

Chapter 51

A Novel Design Approach of Subband Coder and Decoder of Speech Signal Using Log Normal Probability Distribution

Sangita Roy and Sheli Sinha Chaudhuri

Abstract In modern speech communication less bandwidth and low data rate is very essential from portable systems with limited storage capacity. Researchers are concerned with the tradeoff between bandwidth and SNR (Signal-to-Noise), BER (Bit Error Rate). Speech signal can be compressed below 64 Kbps taking care of SNR above 30 dB, and BER below 10^{-5} . Here the authors proposed Log Normal Distribution in the design of Subband Coder and Decoder of Speech Signal taking care of the SNR and BER criterion with data rate 9.3316 Kbps.

Keywords Subband · BER · Probability distribution · SNR

51.1 Introduction

Speech is the basic form of human communication. Speech communication is of immense importance as the speech signal is different from any other sounds. In the last few decades speech communication has been of great concern due to the fast growing technologies. The communication channel is probabilistic in nature. Therefore, measuring parameters are essential to ensure the quality of the speech signal through the channel. The parameters for good speech signal transmission are (i) low bit rate, (ii) more than 30 dB SNR and (iii) BER less than 10^{-5} [1–3]. Speech signals follows some probability density functions (PDF), i.e. Gaussian [4], Rayleigh [5], Log Normal [6], etc. The authors identified Log Normal distribution for characterization of speech signal and compared it with tested speech signal distributions, i.e. standard power spectral density of speech signal, Gaussian

S. Roy (✉)

ECE Department, Narula Institute of Technology, WBUT, Kolkata, India
e-mail: roysangita@gmail.com

S.S. Chaudhuri

ETCE Department, Jadavpur University, Kolkata, India
e-mail: shelism@rediffmail.com

© Springer India 2015

K. Maharatna et al. (eds.), *Computational Advancement in Communication Circuits and Systems*, Lecture Notes in Electrical Engineering 335,
DOI 10.1007/978-81-322-2274-3_51

463

Distribution [7], Rayleigh Distribution [8]. By employing these PDFs (Probability Density Functions) communication channels as well as PSD (power spectral density) of speech signal can be used efficiently. Subband coding (SBC) is a kind of transform coding. A signal is divided into a number of different frequency bands and encodes each one independently. It enables data reduction by discarding information about frequencies which are masked. The result differs from the original signal, but if the discarded information is chosen carefully, the difference will not be noticeable, or more importantly, objectionable [9–12].

51.2 Literature Survey

The paper—“A low-complexity audio data compression technique using subband coding (SBC) and a recursively indexed quantizer (RIQ)” compared SBC and RIQ with conventional coding techniques. The system shows SNR 2–5 dB higher than that of SNRs of other coders of similar computational complexity of wideband audio signals [7]. The basic concept of “Frequency Domain Coding of Speech” methods is to divide the speech into frequency components by a filter bank (subband coding), or by a suitable transform (transform coding), and then encode them using adaptive PCM (Pulse Code Modulation). Three basic factors of the design of coders are: (1) the type of the filter bank or transform, (2) the choice of bit allocation and noise shaping properties and (3) the control of the step-size of the encoders. Short-time analysis/synthesis, practical realizations of subband and transform coding are interpreted within this framework. Spectral estimation, models of speech production, perception and the “side information” can be most efficiently represented and utilized in the design of the coder (particularly the adaptive transform coder) to control the dynamic bit allocation and quantizer step-sizes. Recent developments and examples of the “Vocoder-driven” adaptive transform coder for low bit-rate applications is also discussed [8]. In digital telecommunication systems different signals are processed with different sampling rates, leading to significant errors. In “Subband Coding of Speech Signals Using Decimation and Interpolation”—a structure of a two-channel quadrature mirror filter with low pass filter, high pass filter, decimators and interpolators, is proposed to perform subband coding of speech signals in the digital domain. The performance of the proposed structure is compared with the performance of delta-modulation encoding systems. The results show that the proposed structure significantly reduces error and achieves considerable performance improvement compared to delta-modulation encoding systems [13]. Gaussian Distribution is well suited for describing the Power Spectral Density of Speech Signal. In statistical voice activity detection (VAD) Rayleigh Distribution has been used as the distribution has longer asymmetric tail than Gaussian distribution. MMSEEs (Minimum Mean Square Estimators) for speech enhancement have employed various PDFs, such as Gaussian Distribution, Log Normal Distributions, etc.

