when it takes inputs directly from the input sequence, is termed as input node. Figure 1 shows the basic architecture of a CGPANN node. Inputs $i_1, i_2 ... i_n$ are scaled using weights from weight matrix W defined by Eq. 1.

$$W = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ \vdots \\ W_n \end{bmatrix} \tag{1}$$

The scaled inputs, after summation, are passed through an activation function. This function produces the ultimate output of the system. The combination of input nodes forms an input layer. Layers that take their input(s) from input nodes are termed as intermediate layers and consists of intermediate nodes. The output node(s) lead(s) to the formation of output layer. This layer takes its respected input(s) from either intermediate nodes or from the input layer while giving the system output as shown in Figure 2. Nodes that do not contribute to the system output are termed as Junk nodes.

Fig. 2. Feed Forward CGPANN



In Figure 2, $i_1, i_2 ... i_n$ serves as inputs to the input layer thus the input nodes $1, 2, 3...n$ acts as input nodes. Note that the input node can take more than one input depending upon the CGPANN architecture. Junk nodes are pointed by doted arrows in Figure 2. The output layer that may consist of a single node or multiple nodes, depending upon the architecture

TABLE I
AVERAGE SNR IN dB FOR SINGLE POINT ESTIMATION (RECONSTRUCTED SIGNAL)

S. No.	Number of Nodes						
	50	100	150	200	250	Average	STDEV
Audio Signal 2	37.82	42.98	34.47	41.60	43.36	40.05	3.81
Audio Signal 3	37.62	43.50	34.60	42.00	42.40	40.02	3.77
Audio Signal 4	36.87	21.84	20.96	15.11	19.88	22.93	8.21
Audio Signal 5	32.70	19.41	17.81	22.23	20.88	22.60	5.88
Audio Signal 6	36.80	26.11	20.74	29.31	26.14	27.82	5.89
Audio Signal 7	38.15	25.49	18.71	21.84	26.52	26.14	7.39
Audio Signal 8	44.52	38.64	39.36	40.22	39.61	40.47	2.33
Audio Signal 9	47.17	40.22	40.64	36.28	38.89	**40.64**	4.03
Average	38.96	32.27	28.41	31.07	32.21		

of CGPANN, is providing the network output indicated by O_{sys}.There is only single output layer node (N_{out}) in the CGPANN Model shown in Figure 2. Note that all the weighted inputs are summed and this intermediate value (int_v) is passed through an activation function for evaluating the overall output of the CGPANN neuron. Log-sigmoid is used as the activation function in the proposed model given by Eq. 2.

$$f(x) = \frac{1}{1 + e^{(int_v)}}, \qquad int_v = \sum_{i=1}^{N} [W_i].[I_i] \tag{2}$$

Here, int_v provides the weighted sum of the node input for a particular active node of the CGPANN that is defined by int_v in Eq. 2. W_i is the assigned weight matrix, I_i is the respected inputs matrix to the particular node in a genotype g. If $O_{(g,x,y)}$ is the output of a node x in a genotype g at layer y, then the j_{th} input $I_{(g,x,j)}$ can further be defined by Eq. 3.

$$\begin{aligned} I(g,x,j) = PRG([&I(g,c,1), I(g,c,2), ..., I(g,c,N) \\ &..., I(g,c,K)] : [O(g,c,1), O(g,c,2), \\ &....O(g,c,j-1)]) \end{aligned} \tag{3}$$

The contribution of an individual node (Nq) can be given by Eq. 4.

$$O_{g,c,q} = \frac{1}{1 + e^{(int_v(g,x,y))}} \tag{4}$$

where q is an integer number assigned to a node from the y_{th} layer.

The particular CGPANN uses $1+\lambda$ evolutionary strategy for producing offspring where λ is the number of mutant offspring to be produced in a single generation. The fittest genotype is selected based on a specific fitness criteria and is subjected to mutation.

IV. EXPERIMENTAL SETUP

To evaluate the proposed algorithm, CGPANN has been applied for the estimation of missing data values in audio time series data. The audio signal under consideration is distorted because of long gaps resulting from severe impulsive noise that occurs frequently at random intervals. Our assumption is that due to noise, signal sample is completely lost and is represented by zero in the experiments. Simulations were performed on the audio time series data in which the signal

TABLE II
MAPE VALUES FOR SINGLE POINT ESTIMATION (RECONSTRUCTED SIGNAL)

