Journal of Image Processing and Artificial Intelligence
e-ISSN: 2581-3803
Volume 04, Issue 02
4th National Conference on Advancements in Information Technology

Speech Enhancement by Noise Cancellation: A Review

[1]Shraddha C, [2]Chayadevi M L, [3]M A Anusuya
[1]Research Scholar, Department of CS&E, JSS Academy of Technical Education, Bengaluru
[1]Assistant Professor, Department of CS&E, Vidya Vardhaka College of Engineering, Mysuru
[2]Associate Professor, Department of IS&E, JSS Academy of Technical Education, Bengaluru
[3]Associate Professor, Department of CS&E, JSS Science and Technology University, Mysuru

Abstract

It is observed that recognition rate of speech decreases, with the increase of noise in the background. Noise in the background as a tendency to decay the system's robustness. This paper gives a brief survey on speech enhancements methods using various noise cancellation techniques for different SNR's in a noisy environment.

Keywords-*Noise cancellation, Speech enhancement, Hybrid approach*

INTRODUCTION

The main goal of the researchers working on enhancement of speech is to enhance the quality of speech, which is experimented by applying different algorithms. The outcome of enhancement is looked in for gain in intelligibility and quality of corrupted speech signal which are analysed using various techniques of signal processing. Improving the quality of corrupted speech caused by injection noise to the speech signal has become hot cake in the field of speech enhancement and the result of enhanced speech is used in many applications like Voice over Internet Protocol, in telecommunication, aids for deaf patients. An unwanted signal that obstructs the communication or quantity of another communicating signal is said to be noise. Noise cancellation/reduction is a process of removing noise from a signal. Noise reduction [1] is a very challenging problem. In addition, noise characteristics vary in time. Various types of noise can effect quality of conversation.

Signals which are transmitted through wireless devices always tend to get effected by noise. Noise should be handled carefully in communicating systems while transmitting signal as they produces great extent of errors [2].

Various types of noise, based on source of generation:

a) Acoustic noise- unwanted signal produced due to the motion of the object, collision between them or due to vibration.
b) Thermal noise- unwanted signal produced due to heat. It is caused due to irregular movements of energized elements in conducting material.
c) Electromagnetic noise- it's a noise generated at all frequencies in band and noise due to long distance frequency travel and communication
d) Electrostatic noise- generated by the presence of voltage.
e) Channel distortion, echo and fading- relative motion between transmitter and receiver, Causes fluctuation in the received signal and signal to fade. Hence faded signal strength becomes weak and quality of signal produced degrades.
f) Processing noise – noise resulted from D/A processing of signal.

Various types of noise based on different frequency:

a) White noise- noise which is indeterminist and can't be predicted in natural way. All frequencies are with even strength.

b) Band- limited noise- unwanted signal restricted to bandwidth range.
c) Narrowband noise- a signal of 60 Hz produced by electric device.
d) Coloured noise- unwanted signal noise with irregular frequency.
e) Impulsive noise- spontaneous type and generates for short duration signal.
f) Transient noise- spontaneous type, generates for short duration signal and noise pulse are broad in nature.

The pulsating surface caused due to pressure wave is the key indicator of noise. Energy of the sound cannot be measured properly, but pressure of sound can be measured. Acoustic energy is surrogate of sound pressure. Human can bear wanted pleasant sound pressure of 0.00002 Pa and unwanted signal of 200 Pa.

Quantification of sound is made through log-based scale called Decibel (dB) and is measured as Decibel = 10 Log (acoustic energy/ reference energy), where the reference energy is measured as 10*bel= decibel (dB). Sound is compressed to a range of 0–140 db. Scale starts at zero when sound pressure equals the threshold of human hearing. Sound energy measured with the help of sound pressure as dB= 20 log (p/p_o).

