M-CHANNEL MULTIPLE DESCRIPTION CODING BASED ON UNIFORMLY OFFSET QUANTIZERS WITH OPTIMAL DEADZONE

Lili Meng[1], *Jie Liang*[2], *Yao Zhao*[1], *Huihui Bai*[1], *Chunyu Lin*[1] *and André Kaup*[3]

[1] Institute of Information Science, Beijing Jiaotong University, Beijing, China
[2]School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada
[3]University of Erlangen-Nuremberg, Erlangen, Germany

ABSTRACT

This paper proposes an improved source-splitting-based two-rate M-channel multiple description coding scheme, where the source is split into M subsets. In each description, one subset is coded at a high rate, and others are predictively coded at a low rate. Uniform offsets among low-rate quantizers of different descriptions are achieved by employing unequal deadzones and by quantizing the predictions. When several descriptions are received, the optimal reconstruction of each subset is achieved by finding the intersection of all received quantization bins. The closed-form expression of the expected distortion is obtained. The proposed scheme is applied to lapped transform-based multiple description image coding and achieves improved performance. The optimal deadzone selection and its impact are also given in this paper.

***Index Terms*—** Multiple description coding, predictive coding, deadzone quantization, random quantization

1. INTRODUCTION AND RELATIONSHIP TO PRIOR WORK

Transmission loss in communication networks is an important factor that needs to be considered when designing the system. Multiple description coding (MDC) addresses this by sending several descriptions of the source such that the reconstruction quality improves with the number of received descriptions. The MDC with two descriptions (or channels) have been studied extensively. In this paper, we focus on MDC with M descriptions, where $M > 2$, which is more useful in practice.

A popular approach to generate MDC is based on source splitting, which is indeed the earliest method [1]. In [2], an M-channel MDC method called RD-MDC is proposed for JPEG2000-based image coding, where each JPEG2000 code-block is encoded at two rates. The higher-rate coded code-blocks are divided into M subsets and are assigned to M descriptions. Each description also carries the lower-rate codings of the remaining code-blocks. In [3], the two-rate coding in [2] is generalized to multiple rates, but its complexity increases rapidly with M.

In [4], a more efficient predictive coding-based M-channel MDC scheme is proposed using two-rate coding and staggered quantization (TRPCSQ), where the low-rate coding is made mutually refinable using staggered uniform quantizers. To preserve uniform staggering, the prediction is also quantized. However, the application of staggered quantizers creates asymmetric quantization bins with respect to zero, which compromises the coding efficiency of the scheme.

In [5], a prediction compensated MDC (PCMDC) scheme is developed for $M = 2$, where the source is partitioned into two subsets, and each subset is encoded as the base layer of one description. Each description also encodes the prediction residual of the other subset. In [6], a three-layer MDC (TLMDC) scheme is developed, which generalizes the PCMDC to $M > 2$ via sequential prediction. When more than two low-rate reconstructions of a subset are available, their average is used as the final reconstruction. A third layer is also added to refine the low-rate-coded subsets when only one description is lost, which is the dominant error scenario when the channel loss probability is low. The performance of TLMDC is shown to be better than TRPCSQ.

In this paper, we propose an M-channel MDC scheme that improves both the TRPCSQ [4] and TLMDC [6]. As in TLMDC, this new method uses two rate predictive coding and sequential prediction. As in TRPCSQ, it uses uniform offsets among different low-rate quantizers, which also requires the quantization of the predictions. However, different from TRPCSQ, the uniform offsets are achieved by employing unequal deadzone sizes in different quantizers, which avoids the asymmetric quantizer problem in TRPCSQ. In addition, when both high-rate and low-rate coding of a subset are available, joint de-quantization is also applied.

