ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2010, Vol.39, No.4

A SWITCHED-ADAPTIVE QUANTIZATION TECHNIQUE USING µ-LAW QUANTIZERS

Zoran Peric, Jelena Nikolic, Aleksandar Mosic, Stefan Panic

Faculty of Electronic Engineering, Nis, Serbia
e-mail: jelena.nikolic@elfak.ni.ac.rs

Abstract. In this paper, we propose a switched-adaptive quantization technique suitable for high quality quantization of signals which, as well as speech signals, have short-term statistics modeled by Gaussian probability density function. Particularly, from the set of the 2k disposable quantizers, the technique we propose performs a two-stage selection of the quantizer which is designed for the signal statistic nearest to the one of the signal to be processed. In the first selection stage, the choice of the quantizer from the set of k disposable quantizers is managed by the switched technique. Further, in the second selection stage, the choice is made between the restricted and the unrestricted quantizers, regarding the fact whether the maximum amplitude of the samples within the current frame is lower or higher than the upper support region threshold of the restricted quantizer, respectively. In order to properly define the frame length, in this paper we are introducing the rate-quality compromise criterion, along which we are studying the gain in the signal to quantization noise ratio and in the compression, that are achievable with the proposed technique in reference to the G.711 recommendation and another technique we have recently proposed.

Keywords: μ-law quantization, signal to quantization noise ratio, switched-adaptive quantization technique.

1. Introduction

To achieve high-quality quantized speech signals, the contemporary public switched telephone networks utilize the log-companded 8-bit Pulse Code Modulation (PCM) proposed by the G.711 recommendation [2]. Although a great number of quantizers have been developed to provide an additional enhancement of the speech signal quality [3], [5]-[7], [9]-[12], especially for the VoIP applications where identical PCM format is used, there is still a need to continue the research in this field. The quality of a quantized signal is generally influenced by the width of a quantizer's support region and the number of quantization levels [1], [7]. The importance of the suitable support region choice was pointed out in [8], where the logarithmic decrease of the support region threshold with the number of quantization levels was also ascertained. The shrinkage of the quantizer's support region, i.e. of its granular region, for a fixed number of the quantization levels, causes a reduction of the granular distortion, while at the same time possibly resulting in an unwanted increase of the overload distortion [7]. To overcome the distinguished problem, this paper proposes a novel switched-adaptive quantization technique that performs a two-stage selection of the quantizer which is designed for the signal statistic nearest to the one of the signal to be processed. The presented technique actually upgrades the model of the switched log-compandor, proposed in [11], with a kind of support region adaptation. Particularly, in the second stage of the quantizer selection, we propose the use of the restricted compandor, when all the amplitudes of the samples within the frame belong to the restricted compandor support region. Otherwise, we propose the selection of the unrestricted compandor having a wider support region. In order to provide the smallest possible distortion and a more frequent selection of the restricted compandor, which introduces only the granular distortion, we propose the determination of its support region thresholds to satisfy the total distortion minimization criterion. In the third section of the paper, we demonstrate that with a more preferable selection of the restricted scalar compandor having smaller support region than the unrestricted one, the proposed technique provides the gain in the signal to quantization noise ratio of about 2 dB in reference to the technique proposed in [11], where only the unrestricted μ-law compandors are at disposal. As well as in [5], [11], due to the time varying characteristics of speech signals, in this paper we have decided to utilize the frame-by-frame manner to process the input signal. Unlike [5], where the design of the scalar compandors followed the optimum compression law for the assumed Laplacian probability density function of the input signal, here, as in [11], we consider μ-law quantization, while assuming Gaussian probability density function, which is a more realistic probability density function for the short-term statistics of speech signals. Finally, since the choice of the suitable frame length is not a simple task, in this paper we have decided to introduce a rate-quality compromise criterion. According to this criterion we have ascertained the compromise frame length along with the appropriate signal to quantization noise ratio characteristic in the wide range of the frame variances, for the proposed model of the switched-adaptive μ-law quantizer.