51.3 Basic Principles of the Proposed System Model [8]

51.3.1 Design Procedure for Subband Coding for Speech Signal

The Power Spectral Density (PSD) of a voice signal has been considered to be restricted to 3.5 kHz only, Power Spectral Density to be in watt/Hz or dB (Fig. 51.1).

In this figure frequency axis is divided into a number of subbands (say $0-f_1$, f_1-f_2 , f_2-f_3 , f_3-f_4 , etc.). The frequency band ($0-f_1$) is baseband signal, whereas (f_1-f_2), (f_2-f_3), (f_3-f_4), etc. are bandpass signals. Each band will be translated into baseband by multiplying with the lowest frequency component of the said subband. Here seven subbands have been considered (Fig. 51.2).

Fig. 51.1 PSD of speech signal

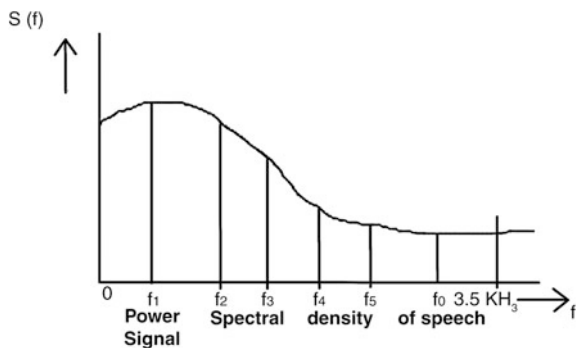
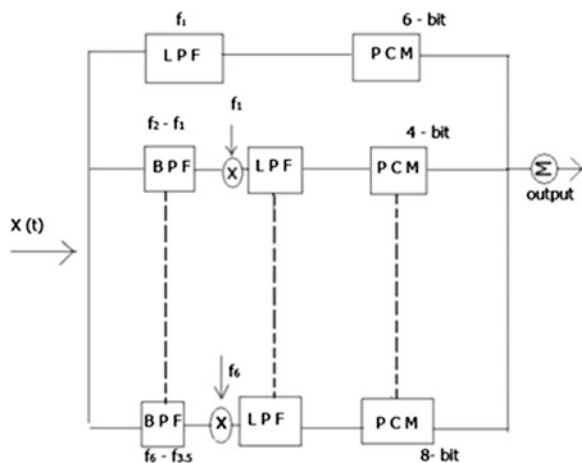


Fig. 51.2 Block diagram of subband coder



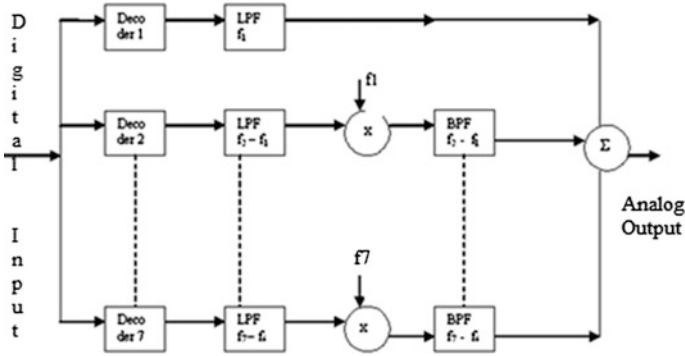


Fig. 51.3 Block diagram of subband receiver

The transmitter consists of one LPF and six BPFs. All BPFs outputs are multiplied by the lowest frequency component of those bands at the multiplier block. Then outputs are PCM and then added by summer. Finally the summed output is put into channel (Fig. 51.3).

At the receiver signals are decoded by seven decoders. Then each signal is passed through LPF of cut-off frequency f_1, f_2-f_1, f_3-f_2 , etc. From the second to the seventh signal outputs are multiplied by their respective lowest frequency components and then passed through BPFs of f_2-f_1, f_3-f_2 , etc. Then the outputs are summed up to get a replica of the original signal.

51.4 Proposed Method with Log Normal Distribution PSD

Speech Coding follows different probability distributions. Authors have already worked with Gaussian and Raleigh Distributions [7, 8]. Here, they have chosen Log Normal Distribution and followed the same procedure as earlier. The results are shown below.