Although staggered quantizers and unequal deadzone quantizers have been used in various MD schemes, their theoretical and image coding performances have not been

This work was supported in part by China Scholarship Council, the Natural Sciences and Engineering Research Council (NSERC) of Canada under grants RGPIN312262, STPGP350740, STPGP380875, the 973 Program (2012CB316401), the National Natural Science Foundation of China (61025013, K12JB00200), the Program for Changjiang Scholars and Innovative Research Team in University. e-mails: mengll_83@hotmail.com, jiel@sfu.ca, {yzhao,hhbai,cylin}@bjtu.edu.cn, kaup@lnt.de.

systematically studied and compared to each other, especially for $M > 2$. For example, in [7], staggered quantizers are used to improve the central decoder of the two-description RD-MDC, but theoretical analysis is only derived for the special case when the low-rate quantizer stepsize is an integer multiple of the high-rate one. In [8], both unequal quantizations and unequal deadzones are suggested for MD video coding, but no theoretical analysis is given.

In this paper, based on TRPCSQ and the random quantization theory developed in [9], we obtain the closed-form expression of the expected distortion of the proposed method. This enables optimization of some system parameters, for example, the optimal transform. The proposed scheme is applied to lapped transform-based MD image coding. Experimental results show that the proposed scheme achieves better performance than TRPCSQ and TLMDC. We also study the problem of finding the optimal deadzones for the low-rate quantizers, and demonstrate its impact in image coding.

2. SYSTEM DESCRIPTION AND PERFORMANCE ANALYSIS

In this section, we describe the proposed uniformly offset quantizer-based M-channel MDC scheme using, and derive the closed-form expression of its expected distortion.

2.1. System Description

In TRPCSQ, the uniform offsets among different low-rate coded quantizers are achieved by shifting the bins of a uniform quantizer by multiple of $q_1/(M-1)$ in different descriptions. In addition, the prediction $\bar{x}_i$ is also quantized by a uniform quantizer with the same stepsize q_1. Therefore during the reconstruction, after shifting the quantization partition according to the reconstructed prediction, the uniform offsets among different quantizers are always maintained. However, the drawback of TRPCSQ is that shifting the bins of a uniform quantizer leads to asymmetric bins with respect to zero, which reduce the coding efficiency, especially at low rates.

In this paper, instead of shifting the bins of a quantizer by different amounts, we generate the initial uniformly offset quantizers by adopting quantizers of unequal deadzone sizes. We denote the proposed method multiple description coding with uniformly offset quantizers (MDUOQ).

In the i-th description of MDUOQ, the i-th subset is still encoded by a quantizer of stepsize q_0. Any other subset $j \neq i$ is first sequentially predicted from previously reconstructed subsets in the same description. The prediction is then quantized by a uniform quantizer with stepsize q_1, as in TRPCSQ. After that, the reconstructed prediction is used to obtain the prediction residual, and the residual is finally quantized by a deadzone quantizer with deadzone size of $2(\delta + \frac{l}{M-1})q_1$, where $2\delta q_1$ is the smallest deadzone size, and $l = (j - i - 1) \bmod M$. As a result, across all the M descriptions, each subset is predictively coded by $M-1$ low-rate quantizers with deadzones of $2(\delta + \frac{l}{M-1})q_1$, $l = 0, \ldots, M-2$, respectively; hence this also creates a uniform offset of $q_1/(M-1)$ among neighboring quantizers. At the decoder, we can refine the reconstruction by finding the intersection of all received high-rate and low-rate quantization bins.

However, different from TRPCSQ, the uniform offsets among low-rate quantizers are not always preserved in MDUOQ, due to the use of predictive coding and the different sizes of the deadzone and other reqular bins. That is, some joint quantization bin boundaries will be changed after adding the reconstructed prediction. For example, if a quantizer with deadzone $2\delta\, q_1$ is shifted to the right by prediction kq_1 ($k > 0$), it can be verified that k quantization bin boundaries within $[(1-\delta)q_1, (k-\delta)q_1]$ will generally have different (non-optimal) values from the original (unshifted) deadzone quantizer, but other bin boundaries still have the same values as the original quantizer. Similarly, if the prediction is negative ($k < 0$), only the k boundaries in $[(k+\delta)q_1, (-1+\delta)q_1]$ will be affected.