Signal should be converted from analog to digital before processing it. Once processed various noise cancellation techniques are applied to improve quality of speech.

a) LINEAR FILTERING OF DIGITAL SIGNAL- it's the basic technique. Linear filtering encompasses signal processing in a time domain, reflected in a change of source signal spectrum content. Linear digital filters consist of two types: Finite Impulse Response filters – FIR filters and Infinite Impulse Response filters – IIR filters.
b) TIME DOMAIN NOISE CANCELLATION- Noise which occurs in the same frequency band as speech can be removed based on its statistical properties. The simplest separation and reduction of stationary noise is possible with the FIR filter in the time domain.
c) NOISE CANCELATION IN FREQUENCY AND SPATIAL DOMAIN- Filtering of unwanted signal done mainly in frequency domain. In spectral filtering technique, signals are analysed, unwanted signals are removed and then finally signals are synthesised. Here filtering is applied on the frequencies whose power is below a certain threshold also called noise floor.
d) NOISE CANCELLATION USING FILTERS- Parameters are automatically adjusted in adaptive filters which helps in removing or in supressing the noise through Adaptive Noise Canceller technique. It uses input obtained from sensors placed in the noise generating fields where the strength of the signal detects to be weak or signal is undetectable. Later which the filter will detect the signal and remove the unwanted signal.
e) RESCURSIVE FILTERS FOR NOISE REDUCTION- recursive filters estimate clean speech signals based on filtering results from a previous frame and noise in the present frame.

Noise Cancellation for Speech Enhancement

Quality of speech degrades with the presence of noise. Once speech is extracted, it's detected for the presence of noise and it's removed. Cancellation of the noise at early stage helps in obtaining only wanted speech, memory utilization and hence enhancing the quality of speech will be easily.

Following are the some of the Noisy

datasets:-NOIZEUS, TIMIT, SPINE (speech in noisy environment), MUSAN (music speech and noise), AURORA, Nosie-92, Chi-ME, AURORA 4, AURORA2, Voice bank, NOISEX and ACE Challenge database of AIR's& noise.

Literature Survey (year vice)

1990-2000 In 90'sthere was drift in research from recognition and enhancement of speech to enhancement of speech by noise cancellation for improving the quality of speech. Dirk Van ompernolle [3] used Switching Adaptive Filters to cancel the noise which showed significant improvement at 10db SNR. A.G. MAHER, R.W. KING and Steve [4] [5] compared various noise cancellation methods of which spectral subtraction showed significant improvement with SNR at10db.Pascal Scalart [6] proposed adaptation algorithm resulted in noise reduction and echo noise control. Dionysis E. Tsoukalas, John N. Mourjopoulos and George Kokkinakis [7] used optimal nonlinear filtering of the short-time spectral amplitude (STSA) showed significant intelligibility gains up to 40%. Nathalie Virag [8] used Single channel subtractive-type improved quality of speech at low SNR's. In this duration research in sighted to enhance the speech by noise cancellation/DE noising the speech.

2001-10 During this period, research opened a gate way for speech enhancement with hybrid approaches to improve quality of speech. Ning Ma, Muvtin Bouchard, and Rujk A.Goubran [9] combined noisy speech model and kalman filtering with masking properties reported no delay and produced PESQ(Perceptual Evaluation of Speech Quality) scores of 0.1 to 0.2 better than the standard Kalman filtering method. Chang Huai YOU, So0 Ngee KOH and Susantorahardja [10] used kalman filtering with masking method performed well for weak speech spectral components. Gwo-Hwa Ju and Lin-Shan Lee [11] combined GSVD- and PCGSVD approach improved speech quality, intelligibility and recognition accuracy for the noise is stationary or nonstationary. Suman Senapati, Sandipan Chakroborty, Goutam Saha [12] used Log Gabor Wavelet yield a higher improvement in Segmental SNR (S-SNR) and also lower Log Spectral Distortion (LSD). Huijun Ding and Ing Yann Soon, Soo Nee Koh, Chai Kiat Yeo [13] used a hybrid Wiener spectrogram filter (HWSF) for effective noise reduction resulted in pleasant sounding speech for human listeners. Nima Yousefian, Ahmad Akbari, Mohsen Rahmani [14] used PLD (Power Level Difference) based speech enhancement method showed effective dealing with non-stationary noise, good performance in case of nearby microphones and independence from the time delay estimation between the received signals. Ching-Ta Lu, Kun-Fu Tseng [15] Integrated intra-frame masking properties of the human ears and the inter-frame SNR variation to adapt the gain factor of a sub band approach removed residual noise. Teddy Surya Gunawan, Eliatham Ambikairajah, Julien Epps [16] proposed forward masking model at 0 dB SNR achieves 30% improvement in Δ PESQ. Kuldip Paliwal, Kamil WO 'jcicki' and Belinda Schwerin [17] proposed spectral subtraction algorithm in the modulation domain achieves better noise suppression. M.A.Anusuya and S.K.Katti [18] attempts to provide a comprehensive survey of research on speech and insight about noisy speech recognition. In this duration research in sighted with enhancing the low SNR speech.

