

A Novel Approach of Speech Enhancement by Implementing Cellular Automata Algorithm

Mahesh B. Adhav
PG Student, Electronics Engineering,
GCOE, Amravati, India

Jagdish D. Kene
Associate Professor,
Electronics Engineering,
GCOE, Amravati, India

ABSTRACT

The main purpose of speech enhancement technique is to eliminate the background noise from speech signal to improve the quality of the speech signal. Speech signal is often corrupted by additive background noise like train noise, market noise etc. In such noisy environment listening at the end user is very difficult. This paper presents speech enhancement method using Cellular Automata (CA) algorithm. Based on Peak Signal to Noise Ratio and Mean Square Error criterion the system evaluate the amount of surrounding noise present in speech. The proposed algorithm helps to improve the quality of speech signal at the listener end.

Keywords

Keywords Speech Enhancement, Cellular Automata, PSNR, MMSE.

1. INTRODUCTION

We ask that authors follow some simple guidelines. Speech communication takes place everywhere like in classroom, market, home, offices, street, bus-stop, etc. Communication between speaker and listener is generally corrupted by background noise, which creates problem for listener. However, when the conversation takes place through the telephone disturbances are more irritating because the brain will not get the extra visual [5]. The speech signal transmitted to the other party is picked up by a microphone at receiver. The microphone signal contains both speech as well as noise at some ratio known as speech to noise ratio, which depend on position of microphone with respect to the speaker's mouth. To increase the SNR and to allow for the listener to grasp the speech clearly, a speech enhancement method should be applied. Therefore from last several years a lot of work is done on speech recognition, speaker verification and speech enhancement. The main task in speech enhancement is separating speech signal from noise, for that characteristics of speech signal and noise signal to be known, which is difficult task. As noise comes from different sources so it is impossible to model them. Schroeder was the first author to propose speech enhancement technique based on analog implementation of the spectral magnitude subtraction method [3].

The remaining part of this paper is organized as follows; section II contain previous work related to speech enhancement, details of proposed method is presented in section III, section IV describes simulation results finally, our concluding are presented in last Section .

2. RELATE WORK

Since few decades, number of speech enhancement techniques are proposed, among all of them denoising of speech signal by using Wavelet based transform is widely useful in mobile system, in this technique various parameters such as the

wavelet basis, the layer of decomposition, threshold value, and the threshold function were studied [2]. The simulation results show by selecting the optimum values of the parameters mentioned above can not only remove the background noise from speech signals but also helps in maintaining the step and break information. Another wavelet based method is using super soft thresholding in which super soft thresholding technique is used to decomposed signal to remove noise [9]. Generally wavelet-based algorithm requires detection of unvoiced signal as it affects the performance, but super soft thresholding does not require any voiced/unvoiced detection method. In case of spectral subtraction (SS) method and their further versions [4], [6], estimated noise signal is subtracted from noisy speech signal to obtain clean signal. For this method it is assume that noise signal is stationary or slowly changing with time. Minimum mean square error (MMSE) estimators [7] and its updated versions [8], these methods do not require any prior knowledge about the speech and noise signals, nor any training stage beforehand, so they are highly flexible and allow implementation in various situations. However, these algorithms usually assume that the noise is stationary and are thus not good at dealing with non-stationary noise types, especially under low signal-to-noise (SNR) conditions. Mean square error (MSE) should be as minimum as possible to minimize the MSE adaptive least mean square algorithm is proposed which show that excess mean square error is improved up to 20 dB in different noisy environment [10].

3. PROPOSED METHOD

Cellular Automata (CA) were originally proposed by John von Neumann as formal models of self-reproducing organisms [1]. CA generally used because of simplicity and parallel processing. CA has ability to perform complex operation and model behavior of complex system in nature. It can be used in number of applications like mathematics, natural science, microstructure, computer science, biology, etc. CA has different forms like dynamic automata, tessellation automata, iterative automata, 1-dimensional automata, 2-dimensional automata, linear cellular automata, additive cellular automata. In proposed work, combination of iterative array and dynamic Cellular Automata is being implemented because in one iteration the weights are being updated and dynamically changes with respect to time. In this models each elementary device, a cell or site, can take some values and is updated at intervals according to a rule which expresses the actual value from its preceding value and that of its neighbors [12].

Speech enhancement using CA algorithm has two phase first is training mode in which speech signal without noise is given input to the system and second phase is testing mode in which speech signal corrupted by adaptive background noise is applied to second part of system as shown in fig. 1. Here

system adds some random noise to original signal, this signal along with original signal applied to CA block.