2. Novel Switched-Adaptive Quantization Technique

The quantization technique developed in this paper processes the input signal on a frame-by-frame manner. As illustrated in Figure 1, the buffering of the j-th frame, which contains M samples denoted by $x_{(j-1)M+i}$, i = 1, ..., M, is followed by an estimation of the frame variance σ_j^2

$$\sigma_j^2 = \frac{1}{M}\sum_{i=1}^{M} x_{(j-1)M+i}^2 . \qquad (1)$$

Since the disposable quantizers are designed for the variances $(\hat{\sigma}_p^{disp})^2$, p=1, …, k, which are log-uniformly distributed in the range of the frame variances B=20log($\sigma_{\max}/\sigma_{\min}$)

$$20\log(\hat{\sigma}_p^{disp}) = 20\log(\sigma_{\min}) + \frac{(2p-1)B}{2k}, \quad p = 1,...,k . \qquad (2)$$

The log-uniform quantization of the estimated frame variance σ_j^2 is actually performed in the manner of the switched quantization technique. For instance, considering the set of variances $(\hat{\sigma}_p^{disp})^2$, p=1,…,k, for which the disposable quantizers are designed, one can ascertain the s-th variance $(\hat{\sigma}_s^{disp})^2$ as the nearest variance to the current frame variance σ_j^2. $(\hat{\sigma}_s^{disp})^2$ therefore, stands not only for the quantized value of σ_j^2, $\hat{\sigma}_j^2 = (\hat{\sigma}_s^{disp})^2$, but it also provides the first selection stage of the two-stage encoder selection process. Accordingly, the second stage of the encoder selection process is managed by the s-th encoder selector. Since it selects one of the two encoders at disposal, Enc. s,0 or Enc. s,1, performing a kind of a support region adaptation, we have named the technique we propose the switched-adaptive technique. The s-th encoder selector compares the estimated maximum amplitude of the samples within the current frame x_j^{max} with a socalled decision threshold x_d, and, depending on the comparison result, makes a choice between the two encoders at disposal, both designed for the variance $\hat{\sigma}_j^2 = (\hat{\sigma}_s^{disp})^2$. The threshold x_d is chosen to be equal to the upper support region threshold of the s-th restricted quantizer $x_{s,0}^{max}=x_d=c_0\hat{\sigma}_j$, $c_0=c_{s,0}=c_{p,0}$, p=1,…,k, where c_0 denotes the loading factor of the restricted quantizers at disposal.

