MULTIPLE DESCRIPTION CODING WITH RANDOMLY OFFSET QUANTIZERS

Lili Meng[1], *Jie Liang*[2], *Upul Samarawickrama*[3], *Yao Zhao*[1], *Huihui Bai*[1] *and André Kaup*[4]

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

ABSTRACT

A multiple description coding scheme based on prediction-induced randomly offset quantizers is proposed, where each description encodes one source subset with a small quantization stepsize, and other subsets are predictively coded with a large quantization stepsize. Due to the prediction, the quantization bins that a coefficient belongs to in different descriptions are randomly overlapped with each others. The optimal reconstruction is obtained by finding the intersection of all received quantization bins. Using the recently developed random quantization theory, the closed-form expression of the expected distortion is obtained. The proposed scheme is then applied to lapped transform-based multiple-description image coding, and an iterative optimization scheme is developed to find the optimal lapped transform. Experimental results show that the proposed scheme achieves better performance than other methods in this category.

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

1. INTRODUCTION

Multiple description coding (MDC) addresses the packet losses in a communication network by sending M ($M \geq 2$) descriptions of the source such that the reconstruction quality improves with the number of received descriptions.

In [1], a multiple description scalar quantizer (MDSQ) method is developed. However, the MDSQ requires complicated index assignment. In [2], a two-stage modified MDSQ (MMDSQ) that is asymptotically optimal is designed, where two staggered uniform scalar quantizers are used to generate the first layer bits of each description, respectively. When both descriptions are received, another uniform scalar quantizer is used to further partition the joint bins of the two staggered quantizers. The output of the second-layer quantizer is evenly split into the two descriptions.

Source splitting is another method to generate multiple descriptions. The earliest example appeared in [3], where each description includes half of the input samples, which are used to predict the missing samples if only one description is received. In [4], transform coding is used, and each description includes one subset at high rate and another subset at low rate. In [5], the two-rate approach is applied to JPEG 2000-based M-description coding under the name RD-MDC.

In [6], a prediction compensated MDC (PCMDC) scheme is developed for $M = 2$. Instead of encoding another subset directly at a low rate, its prediction residual is encoded, which is more efficient than [4,5]. In [7], a M-channel MDC scheme is proposed using two-rate predictive coding and staggered quantization (TRPCSQ), where the low-rate coding is made mutually refinable using uniformly staggered quantizers. To preserve uniform staggering, the prediction is also quantized. In [8], 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.

In this paper, we propose an improved MDC scheme for MDLTPC, TRPCSQ and TLMDC in [6–8]. As in TLMDC in [8], the new method uses two-rate predictive coding and sequential prediction. However, instead of simply averaging the low-rate reconstructions from different descriptions, an improved reconstruction is obtained by finding the intersection of all received quantization bins, which have random offsets due to the predictions. Moreover, different from [6–8], the reconstructions of the high-rate coded subsets are also refined by the refined low-rate reconstruction.

Although staggered quantizers have been used in various MD schemes, their theoretical and image coding performances have not been systematically studied, especially for $M > 2$. For example, in [9], 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 [10], a total variation-based optimization method is developed to get better MDC decoder, where the intersection of all received quantization bins is used as an optimization constraint.

In this paper, based on the random quantization theory recently developed in [11], we obtain the closed-form expression of the theoretical performance of the proposed method for any value of M. We then apply the method to lapped transform-based MD image coding, and an iterative optimization scheme is developed to find the optimal lapped transform. Experimental results show that the proposed scheme achieves better performance than [6–8].

2. SYSTEM DESCRIPTION AND PERFORMANCE ANALYSIS

In this section, we describe the framework of the proposed MDC scheme, based on prediction-induced randomly offset quantizers

This work was supported in part by China Scholarship Council, the Natural Sciences and Engineering Research Council (NSERC) of Canada under grants RGPIN312262, STPGP350740 and STPGP380875, Alexander von Humboldt Foundation and the National Science Foundation of China for Distinguished Young Scholars (No.61025013). e-mails: mengll_83@hotmail.com, jiel@sfu.ca, upsamara@microsoft.com, {yzhao,hhbai}@bjtu.edu.cn, kaup@lnt.de.