The problem above can be avoided by not using predictive coding, that is, quantizing the coefficients directly using different deadzone quantizers, but this would lose the coding efficiency of predictive coding. Our experimental results show that it is still beneficial to keep the predictive coding.

2.2. The Expected Distortion of the System

We next derive the closed-form expression of the expected distortion of the system, which allows us to optimize the system parameters, such as the transform and bit allocation, as in [4–6]. However, finding the closed-form expression of the expected distortion of MDUOQ is very challenging, due to the lack of closed-form R-D formula for deadzone quantizers, as well as the nonuniform offsets of some bins as described above, which are also caused by the different deadzones. Nevertheless, if we ignore the impacts of these nonuniform offsets, the performance of the joint dequantization of the low-rate quantizers in MDUOQ can be approximately obtained.

The expected distortion can be written as

$$D = \sum_{k=0}^{M} p_k D_k, \tag{1}$$

where $p_k = \binom{M}{k} p^{M-k}(1-p)^k$ is the probability of receiving k descriptions, and D_k is the corresponding expected distortion. When $k = 0$, D_k is simply the variance of the input.

Let R_0 and R_1 (bits/sample) be the average bit rate of the high-rate-coded and low-rate-coded subsets, respectively. Assume the overall bit rate constraint is R bits/sample/description, *i.e.*, $\frac{1}{M}(R_0 + (M-1)R_1) = R$.

In the proposed MDUOQ scheme, when k descriptions are available, k out of M subsets will be reconstructed from both high-rate and low-rate coding, and the rest will be jointly reconstructed from low-rate coding. We assume the quantization errors of different blocks are uncorrelated, and their contributions to the reconstruction error are additive. Therefore D_k can be written as

$$D_k = \frac{1}{M}\ (kD_{0,k} + (M-k)D_{1,k}), \qquad (2)$$

where $D_{0,k}$ is the average reconstruction error of subsets with one high-rate and $k-1$ low-rate codings. $D_{1,k}$ is the average reconstruction error of subsets with k low-rate codings.

We first find the expression of $D_{1,k}$, which is approximately given by Eq. (19) in [4].

$$D_{1,k} = \frac{q_1^2}{12} S_{1,k}, \qquad (3)$$

where

$$S_{1,k} = \frac{1}{(M-1)^2\binom{M-1}{k}} \sum_{l=1}^{M-k} \binom{M-2-l}{k-2} l^3. \qquad (4)$$

This joint de-quantization can be viewed as equivalent to a quantizer with reduced stepsize

$$q'_{1,k} = q_1\sqrt{S_{1,k}}. \qquad (5)$$

We next represent $D_{1,k}$ in term of R_1, the average rate of the low-rate coded subsets. Let the rate and entropy of each residual subset be $R_{1,i}$ and $h_{1,i}$, $i = 1, \ldots, M-1$. Assuming the rate is high and entropy coding is applied to encode the quantized coefficients, their relationship with q_1 is [10]

$$R_{1,i} = h_{1,i} - \log_2 q_1 = \frac{1}{2}\log_2(2\pi e\sigma_{1,i}^2) - \log_2 q_1, \qquad (6)$$

where we assume all the data are Gaussian, and $\sigma_{1,i}^2$ is the variance of the residual of the i-th subset. Note that although Laplacian model is more accurate in practice [11], Gaussian model is necessary here to facilitate the optimization of the transform, as in [4–6].