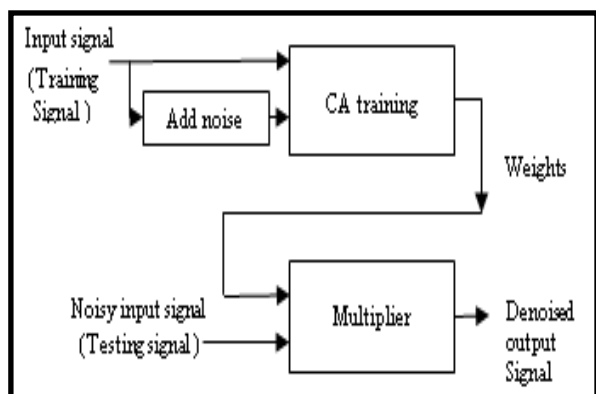


Fig.1 Block diagram of Speech enhancement using CA algorithm

Cellular Automata training block calculate weights of both noisy signal and original signal as define in equation (1).

$$W_{n+1} = W_n + (SF * error * noisy\ sound) * 1/SNR \dots\dots (1)$$

Where SF= Speed up factor, which decide speed of operation of system and inversely proportional to noise in speech signal. Increase in noise in speech consumes more time to denoise it.

$$SF = \lfloor 0 * c \left(\frac{1}{1+a*s^b} \right)$$

SF a (1/noise)

Where, c, a, b – Constants.

s- Speech signal

In second phase of system, the multiplier circuit is used which is fed in terms of noisy speech signal and weights of original signal. This multiplier generates output which represent enhanced or noiseless signal.

3.1 Mean Square Error (MSE)

Mean square error is dominant quantitative performance metric in the field of signal processing. It is measure of signal fidelity, which means it will compare two signals and output value describe the degree of similarity [10].

Consider two signal $A = \{A_i | i= 1, 2, 3, \dots, N\}$ is enhance speech signal with noise, $B = \{B_i | i=1, 2, 3, \dots, N\}$ is training signal without any noise, here A, B both are finite length discrete signals. Then MSE between these two signals can be defined in equation (2), for better operation value of MSE should be as low as possible.

$$MSE = \frac{1}{N} \sum_{i=0}^N (A_i - B_i)^2 \dots\dots\dots (2)$$

3.2 Peak Signal to Noise Ratio (PSNR)

PSNR is second important parameter to measure in order to study speech enhancement. It is define as ratio of signal power to noise power. It is generally express in dB. For achieving better performance of system the value of PSNR should be maintain in higher range. PSNR can be define mathematically as in equation (3).

$$PSNR = 10 * \log_{10} (1/\sqrt{MSE}) \dots\dots\dots (3)$$

4. SIMULATION RESULTS

CA algorithm is tested on spoken English sentence at different environment like in presence of white noise, babble noise, car

noise and simulated using MATLAB. Microphone in mobile is used to capture speech signal and it is stored in .wav format. Fig.2 show waveforms of input white noise at 10 dB SNR. Fig. 2(a) is our training signal without any noise contamination. System add some random noise to this training signal which is shown in Fig. 2 (b). CA algorithm process this random noisy signal and gives denoise signal known as CA corrected signal shown in Fig. 2 (c). Input speech signal with adaptive background noise is known as testing signal shown in Fig. 2 (d). Finally enhanced speech signal is obtain at output shown in Fig. 2 (e).

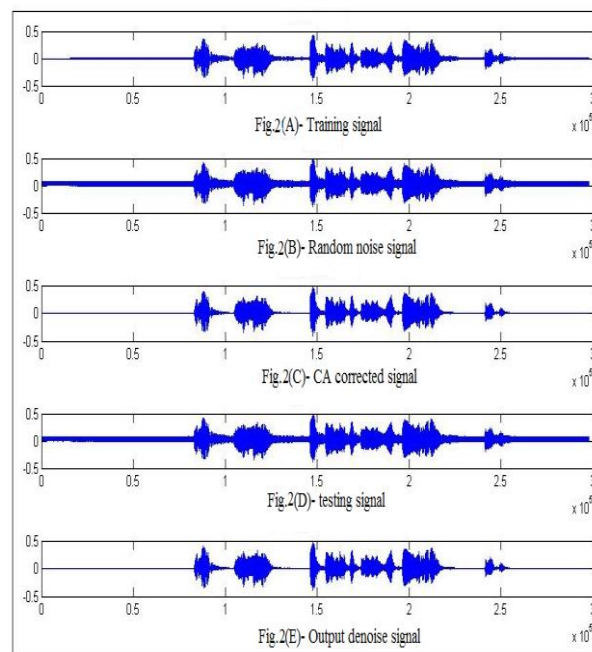


Fig. 2 Simulation result of CA algorithm for white noise at 10 dB SNR input.

