JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Split Vector Quantization Using LBG Algorithm

[1]D. Vasudha Rani, [2]G. MadhuSudhan Reddy, [3]B. Harika, [4]B. Vamsi Krishna

[1]Student, [2]Student, [3]Student. [4]Student
Electronics and Communication Engineering,
R.V.R & J.C. College of Engineering, Guntur, India

Abstract- Speech compression is an essential task in modern communication systems, where the efficient utilization of available bandwidth is pivotal. In this context, the Linde-Buzo-Gray (LBG) algorithm has been widely adopted for vector quantization-based speech compression. The LBG algorithm is a computationally efficient and straightforward algorithm that can efficiently compress speech signals while maintaining good speech quality. This paper explains the implementation's steps of the LBG algorithm to compress and code speech signal directly from raw input speech.

Keywords - **Speech compression, vector quantization, Linde-Buzo-Gray (LBG) algorithm, Codebook Generation**

Introduction

Speech compression is a fundamental process in modern communication systems that involves reducing the amount of data required to represent a speech signal while maintaining acceptable speech quality. This process is critical in scenarios where the available bandwidth is limited, such as in wireless communication systems or internet-based voice communication. One of the most popular methods for speech compression is vector quantization.

I. VECTOR QUANTIZATION

A finite set of scalar or vector quantities used to map an infinite set of scalar or vector quantities are known as quantization. The lossy compression method known as vector quantization divides the voice signal into tiny, non-overlapping vectors. The signal is then encoded using a codebook, which is a collection of code vectors, for each vector. Finding the optimum code vector in the codebook to represent each vector in the speech signal is the fundamental tenet of vector quantization. By reducing the distortion between the original signal and the reconstructed signal, the signal produced by decoding the codebook—this is accomplished. One of the most popular algorithms for generating the codebook is the Linde-Buzo-Gray (LBG) algorithm, which is an iterative clustering algorithm that partitions the training data into a predetermined number of clusters.

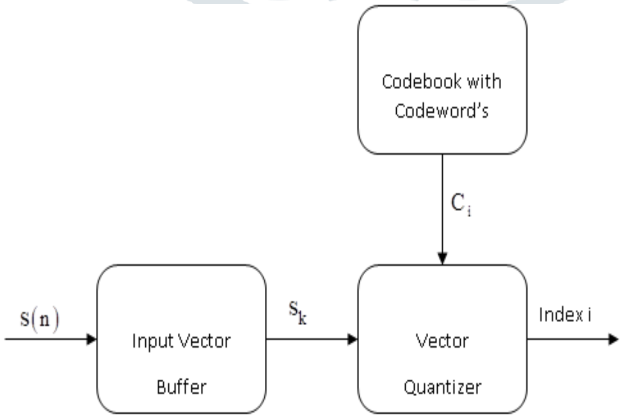

Figure-1. Vector Quantization

II. CODEBOOK GENERATION

The creation of codebooks is necessary for the vector quantization of speech signals. The Linde, Buzo, and Gray (LBG) algorithm is an iterative design process used to create codebooks. A training sequence is used as the LBG algorithm's input. The concatenation of a number of LSF vectors obtained from individuals belonging to various ages and racial groupings makes up the training sequence. Background noise must not exist in the voice signals utilized to create the training sequence. The speech signals can be captured in open spaces, computer rooms, and soundproof booths. In this project, computer rooms are used to record voice signals. speech coding and speech recognition can utilize speech databases like the TIMIT database in practice.

The creation of an initial codebook, which is the centroid or mean acquired from the training sequence, is a prerequisite for the codebook generation utilizing the LBG algorithm. Using the splitting approach, the acquired centroid is then divided into two centroid-like objects or codewords. These two codewords are divided into four using the iterative LBG algorithm, eight by four, and so on until the codebook has the necessary number of codewords.

III. LBG ALGORITHM

The Linde-Buzo-Gray (LBG) algorithm, also referred to as a well-liked clustering technique, is employed in signal processing and data compression.

A recursive method described below effectively implements the LBG algorithm:

1. To begin with, the LBG algorithm needs a training set of LSF parameters as an input for the codebook generation process. The training sequence is derived from a collection of speech samples that were captured in a computer room from various groups of people.
2. Define "R" as the training sequence's region.
3. Obtain an initial codebook from the training sequence; this first codebook should be "C," which is the centroid or mean of the training sequence.
4. Split the initial codebook C into a set of codewords C_n^+ and C_n^- where

$$C_n^+ = C(1+\varepsilon)$$

$$C_n^- = C(1-\varepsilon)$$

$\varepsilon = 0.01$ is the minimum error to be obtained between old and new codewords.

5. Determine the difference between each codeword and the training sequence, and let the result be "D".
6. Based on the difference "D" between the training sequence and the codewords, divide the training sequence into two regions, R1 and R2. The training vectors closer to falls in the regions R1 and R2 are the training vectors, respectively.
7. Make TV1 and TV2 for the training sequence vectors that fall in the regions R1 and R2, respectively.
8. Discover TV1 and TV2's new centroid or mean. Let CR1 and CR2 be the new centroids.
9. Replace the old centroids C_n^+ and C_n^- by the new centroids C_{R1} and C_{R2}.
10. Repeat steps 5 to 10 until $\frac{D^1 - D}{D} < \varepsilon$.
11. Continue to repeat steps 4 through 11 until the codebook has the necessary amount of codewords.

where $N = 2^b$ represents the number of codewords in the codebook and 'b' represents the number of bits used for codebook generation, D^1 represents the difference between the training sequence and the new codewords and 'D' represents the difference between the training sequence and the old codewords.