2011-18 In 18's research work was more into enhancement of speech by hybrid approach. Krishnamoorthy and S.R.M. Prasanna [19] combined linear prediction

(LP) residual weighting in the time domain and spectral processing in the frequency domain provided better noise suppression and enhancement. M. Nandini Kala and N Nirmal Singh [20] proposed particle filter resulted in less intrusive background noise. Rongshan Yu [21] proposed soft audible noise masking method incorporating a psychoacoustic model to determine the exact amount of audible noise to be suppressed resulted in better noise suppression. Kit Yan Chana, Pei Chee Yonga, Sven Nordholma, Cedric K and Yiub, Hak Keung Lam [22] proposed Hybrid ANFIS and sigmoid filter showed improvement in terms of recognition accuracy and computational time can be achieved by the hybrid noise suppression filter under various noisy environments in factories. Anil Garg, O.P. Sahu [23] proposed optimized mask filter result showed significant noise suppression and improve in the quality of speech. Ummidala Santosh Kumar and Dr. G. Manmadha Rao [24] combined Digital Audio Effects with Improved Adaptive Kalman Filter showed additive color noise was found to be better in terms of Signal-to-Noise ratio and intelligibility. Wahbi Nabi, oureddine Aloui and Adnane Cherif [25] proposed exploitation of the coherence function and the quadrature mirror filter banks showed effectiveness in term of speech quality and intelligibility when the speech is corrupted by different noise types. Samir Ouelhaa, AbdeljalilA¨ıssa-El-Bey and Boualem Boashash [26] proposed human physiology based time-frequency (TF) representation(HPTF) using Mel filter banks results in an improvement of up to 4.72 dB with respect to SNR, 2.79 w.r.t SSNR and 4.72 dB w.r.t MSE for a speech database signals corrupted with four different noises. In this duration through hybrid approach speech was processed for pleasant listening.

Table 1: *Summary on Recent Advancements in Noise Cancellation Techniques*

Sl.no	Year	Tittle	Authors	Methods	SNR	Dataset And Result	Remarks
1	2012	*A Neural Network Based Effective Quality of Speech Enhancement using Monte Carlo Method*	M. Nandini Kala	particle filter	10 to 40 dB	Noisex-92 less intrusive background noise	Better performance at high SNR's
2	2014	*A hybrid noise suppression filter for accuracy enhancement of commercial speech recognizers in varying noisy conditions*	Rongshan Yu	Hybrid ANFIS and sigmoid filter	5,10, 15 dB	Noisex-92 better noise suppression	only stationary noise can be filtered
5	2016 .	*Speech Enhancement Using Combination of Digital Audio effects with Kalman Filter*	Ummidala Santosh Kumar	Kumar and Dr. G. Manmadha Rao combined Digital Audio Effects with Improved Adaptive Kalman Filter	0,5,10, 15 dB	Noizeus additive color noise was found to be better in terms of Signal-to-Noise ratio and intelligibility	Could have Experimented for other noise types
6	2017	*An improved speech enhancement algorithm for dual-channel mobile phones using wavelet and genetic algorithm*	Wahbi Nabi	exploitation of the coherence function and the quadrature mirror filter banks	-3dB	effectiveness in term of speech quality and intelligibility when the speech is corrupted by different noise types	Experimented for only 2cm distanced microphone
7	2018	*An improved time-frequency noise reduction method using a psycho-acoustic Mel model*	Samir Ouelhaa	human physiology based time-frequency (TF) representation (HPTF) using Mel filter banks	0, 3, 5, 8, 10 and 15 dB	NOIZEUS improvement of up to 4.72 dB with respect to SNR, 2.79 w.r.t SSNR and 4.72 dB w.r.t MSE for a speech database signals corrupted with four different noises	quantify the improvement in terms of classification, measure the quality of extracted features from the DHPTF representation

