Improving the Switched Split Vector Quantization Technique using a Joint Source Channel Coding Approach

Tran Ngoc Tuan*, Nguyen Quoc Trung, Tran Hai Nam

Hanoi University of Science and Technology, No. 1, Dai Co Viet Str., Hai Ba Trung, Ha Noi, Viet Nam
Received: June 06, 2016; Accepted: November 03, 2017

Abstract

This paper deals with enhancing the error resilient of the Switched Split Vector Quantization (SSVQ) techniques by adopting the optimal Index Assignment approach, a Joint Source-Channel coding method. SSVQ is one of the latest structured vector quantization schemes and it has several advantages over other schemes. The new method proposed in this paper can improve the SSVQ encoder without the addition of extra bits and coding complexity. In addition, the application of the new method in speech coding is also investigated in this paper. The effectiveness of IA-SSVQ method is validated by comparing it with other methods through simulations.

Keywords: Joint Source-Channel coding, Vector Quantization, Index Assignment, Switched Split Vector Quantization.

1. Introduction

Signal coding has played a significant role in the success of digital communication, in which, the fundamental operation is quantization. Vector quantization (VQ) are known to theoretically achieve the lowest distortion, at a given rate and dimension, of any quantization scheme [1,2]. In practice, VQ is widely-used for low bit-rate coding of analog signals, especially highly correlated sources.

An optimal vector quantizer operates using a single large codebook with no constraints imposed on its structure. However, the VQs using large codebook are impractical, because the memory and computational requirement for VQ encoding is prohibitively high and the training process takes too much time. Several structurally constrained VQ schemes have been developed [1], which reduce the complexity of implementation with moderate loss of quantization performance. Switched Split Vector Quantization (SSVQ) [3,4] is one of the latest structured vector quantization schemes and it is further explored in [5,6] to show its competitive performance advantage over other VQ methods.

As most compression methods, the quality of reconstructed signal rapidly deteriorates when the channel noise is introduced. In order to protect against channel errors, the traditional approach is to increase the bit-rate for channel coding. Joint source-channel coding (JSCC) is an alternative that provides a technique to mitigate channel errors without an increase of the bit-rate. This paper deals with enhancing the error resilient of the SSVQ technique by using JSCC approach.

In the past, several methods based on JSCC technique were proposed for improving the VQ coder robustness for transmission over noisy channel. In order to improve the SSVQ method, the Channel Optimized Switched Split Vector Quantization (COSSVQ) method was proposed [7], which is based on Channel Optimized Vector Quantization approach (COVQ) [8]. In this approach, the channel statistical distribution is taken into account during both the source quantization and the codebook design. However, it requires long training time and its performance is usually degraded when the channel quality is high.

In this paper, a method based on Index Assignment approach [9] is developed to improve the error resilience of the coder using SSVQ technique. Different from COSSVQ method, the proposed method does not sacrifice any performance for the better channel and does not add any complexity to the encoder. This approach is implemented simply by rearranging the codebooks in the optimized order, therefore it can be used for improving the existing SSVQ systems with no need to redesign the coder. In addition, the application of this method in speech coding is also investigated in this work and the performance of the proposed method is validated through experiments in Section 5.

---

* Corresponding author: Tel.: (+84) 912.466.789
Email: tuan.tranngoc@hust.edu.vn

2. Switched Split Vector Quantization and the Index Assignment problem.

2.1. Vector Quantization.

When a set of discrete-time amplitude values is quantized jointly as a single vector, the process is known as Vector Quantization (VQ) or block quantization [1]. A vector quantizer $Q: \Re^K \rightarrow \mathbf{C}$ maps a continuous source vector $\mathbf{x} \in \Re^K$ to a codevector $\mathbf{c}_i \in \mathbf{C}$ by the nearest neighbour rule. The codebook $\mathbf{C} = \{ \mathbf{c}_i; 1 \le i \le N \}$ is the set of K-dimensional codevectors. The output of the vector quantizer is the index i of the codevector $\mathbf{c}_i$ which satisfies:

$$i = \underset{k}{\operatorname{argmin}}\, d\left(\mathbf{x}, \mathbf{c}_k\right) \quad (1)$$

where $d(\mathbf{x}, \mathbf{c}_k)$ is the nonnegative distance between two vectors. A common distortion measure is the squared Euclidean distance (SED), given by:

$$d\left(\mathbf{x}, \mathbf{y}\right) = \sum_{i=1}^{K} \left(x_i - y_i\right)^2 \quad (2)$$

