Multi Switched Split Vector Quantizer

M. Satya Sai Ram, P. Siddaiah, and M. Madhavi Latha

Abstract—Vector quantization is a powerful tool for speech coding applications. This paper deals with LPC Coding of speech signals which uses a new technique called Multi Switched Split Vector Quantization, This is a hybrid of two product code vector quantization techniques namely the Multi stage vector quantization technique, and Switched split vector quantization technique,. Multi Switched Split Vector Quantization technique quantizes the linear predictive coefficients in terms of line spectral frequencies. From results it is proved that Multi Switched Split Vector Quantization provides better trade off between bitrate and spectral distortion performance, computational complexity and memory requirements when compared to Switched Split Vector Quantization, Multi stage vector quantization, and Split Vector Quantization techniques. By employing the switching technique at each stage of the vector quantizer the spectral distortion, computational complexity and memory requirements were greatly reduced. Spectral distortion was measured in dB, Computational complexity was measured in floating point operations (flops), and memory requirements was measured in (floats).

Keywords—Unconstrained vector quantization, Linear predictive Coding, Split vector quantization, Multi stage vector quantization, Switched Split vector quantization, Line Spectral Frequencies.

I. INTRODUCTION

MOST forms of speech coding techniques are based on lossy algorithms. Lossy algorithms are considered acceptable when encoding speech because the loss in quality is undetectable by the human ear.Uncompressed speech is usually transmitted at 64 kb/s, using 8 bits/sample and at a rate of 8 KHZ for sampling. Any bit rate below 64 kb/s is considered as compression. This paper deals with a lossy compression technique called Linear predictive coding [1]-[3] which uses a vector quantization technique[4]-[6] called Multi switched split vector quantization. Multi switched split vector quantization technique is a hybrid of Multi stage vector quantization technique (MSVQ) [7]-[9], and Switched split vector quantization technique (SSVQ) [7]. As quality, complexity and memory requirements have a direct impact on marketability and cost of the under laying products or services, the performance of MSSVQ is evaluated by using the spectral distortion, computational complexity and memory requirements. In MSSVQ a number of Multi stage vector quantizers are connected in cascade where the difference between the input vector and quantized vector of one stage is

M. Satya Sai Ram is with Department of ECE, R. V. R. & J. C. College of Engineering, Guntur, 522019, A.P., India (e-mail: m_satyasairam@yahoo.co.in).

P. Siddaiah is with Department of ECE, K.L College of Engineering, Guntur-522502, A. P., India (e-mail: siddaiah_p@yahoo.com).

M. Madhavi Latha is with Department of ECE J.N.T.U College of Engineering Hyderabad, 500072, A.P, India (e--mail: mlmakkena@yahoo.com).

fed as an input to the next stage. Each stage of MSVQ employs SSVQ using soft decision in which a number of codebooks are connected in parallel. In this paper two codebooks connected in parallel are taken so as to maintain a tradeoff between the switch bits and the number of codebooks to be searched at each stage of Multi stage vector quantizer i.e, when two codebooks are connected in parallel with soft decision scheme the input vector is quantized by using the two codebooks connected in parallel, where as with hard decision scheme the input vector is quantized in only one of the two codebooks. As only one bit is required for the two switches with both the soft and hard decision schemes, there can be an improvement in spectral distortion performance with soft decision scheme for the same computational complexity and memory requirements when compared to hard decision scheme, so this paper deals with MSSVQ using a soft decision scheme.

MSSVQ algorithm mainly consists of the following steps:
  a) Select a switch
  b) Generate the codebook from the training sequence
  c) Obtain the quantized vectors from from the codebook generated
  d) Extract the new trained sequence from the old and quantized training sequence.
  e) Repeat steps a to d for all the switches of a stage.
  f) Obtain the approximate vector at each stage
  g) Repeat steps a to f for the required number of stages
  h) Finally obtain the approximate of the input vector by summing the approximate vectors at each stage.

The aim of this article is to provide a general review of MSSVQ, and to compare its performance with other existing product code vector quantization techniques. The practical limitations, regarding computational complexity and memory requirements as a function of bit rate are discussed. The spectral distortion performance of MSSVQ is evaluated in LSF parameter quantization [10]-[12] for narrow band speech coding.