(MDROQ), and derive the closed-form expression of its expected distortion.

2.1. System Description

In the proposed MDROQ method, to get M descriptions, the input correlated samples are conceptually partitioned into M subsets. In the encoder, each description encodes one subset of the source with a smaller quantization stepsize q_0. All other subsets are sequentially predicted from the previously encoded subsets in the same description, and the prediction residuals are encoded with a larger quantization stepsize q_1. This is the same as the sequential prediction in TLMDC [8], but is simpler than TRPCSQ [7].

The main difference of the proposed scheme from [6–8] is to jointly reconstruct each sample from all received descriptions, based on the intersection of all received quantization bins.

Assume x is quantized with stepsize q_0 in Description 0. Let $\bar{x}_i$ be the prediction of x in the i-th description from previously reconstructed samples in the same description, and e_i the corresponding prediction residual for x. e_i is quantized with stepsize q_1, and the reconstructed residual is denoted by $\hat{e}_i$. The reconstruction of x in the i-th description is thus:

$$\hat{x}_i = \bar{x}_i + \hat{e}_i. \tag{1}$$

As a result, the quantization of e_i induces a bin partition for x with the same stepsize q_1, but the bins are shifted from that of e_i by $\bar{x}_i$.

When multiple predictive codings of x are received, the bins that x belongs to in these descriptions will have some random offsets, caused by the different prediction values $\bar{x}_i$ in different descriptions, because they use different references for prediction. Clearly a refined reconstruction can be obtained if we find the intersection of all these bins, and then find the optimal reconstruction of the intersection.

Similarly, if both high-rate and low-rate coded versions of x are received, we can also refine the reconstruction of the high-rate coding by finding the random intersection of all received high-rate and low-rate quantization bins. The random offsets among different quantizers are also caused by the prediction of the low-rate coding. The benefit of this operation increases with the redundancy of the scheme, as q_1 will be closer to q_0. Note that the refinement of low-rate coding can only be used when $M \geq 3$, whereas the refinement of the high-rate coding can be used for $M \geq 2$. Therefore for two-description coding, only the latter is applicable.

2.2. General Formula of the Expected Distortion

We next derive the closed-form expression of the expected distortion of the proposed MDROQ scheme, which can be written as

$$D = \sum_{k=0}^{M} p_k D_k, \tag{2}$$

where $p_k = \binom{M}{k} p^{M-k}(1-p)^k$ is the probability of received 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 MDROQ 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

$$D_k = \frac{1}{M}\left(kD_{0,k} + (M-k)D_{1,k}\right), \tag{3}$$

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 $D_{1,k}$. This is quite challenging due to the random length of the intersection of the k low-rate quantization bins that a signal belongs to. We start by finding the leftmost bin of these k bins, and project the lower ends of other $k-1$ bins to the leftmost bin. This gives $k-1$ randomly chosen thresholds in the leftmost bin, which partition the bin into k intervals. The signal is in the rightmost interval, which is the intersection of all k bins. Therefore the reconstruction error is $D_{1,k} = E[U^2]/12$, where U is the length of the rightmost interval.

To find $E[U^2]$, note that this problem is similar to the random quantization problem recently studied by Goyal in [11], where a signal uniformly distributed in $[0, q_1)$ is quantized by a k-level quantizer with $k-1$ randomly selected thresholds in $[0, q_1)$. The difference is that in our problem, the signal is always in the rightmost interval, whereas it can be in any interval in [11]. However, as pointed out in [11], when the thresholds are uniformly and independently chosen from $[0, q_1)$, all the intervals have the same distribution. Therefore using order statistics theory, the expected distortion is found to be (Eq. (4) in [11])

$$D_{1,k} = E[U^2]/12 = \frac{q_1^2}{2(k+1)(k+2)} = \frac{q_1^2}{12} S_{1,k}, \tag{4}$$

where

$$S_{1,k} = \frac{6}{(k+1)(k+2)}. \tag{5}$$

Clearly, the more quantizers are available, the smaller the distortion will be. This joint de-quantization can be viewed as equivalent to a quantizer with reduced stepsize

$$q'_{1,k} = q_1 \sqrt{S_{1,k}}. \tag{6}$$

For comparison purpose, a k-level uniform scalar quantizer for a source in $[0, q_1]$ has an expected distortion of $q_1^2/(12k^2)$. Therefore the random quantizer is worse than the uniform quantizer by a factor of $6k^2/((k+1)(k+2))$, which approaches 6 as k increases [11].