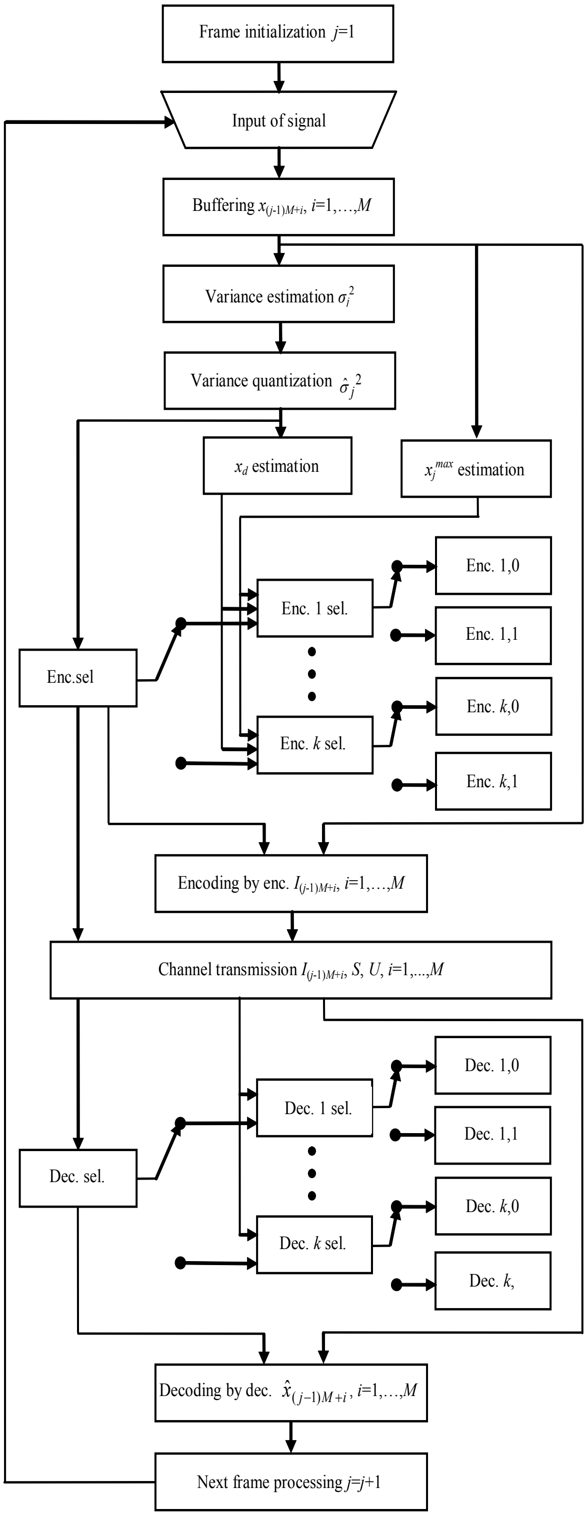

Figure 1. The switched-adaptive quantization technique algorithm

If it holds $|x_j^{max}|\le x_d$, the encoder denoted by Enc. s,0 is selected. Otherwise, the proposed algorithm selects the encoder denoted by Enc. s,1. As the encoder Enc. s,1 is designated for encoding of the unrestricted signals, it is designed for the higher upper support region threshold $x_{s,1}^{max}=c_1\hat{\sigma}_j$, $c_1=c_{s,1}=c_{p,1}>c_0$, p=1, …, k, where c_1 denotes the loading factor of the unrestricted quantizers at disposal. When the encoder is selected and the encoding procedure is performed, the generated index $I_{(j-1)M+i}$, i=1,...,M, is then transmitted along with the side information. Depending on the transmitted side information S and U, a two-stage decoder selection is firstly made between k decoders at disposal, and after that betwen the two decoders, Dec. s, 0 and Dec. s,1, which refer to the s-th restricted and the unrestricted quantizer, respectively. Finally, the quantized samples $\hat{x}_{(j-1)M+i}$, i=1,...,M, are obtained by decoding the received index $I_{(j-1)M+i}$, i=1,...,M, with the selected decoder.

Let us now define the bit rate R corresponding to the proposed algorithm

$$R = R_N + \frac{(R_s + R_u)}{M}. \tag{3}$$

Take a notice of the fact that by $R_N=\log_2 N$ we denote the number of bits per sample required for the N-level μ-law quantization. Further, by $R_s=\log_2 k$ and R_u=1 we denote the number of bits per frame having length M that is required for transmission of the side information S and U, respectively. In accordance with the described algorithm, the total distortion introduced by the quantization is:

$$D = D_{s,0}{}^{g} P_{f0} + (D_{s,1}{}^{g} + D_{s,1}{}^{o}) P_{f1}, \tag{4}$$

where P_{f0} and P_{f1} denote the probabilities of the restricted and the unrestricted quantizer selection, respectively. As the restricted quantizer is selected when all the samples within the frame belong to its support region, it introduces only the granular distortion $D_{s,0}{}^{g}$, annulling in such a way the unwanted overload distortion $D_{s,0}{}^{o}$=0. On the other hand, the unrestricted quantizer is selected when x_j^{max} overreaches x_d and, therefore, it can generally introduce both the granular and the overload distortion, denoted by $D_{s,1}{}^{g}$ and $D_{s,1}{}^{o}$, respectively. For the assumed Gaussian probability density function [7]

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2}), \tag{5}$$

and μ-law quantization, we can derive the corresponding distortions

$$D_{s,l}{}^{g} = \left[q_{s,l}{}^2 + \sqrt{\frac{2}{\pi}} q_{s,l}\mu_l \left(2 - (2+\mu_l)\exp\left(-\frac{q_{s,l}{}^2}{2} \right) \right) + \mu_l^2 erf\left(\frac{q_{s,l}}{\sqrt{2}} \right) \right] \times \sigma^2 \frac{\ln^2(1+\mu_l)}{3N^2 \mu_l{}^2},$$

$$q_{s,l} = \frac{c_l \hat{\sigma}_s^{disp}}{\sigma}, \; l = 0,1, \tag{6}$$

$$D_{s,1}{}^{o} = \sigma^2 \left[\left(1 + q_{s,1}{}^2\right)\left(1 - erf\left(\frac{q_{s,1}}{\sqrt{2}}\right)\right) - \sqrt{\frac{2}{\pi}} q_{s,1} \exp\left(-\frac{q_{s,1}{}^2}{2} \right) \right], \tag{7}$$

and probabilities

$$P_{f0} = \left(\int_{-x_d}^{x_d} p(x) dx \right)^M = \left(erf\left(\frac{c_0}{\sqrt{2}} \right) \right)^M, \tag{8}$$