Fig.1 shows the principle of VQ. Only the index i is transmitted over the channel to the receiver. Upon receiving i correctly, the VQ decoder can reconstruct $\mathbf{x}$ to $\mathbf{c}_i$ by a simple table lookup operation.
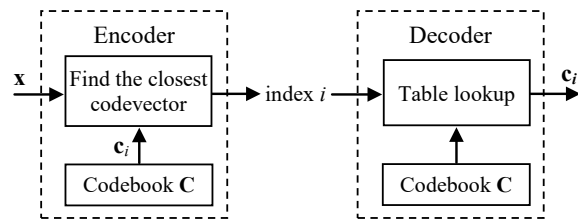

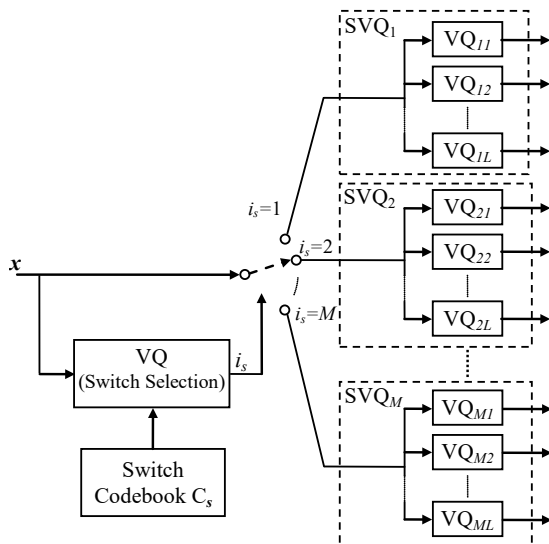**Fig. 1.** Principle of vector quantization

Fig. 2. Block diagram of a SSVQ encoder

The codebook design process is also known as training the codebook. A widely used algorithm for VQ codebook design is the Linde-Buzo-Gray (LBG) algorithm [10].

2.2. Switched Split Vector Quantization.

SSVQ is a hybrid of Switch Vector Quantization and Split Vector Quantization. In this scheme, the vector space is divided into non-overlapping switching regions and a separate Split Vector Quantizer (SVQ) [11] is designed for each region. The SVQ divides vectors into subvectors of lesser dimension and they are then quantized using independent codebooks. An L-part K-dimension SVQ is composed of L classical VQs of smaller sizes and dimension of $K_1, K_2, ..., K_L$.

The block diagram of a Switched Split Vector Quantizer is shown in Fig.2. Each vector to be quantized is first switched to one of the M possible directions based on the nearest-neighbour criterion, using the switch VQ codebook $\mathbf{C}_s$.

$$i_s = \underset{i}{\operatorname{argmin}}\, d\left(\mathbf{x}, \mathbf{c}_{si}\right) \quad (3)$$

Next, the vector will be quantized using the corresponding L-part SVQ. Therefore, the SSVQ coder transmits to the decoder an index i composed $L+1$ concatenated binary indices. The first index i_s indicates the switch direction and the remaining L indices $i_1, i_2, ..., i_L$ are provided by the corresponding local SVQ_{i_s}.

2.3. Index Assignment for Vector Quantization.

The effect of channel errors is to cause errors in the received indices which can result in significant distortion in decoded vectors. Let $P_a(i)$ denote the a priori probability of codevector $\mathbf{c}_i$, The IA function π is a permutation of the integers $\{0,1,...,N\text{-}1\}$ and $\pi(i)$ assigns an index to codevector $\mathbf{c}_i$. The overall distortion caused by channel noise is:

$$D\left(\pi\right) = \sum_{i=1}^{N} P_a(i) \sum_{j=1}^{N} P_c\left(\pi(j), \pi(i)\right) d\left(\mathbf{c}_i, \mathbf{c}_j\right) \quad (4)$$

In case of binary symmetric channel (BSC) with bit error rate (BER) ε, the codeword transition probability $P_C(i,j)$ is given by:

$$P_C(i,j) = \varepsilon^{h(i,j)}(1 - \varepsilon)^{n - h(i,j)} \quad (5)$$

where $h(i,j)$ denote the Hamming distance (number of bit differences) between i and j.

Different IAs affect the overall distortion $D(\pi)$ in case of channel error, so the IA problem is to find the optimal IA solution π which minimize $D(\pi)$. There are $N!$ possibilities to order N codewords, and to find an optimal solution for codebooks larger than 32 entries is practically impossible. For this reason, a

number of different IA approximate solutions have been proposed [9,12,13].