We next represent $D_{1,k}$ in term of R_1, the average rate of the low-rate coded subsets. Note that although our problem has the same distortion as the random quantizer in [11], the rate formula in Eq. (2) of [11] is not applicable here, as it is the rate of encoding the index of a single quantizer with random thresholds, whereas R_1 in our system is the average rate of regular uniform quantizers.

Let the rate and entropy of each residual subset be $R_{1,i}$ and $h_{1,i}$, $i = 1, \ldots, M-1$. When the rate is high and entropy coding is used, their relationship with q_1 is

$$R_{1,i} = h_{1,i} - \log_2 q_1 = \frac{1}{2}\log_2(2\pi e \sigma_{1,i}^2) - \log_2 q_1, \tag{7}$$

where we assume all the data are Gaussian, and $\sigma_{1,i}^2$ is the variance of the residual of the i-th subset.

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

$$R_1 = \frac{1}{M-1} \sum_{i=1}^{M-1} R_{1,i}. \tag{8}$$

We can then represent q_1 by

$$q_1 = \sqrt{2\pi e}\left(\prod_{i=1}^{M-1}\sigma_{1,i}\right)^{\frac{1}{M-1}} 2^{-R_1} \triangleq \sqrt{2\pi e}\,\bar{\sigma}_1 2^{-R_1}, \quad (9)$$

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

Therefore the distortion in (4) becomes

$$D_{1,k} = \frac{2\pi e}{12} S_{1,k}\,\bar{\sigma}_1^2\, 2^{-2R_1}. \quad (10)$$

Next, we find the expression of $D_{0,k}$ in (3), *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 the 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 (6). 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 of q_0 and $q'_{1,k-1}$ respectively. The expected distortion of such a random quantization is also studied in [11] and is given by Eq. (8) in it:

$$D_{0,k} = \frac{1}{12} q_0^2 \frac{q'_{1,k-1} - \frac{3}{4}q_0}{q'_{1,k-1} - \frac{1}{2}q_0} \triangleq \frac{1}{12} q_0^2 S_{0,k}, \quad (11)$$

where $S_{0,1} = 1$, this means only receiving the high-rate code.

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

$$S_{0,k} = \frac{\sqrt{S_{1,k-1}}\,\bar{\sigma}_1 2^{-R_1} - \frac{3}{4}\sigma_0 2^{-R_0}}{\sqrt{S_{1,k-1}}\,\bar{\sigma}_1 2^{-R_1} - \frac{1}{2}\sigma_0 2^{-R_0}}, \quad (12)$$

$$D_{0,k} = \frac{2\pi e}{12} S_{0,k}\sigma_0^2 2^{-2R_0}, \quad (13)$$

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

Plugging $D_{0,k}$ and $D_{1,k}$ into (3) and (2), the general expression of the expected distortion becomes

$$\begin{aligned} D = {} & \frac{2\pi e}{12}\left(\sum_{k=1}^{M}\frac{kp_k}{M}S_{0,k}\right)\sigma_0^2 2^{-2R_0} \\ & + \frac{2\pi e}{12}\left(\sum_{k=1}^{M}\frac{(M-k)p_k}{M}S_{1,k}\right)\bar{\sigma}_1^2 2^{-2R_1} \; + \; p_0 D_0 \\ \triangleq {} & \frac{2\pi e}{12}\bar{S}_0\sigma_0^2 2^{-2R_0} + \frac{2\pi e}{12}\bar{S}_1\bar{\sigma}_1^2 2^{-2R_1} \; + \; p_0 D_0, \end{aligned} \quad (14)$$

where D_0 is the variance of the input signal.