These noisy speech examples are at different input SNRs equal to 1dB, 5dB, 10dB and 15dB and signal with improved SNR is obtained at output.

Table 1: Performance measure of CA based on SNR for white noise

Input SNR (dB)	Enhanced SNR (dB)
	CA algorithm
1	9.58
5	38.50
10	51.73
15	59.57

Table 2: Performance measure of CA based on MSE for white noise

Input SNR (dB)	Mean Square Error
	CA algorithm
1	0.33
5	0.011
10	0.0025
15	0.0010

Speech babble is one of the most challenging noise interference for all speech systems [14]. Listener suffers from babble noise when user is in crowd or multi talkers are present. The same speech sample is tested for babble noisy environment and CA algorithm gives result as shown in fig. 3.

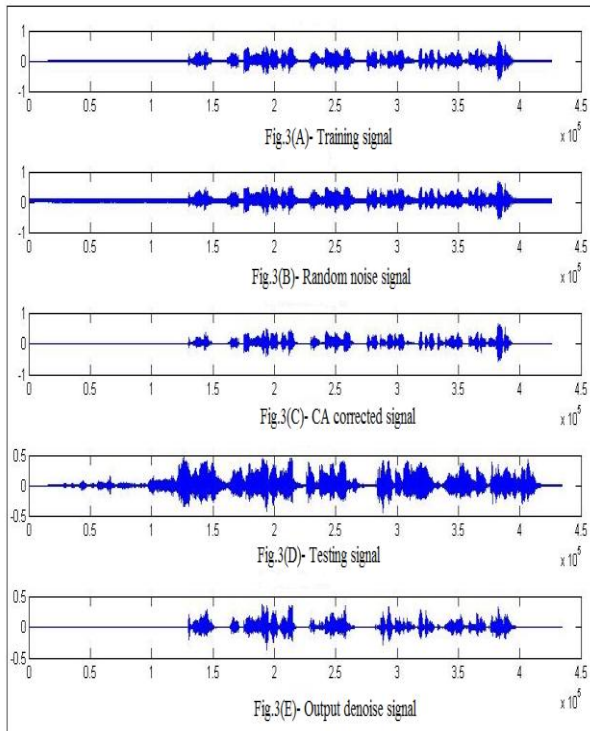


Fig. 3 Simulation result of CA algorithm for babble noise at 10 dB SNR input.

Table 3: Performance measure of CA based on SNR for babble noise

Input SNR (dB)	Enhanced SNR (dB)
	CA algorithm
1	48.49
5	51.12
10	55.52
15	58.36

Table 4: Performance measure of CA based on MSE for babble noise

Input SNR (dB)	Mean Square Error
	CA algorithm
1	0.0032
5	0.0026
10	0.0018
15	0.0012

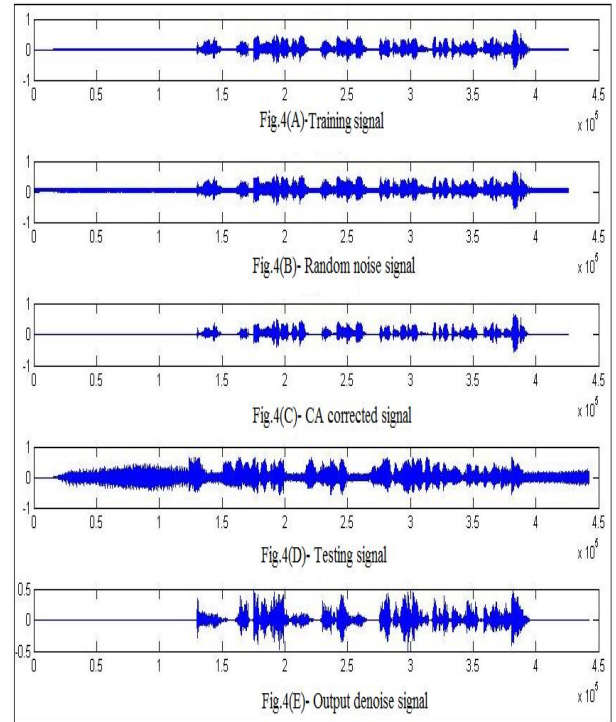


Fig.4 Simulation result of CA algorithm for car noise at 10 dB SNR input.