51.4.1 Mathematical Validation Using MATLAB Simulation

(See Figs. 51.4, 51.5 and 51.6; Tables 51.1 and 51.2)

Fig. 51.4 Cumulative data rate versus frequency

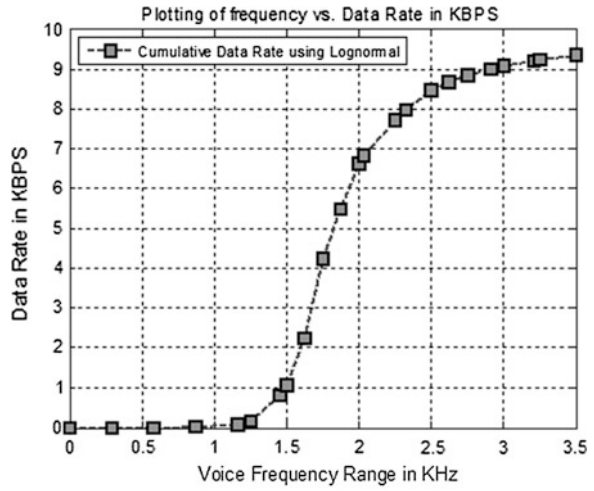


Fig. 51.5 SNR versus frequency

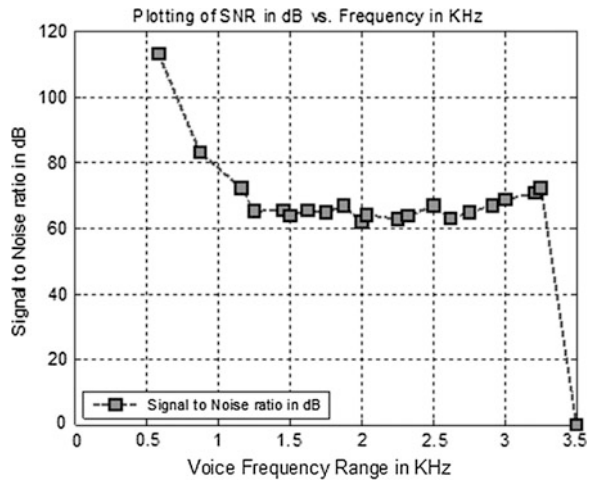


Fig. 51.6 Probability of bit error

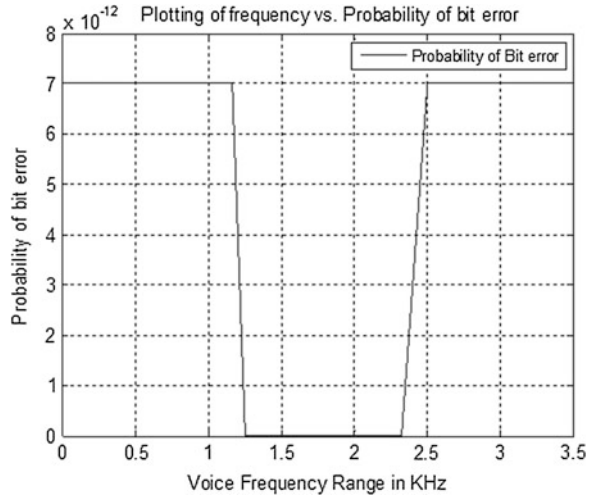


Table 51.1 Data Rate, SNR_{min}, BER of Log Normal Distribution

Nature of speech	Bit rate (Kbps)	SNR _{min} (dB)	Probability of bit error
Log normal distribution	9.3316	61.9169	10 ⁻⁵

Table 51.2 Comparative list of different distribution with their data rates

Sl no.	Name	Data rate (Kbps)	SNR _{min} (dB)	Probability of bit error
1	Existing data rate [7, 8]	64	49	Below 10 ⁻⁵
2	Staircase approximation [7, 8]	19.5	49	Below 10 ⁻⁵
3	Gaussian distribution [7, 8]	12.0128	56.5828	Below 10 ⁻⁵
4	Gaussian with perception criterion [8, 13–15]	11.0959	48.9566	Below 10 ⁻⁵
5	Gaussian with reduced bit allocation	10.9494	48.9566	Below 10 ⁻⁵
6	Gaussian with more reduced bit allocation	10.7175	48.9566	Below 10 ⁻⁵
7	Gaussian with more reduced bit allocation	10.4867	48.9566	Below 10 ⁻⁵
8	The BELL system technical journal, october 1976 [14]	9.6	–	–
9	Rayleigh distribution [8]	9.4875	68.99	Below 10 ⁻⁵
10	Log normal distribution	9.3316	61.9169	Below 10 ⁻⁵

