International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Liner Prediction and DCT based De-noising of Speech Signal

Diksha Sharma[1], Rupinder Kaur[2]

[1, 2]DIET, Kharar, Punjab, India

Abstract: ***The proposed algorithm adopts LP residual as one of the sparse representation of speech, considering it is feasible and advantageous, as analysed in the previous section. To make full use of the sparsity of speech, DCT coefficients are also included to contribute as a measurement. The proposed algorithm aims to recover the clean speech, whose LP residual and DCT coefficients are both sparse, via solving an optimization problem under a series of constraints.***

Keywords: DCT. Linear Prediction, Speech Enhancement

1. Introduction

Speech processing is a diversified area of research, primarily due to the fact that speech processing will always be needed in such important areas as telecommunication technologies. For example, speech processing in real-time and transparent speech transmission sets many challenges for researchers. The requirements for the quality of processed speech are increasing constantly. Since speech has many aspects, distinguishing between different utterances is not enough; all the other information related to e.g., identification of the speaker, the speaker's disposition or other emphases included in the speech signal should also be delivered to the receiver. In phonetics, the characteristics of speech, including intonation, enunciation and length of the phoneme, i.e., quantitative, are all called the prosody of the speech. Depending on the application and the required quality, all the perceptually important characteristics of speech must be maintained during processing. Future wireless or mobile communication applications, for example, will no longer tolerate speech quality which is inferior to the toll quality.

2. Related Works

Yong Xu, Jun Du, Li-Rong Dai, and Chin-Hui Lee; (2014) "An Experimental Study on Speech Enhancement Based on Deep Neural Networks," Here we presents a regression-based speech enhancement framework using deep neural networks (DNNs) with a multiple-layer deep architecture [1]

Shenoy, R.R. ; Seelamantula, C.S.(2014) Frequency domain linear prediction based on temporal analysis. This states that Frequency-domain linear prediction (FDLP) is widely used in speech coding for modeling envelopes of transients signals, such as voiced and unvoiced stops, plosives, etc. FDLP fits an auto regressive model to the discrete cosine transform (DCT) coefficients of a sequence. [2]

Ante Juki´c1, Toon van Waterschoot2, Timo Gerkmann1, Simon Doclo1(2014) Speech Dereverbersation with Multi-Channels Linear Prediction And Sparse Priors For The Desired SIGNAL. This states that the quality of recorded speech signals can be substantially affected by room reverberation. [3]

Kumar, N. ; Van Segbroeck, M. ; Audhkhasi, K. ; Drotar, P. ;Narayanan, S.S. (2014) Fusion of diverse de-noising systems for robust automatic speech recognition. This states that We present a framework for combining different de-noising front-ends for robust speech enhancement for recognition in noisy conditions. [4]

Peddinti, V. ; Hermansky, H.(2013) Filter-bank optimization for Frequency Domain This states that The sub-band Frequency Domain Linear Prediction (FDLP) technique estimates autoregressive models of Hilbert envelopes of sub band signals, from segments of discrete cosine transform (DCT) of a speech signal, using windows. [5]

Amro, I. Higher Compression Rates for Code Excited Linear Prediction Coding Using Lossless Compression (2013).This states that In this paper, we exploit the Hamming Correction Code Compressor (HCDC) Code Excited Linear Prediction frame's Parameter. These parameters includes Linear Prediction Coefficients, [6]

S.C. Shekokar, M. B. Mali(2013) Speech Enhancement Using DCT. This states that The speech enhancement problem comprises of various problems characterized by the type of noise source, the nature of interaction between speech and noise, the number of sensor signals (microphone outputs) available for enhancement and the nature of the speech application. [7]

Meng Guo ; Elmedyb, T.B. ; Jensen, S.H. ; Jensen, J.(2010) Analysis of adaptive feedback and echo cancelation algorithms in a general multiple-microphone and single-loudspeaker system. This states that In this paper, we analyze a general multiple-microphone and single-loudspeaker system, where an adaptive algorithm is used to cancel acoustic feedback/echo and a beam former processes the feedback/echo canceled signals. [8]

Wen Jin Xin Liu ; Scordilis, M.S. ; Lu Han (2009) Speech Enhancement Using Harmonic Emphasising Adaptive Comb Filtering. This states that An enhancement method for single-channel speech degraded by additive noise is proposed.[9]

Zhimin Xiang and Yuantao GuAdaptive; (2013) "Speech Enhancement Using Sparse Prior Information," In recent years, sparse representation is adopted to improve the quality of noise corrupted speech. [10]

Yongjun He ,Jiqing Han ; Shiwen Deng ; Tieran Zheng ; Guibin Zheng; (2012) "A Solution to Residual Noise in Speech De-noising with Sparse Representation," Sparse representation has been extensively investigated in signal processing community. [11]

.