R_1 is the average of all $R_{1,i}$'s. That is,

$$R_1 = \frac{1}{M-1}\sum_{i=1}^{M-1} R_{1,i}. \qquad (7)$$

We can then represent q_1 by

$$q_1 = \sqrt{2\pi e}\left(\prod_{i=1}^{M-1}\sigma_{1,i}\right)^{\frac{1}{M-1}} 2^{-R_1} \triangleq \sqrt{2\pi e}\,\bar{\sigma}_1 2^{-R_1}, \qquad (8)$$

where $\bar{\sigma}_1$ is the geometric mean of all $\sigma_{1,i}$'s.

Therefore the distortion in (3) becomes

$$D_{1,k} = \frac{2\pi e}{12} S_{1,k}\,\bar{\sigma}_1^2\, 2^{-2R_1}. \qquad (9)$$

Next, we find the expression of $D_{0,k}$ in (2), *i.e.*, the average reconstruction error of subsets with one high-rate and $k-1$ low-rate codings. When $k = 1$, the distortion is simply given by high-rate coding. When $k > 1$, we first find the refined quantization bin from the $k-1$ low-rate codings. Let $q'_{1,k-1}$ be the final quantization step, as in (5). We next combine the refined low-rate quantization bin and the quantization of the high-rate coding to find the final quantization bin. This is equivalent to joint de-quantization from two randomly staggered uniform quantizers with stepsize q_0 and $q'_{1,k-1}$ respectively. The expected distortion of such a random quantization is studied in [9] and is given by Eq. (8) in it:

$$D_{0,k} = \frac{1}{12}q_0^2\frac{q'_{1,k-1} - \frac{3}{4}q_0}{q'_{1,k-1} - \frac{1}{2}q_0} \triangleq \frac{1}{12}q_0^2 S_{0,k}, \qquad (10)$$

where we define $S_{0,1} = 1$. This is equivalent to reducing the stepsize of the high-rate quantizer from q_0 to $q'_{0,k} = q_0\sqrt{S_{0,k}}$.

In terms of R_0 and R_1, $S_{0,k}$ and $D_{0,k}$ can be written as:

$$S_{0,k} = \frac{\sqrt{S_{1,k-1}}\,\bar{\sigma}_1 2^{-R_1} - \frac{3}{4}\sigma_0 2^{-R_0}}{\sqrt{S_{1,k-1}}\,\bar{\sigma}_1 2^{-R_1} - \frac{1}{2}\sigma_0 2^{-R_0}}, \qquad (11)$$

$$D_{0,k} = \frac{2\pi e}{12} S_{0,k}\sigma_0^2 2^{-2R_0}, \qquad (12)$$

where σ_0^2 is the entropy power of the signal at high-rate coding.

Plugging $D_{0,k}$ and $D_{1,k}$ into (2) and (1), the general expression of the expected distortion becomes

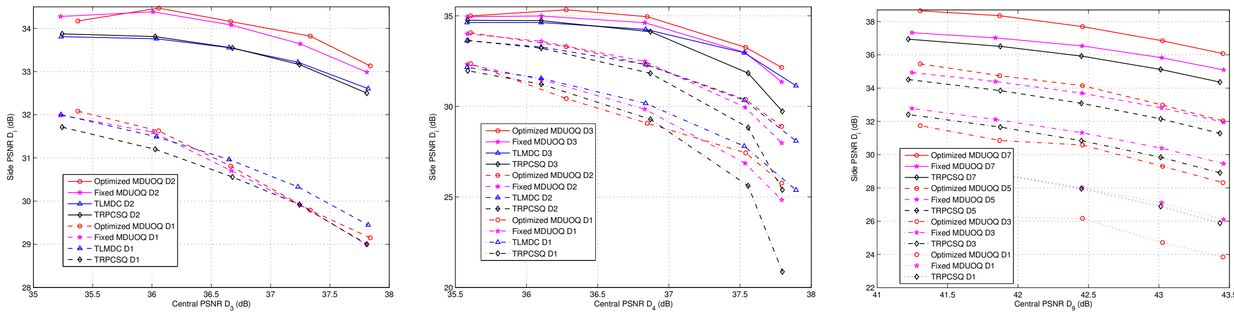
$$D = \frac{2\pi e}{12}\left(\sum_{k=1}^{M}\frac{kp_k}{M}S_{0,k}\right)\sigma_0^2 2^{-2R_0} + \frac{2\pi e}{12}\left(\sum_{k=1}^{M}\frac{(M-k)p_k}{M}S_{1,k}\right)\bar{\sigma}_1^2 2^{-2R_1} + p_0 D_0, \qquad (13)$$

where D_0 is the variance of the input signal.