II. MULTI SWITCHED SPLIT VECTOR QUANTIZATION

The basic idea of MSSVQ is to use p stages, m switches and s splits, and its goal is to reduce the spectral distortion, computational complexity, and memory requirements this is achieved by the use of two product code vector quantization techniques MSVQ and SSVQ. SSVQ is a hybrid of switch vector quantization and split vector quantization [13]-[14] techniques. The use of split vector quantizer makes the less availability of bits at each split of the vector quantizer as a result the complexity and memory requirements were greatly

reduced but the dependencies that exists across the dimensions (splits) of a vector will be lost as a result the spectral distortion will be slightly increased. With the use of Multi stage vector quantizer the number of bits used for quantization will be divided at each stage of the vector quantizer as a result the complexity and memory requirements can be reduced greatly, likewise the use of a switch vector quantizer exploits the correlation that exists across all dimensions of a vector quantizer.

At each stage of SVQ the 10-dimmensional LSF vector is split into 3 parts of 3, 3, 4 divisions respectively. During codebook generation bits are allocated depending on the frequency of the LSFs. Preference is given to high frequency LSFs, when the number of bits is not divisible by 3. For a particular switch the generation of codebooks at different stages is shown in Fig. 1.
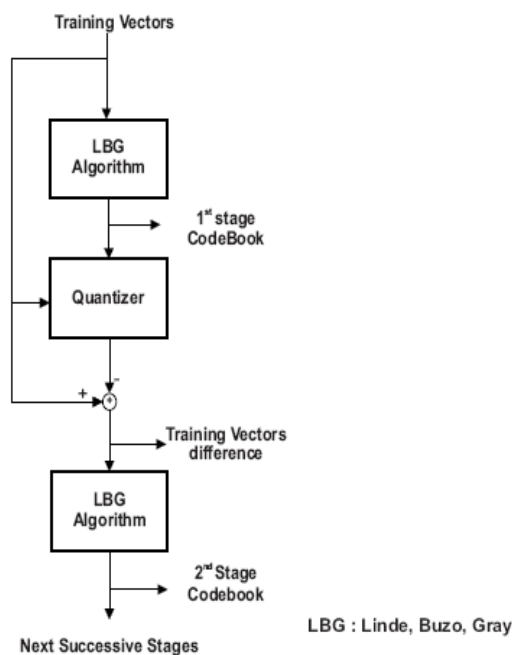

Fig. 1 Codebook Generation at different stages

> Initially the codebook at the first stage is generated by using the Linde, Buzo and Gray (LBG) [15] algorithm with the training vectors set as an input.
> Secondly the training difference vectors are extracted from the input training vectors set and the quantized training vectors of the first stage.
> Finally the training difference vectors are used to generate the codebook of the second stage.

This procedure is continued for the required number of stages and the number of codebooks to be generated will be equal to the number of stages used for quantization.

An p x m x s MSSVQ is shown in Fig. 2, where p corresponds to the number of stages, m corresponds to the number of switches, and s corresponds to the number of splits.

> Each input vector x that is to be quantized is applied to SSVQ at the first stage so as to obtain the approximate vectors at each codebook of the first stage.
> Extract the approximate vector with minimum distortion from the set of approximate vectors at the first stage i,e. $\hat{x}_1 = Q[x_1]$.
> Compute the error vector resulting at the first stage of quantization and let the error vector be, $e_1 = x_1 - \hat{x}_1$.
> The error vector at the first stage is given as an input to the second stage so as to obtain the quantized version of the error vector $\hat{e}_1 = Q[e_1]$.
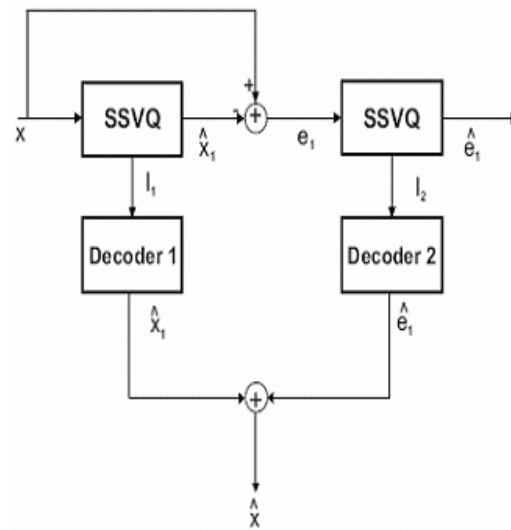
(I_i denotes the Index of Ith quantizer)
SSVQ: Switched Split Vector Quantization

