International Journal of Electrical and Electronics Engineers
Vol. No.8 Issue 01, January-June 2016
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

STATISTICAL BASED SPEECH ENHANCEMENT APPROACHES: A BRIEF REVIEW

Aditi Kapoor[1], Dr.Geeta Kaushik[2], Anil Garg[3]

[1]PG Scholar, [2, 3]Associate Professor, ECE Department, Maharishi Markandeshwar University,(India)

ABSTRACT

Speech enhancement is a sustained standing issue with various applications like automatic recognition, coding of speech signals and hearing aids. Single channel speech enhancement approach is used for enhancement of effective speech depraved by additive backdrop noises. The backdrop noise can have the conflicting impact on our capability to converse without interruption or fluently in very noisy environment, like busy streets. In this paper a review on different approaches and advantage of adaptive filter is discussed in order to improve quality and to increase intelligibility.

Keywords: Adaptive filter, Automatic recognition, Coding of speech signals, Hearing aid, Intelligibility, Quality.

I. INTRODUCTION

Speech signals are extensively used signals among humans, to transfer messages. Speech processing systems are pre-owned in a wide range of applications such as speech recognition for automatic information, speech coding for communications and regarding aids to hearing impaired persons speech pre-processing is used. These systems are described under the acceptance that corruptive backdrop noises are absent. In a noisy situation, speech enhancement is recommended to improve the execution of these systems [1].To describe algorithms a word is used that is called speech enhancement, which can be used to increase intelligibility, improve the quality, improves the performance of the voice communication systems and decrease the hearing fatigue of noisy speech [1]. For concluding speech intelligibility considering normal-hearing and hearing-impaired auditor various index are mentioned in this paper.A simple adjustment to the AI measure explaining for non-linear processing in the existence of additive noise is explained also speech intelligibility index is described to determine whether clinically obtainable measure of audibility is more précised than the pure-tone average at concluding the lexical abilities concerning children wear hearing aids [2].

II. LITERATURE REVIEW

This section explains a short description regarding the normal and impaired listeners and the various algorithms used for speech enhancement in order to improve the quality and to increase the intelligibility.

2.1 Complicationin Considering Speech In Noise: How Much Poor Are The Hearing Impaired?

The issue investigated first that:how much normal-impaired auditoris bad than normal in their ability to figure out speech in noise? This is often evaluated by supposing the speech-to-noise ratio recommended achieving a certain amount of intelligibility, such as 50% precise. This ratio is termed the speech receptionthreshold (SRT) as well as it is usually expressed in dBthe models flop to accurately conclude speech intelligibility (SI) along nonlinearly processed noisy speech [2]. Latest studies conclude SI for normal-hearing auditor situated on signal-to-noise ratio measure with envelope domain, with framework about sEPSM. Conclusion shows that with only modeling the failure of audibility, the model cannot explain for SRT with hearing impaired (HI) people along stationary noise correlated to normalhearing (NH). Though by only creating the loss of audibility, affecting model cannot detail for higher SRT concerning HI persons in stationary noise in comparison to NH [2]. Also the speech intelligibility index is used to determine whether clinically obtainable measure of audibility is more précised than the pure-tone average at concluding the lexical abilities concerning children wear hearing aids. (CHA) children with hearing loss determine poorer execution thanchildren with normal hearing (CNH) on nonword repetition; tests of word and receptive vocabulary. Word andreceptive vocabulary and nonword repetition are predicted stronger by aided SII rather than Pure- tone average(PTA) [2]. Aided speech intelligibility index(SII) persist a significant predictor of receptive vocabulary and nonword repetition after accounting for PTA. Unlike PTAaided SII consolidate hearing aid amplification characteristics as well as frequency weightings and may produce a more valid assessment of the child's access along with ability to grasp from auditory input in real-world environments [3].

2.2 Explanation Of Difficulty In Considering Speech In Noise

The articulationindex (AI)produces a way of evaluating the effect of audibility on speech intelligibility.The link among audibility and hearing aid capability was tested in cohort of sufferer who achieves hearing aids through veteran's administration.Part of capability comprised with two hearing specific survey and self-reported. Valuation of global satisfaction and hearing aid adherence.The elementary target was to determine direct communication between enhanced audibility measured with AI and comprehensive capability of hearing aid.As shortcut AI approach used to fit hearing aids author needed to resolve whether AI methods layout for use in clinic provide ditto report as the more conventional and stagnant AI calculation [4]. Throughout the hearing aid fitting the desiderate gain the individual frequency was calculated for any party using national acoustic laboratories. The frequency gain feedback of hearing gain was accommodating to contribute the finest available match to choose gain for a conversational level input signal. Basically AI is determined by counting up number of dB by that speech peaks beat threshold at individual frequency with the condition that speech range at each frequency will not beat 30 dB .To determine the progress in usable speech information with amplification it has been determined the change in AI among aided and unaided action for each ear. Profit is determined by subtracting aided score by unaided score. Sufferer who achieves bigger speech audibility among hearing aids release they are limited able to recognize speech in background noise. This is apparently because all sufferers were capable with linear hearing aids that magnify both speech and noise[4].