3. OPTIMIZATION OF THE DEADZONES

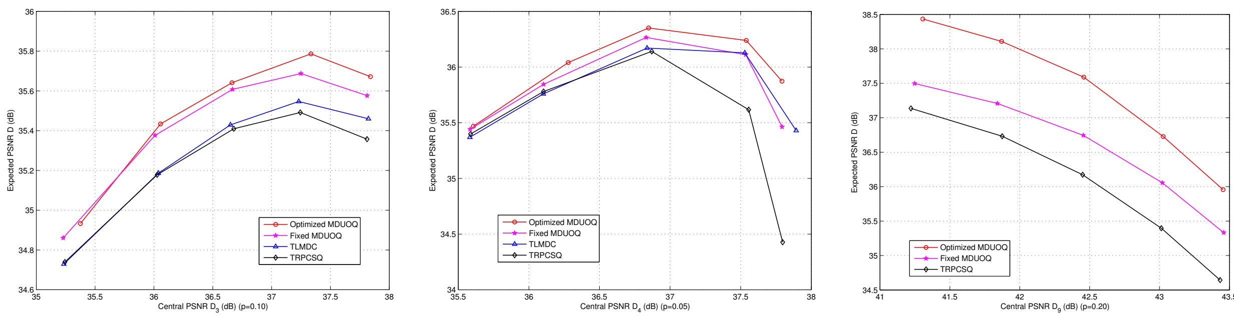
As described in Sec. 2.1, the smallest deadzone size $2\delta\, q_1$ determines all other deadzone sizes in the proposed scheme. Therefore the choice of δ is an important factor that affects the performance of the system. A simple solution is to always fix δ. Another way is to find the optimal δ^* for each case. An important question is how much is the difference between the two approaches. In this paper, we use a simple method to search the optimal δ^*. The two approaches will be compared in the next section.

To find the optimal δ^* for the given input signal with target bit rate R^*, loss probability p, and q_0, the algorithm loops through different δ with a step size of 0.05. For the given



(a) (b) (c)

Fig. 1. The side PSNRs vs. central PSNR tradeoff. (a) Boat: $M = 3$, total rate 1.0 bpp. (b) Peppers: $M = 4$, total rate 1.0 bpp. (c) Lena: $M = 9$, total rate 2.0 bpp.



(a) (b) (c)

Fig. 2. The expected PSNR vs. central PSNR. (a) Boat: $M = 3$, total rate 1.0 bpp, $p = 0.10$. (b) Peppers: $M = 4$, total rate 1.0 bpp, $p = 0.05$. (c) Lena: $M = 9$, total rate 2.0 bpp, $p = 0.20$.

q_0 and δ, it adjust q_1 until the bit rate meets the target. The expected distortions corresponding to all δ are calculated, and the optimal deadzone for the given q_0 can be found from them.

4. EXPERIMENTAL RESULTS

We now compare the performances of the proposed MDUOQ with TRPCSQ [4] and TLMDC [6] in MD image coding. As in TRPCSQ and TLMDC, the time domain lapped transform-based image coding in [12] is used. For the proposed method with the fixed minimal deadzone size, experimental results with some typical images show that $\delta = 0.6$ gives the best overall result. This is used in the following tests.

Fig. 1 compares performances of four methods: MDUOQ with optimal deadzone, MDUOQ with fixed deadzone, TRPCSQ and TLMDC. Three images with different numbers of descriptions are tested. The figures show the tradeoff between the central PSNR D_M and side PSNR D_i. For $M = 9$, not all D_i are displayed to avoid too crowded figure.