Fig. 2 Block Diagram of MSSVQ

This process is continued for the required number of stages. Finally the decoder takes the indices, I_i, from each stage and adds the quantized vectors at each stage so as to obtain the reconstructed vector $\hat{x}$ given by $\hat{x} = Q[x_1] + Q[e_1] + Q[e_2] + \ldots$. Where $Q[x_1]$ is the quantized input vector at the first stage, $Q[e_1]$ is the quantized error vector at the second stage and $Q[e_2]$ is the quantized error vector at the third stage and so on.. As this process involves the quantization of the error vectors and summing of the error vectors with the approximate vector at the first stage the spectral distortion performance can be greatly improved when compared to SSVQ and SVQ.

III. COMPLEXITY AND MEMORY REQUIREMENTS

The computational complexity of a Switch vector quantizer is given by:

$$Complexity_{switch} = 4n2^{b_m} - 1 \qquad (1)$$

The computational complexity of a Split vector quantizer, is given by:

$$Complexity_{SVQ} = \sum_{i=1}^{s} (4n_i 2^{b_i} - 1) \qquad (2)$$

The computational complexity of a Multistage vector quantizer is given by:

$$Complexity_{MSVQ} = \sum_{i=1}^{m} (4n 2^{b_i} - 1) \qquad (3)$$

The computational complexity of a Switched split vector quantizer is given by:

$$Complexity_{SSVQ} = (4n2^{b_m} - 1) + \sum_{i=1}^{s} (4n_i 2^{b_i} - 1) \qquad (4)$$

The computational complexity of a Multi switched split vector quantizer is given by:

$$Complexity_{MSSVQ} = p(4n2^{b_m} - 1) + \sum_{j=1}^{p} [\sum_{i=1}^{s} (4n_i 2^{b_{ij}} - 1)] \qquad (5)$$

The memory requirements of a Split vector quantizer is given by:

$$Memory_{SVQ} = n2^{b_m} \qquad (6)$$

The memory requirements of a Multistage vector quantizer is given by:

$$Memory_{MSVQ} = \sum_{i=1}^{m} n 2^{b_i} \qquad (7)$$

The memory requirements of a Switched split vector quantizer is given by

$$Memory_{SSVQ} = n2^{b_m} + 2^{b_m} \sum_{i=1}^{s} n_i 2^{b_i} \qquad (8)$$

The memory requirements of a Multi switched split vector quantizer is given by:

$$Memory_{MSSVQ} = pn2^{b_m} + 2^{b_m} \sum_{j=1}^{p} [\sum_{i=1}^{s} n_i 2^{b_{ij}}] \qquad (9)$$

where n is the dimension of the vector

b_m is the number of bits allocated to the switch vector quantizer

$m = 2^{b_m}$ is the number of switching directions,

p is the number of stages

s is the number of splits.

IV. Spectral Distortion

In order to objectively measure the distortion between a coded and uncoded LPC parameter vector, the spectral distortion is often used in narrow band speech coding. For the ith frame the spectral distortion (in dB), SD_i, [16] is defined as:

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

Where F_S is the sampling frequency and $x_i(f)$ and $\hat{x}_i(f)$ are the LPC power spectra of the uncoded and coded i^{th} frame, respectively. f is the frequency in Hz, and the frequency range is given by f_1 and f_2. the frequency range used in practice is 0-4000Hz. The average spectral distortion SD [10] is given by

$$SD = \frac{1}{N} \sum_{n=1}^{N} SD_i \qquad (11)$$

The conditions for transparent speech from narrowband LPC parameter quantization are.

➤ The average spectral distortion (SD) must be less than or equal to 1dB.
➤ There must be no outlier frames having a spectral distortion grater than 4dB.
➤ The no of outlier frames between 2 to 4dB must be less than 2%.

