Shape - Gain Vector Quantization of Speech Signals

[1]Yogesh Kumar Mishra, [2]Jai Prakash Saini
[1]Associate Professor, [2]Director, NSIT, New Delhi
[1]Electronics Engineering Department,
[1]Kamla Nehru Institute of Technology, Sultanpur, India

***Abstract*—This paper presents the implementation of shape-gain vector quantization which a type product code vector quantization. The shape and the gain parameters are particularly suitable for speech signals which have large dynamic range. The implementation results is shows that optimum signal to noise ratio is achieved with codebook dimension of ten.**

***Index Terms*—Shape-Gain, Vector Quantization, Speech Signal.**

I. Introduction

The speech signal have large dynamic range of power levels. This property of speech signal makes it suitable for Shape Gain Vector Quantization. Shape gain vector quantization can be viewed as a type of product code involving shape and gain [1]. In this case some statistical parameter such as average or root mean square of speech vector is extracted which is named as gain and treated as scaling factor [2]. The speech vector normalized by scaling factor is called shape. The very definition of shape gain vector quantization suggests that in the situation when the probability distribution function of shape is independent of the gain, the shape gain vector quantization is most suitable.

II. Shape Gain Vector Quantization of Speech Signal

Consider a speech signal represented in terms a vector as $X = [x_1, x_2\ , x_3\ ….. x_n]$ and root mean square value of the speech vector is denoted by g then

$$g = [\sum_i^n {x_i}^2]^{1/2} \tag{1}$$

and shape S is given by

$$S = \frac{X}{g}; \qquad g \neq 0. \tag{2}$$

To find optimality of quantizer various distortion measures are framed as performance parameters, which includes mean square error, weighted mean square error and linear prediction distortion measures [3][4][5].

The mean square error parameter performance criterion yields

$$(d, \hat{g}\,\hat{S}) = \left\|X - \hat{g}\,\hat{S}\ \right\|^2, \tag{3}$$

where $\hat{g}$ and $\hat{S}$ are quantized values of g and S respectively. The above equation (3) can be written as

$$(d, \hat{g}\,\hat{S}) = \|X\|^2 + \hat{g}^2 + 2\,\hat{g}\,X^t\,\hat{S}$$
$$= \|X\|^2 + [\ \hat{g} - (X^t\,\hat{S})\,]^2 - (X^t\,\hat{S}\,)^2 \tag{4}$$

The above equation (4) shows that mean square error is minimized by first selecting the space vector that maximizes the last term $X^t\hat{S}$ and then minimizing the second term $\hat{g}$ - $(X^t\hat{S})^2$.

III. Codebook Design

One way to design the shape gain vector quantizer shown in **Fig. 1** which based on equation (4). The code (i, j) is generated from gain and shape codebooks from encoder unit ant transmitted to decoder. The various steps are enumerated as

Step 1

Let us consider the a region ***R*** given by $R = \{R_{ij};\ i = 1, 2, 3 ... N_g;\ \ j = 1, 2, 3 ... N_s\}$ having N_g X N_s partitions. Initialize the shape code book with $C_s = \{\hat{S}_j;\ \ j = 1, 2, 3 ... N_s\}$ and gain codebook with $C_g = \{g_i;\ \ i = 1, 2, 3 ... N_g\}$. If the input vector $X \in R_{ij}$, then x is mapped into (i, j) and corresponding reproduction vector is given by $g(x) = g_i$ and $S(x) = S_j$.

Step 2

Define the gain partition $G = \{G_i\,;\ i = 1, 2, 3 ... N_g\}$ by $G_i = \bigcup_{j=1}^{N_s} R_{ij}$. G_i being the input vector space which maps into a the gain code word $\hat{g}_i$. Also define the shape partition $A = \{A_j;\ j = 1, 2, 3, \ldots N_s \backslash\}$ by $A_j = \bigcup_{i=1}^{N_g} R_{ij}$. A_j being input vector space which maps into the shape code word S_j.

Step 3

Let us denote the average distortion by $D(C_g, C_s, R) = E_d(X, g(x)S(x))$. Optimum distortion for a given set of shape and gain code vector $(C_s, C_g\ )$, $R\ (C_s, C_g)$ denote the minimum distortion. Mathematically $D(C_g, C_s, R) \geq D(C_g, C_s, R(C_s, C_g))$. Therefore shape and gain is minimized to the average distortion for a fixed pair of code word.