3. The proposed IA-SSVQ method.

In order to improve the robustness of the SSVQ coders, we adopt an JSCC approach carried out by the IA method and develop a new method named IA-SSVQ. The switch codebook $\mathbf{C}_S$ need to be reassigned in the optimized order provided by an IA algorithm and the order of SVQs is also rearranged according to the new order of codevectors in $\mathbf{C}_S$. Next, continue using the IA algorithm to find the optimal IA for each codebook of local SVQs and rearranging them in such optimized order.

The scheme for designing a M-switch IA-SSVQ with the training set $\mathbf{S}$ of length n_s is described below:

- Train the M-length switch codebook $\mathbf{C}_s$ from $\mathbf{S}$.
- Corresponding to M vectors $\mathbf{c}_{s1}, \mathbf{c}_{s2}, ... \mathbf{c}_{sM}$ in $\mathbf{C}_s$, partition $\mathbf{S}$ into M non-overlapping cells $\mathbf{R}_1, \mathbf{R}_2 ... , \mathbf{R}_M$ of length $n_{s1}, n_{s2}, ..., n_{sM}$.
- Train codebooks of the M local SVQs. (The SVQ$_i$ is trained using the training set $\mathbf{R}_i$).
- Find the optimal IA solution of $\mathbf{C}_S$ by using an IA algorithm with a priori probability of vector $\mathbf{c}_{si}$ given by $P_a(i) = n_{si}/n_s$ $(1 \le i \le M)$.
- Permute $\mathbf{C}_S$ by the optimized IA solution and rearrange the order of SVQs according to the new positions of vectors in $\mathbf{C}_S$.
- Apply IA method to rearrange all sub codebooks of the M local SVQs in the optimized order.

In the case of upgrading the existing system, only the last 3 steps need to be executed.

4. Application of IA-SSVQ in speech coding.

Most low bit rate speech coders employ the linear predictive coding (LPC) model [14] in which the short-term spectral is approximated by the all-pole filter whose transfer function is $H_{LPC}(z) = 1/A(z)$ and $A(z)$ is an inverse filter, given by:

$$A(z) = 1 + \sum_{i=1}^{p} a_i z^{-i} \qquad (6)$$

The order p is typically set to 10 for narrowband speech coders and to 16 for wideband speech coders. The quantization of LPC coefficients $\{a_i\}_{i=1}^{p}$ play a major role in the overall bit-rate and preserving the quality of the reconstructed speech.

In order to evaluate the performance of a LPC quantizer, the most popular approach is the spectral distortion (SD). For the i-th frame, the SD$_i$ in Decibel, defined as [11]:

$$SD_i = \sqrt{\frac{1}{n_1 - n_0} \sum_{n=n_0}^{n_1-1} \left[10 \log_{10} \frac{S\left(e^{j2\pi n/N}\right)}{\hat{S}\left(e^{j2\pi n/N}\right)} \right]^2} \qquad (7)$$

where $S(e^{j2\pi n/N})$ and $\hat{S}(e^{j2\pi n/N})$ are the original and quantized power spectrum of the LPC filter corresponding to the i-th frame of speech signal. The requirements usually considered necessary to achieve good quality speech are [11]: The average distortion is about 1dB, the number of outlier frames having SD in the range 2-4dB is less than 2% and no outlier frame having SD larger than 4dB.

In practice, the LPC coefficients are not directly quantized because they have poor quantization properties. Line Spectral Frequency (LSF) [15] has become the major representation of LPC coefficients because of its excellent properties in terms of model filter stability and robust quantization. The LSFs are defined as the roots of the following polynomials:

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1})$$
$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1}) \qquad (8)$$

All roots of $P(z)$ and $Q(z)$ are located on the unit circle of the z-plane and are interlaced with each other so that LSFs are in ascending order.

To further improve the performance of the coder, the weighted Euclidean distance (WED) may be used instead of SED as distortion measure for LSF vectors. The WED $d(\mathbf{f}, \hat{\mathbf{f}})$ between the original and quantized LSF vectors is given by [11]:

$$d\left(\mathbf{f}, \hat{\mathbf{f}}\right) = \sum_{i=1}^{p} \left[w_i \left(f_i - \hat{f}_i \right) \right]^2 \qquad (9)$$

where w_i is the spectral weight corresponding to the i-th LSF:

$$w_i = \left[\left| H\left(f_i\right) \right|^2 \right]^r \qquad (10)$$

where $|H(f_i)|^2$ is the LPC power spectrum at frequency f_i and r is an empirical constant determined experimentally. A value of $r = 0.15$ has been found satisfactory [11].

Due to the high correlation property of LSFs, VQ of them is most suitable for low bitrate but high quality quantization. SSVQ which has been studied recently is an effective structurally constrained VQ method for quantizing LSF coefficients and has many advantages over other VQ techniques[5,6]. Therefore, using IA-SSVQ method can improve the robustness of the speech coder and the effectiveness of this method is confirmed by experiment in Section 5.