V. Results

Tables I, II, III, IV and V shows the spectral distortion (dB), computational complexity (kflops/frame), and memory requirements (ROM) at various bit rates for a 3- part split vector quantizer, 3-stage multistage vector quantizer, 2-switch, 3- part switched split vector quantizer, 2-stage, 2-switch, 3-part multi switched split vector quantizer and 3-stage, 2-switch, 3-part multi switched split vector quantizer . From tables 1 to 5 and from Fig's 3 & 4 it is observed that 3-stage, 2-switch, 3-part MSSVQ has better spectral distortion performance, less computational complexity, and memory requirements when compared to 3-part SVQ, 3-stage MSVQ, and 2-switch, 3-part SSVQ. From Fig's 5 & 6 it is observed that for 2-stage, 2-switch, 3-part MSSVQ, and 3-stage, 2-switch, 3-part MSSVQ as the number of stages increases the computational complexity, and memory requirements decreases as the number of bits/stage decreases from 24 to16 bits/frame of quantization and from 15 to 9 bits/frame the case is reversed. For 3-part SVQ transparency in quantization is achieved at 24 bits/frame, for 3-stage MSVQ transparency is achieved at 22 bits/frame, for 2-switch, 3- part SSVQ transparency is achieved at 22 bits/frame, and for 2-stage, 2-switch, 3-part MSSVQ, and 3-stage, 2-switch, 3-part MSSVQ transparency is achieved at 21 bits/frame. From the results it is proved that MSSVQ has better spectral distortion performance, less computational complexity and memory requirements when compared to all the above mentioned product code vector quantization techniques. From tables 6 to 8 gives the number of unstable frames for a given bit rate the unstability is due to the independent quantization of the sub vectors in SVQ, SSVQ and MSSVQ and it can be observed that the number of unstable frames increases with the decrease in bitrate.
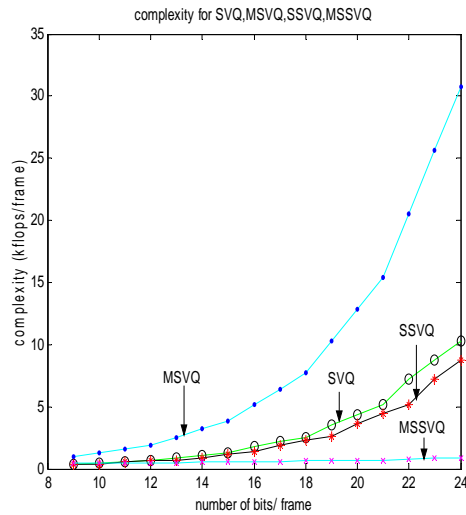
Fig. 3 Complexity for 3-part SVQ, 3-stage MSVQ, 2-switch 3-part SSVQ, and 3-stage 2-switch 3-part MSSVQ at various bit rates
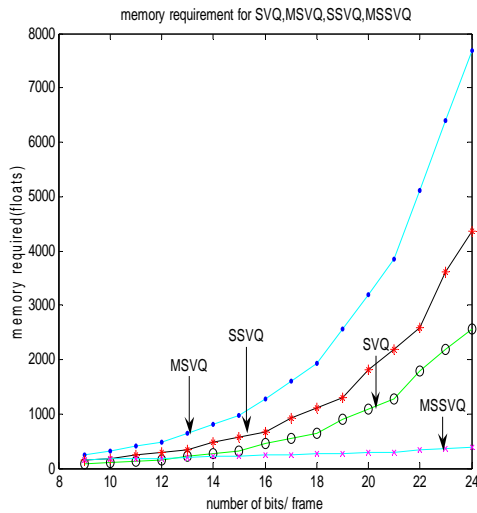