2.3 Concept Of Synthetic Speech

Basicallymodification to synthetic and natural speech which desire improving intelligibility in noise. The present study analyze the benefits of speech modification algorithm in an extensive speech intelligibility evaluation also quantifies the equivalent intensity variation, describe as the amount in decibels so unmodified speech would essential to be adjusted by procedure in order to accomplish the related intelligibility as modified speech. As comparing plain natural speech with synthetic speech the synthetic speech in noise will always remain less intelligible, but modified synthetic speech decreased thisdeficit by significant amount [5]. For decomposing the standard digital signal processing this technique Minimal-Pair ABX is proposed for the evaluation of speech illustration in the low resource setting or zero. This framework need same-different word discrimination function that does not lean on training a classifier, nor phonetically labeled data. This framework is applied for computing PLP and MFC coefficients in order to decompose the standard signal processing pipelines [6].

2.4 Extending The Articulation Index For Non Linear Distortion

A simple adjustment to the AI measure explaining for non-linear processing in the existence of additive noise. The adjustment was based on subsequent two observations. First the input SNR in limited band cannot be enhanced following any mode of non-linear processing. Secondly for the effect of non-linear or suppression effect on the target envelope when $\hat{S} < S$ than when $\hat{S} > S$ the output envelope $\hat{S}$ reflects more reliable [7].

2.5 Effects Of Compression On Intelligibility And Sound Quality

A theoretical scheme which is used to evaluate potential aspect that can impact intelligibility of processed speech. It is consider if distortions are perfectly restrained than considerable gain in intelligibility can accomplished. To test this consideration author presented composed speech distortions with processed speech in which intelligibility are conducted with human listeners. The 7 normal hearing listeners cooperate in a total 32 conditions. For individual SNR level processing action comprised speech processing accepting three contrasting speech enhancement algorithms with 1) no constraints imposed, 2) Region II constraints,3) Region I constraints, 4) Region I and II constraints, and 5)Region III constraints. According to the test it demonstrate that region I and region I+II constraints are highest tough in terms of resigned constantly huge gain in intelligibility separate the speech enhancement algorithmused [8]. From the three enhancement algorithms tested the wiener algorithm is approved when commanding region I and region I+II constraints .as this algorithm allows the biggest gains in intelligibility for both SNR levels tested. Comparatively concentrate on minimizing a squared error criterion one can concentrate rather on minimizing given criterion subject to proposed constraints.SE algorithm used to be construct so as to enlarge a metric (e.g. SNR ESI, AI) that is familiar to correspond highly with speech intelligibility.For reason SE algorithm need to perform to enlarge a corrupt square error cost function as done by better statistical model based algorithms. Algorithms that enlarge SNRESI metric are possible to implement substantial gains in intelligibility [8].

2.6 Prediction of Intelligibilty

The speech is situated on envelope power spectrum model (sEPSM) surmise (SNR_{env}) ratio succeeding modulation-frequencyparticularprocessing. The sEPSM is confined to conditions along stationary interferers, as long as the long-term integration about envelope power moreover cannot account for elevated intelligibility typically achieved with fluctuating maskers. Presently a multi-resolution version regarding sEPSM is conferred where SNR_{env} is approximated in temporal segments along modulation filter vulnerable duration. The mr-sEPSM(multi-resolution version) framework consult single decision metric, $SNR_{env,}$approximated from normalized variance about envelope fluctuations about noisy speech along with noise alone. The span of

International Journal of Electrical and Electronics Engineers
Vol. No.8 Issue 01, January-June 2016
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