S. No.	Number of Nodes						
	50	100	150	200	250	Average	STDEV
Audio Signal 2	6.94	5.10	9.17	6.91	5.38	6.70	0.09
Audio Signal 3	6.56	4.14	9.08	6.25	4.86	**6.18**	0.09
Audio Signal 4	8.29	6.58	8.02	18.30	8.42	9.92	0.09
Audio Signal 5	11.16	8.57	11.11	8.95	7.22	9.40	0.11
Audio Signal 6	8.01	6.09	11.28	10.26	7.06	8.54	0.13
Audio Signal 7	8.94	7.63	12.09	13.26	8.46	10.08	0.14
Audio Signal 8	7.38	9.03	8.30	9.30	8.85	8.57	0.09
Audio Signal 9	5.85	7.27	7.02	10.85	8.24	7.84	0.09
Average	7.89	**6.80**	9.51	10.51	7.31		

TABLE III
AVERAGE SNR IN dB FOR MULTI POINT ESTIMATION (RECONSTRUCTED SIGNAL)

S. No.	Number of Nodes						
	50	100	150	200	250	Average	STDEV
Audio Signal 2	20.00	20.45	20.62	21.31	21.14	20.70	0.53
Audio Signal 3	27.15	25.81	25.81	25.72	26.02	26.10	0.60
Audio Signal 4	29.65	29.37	28.82	25.92	27.77	28.30	1.52
Audio Signal 5	23.23	23.74	24.01	25.95	24.82	24.35	1.06
Audio Signal 6	25.25	25.75	25.66	25.53	25.79	25.60	0.22
Audio Signal 7	23.71	24.03	23.95	24.05	24.15	23.98	0.17
Audio Signal 8	28.02	28.60	28.65	28.77	29.14	**28.64**	0.40
Audio Signal 9	27.19	27.25	27.23	27.57	27.11	27.27	0.18
Average	25.53	25.62	25.59	25.60	25.74		

TABLE IV
MAPE VALUES FOR MULTI POINT ESTIMATION (RECONSTRUCTED SIGNAL)

S. No.	Number of Nodes						
	50	100	150	200	250	Average	STDEV
Audio Signal 2	11.11	10.56	10.41	9.71	9.87	10.33	0.56
Audio Signal 3	6.12	5.22	5.24	5.41	5.13	**5.42**	0.40
Audio Signal 4	12.17	12.37	12.30	12.15	12.43	12.28	0.12
Audio Signal 5	11.25	11.26	11.09	9.88	10.88	10.87	0.58
Audio Signal 6	9.35	9.55	9.61	10.07	9.78	9.67	0.27
Audio Signal 7	11.17	11.46	11.43	11.09	11.61	11.35	0.22
Audio Signal 8	6.58	6.44	6.40	12.66	6.33	7.68	2.79
Audio Signal 9	6.89	6.77	6.77	6.83	6.71	6.80	0.07
Average	9.33	9.20	9.16	9.73	**9.09**		

TABLE V
AVERAGE SNR IN dB OF RECONSTRUCTED SIGNAL HAVING VARIABLE NOISE

S. No.	Number of Nodes						
	50	100	150	200	250	Average	STDEV
Audio Signal 2	25.38	25.86	26.02	26.90	26.51	26.14	0.59
Audio Signal 3	31.76	30.96	30.95	30.74	31.18	31.12	0.39
Audio Signal 4	28.66	29.18	29.58	34.46	30.68	30.51	2.33
Audio Signal 5	28.46	28.95	29.21	31.51	29.947	29.61	1.19
Audio Signal 6	30.59	31.08	31.04	31.04	31.24	31.00	0.24
Audio Signal 7	28.09	28.40	28.42	28.58	28.52	28.40	0.19
Audio Signal 8	33.66	34.81	34.66	34.80	34.51	**34.49**	0.48
Audio Signal 9	29.80	30.00	30.00	30.31	30.09	30.04	0.18
Average	29.55	29.91	29.98	31.04	30.33		

is artificially corrupted. We have considered three scenarios explained in subsections A, B and C. The experiments done in this research is based on the concept of triggered jumping window mechanism used for Single Point Estimation (SPE) and moving/sliding window mechanism [24] for Multi Point Estimation (MPE). At the start of simulation experiments, the audio signal in time domain is artificially degraded, originally sampled at standard rate of 44.1 kHz. We limited our experiments to the time domain analysis and showed the reconstruction model efficiency in various scenarios.