Fig. 4 Memory requirements for 3-part SVQ, 3-stage MSVQ, 2-switch 3-part SSVQ, and 3-stage 2-switch 3-part MSSVQ at various bit rates
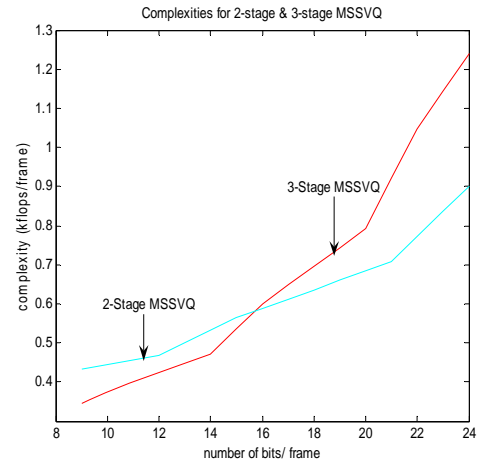

Fig. 5 Complexity for 2–stage 2-switch 3-part and 3–stage 2-switch 3-part MSSVQ at various bit rates
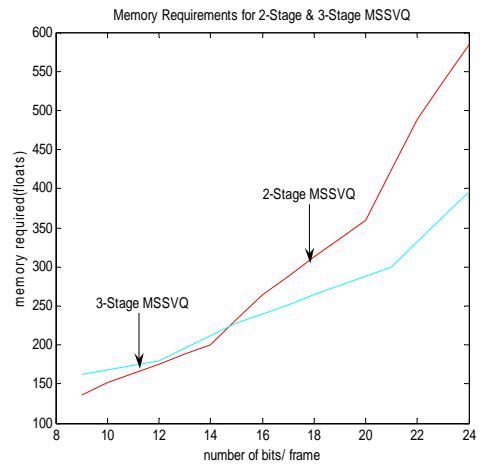

Fig. 6 Memory requirements for for 2–stage 2-switch 3-part and 3–stage 2-switch 3-part MSSVQ at various bit rates
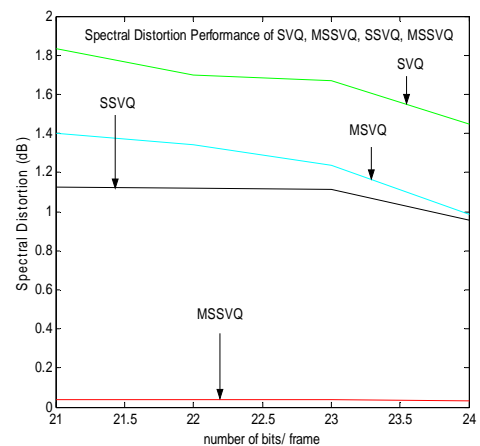

Fig. 7 Spectral Distortion Performance for 3-part SVQ, 3-stage MSVQ, 2-switch, 3-part SSVQ, and 3-stage, 2-switch, 3-part MSSVQ at various bit rates

TABLE I
SPECTRAL DISTORTION, COMPLEXITY, AND MEMORY REQUIREMENTS FOR 3-PART SPLIT VECTOR QUANTIZATION TECHNIQUE

Bits / frame	SD(dB)	2-4 dB	>4dB	Complexity (kflops/frame)	ROM (floats)
24(8+8+8)	1.45	0.43	0	10.237	2560
23(7+8+8)	1.67	0.94	0	8.701	2176
22(7+7+8)	1.701	0.78	0.1	7.165	1792
21(7+7+7)	1.831	2.46	0.2	5.117	1280

TABLE II
SPECTRAL DISTORTION, COMPLEXITY, AND MEMORY REQUIREMENTS FOR 3-STAGE MULTI STAGE VECTOR QUANTIZATION TECHNIQUE

Bits / frame	SD(dB)	2-4 dB	>4dB	Complexity (kflops/frame)	ROM (floats)
24(8+8+8)	0.984	1.38	0	30.717	7680
23(7+8+8)	1.238	1.2	0.1	25.597	6400
22(7+7+8)	1.345	0.85	0.13	20.477	5120
21(7+7+7)	1.4	1.08	0.3	15.357	3840

TABLE III
SPECTRAL DISTORTION, COMPLEXITY, AND MEMORY REQUIREMENTS FOR 2-SWITCH 3-PART SWITCHED SPLIT VECTOR QUANTIZATION TECHNIQUE