It can be seen from the figure that when $M = 3$ and $M = 4$, the proposed method with the optimal deadzone has similar performance to the fixed deadzone method. However, when $M = 9$, using the optimal deadzone can improve more than 1 dB when the number of received descriptions is large.

Compared to TRPCSQ and TLMDC, it can be seen from Fig. 1 that for the same central PSNR, the side PSNR of the MDUOQ is better when more descriptions are available. When very few descriptions are received, TLMDC and TRPCSQ sometimes can be better than MDUOQ, but since the probability of these cases is usually lower than that of receiving more descriptions, the proposed method will have better expected performance. This can be verified from Fig. 2, which shows the expected distortion (in PSNR) under different loss probability p for different central PSNR. These figures can be used to identify the optimal bit allocation for the given loss probability p. Note that TLMDC is not shown in Fig. 2 (c), because TLMDC is mainly designed for small p.

5. CONCLUSION

We develop an improved MDC scheme that uses unequal deadzones to generate uniformly offset quantizers. The theoretical performance is obtained via random quantization theory. Experimental results show that it outperforms existing staggered quantizer method that shifts the bins of a uniform quantizer. We also study the optimal deadzone, and found that the gain of optimal deadzone is significant when the number of descriptions is large. An open problem is how to find a fast algorithm to determine the optimal deadzone for given image, bit rate and channel condition.

6. REFERENCES

[1] N. Jayant, "Subsampling of a DPCM speech channel to provide two self-contained half-rate channels," *Bell Syst. Tech. Jour.*, vol. 60, pp. 501–509, Apr. 1981.

[2] E. Baccaglini, T. Tillo, and G. Olmo, "A flexible R-D-based multiple description scheme for JPEG 2000," *IEEE Signal Process. Lett.*, vol. 14 (3), pp. 197–200, Mar. 2007.

[3] T. Tillo, E. Baccaglini, and G. Olmo, "A flexible multi-rate allocation scheme for balanced multiple description coding applications," *Proc. IEEE 7th Workshop on Multimedia Signal Processing*, pp. 1–4, Nov. 2005.

[4] U. Samarawickrama, J. Liang, and C. Tian, "M-channel multiple description coding with two-rate coding and staggered quantization," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 20 (7), pp. 933–944, Jul. 2010.

[5] G. Sun, U. Samarawickrama, J. Liang, C. Tian, C. Tu, and T. D. Tran, "Multiple description coding with prediction compensation," *IEEE Trans. Image Proc.*, vol. 18 (5), pp. 1037–1047, May 2009.

[6] U. Samarawickrama, J. Liang, and C. Tian, "A three-layer scheme for M-channel multiple description image coding," *Sig. Proc.*, vol. 91 (10), pp. 2277–2289, Oct. 2011.

[7] T. Tillo and G. Olmo, "Improving the performance of multiple description coding based on scalar quantization," *IEEE Signal Process. Lett.*, vol. 15, pp. 329–332, 2008.

[8] S. Zhu, S.-K. Au Yeung, and B. Zeng, "A comparative study of multiple description video coding in P2P: normal MDSQ versus flexible dead-zones," *Proc. IEEE Conf. Multimedia and Expo*, pp. 1118–1121, 2009.

[9] V. K. Goyal, "Scalar quantization with random thresholds," *IEEE Signal Process. Lett.*, vol. 18, pp. 525 – 528, 2011.

[10] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, Boston, 2002.

[11] X. Li, N. Oertel, A. Hutter, and A. Kaup, "Laplace distribution based Lagrangian rate distortion optimization for hybrid video coding," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 19, no. 2, pp. 193–205, 2009.

[12] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-processing," *IEEE Trans. Signal Proc.*, vol. 51 (6), pp. 1557–1571, Jun. 2003.