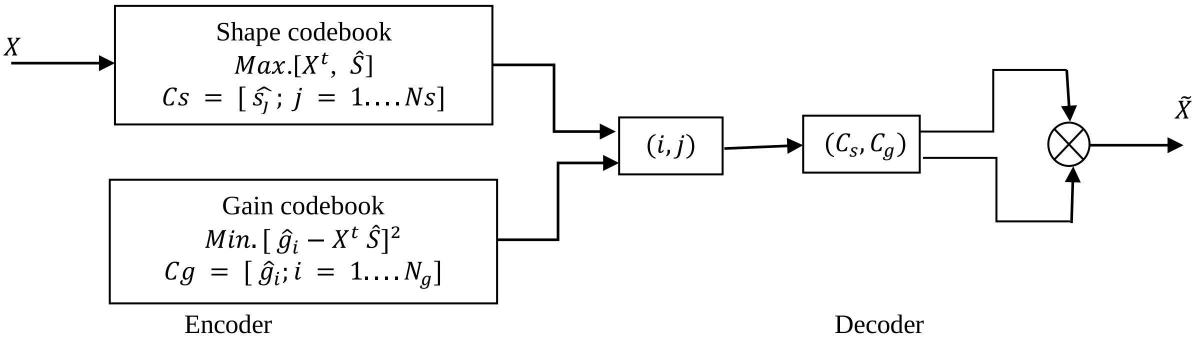

Figure 1: shape gain vector quantizer

IV. SIMULATION RESULT

Signal to noise ratio of the shape gain vector quantization with respect to codebook dimension of a shot term voiced speech sampled at a $8\,KHz$ taking length utterance as $1.5\,ms$ of is shown in the graph as depicted in **Fig. 2**. It has been found that the SNR saturates beyond codebook of dimension ten.

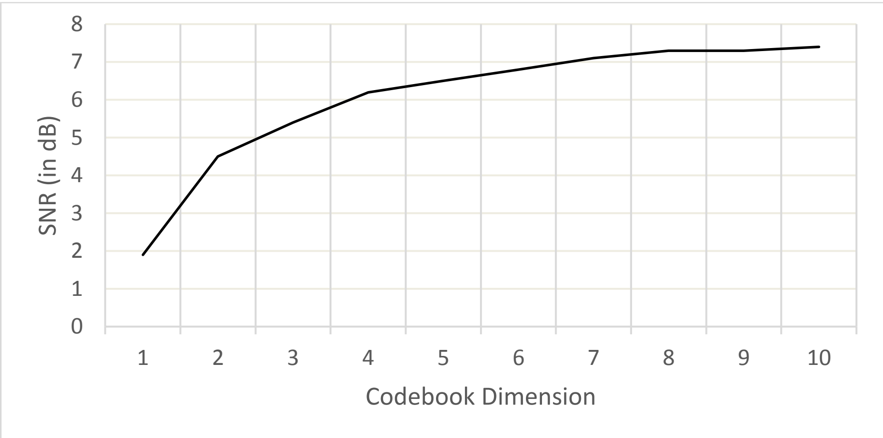

Figure 2: SNR Versus Codebook dimension Plot

V. ACKNOWLEDGMENT

We acknowledge the use VOICEBOX Copyright (C) Mike Brookes 2003 Version: $Id: voicebox.m 9702 2017-04-19 10:48:12Z, which is MATLAB toolbox for speech processing, under GNU General Public License.

REFERENCES

[1] M. J. Sabin, and M. R. Gray,, "Product code vector quantizers for waveform and voice coding," IEEE Trans. Acous., Speech, Signal Processing, vol. ASSP-32, pp. 474-488, June 1984.

[2] A. Buzo, A. H. Gray, Jr., R. M. Gray, and J. D. Markel. "Speech coding based upon vector quantization," IEEE Trans. Acoust. Speech Signal processing, ASSP-28:562-574, Oct. 1980.

[3] A. H. Gray, Jr. and J. D. Markel, "Distance measure for speech processing," IEEE Tans. Acoust., Speech, Signal Processing, vol.ASSP-24, no. 5, pp.380-291, Oct 1976.

[4] T. P. Barnwell, III, "Correlation analysis of subjective and objective measure of speech quality," in proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (Denver, CO. Apr. 1980), pp. 706-709.

[5] V. Viswanathan, W. Russeli, and A. W. F. Huggins, "Objective speech quality evaluation for mediumband and narrowband real time speech coder," in proc. IEEE Int. conf. acoust., Speech, signal Processing (Boston, MA, apr. 1983), pp. 443-546.