Bits / frame	SD(dB)	2-4 dB	>4dB	Complexity (kflops/frame)	ROM (floats)
24(12+12)	0.957	1.06	0	8.78	4372
23(11+12)	1.113	1.29	0.14	7.244	3604
22(11+11)	1.119	0.52	1.3	5.196	2580
21(10+11)	1.127	1.3	0.56	4.428	2196

TABLE IV
SPECTRAL DISTORTION, COMPLEXITY, AND MEMORY REQUIREMENTS FOR A 2-STAGE 2-SWITCH 3-PART MULTI SWITCHED SPLIT VECTOR QUANTIZATION

Bits / frame	SD(dB)	2-4 dB	>4dB	Complexity (kflops/frame)	ROM (floats)
24(12+12)	0.071	0	0	0.9	396
23(11+12)	0.083	0	0	0.836	364
22(7+7+8)	0.92	0	0	0.772	332
21(7+7+7)	0.13	0	0	0.708	300

TABLE V
SPECTRAL DISTORTION, COMPLEXITY, AND MEMORY REQUIREMENTS FOR A 3-STAGE 2-SWITCH 3-PART MULTI SWITCHED SPLIT VECTOR QUANTIZATION

Bits / frame	SD(dB)	2-4 dB	>4dB	Complexity (kflops/frame)	ROM (floats)
24(8+8+8)	0.0322	0	0	0.9	396
23(7+8+8)	0.0381	0	0	0.836	364
22(7+7+8)	0.0373	0	0	0.772	332
21(7+7+7)	0.0377	0	0	0.708	300

TABLE VI
NUMBER OF UNSTABLE FRAMES FOR 2-SWITCH 3- PART SSVQ WITH NO ORDERING CONSTRAINT

Bits / frame	No of switch bits	No of unstable frames
24(8+8+8)	1	19
23(7+8+8)	1	23
22(7+7+8)	1	29
21(7+7+7)	1	32

TABLE VII
NUMBER OF UNSTABLE FRAMES FOR 2-STAGE 2-SWITCH 3-PART MSSVQ WITH NO ORDERING CONSTRAINT

Bits / frame	No of switch bits	No of unstable frames
24(8+8+8)	1	11
23(7+8+8)	1	16
22(7+7+8)	1	19
21(7+7+7)	1	24

TABLE VIII
NUMBER OF UNSTABLE FRAMES FOR 3-STAGE 2-SWITCH 3-PART MSSVQ WITH NO ORDERING CONSTRAINT

Bits / frame	No of switch bits	No of unstable frames
24(8+8+8)	1	9
23(7+8+8)	1	13
22(7+7+8)	1	17
21(7+7+7)	1	21

VI. CONCLUSION

MSSVQ provides better trade-off between bit rate and spectral distortion performance, computational complexity, and memory requirements, when compared to other product code vector quantization schemes like SVQ, MSVQ, and SSVQ. So MSSVQ is proved to be better. The decrease in the computational complexity is due to the less availability of bits at each stage of quantization as the number of stages increases. From Fig. 4 it can be observed that for SSVQ the memory required is high when compared to SVQ. This has been overcomed by MSSVQ where the memory required is less when compared to SVQ, MSVQ, and SSVQ. So MSSVQ is

proved to be better when compared to all the above product code vector quantization techniques.

ACKNOWLEDGMENT

REFERENCES

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

[2] Harma, "A. Linear predictive coding with modified filter structures," IEEE Trans. Speech Audio Process, Vol 9, pp.769-777, Nov 2001.

[3] Viswanathan. R., Makhoul. J, "Quantization properties of transmission parameters in linear predictive systems," IEEE Trans. on Acoustics, Speech Signal Processing,vol 23, no.3, pp. 309-321, June1975.

[4] Gray.R.M., Neuhoff. D.L, "Quantization,". IEEE Trans. Inform. Theory, pp.2325-2383,1998.

[5] Agiomyrgiannakis. Y, Stylianou. Y, "Conditional Vector Quantization for Speech Coding," IEEE Trans. Speech Audio Process, Vol 15, no.2, pp.377-386, Feb 2007.

[6] Bist.A, "Narrowband process quantization," Data Compression Conference, 1995. DCC '95. Proceedings, March 1995.

[7] 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.