IV. SPECTRAL DISTORTION MEASURE

Voice coders use spectral distortion, which is measured in decibels, to determine the quality of the voice stream. (dB). Between the LPC power spectra of the quantized and unquantized speech samples, the spectral distortion is measured. The final value of the spectral distortion is determined by averaging or calculating the mean of the spectral distortion over all frames.

$$SD_i = \sqrt{\frac{1}{(f_2 - f_1)} \int_{f_1}^{f_2} \left[10\log_{10} s_i(f) - 10\log_{10} \hat{s}_i(f)\right]^2 df} \; (dB) \qquad (1)$$

Where $s_i(f)$ and $\hat{s}_i(f)$ are the LPC power spectra of the unquantized and quantized i^{th} frame respectively. The frequency "f" is expressed in Hz, while "f1" indicates the frequency range. For narrowband speech coding, the frequency range in use is 0 to 4000 Hz.

The average or mean of the spectral distortion SD is given by equation

$$SD = \frac{1}{N}\sum_{i=1}^{N} SD_i \qquad (2)$$

The requirements for transparent speech coding are:

1. The spectral distortion (SD) average or mean be less than or equal to 1 DB.
2. No outlier frames with spectral distortion over 4dB are allowed.
3. There must be fewer than 2% of outlier frames between 2 and 4 dB.

V. RESULTS AND DISCUSSION

Frames having average spectral distortion greater than 1dB are considered as outlier frames, there must be no outlier frames having a spectral distortion greater than 4 dB and the number of outlier frames between 2 to 4 dB must be less than 2%. For transparent coding, the average spectral distortion must be less than 1 dB.

Table 1 Spectral distortion for an LBG Split Vector Quantizer

Bits / frame	SD (dB)	Percentage of outliers		Complexity (Kflops/frame)	ROM (Floats)
		2-4 dB	>4dB		
24(8+8+8)	1.41	0.22	0.03	10.23	2560
23(8+8+7)	1.41	0.23	0.03	8.701	2176
22(8+7+7)	1.43	0.24	0.03	7.165	1792
21(7+7+7)	1.91	0.27	0.10	5.117	1280
20(7+7+6)	1.91	0.28	0.10	4.342	1088

VI. CONCLUSION

In this paper, a LBG algorithm based split vector quantization has been proposed for speech compression. The distortion of speech is optimized by employing LBG technique. The algorithm creates centroids efficient codebook design and efficient vector quantization of training vectors.

VII. REFERENCES

[1] 3. Gao Chang, LI Haifeng and Ma Lin, “Content-based compressive sensing for speech signal,” Journal of Signal Procesing, vol. 28, pp. 851-858, June 2012.

[2] Lamia and M.Arnaud, “Towards improving speech detection robustness for speech recognition in adverse conditions,” Speech Communication, vol. 40, pp. 261-266, March 2003.

[3] S. N. Kumar, A. Lenin Fred, H. Ajay Kumar, P. Sebastian Varghese & Ashy V. Daniel Conference paper First Online: 08 February 2019.

[4] Eiben AE, Smit SK (2011) Parameter tuning for configuring and analyzing evolutionary algorithms. Swarm Evol Comput 1(1):19–31.

[5] Saumya Mishra, Shraddha Singh “A Survey Paper on Different Data Compression Techniques”, Indian journal Research Paper, Volume : 6, Issue : 5, May 2016.

[6] J. Lukasaik, C. McElory, E. Chang, “Compression Transparent Low-Level Description of Audio Signals”,IEEE Int. Conf. on Multimedia and Expo, 6 July 2005.

[7] PramilaSrinivasan, Leah H. Jameason, “High Quality Audio Compression Using an Adaptive Wavelet Packet Decomposition and Psychoacoustic Modelling”,IEEE Trans. On Signal Proc., vol. XX, No. Y, 1999.

[8] Daniele Giacobello, Mads G. Christensen, Manohar N. Murthi, Soren H. Jenson, Marc Moonen, “Speech Coding Based on Sparse Linear Prediction”IEEE Trans. Acoustics, speech, and signal processing vol. 15, no. 5, 2010.

[9] M. Ramya, M. Sathyamoorthy, “Speech Coding by Using Subband Coding”,ISBN: 978-1-4675-2248-9, ICCCE 2012, 12 and 13 Apr 2012.

[10] Atal. B.S, “The history of linear prediction,” IEEE Signal Processing Magazine, Vol 23, pp.154-161, March 2006.

[11] Stephen. So, & Paliwal. K. K, “Efficient product code vector quantization using switched split vector quantizer,” Digital Signal Processing journal, Elsevier, Vol 17, pp.138-171, Jan 2007.

[12] H. B. Kekre, Tanuja K. Sarode, “Fast Codebook Generation Algorithm for Color Images using Vector Quantization,” International Journal of Computer Science and Information Technology, Vol. 1, No. 1, pp: 7-12, Jan 2009.