2.3. TDLT based MD Image Coding and Optimization

We next apply the proposed MDROQ scheme to lapped transform-based MD image coding, which has been shown to be a framework with state-of-the-art MDC performance [6–8]. The time-domain lapped transform (TDLT) is used [12], which employs a prefilter at block boundaries before the DCT and a postfilter after the inverse DCT. The TDLT has been adopted by the JPEG XR standard, which is a low-cost alternative to JPEG 2000.

To get MDC, the transformed blocks are partitioned into M subsets after the DCT. In each description, one subset is coded with a higher bit rate R_0 and other subsets are coded with a lower rate R_1.

An attractive feature of the TDLT is that its pre/postfilters can be optimized for different applications. The objective is to find the optimal TDLT prefilter that minimizes the average distortion D in (14), subject to the bit rate constraint of $R_0 + (M-1)R_1 = MR$. In this case, for a given set of pre/post-filters, we can find the optimal bit allocation and the corresponding minimal distortion, then the optimal pre/post-filter can be found by minimizing the distortion. In the optimization, we assume the source is a unit-variance first-order Gauss-Markov source with correlation coefficient $\rho = 0.95$.

A naive way to solve the problem is to define a Lagrangian cost function $\mathcal{L} = D + \lambda(R_0 + (M-1)R_1 - MR)$, and using a numerical optimization program to minimize $\mathcal{L}$. However, the solution is very sensitive to λ.

In this paper, we propose an effective iterative approach to solve the problem. We first let all $S_{0,k} = 1$, *i.e.*, ignoring the refinement of the high-rate quantizers. In this case, since $S_{0,k}$ is no longer a function of R_0 and R_1, the distortion in (14) can be easily minimized by the Lagrangian multiplier method, and the optimal bit allocation is given by

$$\begin{aligned} R_0 &= min\left(MR, R + \frac{M-1}{2M}\log_2\frac{(M-1)\bar{S}_0\sigma_0^2}{\bar{S}_1\bar{\sigma}_1^2}\right), \\ R_1 &= max\left(0, R - \frac{1}{2M}\log_2\frac{(M-1)\bar{S}_0\sigma_0^2}{\bar{S}_1\bar{\sigma}_1^2}\right). \end{aligned} \quad (15)$$

We then use R_0 and R_1 above to calculate the $S_{0,k}$ in (12), which can then be used to update the bit allocation in (15). After that, the distortion in (14) is re-calculated. The iteration terminates when the distortion change is less than a threshold.

This iteration method does not need to select λ. The bit rate constraint is strictly met by (15). In addition, when applied to the lapped transform-based setup, the optimized pre/post-filters are not sensitive to the bit rate R and error probability p. When the block size is 8, the iteration above can always converge in less than five times with an accuracy of 10^{-6} in the expected distortion.

3. EXPERIMENTAL RESULTS

In this section, we compare the performances of MDROQ with MDLTPC, TRPCSQ, and TLMDC in MD image coding using testing images of various characteristics. It has been reported in [6–8] that these methods have better performances than many other methods such as the PCT/GPCT, MMDSQ, RD-MDC, and MDLVQ.

Fig. 1 (a) compares the two-description MDROQ and the MDLTPC in [6] for the image Lena. It can be seen that for the same side PSNR, the central PSNR of the MDROQ can be up to 0.5 dB better than the MDLTPC, due to the refinement of the high-rate quantizer. Alternatively, for the same central PSNR, the side PSNR of MDROQ achieves up to 0.3 dB gain over MDLTPC.

Fig. 1 (b)-(d) compare the relationships between the side PSNR D_i and central PSNR D_M of MDROQ and TRPCSQ [7] for different images when $M = 3$, 4, and 9, respectively. To avoid too crowded figures, only odd i's are shown in Fig. 1 (d). It can be seen that the proposed MDROQ outperform TRPCSQ in almost all cases. Up to 5 dB gain can be obtained when the redundancy is low, *i.e.*, when q_1 is large, which corresponds to the right corners of the curves. This is because the asymmetric quantizers in TRPCSQ is not efficient at low rates. When there is moderate or high redundancy, the gain of side distortion is still up to 0.5 dB.