CONCLUSION

In this paper, we have discussed variety of speech enhancement by noise cancellation technique. Research has opened the gate way for hybrid approach which combines two or more techniques to reduce the noise and enhancing the speech. Further system needs to be designed with an efficient speech enhancement algorithm in pre-processing stage for supressing the noise signals.

ACKNOWLEDGEMENT

The authors remain thankful for Dr. S.K.Katti, for helping us to prepare this manuscript. We thank all the people who have helped us in the preparation of this manuscript.

REFERENCES

1. Urmila Shrawankaret et.al, "*Noise Estimation and Noise Removal Techniques for Speech Recognition in Adverse Environment*", 6th IFIP TC 12 International Conference on Intelligent Information Processing (IIP), DOI: hal-01055058, 2014.
2. Lakshmikanth.S et.al, "*Noise cancellation in Speech Signal Processing-A Review*", International Journal of Advanced research in Computer and Communication Engineering, pp. 5175-5186,ISSN ISSN (Print) : 2319-5940, 2014.
3. Dirk Van Compernolleet et.al, "*Switching Adaptive Filters for Enhancing Noisy and Reverberant Speech from Microphone Array Recordings*", IEEE Conference, pp 833–836, DOI: CH2847-2/90/0000-083 1990.
4. A.G. MAHER et.al, "*A Comparison of Noise Reduction Techniques for Speech Recognition in Telecommunications Environments*", the Institution of Engineers Australia Communications Conference, pp 107–111, 1992.
5. Steve A.Billingset et.al, "*Recurrent radial basis function networks for adaptive noise cancellation*", International Journal of Neural Network, DOI: 10.1016/0893-6080(94)00078-Z, pp. 273-290, 1995.
6. Pascal Scalartet et.al, "*A System for Speech Enhancement in the Context of Hands-Free Radiotelephony with Noise Reduction and Acoustic Echo Cancellation*", International Journal of Speech Communication, pp. 203-214, DOI: 10.1016/S0167-6393(96)00056-8, 1996.
7. Dionysis E. Tsoukalaset et.al, "*Speech Enhancement Based on Audible Noise Suppression*", IEEE transactions on speech and audio processing, pp. 457-514, DOI S 1063-6676(97)07842-5, 1997.
8. Nathalie Virag, "*Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System*", IEEE transactions on speech and audio processing, pp 126–137, DOI: S 1063-6676(99)01630-2, 1999.
9. Ning Ma et.al, "*A Perceptual Kalman Filtering Based Approach for Speech Enhancement*", IEEE conference, DOI: 0-7803-7946-2/03, pp.373-376, 2003.
10. Chang Huai YOU et.al, *"Kalman Filtering Speech Enhancement in Corporation Masking Properties for Mobile Communication in a Car Environment",* IEEE International Conference on Multimedia and Expo, pp. 1343-1346, DOI: 0-7803-8603-5/04/$20, 2004.
11. Gwo-Hwa Ju et.al, *"A Perceptually Constrained GSVD-Based Approach for Enhancing Speech Corrupted by Colored Noise",* IEEE transactions on audio, speech, and language processing, pp. 79–83, DOI: 10.1109/TASL.2006.876868, 2007.
12. Suman Senapatiet et.al, "*Speech Enhancement by Joint Statistical*