TABLE VI
SNR IMPROVEMENT

Proposed Model	Noisy Signal SNR	Reconstructed Signal SNR	SNR Improvement
Single Point Estimation	12 dB	32 dB	**20 dB**
Multi Point Estimation	6 dB	25 dB	**19 dB**
Signal Enhancement (Variable Noise)	14 dB	30 dB	**16 dB**

TABLE VII
COMPARISON WITH OTHER TECHNIQUES

S.No	Technique	SNR Improvement/Gain (dB)	Type of distortion
1	Gabor regression model [9]	10	36.5% missing data
2	Gabor regression model [9]	5.94	37.5% missing data
3	Multi-Resolution Fourier Transform [10]	1	Random noise
4	High frequency resolution speech model [19]	8.38	Speech Enhancement
5	Matched Sign Pursuit [11]	12	Random noise
6	Proposed CGPANN	30	25% missing data
7	Proposed CGPANN	20	50% missing data
8	Improved Least Square Estimation [25]	3	Speech Enhancement
9	Kalman Filter [26]	10	Speech Enhancement
10	Proposed CGPANN	20	25-30% (Random Noise)

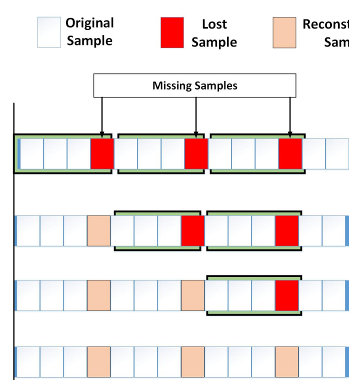


Fig. 3. Triggered Jumping Window Mechanism for Single Point Estimation

A. *Single Point Estimation (SPE)*

In single point estimation, triggered jumping mechanism takes 3 samples as input for predicating the 4th missing signal sample, when 25% data values are missing. Then the window jump to the next iteration and further takes 3 original samples and finds out the 4th sample as shown in Figure 3. Signal sampled at 44.1 kHz is down sampled to 33.075 kHz done by removing 4th sample after every 3 samples. The CGPANN model has shown SNR improvement of 20 dB. The degraded signal SNR is 12 dB and reconstructed signal SNR is 32 dB.

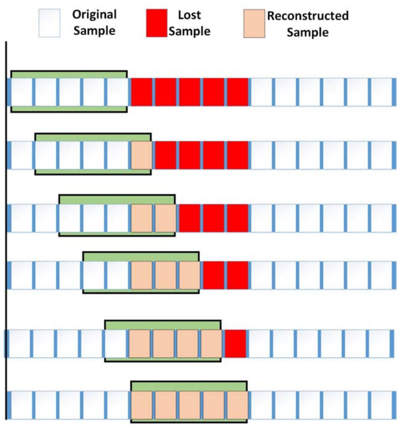


Fig. 4. Sliding/Moving Window Mechanism for Multi Point Estimation

B. Multi Point Estimation (MPE)

Multi point estimation is used when more than one sample is lost during transmission. This problem can be solved by using MPE and reconstructing the original signal from estimated sample values. 50% of samples in an audio signal are replaced by zero to indicate 50% missing data. The signal sampled at a rate of 44.1 kHz was first down sampled to 22.05 kHz by introducing series of gaps of 5 samples after every 5 samples. Amplitude values in time domain are considered and estimated sample is placed back in the sliding window to estimate future values in case of multi-point estimation as shown in Figure 4.

C. Signal Enhancement In Presence Of Random Noise

In the third case, unknown number of samples are removed at random locations hence introducing variable noise. Enhancing the signal quality in the presence of random variable noise is a challenging task. Speech enhancement in the presence of colored noise has been explored by Youshen and Pengyu [25] who have reported SNR improvement of 5 dB to 8 dB using an improved least square estimation. Mariyadasu et al. [26] applied Kalman filtering to a speech signal mixed with random noise and improved SNR from -5 dB to +5 dB. The algorithm presented in this research can recover signal damaged with 10% to 40% random noise.

V. Performance Evaluation Of Reconstruction Algorithm

The proposed method removes noise and enhances the signal quality when 25% to 50% of the data values are missing. Lastly, in the simulations the location of noise is already known and the samples degraded due to noise are completely missing. Algorithm performance is evaluated using signal to noise ratio(SNR) defined as:

$$SNR(input) = 20\, log_{10} \frac{\sum_{i=1}^{k} S_i}{\sum_{t=1}^{m} (S_t - 0)} \tag{5}$$

S_i is undistorted signal samples having k number of the total signal (actual plus target) and S_t is the target signal having m number of missing samples which we wish to estimate in order to reconstruct the whole signal. Signal-to-Noise ratio is the measure of strength of signal divided by unwanted interference. In our case, the unwanted interference is given by the difference between the actual signal samples and estimated samples. At the input side, estimation is not yet done so noise is equal to target signal. This distorted signal acts as input to the proposed system. The output SNR is given by following equation.

$$SNR(output) = 20\, log_{10} \frac{\sum_{i=1}^{k} S_i}{\sum_{t=1}^{m} (S_t - E_t)} \tag{6}$$

where E_t is estimated signal samples. In the output block, noise is the difference between target value and its estimated value. The output SNR is improved because difference (noise) is reduced.