5. Experiments and discussion.

In this section, computational experiments are carried out in Matlab to examine the performance of the IA-SSVQ method and to compare it with the traditional SSVQ and COSSVQ method. These three SSVQ systems with the same selected characteristics quantize and transmit the source over a BSC channel. The sources include a random highly correlated process and sets of speech LSF parameters.

In our experiments, codebooks were generated using LBG algorithm [8] and the SA algorithm [12,13] was applied to find the optimal IA for IA-SSVQ codebooks. The bit error probability used for training IA-SSVQ and COSSVQ codebooks is 0.01.

5.1. Random correlated source.

In this section, the input signal is a first-order Gauss-Markov process with correlation coefficient ρ.

$$x(n) = \rho x(n-1) + w(n) \qquad (11)$$

where $w(n)$ is a zero-mean, unit variance, Gaussian white noise process. In our experiment the value for ρ is 0.9 and the SED (Eq.2) is used as vector distortion measure.

The source is first partitioned into vectors of dimension 8, then these input vectors are quantized by various 16-switch 2-part SSVQ quantizers. The vectors are split into 2 parts with (4,4) division and the bit allocation is (6,6). The performances are evaluated in terms of signal-to-noise ratio (SNR) given by:

$$SNR = 10\log_{10}(\sigma_x/\sigma_n) \qquad (12)$$

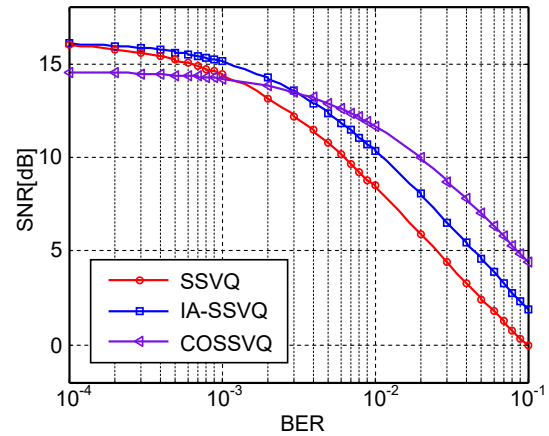where σ_x and σ_n are the signal and noise variances, respectively.

Fig. 3. Performance comparison of SSVQ methods

Fig.3 shows the SNR of system for 3 SSVQ methods against the BER. According to Fig.2, it can be observe that the performance of the IA-SSVQ method outperforms the regular SSVQ method in terms of high SNR. At high BER levels, the COSSVQ method provides better performance compared to IA-SSVQ method, but the IA-SSVQ method is better at low BER.

5.2. LSF Parameters of speech coder.

In this experiment, the TIMIT speech database with a sampling rate of 16kHz [16] was used for training and tesing of the SSVQ. In order to obtain the LSF vectors database, the same preprocessing and LPC analysis of the Adaptive Multirate Wideband speech coder (AMR-WB, ITU-T G.722.2) [17] was used. The training set consists of 644.137 vectors while the testing set contains 235.603 vectors distinct from the training vectors.

In all SSVQ quantizers, the number of switch directions is 32 (m=5) and the 16-dimensional LSF vectors are split into 5 parts with (3,3,3,3,4) division and the bit allocation is (9,8,8,8,8). The WSED was used for measuring the distortion of LSF vectors.

Table 1. Performance comparisons between various 46 bits/frame LSF SSVQ encoders.

BER ε	SSVQ			IA-SSVQ			COSSVQ		
	Average SD (dB)	Outliers % 2-4 dB	>4 dB	Average SD (dB)	Outliers % 2-4 dB	> 4 dB	Average SD (dB)	Outliers % 2-4 dB	> 4 dB
0	0.921	0.499	0.000	0.921	0.499	0.000	0.968	1.499	0.006
0.001	1.077	2.857	1.294	1.003	1.723	0.596	1.035	2.512	0.545
0.002	1.204	4.894	2.455	1.077	2.925	1.129	1.097	3.523	1.029
0.003	1.338	6.691	3.742	1.158	4.125	1.738	1.163	4.505	1.570
0.004	1.461	8.470	4.969	1.234	5.358	2.286	1.227	5.530	2.093
0.005	1.585	10.187	6.173	1.307	6.399	2.857	1.287	6.469	2.586
0.01	2.185	17.265	12.332	1.673	11.679	5.799	1.592	11.170	5.191
0.1	7.887	17.176	79.401	6.011	32.266	55.664	5.316	35.921	49.802