Fig. 1 (b) also includes results of the TLMDC in [8], which is better than TRPCSQ. The TLDMC adds a third layer to improve the quality when $M-1$ descriptions are received. Although the proposed methods do not have the third layer, they can still get better

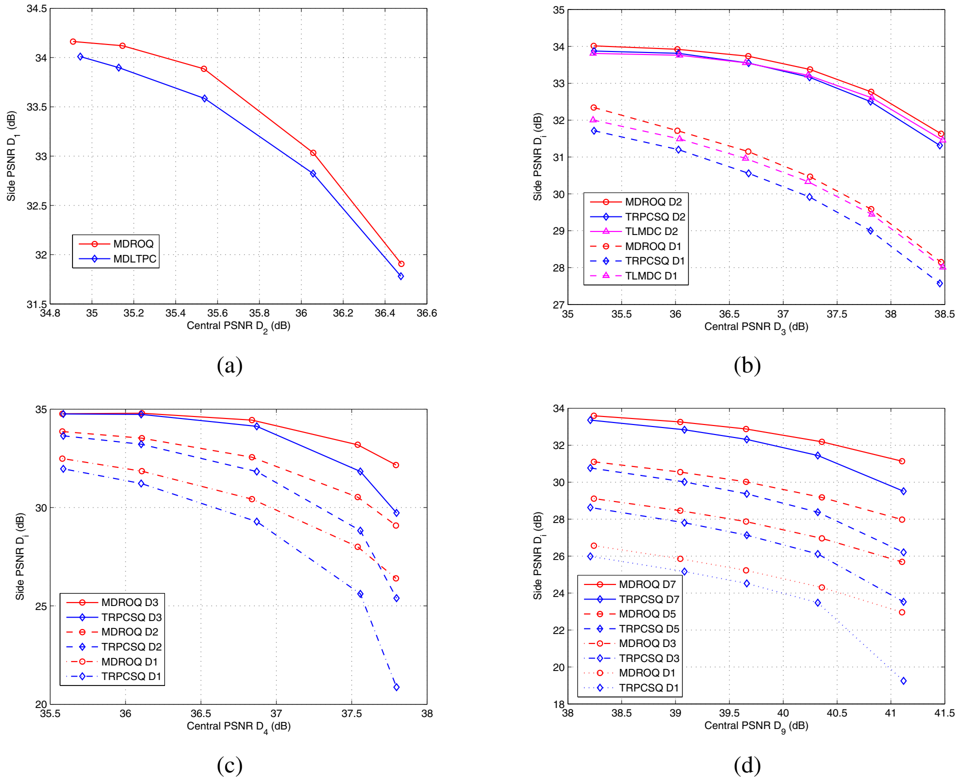

Fig. 1. The side PSNR vs. central PSNR. (a) Lena: $M = 2$, total rate 0.5 bpp. (b) Boat: $M = 3$, total rate 1.0 bpp. (c) Pepper: $M = 4$, total rate 1 bpp . (d) Couple: $M = 9$, total rate 2 bpp.

overall performance than TLMDC, due to improved joint dequantization and the refinement of the low-rate coding.

4. CONCLUSION

A multiple description coding scheme with randomly offset quantizations is proposed. The closed-form expression of the expected distortion of the proposed scheme is obtained based on the random quantization theory in [11]. The proposed scheme is applied to lapped transform-based MD image coding, and better performance than other methods in this category can be achieved. The proposed scheme can also be generalized by using quantizers with different deadzones, and the results will be reported at other places.

5. REFERENCES

[1] V. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 39 (3), pp. 821–834, May 1993.

[2] C. Tian and S. Hemami, "A new class of multiple description scalar quantizer and its application to image coding," *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 329–332, Apr. 2005.

[3] 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.

[4] W. Jiang and A. Ortega, "Multiple description coding via polyphase transform and selective quantization," *Proc. SPIE Conf. on Visual Communications and Image Processing*, vol. 3653, pp. 998–1008, Feb. 1999.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] S. Zhu, S.-K. Au Yeung, B. Zeng, and J. Wu, "Total-variation based picture reconstruction in multiple description image and video coding," *Signal Processing: Image Communication*, vol. 27, no. 2, pp. 126–139, 2012.

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

[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.