Van Hamme, H. ; Cranen, B. ; Boves, L.;(2010) "Compressive Sensing for Missing Data Imputation in Noise Robust Speech Recognition," An effective way to increase the noise robustness of automatic speech recognition is to label noisy speech features as either reliable or unreliable (missing), and to replace (impute) the missing ones by clean speech estimates. [12].

Sigg, C.D., Dikk, T. Buhmann, J.M.; (2010) "Speech Enhancement with Sparse Coding in Learned Dictionaries," The enhancement of speech degraded by non-stationary interferers is a highly relevant and difficult task of many signal processing applications. [13]

Anis Ben Aicha; (2005), "Adaptive Noise Smoothing for Speech Enhancement Filters," Noise estimation is one of the main keys for successful speech enhancement processing. Traditional noise estimators based on inter-frame smoothing of the estimated power spectrum (PSD) of the background noise leads to artificial artifacts on the residual noise well known as 'musical noise'. [14]

Yi Hu,Loizou, P.C (2003). "A Generalized Subspace Approach for Enhancing Speech Corrupted by Colored Noise," A generalized subspace approach is planned for enhancement of speech corrupted by colored noise. [15]

Loizou, Philipos, Kamath, Sunil; (2002) "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," The spectral subtraction method is a well-known noise reduction technique. Most implementations and variations of the basic technique advocate subtraction of the noise spectrum estimate over the entire speech spectrum. [16]

3. Algorithm

The presented work is implemented in following steps:

- Speech Signal Acquisition
- x(n) = $\sum_{k=1}^{p} a_k . x(n-k) + r(n)$
- DCT Coefficients Computation
- Sparse DCT coefficient matrix Generation
- Collecting Non-zero values from sparse dct matrix
- Generation of LP residual Matrix r = A X
- Generation of LP residual Matrix r = A X
- LP residual Computation

The DCT coefficients of the speech signal are extracted using the following formula:

$$y(k) = w(k) \sum_{n=1}^{N} x(n) \cos \frac{\pi(2n-1)(k-1)}{2N}, \qquad k = 1, \ldots, N$$

where

$$w(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k = 1 \\ \sqrt{\frac{2}{N}}, & 2 \le k \le N \end{cases}$$

The residual noise is extracted using linear predictor coefficients and de-noised sound is received after taking the difference of noisy signal and residual noise. The snr before and after is computed in order to measure the performance of the algorithm.

Results and Conclusion

The presented algorithm has been implemented on different audio signal of different frequency. The signal to noise ratio before and after the applying of the proposed algorithm is shown in result table below. The improvement in signal to noise ratio proves the fairness of the presented algorithm when de-noising the audio signal.

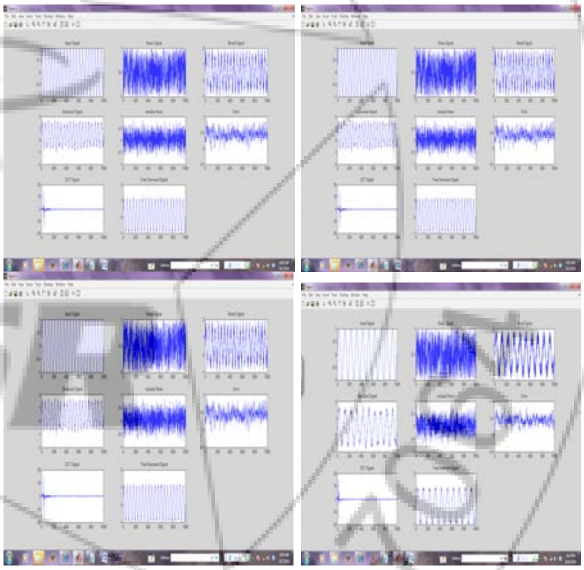
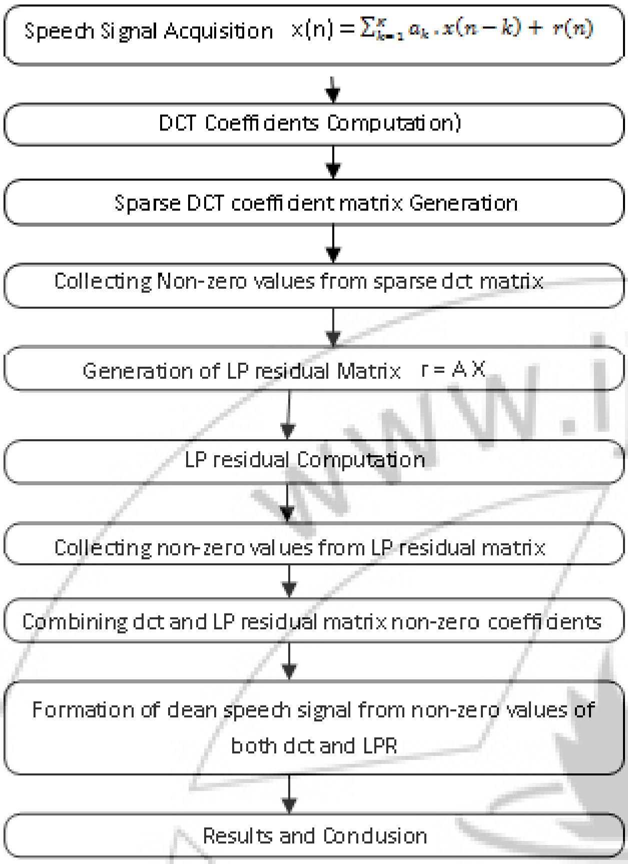