Table 5: Performance measure of CA based on SNR for car noise

Input SNR (dB)	Enhanced SNR (dB)
	CA algorithm
1	39.98
5	41.93
10	44.95
15	46.76

Table 6: Performance measure of CA based on MSE for car noise

Input SNR (dB)	Mean Square Error
	CA algorithm
1	0.0100
5	0.0080
10	0.0056
15	0.0045

From results of Table 1 to Table 6, it is clearly found that CA based speech enhancement algorithm performed well in both low SNR and high SNR environment.

5. CONCLUSION

From the result analysis, we are concluded that improvement in PSNR is achieved up to the remarkable level by implementing Cellular Automata concept with an iterative network. In all three noisy conditions consider for simulation have improvement in both lower and higher ranges. Thus the

Cellular Automata provides effective solution towards performance improvement in mobile communication system, irrespective of surrounding noise.

6. REFERENCES

- [1] Burks, A. W.: Von Neumann's self-reproducing automata. In Burks (1970a).
- [2] L. Du, R. Xu, F. Xu, D. Wang, H. Chen, "Research on Key Parameters of Speech Denoising Algorithm Based on Wavelet Packet Transform," Computer Science and Information Technology (ICCSIT), 3rd IEEE International Conference on, Vol. 6, pp. 551 – 556, July 2010.
- [3] W. Kim, S. Kang, and H. ko "Spectral subtraction Based on phonetic dependency and masking effects" IEEE Proc. vision image signal process, Vol. 14, no.5, pp423-427,2000.
- [4] S. Kamath and P. Loizou, "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," IEEE international conference on acoustics speech and signal processing, Vol.4, pp. 4164–4167, May 2002.
- [5] M. B. Adhav, J. D. Kene, "Effective Study Of Feature Extraction Methods For Speaker Identification," International Journal of Science, Engineering and Technology Research (IJSETR), Vol.4, no.12, pp. 4384-4387, December 2015.
- [6] K. Yamashita, S. Ogata, and T. Shimamura, "Improved spectral subtraction utilizing iterative processing," Electronics and Communications, Japan, vol. 90, no. 4, pp. 39-51, 2007.
- [7] F. Ykhlef, H. Ykhlef, "A smoothed Minimum Mean-Square Error Log-Spectral Amplitude estimator for speech enhancement," Multimedia Computing and Systems (ICMCS), 2014 International Conference on, pp. 246-249, Apr. 2014.
- [8] T. Gerkmann and R. C. Hendriks, "Unbiased MMSE-based noise power estimation with low complexity and low tracking delay," IEEE Trans Audio, Speech, Lang. Process., vol. 20, no. 4, pp. 1383–1393, May 2012.
- [9] R. Santhoshkumar, Dr. B. Kirubagari, "Speech Enhancement Using Super Soft Thresholding in Wavelet Domain," International Journal of Advanced Research in Computer Science and Software Engineering, Vol.5, no.2, pp. 315-320, February 2015.
- [10] S. Wang , Y. Shi, "An Improved speech denoising Algorithm Based on Adaptive Least Mean Square ," International Conference on Industrial and Information Systems, pp. 293 – 296, April 2009.
- [11] R. Martin, "Speech Enhancement Based on Minimum Mean-Square Error Estimation and Supergaussian Priors ," IEEE Trans. on Speech And Audio Processing, Vol. 13, No. 5, pp. 845-856, September 2005.
- [12] H. Aso, N. Honda, "Dynamical characteristics of linear cellular automata, " Journal of Computer and System Sciences, Vol.30, no.3, Pp. 291-317, June 1985.
- [13] Cole, S. N., "Real-time computation by n-dimensional iterative arrays of finite-state machines," IEEE Trans. on Computers, Vol.18, no.4, pp.349–365, April 1969.
- [14] N. Krishnamurthy, J. H. L. Hansen, "Babble Noise: Modeling, Analysis, and Applications," IEEE Trans. On Audio, Speech, And Language Processing, Vol. 17, No. 7, pp. 1394-1407, September 2009.