$$P_{f1} = 1 - P_{f0}, \tag{9}$$

which provide the signal to quantization noise ratio of the proposed quantizer model in line with the basic definition [1], [6], [7]

$$SQNR = 10 \log_{10} \left(\frac{\sigma^2}{D} \right), \tag{10}$$

for the input frame variance σ^2. Although one can expect that further growth of the signal to quantization noise ratio can be achieved by increasing the number of the restricted μ-law quantizers at disposal, in this paper we have kept analyzing the case with two quantizers, the restricted (Enc. s,0, Dec. s,0) and the unrestricted one (Enc. s,1, Dec. s,1), since our aim is to explain the proposed idea in the simplest possible manner.

3. Results and Conclusions

What is presented and discussed in this section are the performances that we have ascertained by applying the proposed technique in quantization of the signals having Gaussian probability density function and a wide range of the frame variances (B=40 dB). We have made a choice of the compromise frame length M^{com} regarding the rate-quality compromise criterion. Namely, it is well known that the frame length must be selected carefully: if it is too small, there may be a large overhead for transmitting the side information, but if the frame length is too large, the selected μ-law compandor may become inadequate for some portions of the frame, leading to large errors in the quantization process [6], [7]. According to the criterion, we introduce here, there is no point in further reduction of the frame length if the unit increase of the number of bits per sample does not result in more than 6 dB increase of signal to quantization noise ratio averaged over the range of the frame variances B. While determing M^{com}, for all of the considered M, we have ascertained the optimum pairs of parameters (c_0^{opt}, μ_0^{opt}) and (c_1^{opt}, μ_1^{opt}) that minimize the total distortion given by (4). By assuming $M=M^{com}$ and the appropriate (c_0^{opt}, μ_0^{opt}) and (c_1^{opt}, μ_1^{opt}), we have ascertained the signal to quantization noise ratio dependences for the bit rates of R_1=7.178 bit/sample and R_2=8.172 bit/sample (see Figures 2 and 3). By comparing the signal to quantization noise ratio characteristic obtained for the R_1 bit rate, with the one defined by the G.711 recommendation in the same variance range [2], we have reveled that the proposed quantizer, along with 0.822 bit/sample compression, provides a more constant and higher level of signal to quantization noise ratio, which can be considered via the gain in the average signal to quantization noise ratio of about 5.7 dB. Similarly, by comparing the signal to quantization

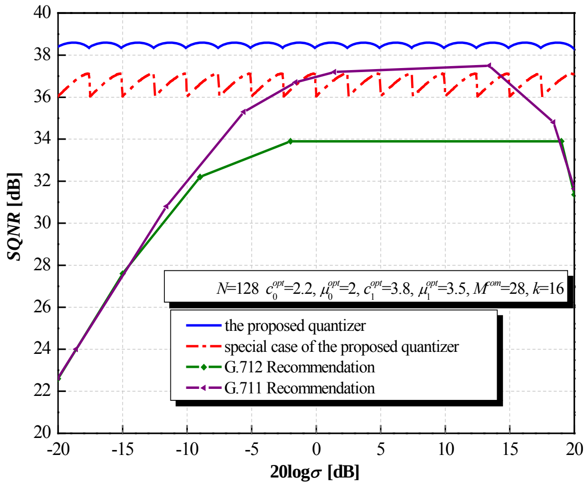

Figure 2. Illustration of the signal to quantization noise ratio characteristics in the wide range of the frame variances B=40 dB for the bit rate of R_1=7.178 bit/sample

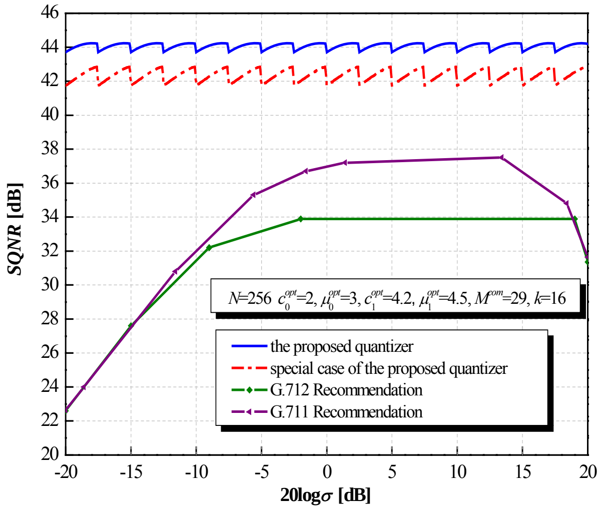

Figure 3. Illustration of the signal to quantization noise ratio characteristics in the wide range of the frame variances B=40 dB for the bit rate of R_2=8.172 bit/sample