51.5 Conclusion

It is evident from the above discussion that both subband coding and existing 64 Kbps line have almost negligible probability of bit error but subband offers lowest data rate bandwidth ever possible. Authors have used different probability distribution for speech coding for validation. Log Normal Distribution shows the least. Therefore it can be deduced that subbanding generates all the possible significant footsteps towards data rate as well as bandwidth savings without losing any significant information and probability of bit error is also least or may be said negligible. PCM requires high bandwidth as well as data rate. But PCM and DM have almost the same SNR up to 30 dB. After 30 dB PCM shows performance-wise better results than DM. It has been shown by Matlab program. If more subbands are used, data rate can be reduced more and more accurate approximation of the original voice signal can be reconstructed. Therefore, authors can conclude that communication engineering will be immensely benefited by using this scheme. There are a lot more distribution support speech signals. These distributions can be simulated and results can be found out.

References

1. L.W. Couch II, *Modern Communication Systems Principles and Applications* (Prentice-Hall of India Private Limited, Delhi, 1995)
2. G.J. Proakis, *Digital Communications*, 4th edn. (Mcgraw-Hill International Edition, New York)
3. H. Taub, D.L. Schilling, *Principles of Communication Systems*, 2nd edn. (Tata Mcgraw-Hill publishing company limited, Noida)
4. I. Tashev, A. Aecero, *Speech Technology Group* (Microsoft Research, Redmond)
5. Y. LI, J. Chen, H. Tan, Voice activity detection under Rayleigh distribution. *J. Electron.* (Springer), **26**(4), 552–556 (2009)
6. M. Suman, T.V. Bhargava, G.P. Teja, K.B.N.P. Kumar, Speech enhancement and recognition compressed speech signal in noisy reverberant conditions, *IRJSP* **2**(2) (2011). ISSN 2249-6505
7. S. Roy, D.B. Gupta, P.K. Banerjee, Studies and implementation of subband coder and decoder of speech signal, in *Proceedings of NCECS* (2012), pp. 8–16
8. S. Roy, D.B. Gupta, S.S Chaudhuri, Studies and implementation of subband coder and decoder of speech signal using raleigh distribution, *LNEE298*, Springer, doi:[10.1007/978-81-322-1817-3_2](https://doi.org/10.1007/978-81-322-1817-3_2)
9. Z. Peric, J. Nikolic, An adaptive waveform coding and its application in speech coding. *Digit. Signal Process.* (Elsevier) **22**, 199–209 (2012)
10. Y.-J. Chen, R.C. Maher, *Sub-band Coding of Audio using recursively indexed quantization* (Department of Electrical Engineering and Center for Communication and Information Science, University of Nebraska, Lincoln, 1995)
11. M.J. Tribolet, R.E. Crochiere, Frequency domain coding of speech. *IEEE Trans. Acoust. Speech Signal Process.* **Assp-27**(5), 550–558 (1979)
12. A.M. Aziz, Subband Coding of Speech Signals Using Decimation and Interpolation, in *13th International Conference ASAT-13*, Military Technical College, KobryElkobbah, Cairo, 26–28 May 2009

13. B. Rivet, L. Girin, C. Jutten, Log—Rayleigh Distribution: a simple and efficient statistical representation of log-spectral coefficients, in *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15 March 2007
14. R.E. Crochiere, S.A. Webber, N. Flanagan, Digital coding of speech in subbands. *BELL Syst. Tech. J.* (1976)
15. E. Zwicker, U. Zwicker, *Psychoacoustics, Facts and Models* (Springer, Berlin, 1990)
16. P.G. Knutson, K. Ramaswamy, J.W. Richardson, Subband adpcm voice encoding and decoding, PCT/US2000/034410, July 2001
17. C.F. Szczutkowski, Subband encoding method and apparatus. EP **0178608**, A2 (1986)