Characterization in the Log Gabor Wavelet Domain", International Journal of Speech Communication, pp. 504–518, DOI: 10.1016/j.specom.2008.03.004, 2008.
13. Huijun Ding et.al, "*A Spectral Filtering Method Based on Hybrid Wiener Filters for Speech enhancement*", International Journal of Speech Communication, pp. 259–267, DOI: 10.1016/j.specom.2008.09.003, 2009.
14. Nima Yousefian et.al, "*Using Power Level Difference for near Field Dual-Microphone Speech Enhancement*", International Journal of Applied Acoustics, pp. 1412–1421, DOI: 10.1016/j.apacoust.2009.04.012, 2009.
15. Ching-Ta Lu et.al, *"Perceptual Speech Enhancement Exploiting Temporal Masking Properties Of Human Auditory System"*, IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, pp. 119–134, DOI: 10.1109/TASL.2006.876868, 2007.
16. Teddy Surya Gunawan et.al, "*Single-Channel Speech Enhancement Using Spectral Subtraction in the Short-Time Modulation Domain*", International Journal of Speech Communication, pp. 450–475, DOI: 10.1016/j.specom.2010.02.004, 2010.
17. Kuldip Paliwal et.al, "*Single-Channel Speech Enhancement Using Spectral Subtraction in the Short-Time Modulation Domain*", International Journal of Speech Communication, pp 450–475, DOI: 10.1016/j.specom.2010.02.004, 2010.
18. M.A.Anusuya and S.K.Katti, "*Speech Recognition by Machine: A Review*", International Journal of Computer Science and Information Security, pp. 181-205, ISSN 1947-5500, 2009.
19. P.Krishnamoorthy and S.R.M.Prasanna, "*Enhancement of Noisy Speech by Temporal and Spectral Processing*", International Journal of Speech Communication, pp. 154–174, DOI: 10.1016/j.neunet.2015.03.008, 2011.
20. M. Nandini Kala et.al, "*A Neural Network Based Effective Quality of Speech Enhancement using Monte Carlo Method*", International Journal of Procedia Engineering, pp. 698 – 707, DOI: 10.1016/j.proeng.2012.06.088, 2012.
21. Rongshan Yu, "*A Hybrid Noise Suppression Filter For Accuracy Enhancement Of Commercial Speech Recognizers In Varying Noisy Conditions*", International Journal of Applied Soft Computing, pp. 132–139, DOI: 10.1016/j.asoc.2013.05.017, 2014.
22. Kit Yan Chanaet et.al, "*Speech Enhancement by Joint Statistical Characterization in the Log Gabor Wavelet Domain*", International Journal of Speech Communication, pp. 504–518, DOI: 10.1016/j.specom.2008.03.004, 2008.
23. Anil Garg and O.P. Sahu, "*Cuckoo Search Based Optimal Mask Generation for Noise Suppression and Enhancement of Speech Signal*", Journal of King Saud University on Computer and Information Sciences, pp. 269–277, DOI: 10.1016/j.jksuci.2014.04.006, 2015.
24. Ummidala Santosh Kumar et.al, "*Speech Enhancement Using Combination of Digital Audio effects with Kalman Filter*", International Journal of Computer Speech & Language, pp. 1208–1211, DOI: 978-1-5090-4620-1/16/, 2016.
25. Wahbi Nabiet et.al, "*An Improved Speech Enhancement Algorithm for Dual-Channel Mobile Phones Using Wavelet and Genetic Algorithm*", International Journal of Computers and Electrical Engineering, pp.1-14, DOI:10.1016/j.compeleceng.2017.01.0

20, 2017.

26. Samir Ouelhaaet et.al, "*An Improved Time-Frequency Noise Reduction Method Using a Psycho-Acoustic Mel Model*", International Journal of Digital Signal Processing, pp.108-8, DOI: 10.1016/j.dsp.2018.04.005, 2018.