Table 1: SNR before and after De-noising

Sr. No.	*Test Signal Freq.*	*SNR Before De-noising (db)*	*SNR After Denoising (db)*
1	5	4.004	8.878
2	10	4.026	8.560
3	20	4.053	8.261
4	30	3.955	8.138
5	40	3.794	8.422

Flow Chart



References

[1] Anis Ben Aicha", *IEEE* "Adaptive Noise Smoothing for Speech Enhancement Filters**"**, vol. 27, no. 2, pp. 113-120, Apr. 1979.enhancing speech corrupted by coloured noise", in *Proc. ICASSP2002*, vol. 4, pp. IV-4164, May 2002

[2] Sigg, T. Dikk and J. M. Buhmann, "Speech enhancement with sparse coding in learned dictionaries", in *Proc. ICASSP2010*, pp. 4758-4761, March 2010.

[3] J. F. Gemmeke, H. Van Hamme, B. Cranen and L. Boves," Compressive sensing for missing data imputation in noise robust speech recognition", *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 272-287, Apr. 2010.

[4] H. Yi and P. C. Loizou, "A generalized subspace approach for enhancing speech corrupted by coloured noise", *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 4, pp. 334-341, July 2003

[5] H. Yongjun, H. Jiqing, D. Shiwen, Z. Tieran and Z. Guibin, "A solution to residual noise in speech de-noising with sparse representation", in *Proc. ICASSP2012*, pp. 4653-4656, March 2012.

[6] Yong Xu, Jun Du, Li-Rong Dai, and Chin-Hui Lee, "An Experimental Study on Speech Enhancement Based on Deep Neural Networks," *IEEE Journal of Selected Topics in Signal Processing, vol.* 21,pp. 65-68 , ISSN.1070-9908 , Jan 2014.

[7] *Zhimin Xiang and Yuantao Gu*Adaptive, "Speech Enhancement Using Sparse Prior Information", Acoustic, ICASSP, IEEE international conference, pp.7025-7029, May 2013.

[8] Shenoy, R.R. ; Seelamantula, C.S.(2014), " Frequency domain linear prediction based on temporal analysis Acoustics", Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on pp.2629 – 2633, may 20014.

[9] Ante Juki´c1, Toon van Waterschoot2, Timo Gerkmann1, Simon Doclo1 (2014), "Speech Dereverbersation with Multi-Channels Linear Prediction and Sparse Priors for the Desired SIGNAL", Hands-free Speech Communication and Microphone Arrays (HSCMA), 2014 4th Joint Workshop pp.23-26.

[10] Kumar, N. ; Van Segbroeck, M. ; Audhkhasi, K. ; Drotar, P. ;Narayanan, S.S. (2014), " Fusion of diverse de-noising systems for robust automatic speech recognition", Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference pp. 5557 – 5561 may 2014.

[11] Peddinti, V. ; Hermansky, H.(2013) ,"Filter-bank optimization for Frequency Domain" , Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference pp**.**7102 – 7106 may 2013.

[12] Amro, I.(2013), "Higher Compression Rates for Code Excited Linear Prediction Coding Using Lossless Compression", DOI: 10.1109/ICCE-Berlin.2013.6698057 IEEE Third International Conference.

[13] S.C. Shekokar, M. B. Mali (2013), "Speech Enhancement Using DCT", International Journal of Engineering and Advanced Technology (IJEAT) Volume-2, ISSN: 2249 – 8958, Issue-5, June 2013.

Author Profile

Ms. Diksha Sharma is pursuing her M.Tech in ECE from DIET, Kharar, Mohali, INDIA. Her field of interest is in digital signal processing based system design and application.