The audio data used in the experimentation is recorded using the speech recorder at a sampling rate of 44,100 samples per second and stored in an uncompressed audio format called Waveform Audio File (WAV). The recorded signal is then divided into 9 equal parts of 2 seconds to get 9 audio files. The audio signal 1 having size of 29.5 KB is chosen randomly for training of the CGPANN system. Each audio signal contains approximately 57000 samples. The testing of the data has been carried out on the remaining 8 audio files. Average SNR(dB) of the reconstructed signal in case of Single Point Estimation (SPE) for different number of nodes is shown in Table I. The audio signal having less variations and less Standard deviation is audio signal 3 which is 3.77 have shown an average SNR of 40.02 dB. The audio signal 4 having more variations (SD=8.21) have shown low SNR of 22.60 dB. A comparison of SNR of CGPANN network have shown that a 50-node neural network gave best result 38.96 dB and the Audio Signal 9 has given good average SNR of 40.64 dB as highlighted in Table 1 .

Table II shows the Mean Absolute Percentage Error (MAPE) [27] values calculated for SPE. The MAPE value of 8.40 is calculated by averaging the different nodes average MAPE values. The single point estimation gives **91.6% accuracy**. From Table 2, it can be seen that system with 100 nodes performs better having average error of 6.80 and Audio Signal 3 shows lowest error of 6.18. Table III shows the results for Multi Point Estimation (MPE) when 50% data values are missing. Keeping number of nodes 250 gave us good average SNR of 25.74 dB and Audio Signal 8 reconstructed with good SNR of 28.64 dB. The lowest average error of 9.09 for MPE is observed in 250-node neural network as shown in Table IV. As a whole Audio Signal 3 reconstruction error is minimum having MAPE of 5.42. The average MAPE is 9.30 calculated from the average MAPE of different nodes in Table IV which proves that proposed system gives **accuracy of 90.7%** for MPE. Furthermore, Table V displays computed results of the SNR values when signal is distorted by random variable noise (10%–40%). The audio signal 8 has shown excellent average result of 34.5 dB, with number of nodes 200 we get best SNR of 31.04 dB. Table VI shows the overall SNR improvement that the proposed system has achieved. The input and output SNR is calculated using the formula given in equations 5 and 6. Output SNR is obtained by taking the average SNR of all nodes and audio signals. The noisy signal has SNR of 12 dB which is lifted up to 32 dB by our proposed system when 25% data is corrupted and SNR gain of 20 dB is recorded. In case of 50% degradation, model indicates SNR improvement of 19 dB by restoring the signal from 6 dB to 25 dB. The input signal distorted by random variable noise has SNR of 14 dB. This low input SNR is lifted up by 16 dB by the CGPANN estimator which correctly estimates the corrupted samples and the reconstructed signal SNR is 30 dB. This is the best result achieved so far in case of random noise for signal enhancement as can be seen in Table VII. Comparison of SNR gain with previously proposed methods is provided in Table VII which clearly shows the CGPANN reconstruction model outperforms rest of the signal restoration algorithms.

VI. Conclusion

The current research aims at devising estimation based reconstruction model which efficiently estimate the lost samples in an audio signal. The performance of proposed CGPANN reconstruction model has been evaluated in different scenarios. First is single point estimation when signal degradation is minimum that is only one sample is estimated using information from past three actual samples. The audio signal is restored with 91.6% accuracy. In the second case, distortion is high as half of the signal is damaged due to noise, CGPANN performs well giving an accuracy of 90.7%. The performance of estimation method was assessed taking into account the SNR improvement from damaged noisy signal to reconstructed signal. The proposed algorithm is a suitable method to restore the missing samples in a signal when transmission noise induced is up to 50%. SNR improvement of 20 dB is recorded for Single Point Estimation and 19 dB for Multi Point Estimation. In real life, the signals vary randomly and are non-stationary. For this purpose, CGPANN is applied for estimating the lost sample in time domain as it is able to overcome the disadvantages in previous method of signal enhancement and gave SNR gain of 16 dB for a signal having random noise.

References

[1] S. Godsill, P. Rayner, and O. Cappé, "Digital audio restoration," in *Applications of digital signal processing to audio and acoustics*. Springer, 2002, pp. 133–194.