We use the common measure of spectral distortion (SD) (Eq.7) [11] to test the LSF quantization performance. In Table 1, the performance both in average SD as well as outlier percentage is depicted for various SSVQ schemes. It can be seen that, the simulation result is similar to the result in Section 5.1. The IA-SSVQ coder provides better performance than the ordinary SSVQ coder in term of low average SD and the number of outlier's frames of SD > 4dB. In comparison with COSSVQ coder, when ε is less than a certain threshold, the performance of IA-SSVQ coder is better and vice versa. In this experiment, the threshold is about 0.004. The reason is the IA-SSVQ and SSVQ codebooks are the same sets, just in different order, so the IA-SSVQ coder preserves the original performance of the SSVQ coder designed for noiseless channel.

6. Conclusion.

In this paper, an efficient and robust structured VQ scheme based on an optimal IA version of the SSVQ technique, namely IA-SSVQ, was developed. The performance of SSVQ methods was investigated for quantizing a random highly correlated source and parameters of the speech coder. The results showed that the IA-SSVQ encoder yields significant improvement over the ordinary SSVQ encoder by providing robustness against channel errors. Although, the performance of COSSVQ scheme is better at high BER, the new scheme has advantage of requiring no increase complexity to the encoder and no sacrifice performance for the better channels. Therefore, the IA-SSVQ can be a good technique for systems transmitting correlated analog signal as well as in speech coder in particular.

References

[1] A. Gersbo and R. Gray, Vector quantization and signal compression, Boston, Ma. Kluwer Academic Publishers, 1992.

[2] T.D. Lookabaugh, R.M. Gray, High-resolution quantization theory and the vector quantizer advantage, IEEE Trans. Inform. Theory 35 (5) (1989) 1020–1033.

[3] S. So, K.K. Paliwal, Efficient vector quantisation of line spectral frequencies using the switched split vector quantiser, Proc. Int. Conf. Spoken Language Processing, Korea, 2004.

[4] S. So, K.K. Paliwal, Switched Split Vector Quantisation of Line Spectral Frequencies for Wideband Speech Coding, INTERSPEECH-2005, Portugal, (2005) 2705-2708

[5] S. So, K.K. Paliwal, Efficient product code vector quantization using switched split vector quantizer, Digital Signal Processing journal, Elsevier, 17(1) (2007) 138-171.

[6] S. So, K. K. Paliwal, A Comparative Study of LPC Parameter Representations and Quantisation Schemes for Wideband Speech Coding, Digital Signal Processing Journal, Elsevier, 17(1) (2007) 114-137.

[7] M. Bouzid, S. Cheraitia, Channel Optimized Switched Split Vector Quantization for Wideband Speech LSF Parameters, Proc. 11th Int. Conf. on Inf. Science, ISSPA2012, Canada, (2012) 1045-1050.

[8] N. Farvadin, A Study of Vector Quantization for Noisy Channels, IEEE Trans. on Inf. Theory, 36(4) (1990) 799-809.

[9] N. Farvardin, V. Vaishampayan, On the performance and complexity of channel-optimized vector quantizers, IEEE Trans. Inf. Theory, 37(1) (1991) 155–160.

[10] Y. Linde, A. Buzo, and R. M. Gray, An algorithm for vector quantization design, IEEE Trans. on Commun., COM-28 (1980) 84-95.

[11] K. K. Paliwal, B. S. Atal, Efficient vector quantization of LPC parameters at 24 bits/frame, IEEE Transactions on Speech and Audio Processing, 1(1) (1993) 3-14.

[12] K. Zeger and A. Gersho, Pseudo-Gray Coding, IEEE Trans. on Commun., 38(12) (1990) 2147-2158.

[13] T.N. Tuan, N.Q. Trung, Improving the Simulated Annealing algorithm for the Index Assignment method to enhance the robustness of communication systems, Vietnamese Journal on Inf. Tech. & Comm.., E-3, 7(11) (2014) 13-20.

[14] A. M. Kondoz, Digital Speech: Coding for Low Bit Rate Communication Systems, 2nd Edition, John Wiley and Sons, 2004.

[15] F. Itakura, Line spectrum representation of linear predictive coefficients of speech signals, J. Acoust. Soc. Amer., 57 (1975) S35.

[16] ITU-T Recommendation G.722.2, Wideband Coding of Speech at Around 16 kb/s Using Adaptive Mutirate Wideband (AMR-WB), 2003.

[17] J. Garofol and al., Darpa TIMIT, Acoustic-Phonetic Continuous Speech Corpus CD-ROM, National Institute of Standards and Technology, NISTIR 493, USA, 1990.