segments used as short- term SNR_{env}estimation was assumed to be reversed associated to center frequencies about modulation bandpass filters in effective model. The mr-sEPSM is determined to account intelligibility accomplish in condition along stationary also fluctuating interferers, also noisy speech distorted through reverberation either spectral subtraction [9]. Another technique known as Hearing –Aid Speech Perception Index(HASPI)is considered for concluding speech intelligibility considering normal-hearing and hearing-impaired auditor. The HASPI is established on an illustrative of auditory periphery that associates refinement due to hearing loss[10]. The index consider the envelope and temporal fine frame output of auditory illustrative for reference signal to output of illustrative for the signal beneath analysis. The auditory illustrative for reference signal is confirmed for normal hearing although the model for test signal associates the peripheral hearing failure. For normal hearing auditor the refined signal is output of normal hearing illustrative having depraved signal as input. HASPI is commence to give definite intelligibility, surmising for a broad domain of signal degradation in addition to speech depraved by noise along with nonlinear distortion, speech refined applying frequency compression, noisy speech refined through noise-suppression algorithm moreover speech where the upraised frequencies are recoveredby output about noise vocoder [10].Also there are various algorithms of speech enhancement which leads to increment of intelligibility ,improves quality, reduction of hearing fatigue connected with noisy speech and also it improves the performance of voice communication system. Adaptive wiener filtering procedure lean on variation of filter transfer function from sample to sample situated on speech signal statistics; the local variance and local mean. It is achieved in time domain comparatively than in frequencydomain in order to hold for time varying aspect of the speech signals. Local variance and estimated local mean of signal are the two in this method that are oppressed [11]. It is simulated that additive noise is of zero mean and have white nature with variance. Recognize short segment of speech signal in which signal is simulated to be stationary. The suggested method correlated to traditional frequency domain wiener filtering wavelet denoising methods and spectral subtraction using other speech quality metrics. The reason for using the adaptive wiener filtering is that as there were some problems to the spectral subtraction and the wiener filtering. Complication with spectral subtraction method is representation of residual noise called musical noise. At the time of silence period spectral subtraction does not attenuate noise enough. Favorable part that is distorted of original speech is one of the complications with wiener filtering. That's why the author suggested adaptive wiener filtering as it has finest completion in comparisonto other speech enhancement methods. The suggested filter has advantage that it works as a single input when it is relying only on noisy signal [11].

III. CONCLUSION

In this paper overview of speech enhancement is discussed along with the classification of techniques in speech enhancement. Also the Statistical based techniques along with their properties and a limitation has been explained. Comparison between classical and Bayesian estimators and analysis of SI index on different approaches has been explained. Also studied the variation of filter transfer function from sample to sample situated on speech signal statistics; the local variance and local mean for an adaptive wiener filtering approach. This paper concludes that adaptive filter improve SI index better than other approaches.

REFERENCES

[1] Rowden, Chris, speech processing(Essex series in telecommunication and information system, 1992).

[2] ChristophScheidiger, SorenJorgensen, TorstenDau, Modeling Speech Intelligibility in Hearing ImpairedListeners, European Acoustics Association,100(3),2014.

[3] Derek J. Stiles, Ruth A.Bentler ,andKarlaK.McGregora, The Speech Intelligibility Index and the Pure-ToneAverage as Predictors of Lexical Ability in Children Fit With Hearing Aids, Journal of Speech, Language, and Hearing Research , 55(3), 2012, 764–778.

[4] Pamela E. Souza, PhD; Bevan Yueh, MD; Margaret Sarubbi, MS; Carl F. Loovis, Fitting hearing aids with theArticulation Index: Impact on hearing aid effectiveness, Journal of Rehabilitation Research and Development,37(4), 2000, 473–481.

[5] Martin Cooke, Catherine Mayo,CassiaValentini-Botinhao ,YannisStylianou, Bastian Sauert , Yan Tang, Evaluating the intelligibility benefit of speech modifications in known noise conditions, speech communication, 55(4),2013, 572–585.

[6] Thomas Schatz, VijayadityaPeddinti, Francis Bach, Aren Jansen, HynekHermansky, Emmanuel Dupoux, Evaluating speech features with the Minimal-Pair ABX task: Analysis of the classical MFC/PLP pipeline, 88, HAL Id: hal-00918599, 2013, 1-5.

[7] Philipos C. Loizou, Extending the articulation index to account for non-linear distortions introduced by noise-suppression algorithms, Acoustical Society of America, 130(2), 2011, 986–995.

[8] Philipos C. Loizou, Gibak Kim, Reasons why current speech-enhancement algorithms do not improve speech intelligibility and suggested solutions, IEEE Trans Audio Speech Lang Processing, 19(1), 2011, 47–56.

[9] SørenJørgensena, Stephan D.Ewert, TorstenDau, A multi-resolution envelope-power based model for speech intelligibility, J. Acoustical Society of. America, 134(1),2013, 436–446.

[10] James M. Kates ,Kathryn H. Arehart, The Hearing-Aid Speech Perception Index, in Speech Communication,65, 2014,75-93.

[11] Marwa A. Abd El-Fattah, Speech enhancement with an adaptive Wiener filter, International Journal of Speech Technology, 17(1), 2014, 53–64.