[8] Krishnan. V, Anderson. D.V, Truong. K.K, "Optimal multistage vector quantization of LPC parameters over noisy channels," IEEE Trans. Speech Audio Process, Vol 12, no.1, pp.1-8, Jan 2004.

[9] Biing-Hwang Juang Gray, A., Jr. "Multiple stage vector quantization for speech coding". IEEE Conference on Acoustics, Speech Signal Processing, vol 7, pp. 597-600, May 1982.

[10] Soong. F, Juang. B, "Line spectrum pair (LSP) and speech data compression," IEEE Conference on Acoustics, Speech Signal Processing, vol 9, no.1, pp. 37-40, Mar 1984.

[11] Bastiaan Kleijn. W, Tom Backstrom, & Paavo Alku. "On Line Spectral Frequencies," IEEE Signal Processing Letters, Vol.10, no.3, 2003.

[12] P. Kabal, and P. Rama Chandran. "The Computation of Line Spectral Frequencies Using Chebyshev polynomials," IEEE Trans. on Acoustics, Speech Signal Processing, Vol 34, no.6, pp. 1419-1426, 1986.

[13] Paliwal. K.K, Atal. B.S, "Efficient vector quantization of LPC Parameters at 24 bits/frame," IEEE Trans. Speech Audio Process, pp.3-14,1993.

[14] Chatterjee. S, Sreenivas. T.V, "Conditional PDF-Based Split Vector Quantization of Wideband LSF Parameters," IEEE Signal Processing Letters, Vol.14, no.9, Sept 2007.

[15] Linde .Y, Buzo. A., & Gray. R.M, "An Algorithm for Vector Quantizer Design," IEEE Trans.Commun, 28, pp. 84-95, Jan.1980.

[16] Sara Grassi., "Optimized Implementation of Speech Processing Algorithms," Electronics and Signal Processing Laboratory, Institute of Micro Technology, University of Neuchatel, Breguet 2, CH-2000 Neuchatel, Switzerland, 1988.

[17] Thomas F. Quatieri, Discrete-Time Speech Signal Processing, Pearson Education Inc, 2002.

[18] L.R.Rabiner and R.W.Schafer, Digital Processing of Speech Signals, Prentice Hall, Englewood Cliffs, 1978.

M. Satya Sai Ram obtained B.Tech degree in Electronics and Communication Engineering from Nagarjuna University, Guntur in 2003. He received his M.Tech degree from Nagarjuna University, Guntur in 2005. He started his career as a lecturer at R.V.R & J.C. College of Engineering in 2005. At present M.Satya Sai Ram is working as an Sr.Lecturer in the department of Electronics and Communication Engineering, R.V.R & J.C. College of Engineering and actively involved in research and guiding students in the area of Speech & Signal Processing,. He has taught a wide variety of courses for UG students and guided several projects. He has published more than 5 papers in International Conferences and Journals.

P. Siddaiah obtained B.Tech degree in Electronics and Communication Engineering from JNTU college of Engineering in 1988. He received his M.Tech degree from SV University, Tirupathi. He did his PhD program in JNTU, Hyderabad. He is the chief Investigator for several outsourcing project sponsored by Defense organizations and AICTE. He started his career as lecturer at SV University in 1993. At present Dr P. Siddaiah is working as an Professor & HOD in the department of Electronics and Communication Engineering, KL College of Engineering and actively involved in research and guiding students in the area of Antennas, Speech & Signal Processing,.. He has taught a wide variety of courses for UG & PG students and guided several projects. Several members pursuing their PhD degree under guidance. He has published several papers in National and International Journals and Conferences. He is the life member of FIETE, IE, and MISTE.

M. Madhavi Latha graduated in B. Tech from NU in 1986, Post Graduation in M.Tech from JNTU in 1993 and Ph. D from JNTU in 2002. She has been actively involved in research and guiding students in the area of Signal & Image Processing, VLSI (Mixed Signal design) and hardware implementation of Speech CODECs. She has published more than 17 papers in National/ International Conferences and Journals. Currently, she has been working as Professor in ECE, JNTU College of Engineering, Hyderabad, Andhra Pradesh. She is the life member of FIETE, and MISTE.