noise ratio characteristics defined by the G.711 recommendation with the one obtained for the R_2 bit rate, which is about 0.172 bit/sample higher than the PCM bit rate, we have ascertained an even greater gain in the average signal to quantization noise ratio of about 11.25 dB. Moreover, in both figures, Figure 2 and Figure 3, we have shown the signal to quantization noise ratio characteristics achievable by the proposed quantizer and its special case corresponding to the quantizer provided in [11]. According to these characteristics we have determined that the quantizer proposed in this paper along with a slight increase of the bit rate of $1/M^{com}$, provides the gains of about 1.8 dB and 1.7 dB, for the bit rates R_1=7.178 bit/sample and R_2=8.172 bit/sample, respectively. Additionaly, from the Figs. 2 and 3 one can conclude that in the considered range of the frame variances B, the quantizer proposed in [11], when designed for the R_1 bit rate and the compromise frame length M^{com}, can not completely overreach G.711 recommendation. However, it is not an issue in case of M=10, as it was assumed in [11], since in such a case we have revealed an increase in the average signal to quantization noise ratio of about 0.6 dB in comparision to the case when M=M^{com}, while the consequence is an increase of the R_1 bit rate of 0.322 bit/sample. Finally, since we have ascertained that the proposed quantzer satisfies the G.712 recommendation [4] in the considered range of the frame variances B, one can believe that it will find practical implementation in the high quality quantization of signals, which, as well as speech signals, have short-term statistics modeled by the Gaussian probability density function.

References

[1] **A. Gersho, R.M. Gray.** Vector Quantization and Signal Compression. *Kluwer Academic Publishers, Boston, Dordrecht, London*, 1992.

[2] **ITU-T, Recommendation G.711.** Pulse Code Modulation (PCM) of Voice Frequencies. *International Telecommunication Union, Geneva, Switzerland*, 1988.

[3] **ITU-T, Recommendation G.711.1.** Wideband Embedded Extension for G.711 Pulse Code Modulation. *International Telecommunication Union, Geneva, Switzerland*, 2008.

[4] **ITU-T, Recommendation G.712.** Transmission Performance Characteristics of Pulse Code Modulation channels. *International Telecommunication Union, Geneva, Switzerland*, 2001.

[5] **J. Nikolic, Z. Peric.** Lloyd-Max's Algorithm Implementation in Speech Coding Algorithm Based on Forward Adaptive Technique. *Informatica*, *Vol.*19, *No.*2, *June* 2008, 255-270.

[6] **L. Hanzo, C. Somerville, J. Woodard.** Voice and Audio Compression for Wireless Communications. *John Wiley & Sons, IEEE Press, October*, 2007.

[7] **N.S. Jayant, P. Noll.** Digital Coding Of Waveforms, Principles and Applications to Speech and Video. *Prentice Hall, New Jersey*, 1984, 115-251.

[8] **S. Na.** On the Support of Fixed-Rate Minimum Mean-Squared Error Scalar Quantizers for a Laplacian Source. *IEEE Transactions on Information Theory*, *Vol.*50, *No.*5, *May* 2004, 937-944.

[9] **W.C. Chu.** Speech Coding Algorithms: Foundation and Evolution of Standardized Coders. *John Wiley & Sons, New Jersey*, 2003.

[10] **Y. Hiwasaki et al.** G.711.1: A Wideband Extension to ITU-T G.711. *Proc. of EUSIPCO* '08, 2008.

[11] **Z. Peric, A. Mosic, S. Panic.** Coding Algorithm based on Loss Compression using Scalar Quantization Switching Technique and Logarithmic Companding. *Journal of Information Science and Engineering*, *Vol.*26, *No.*3, *May* 2010, 967-976.

[12] **Z. Peric, J. Nikolic, Z. Eskic, S. Krstic, N. Markovic.** Design of Novel Scalar Quantizer Model for Gaussian Source. *Information Technology and Control*, *Vol.*37, *No.*4, 2008, 321-325.

Received August 2010.