[2] U. Hammarqvist, "Audio editing in the time-frequency domain using the gabor wavelet transform," 2011.

[3] L. Oudre, "Interpolation of missing samples in audio signals based on autoregressive modeling," *Image Processing On Line. Draft*, 2014.

[4] M. Bartkowiak and B. Latanowicz, "Mitigation of long gaps in music using hybrid sinusoidal+ noise model with context adaptation," in *Signals and Electronic Systems (ICSES), 2010 International Conference on*. IEEE, 2010, pp. 435–438.

[5] C. Perkins, O. Hodson, and V. Hardman, "A survey of packet loss recovery techniques for streaming audio," *IEEE network*, vol. 12, no. 5, pp. 40–48, 1998.

[6] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, pp. 345–356, 2006.

[7] P. J. Wolfe and S. J. Godsill, "Interpolation of missing data values for audio signal restoration using a gabor regression model," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, vol. 5. IEEE, 2005, pp. v–517.

[8] S. Godsill, P. Rayner, and O. Cappé, "Digital audio restoration," in *Applications of digital signal processing to audio and acoustics*. Springer, 2002, pp. 133–194.

[9] P. J. Wolfe and S. J. Godsill, "A gabor regression scheme for audio signal analysis," in *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on.* IEEE, 2003, pp. 103–106.

[10] H. R. R. Scott and R. Wilson, "A multiresolution audio restoration algorithm," in *Applications of Signal Processing to Audio and Acoustics, 1995., IEEE ASSP Workshop on*. IEEE, 1995, pp. 151–154.

[11] P. T. Boufounos, "Greedy sparse signal reconstruction from sign measurements," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*. IEEE, 2009, pp. 1305–1309.

[12] L. Oudre, "Automatic detection and removal of impulsive noise in audio signals," *Image Processing On Line*, vol. 5, pp. 267–281, 2015.

[13] C. F. Stallmann and A. P. Engelbrecht, "Gramophone noise detection and reconstruction using time delay artificial neural networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 6, pp. 893–905, 2017.

[14] C. F. Stallmann, "Digital audio restoration of gramophone records," Ph.D. dissertation, 2015.

[15] S. Maymon and A. V. Oppenheim, "Sinc interpolation of nonuniform samples," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4745–4758, 2011.

[16] J. Selva, "Functionally weighted lagrange interpolation of band-limited signals from nonuniform samples," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 168–181, 2009.

[17] A. Griffin, T. Hirvonen, A. Mouchtaris, and P. Tsakalides, "Encoding the sinusoidal model of an audio signal using compressed sensing," in *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*. IEEE, 2009, pp. 153–156.

[18] S. Meignen, T. Oberlin, P. Depalle, P. Flandrin, and S. McLaughlin, "Adaptive multimode signal reconstruction from time–frequency representations," *Phil. Trans. R. Soc. A*, vol. 374, no. 2065, p. 20150205, 2016.

[19] T. Kristjansson and J. Hershey, "High resolution signal reconstruction," in *Automatic Speech Recognition and Understanding, 2003. ASRU'03. 2003 IEEE Workshop on*. IEEE, 2003, pp. 291–296.

[20] A. Uncini, "Audio signal processing by neural networks," *Neurocomputing*, vol. 55, no. 3, pp. 593–625, 2003.

[21] J. F. Miller and P. Thomson, "Cartesian genetic programming," in *European Conference on Genetic Programming*. Springer, 2000, pp. 121–132.

[22] M. M. Khan, A. M. Ahmad, G. M. Khan, and J. F. Miller, "Fast learning neural networks using cartesian genetic programming," *Neurocomputing*, vol. 121, pp. 274–289, 2013.

[23] G. M. Khan and R. Arshad, "Electricity peak load forecasting using cartesian genetic programming based neuro evolutionary techniques," *International Journal of Computational Intelligence Systems*, vol. 9, no. 2, pp. 376–395, 2016.

[24] P. Stavroulakis, *Neuro-fuzzy and fuzzy-neural applications in telecommunications*. Springer, 2004.

[25] Y. Xia and P. Wang, "Speech enhancement in presence of colored noise using an improved least square estimation," 2013.

[26] M. Mathe, S. P. Nandyala, and T. K. Kumar, "Speech enhancement using kalman filter for white, random and color noise," in *Devices, Circuits and Systems (ICDCS), 2012 International Conference on*. IEEE, 2012, pp. 195–198.

[27] M. Sonmez, A. P. Akgüngör, and S. Bektaş, "Estimating transportation energy demand in turkey using the artificial bee colony algorithm," *Energy*, vol. 122, pp. 301–310, 2017.