

# Multiple Description Coding With Randomly and Uniformly Offset Quantizers

Lili Meng, Jie Liang, *Senior Member, IEEE*, Upul Samarawickrama, Yao Zhao, *Senior Member, IEEE*, Huihui Bai, and André Kaup, *Fellow, IEEE*

**Abstract**—In this paper, two multiple description coding schemes are developed, based on prediction-induced randomly offset quantizers and unequal-deadzone-induced near-uniformly offset quantizers, respectively. In both schemes, each description encodes one source subset with a small quantization stepsize, and other subsets are predictively coded with a large quantization stepsize. In the first method, due to predictive coding, the quantization bins that a coefficient belongs to in different descriptions are randomly overlapped. The optimal reconstruction is obtained by finding the intersection of all received bins. In the second method, joint dequantization is also used, but near-uniform offsets are created among different low-rate quantizers by quantizing the predictions and by employing unequal deadzones. By generalizing the recently developed random quantization theory, the closed-form expression of the expected distortion is obtained for the first method, and a lower bound is obtained for the second method. The schemes are then applied to lapped transform-based multiple description image coding. The closed-form expressions enable the optimization of the lapped transform. An iterative algorithm is also developed to facilitate the optimization. Theoretical analyzes and image coding results show that both schemes achieve better performance than other methods in this category.

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

Manuscript received August 19, 2012; revised April 30, 2013 and August 12, 2013; accepted October 23, 2013. Date of publication November 5, 2013; date of current version December 17, 2013. This work was supported in part by the China (973 Program) under Grant 2012CB316400, in part by the China Natural Science Funds for Distinguished Young Scholar under Grant 61025013, in part by the National NSF of China under Grants 61210006, 61202240, and 61272051, in part by the Program for Changjiang Scholars and Innovative Research Team in University, in part by the China Scholarship Council, in part by the NSERC of Canada under Grants RGPIN312262, STPGP350740, and STPGP380875, and in part by Alexander von Humboldt Foundation of Germany. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. David S. Taubman.

L. Meng was with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China. She is now with the School of Information Science and Engineering, Shandong Normal University, Jinan 730050, China (e-mail: mengll83@hotmail.com).

J. Liang is with the School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada (e-mail: jiel@sfu.ca).

U. Samarawickrama is with Microsoft Corporation, Redmond, WA 98052 USA (e-mail: upsamara@microsoft.com).

Y. Zhao is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the State Key Laboratory of Rail Traffic Control and Safety, Beijing 100044, China (e-mail: yzhao@bjtu.edu.cn).

H. Bai is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100011, China (e-mail: hhhbai@bjtu.edu.cn).

A. Kaup is with the Department of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Erlangen 91058, Germany (e-mail: kaup@int.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2013.2288928

## I. INTRODUCTION

**M**ULTIPLE description coding (MDC) mitigates the impact of packet losses during transmission by sending  $M$  ( $M \geq 2$ ) descriptions of the source such that the reconstruction quality improves with the number of received descriptions [1].

In [2], a multiple description scalar quantizer (MDSQ) method is developed, which is asymptotically optimal at high rates [3]. Many extensions of MDSQ have been proposed, *e.g.*, scalable MDSQ [4], [5]. However, the MDSQ requires complicated index assignment. In [6], a two-stage modified MDSQ (MMDSQ) that is also 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. An improvement of the MMDSQ is reported in [7].

In [8], the MDSQ is extended to more than two channels via a combinatorial optimization approach. Another extension of the MDSQ is proposed in [9], which has multiple stages and each stage refines the preceding stages. However, both the methods in [8], [9] become quite complicated as  $M$  increases.

Lattice vector quantization-based MDC (MDLVQ) methods are studied in [10], [11], where  $M$  descriptions are generated by uniquely assigning each point in a finer central lattice to  $M$  points in a sublattice. These methods also involve a complicated index assignment problem.

Source splitting is another approach to generate multiple descriptions. One of the earliest designs appeared in [12], where the source samples are split into even-indexed and odd-indexed subsets, and each subset is encoded into one description. If one description is lost, the missing data in it are predicted from their neighbors in the other description, but the performance in this case approaches an asymptote at high rates because of the prediction error [1].

In [13], transform coding is used and the transform coefficients are also split into two subsets. Each subset is quantized to generate the base layer of one description. Each description also includes a coarsely quantized version of the other subset, which is used when the other description is lost. The optimal bit allocation between the two descriptions is studied.

In [14], a RD-MDC method is developed by generalizing [13] to JPEG 2000 for two-description coding, where each JPEG 2000 code-block is coded at two rates, one in each

description. The rate allocation is determined by Lagrangian optimization. The RD-MDC is extended in [15] to  $M$  channels ( $M > 2$ ), where each JPEG 2000 code-block is still 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 [16], a multi-rate method with  $M - 1$  degrees of freedom is developed, which generalizes the two-rate method in [15]. In this method, each subset of the source is coded at  $M$  different rates, one for each description.

The pairwise correlating transform (PCT) [17] introduces controlled redundancy before splitting the data, and the redundancy is adjusted by a set of  $2 \times 2$  correlating transforms. If one coefficient is lost, it is estimated from its counterpart in the other description. However, PCT does not perform well at high rates due to the prediction residual, similar to [12]. In [18], a generalized PCT (GPCT) is proposed to encode the prediction residual in each description, but no image coding result is reported. The PCT can be generalized to introduce correlation among more than two coefficients [19].

The erasure correcting codes can also be used to general multiple descriptions by providing unequal loss protections (ULP) to different layers of a scalable source code [20]. However, the method only has good performance when at least  $n$  descriptions are received, for some pre-specified  $n$ .

In [21], a multiple description lapped transform with prediction compensation (MDLTPC) is developed for two-description coding, 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 as the enhancement layer. The application of the MDLTPC in MD image coding achieves better performances than MMDSQ, RD-MDC, PCT, and GPCT.

In [22], another  $M$ -channel MDC scheme is proposed using two-rate predictive coding and staggered quantization (TRPCSQ), where the  $M - 1$  lower-rate quantizations for each coefficient are designed to be uniformly staggered using two operations: (1) The  $M - 1$  lower-rate quantizers for its prediction residuals in different descriptions are uniformly shifted to each other. (2) The corresponding predictions in different descriptions are also quantized by a uniform quantizer with the same quantization step. The TRPCSQ achieves better performance than MDLVQ and RD-MDC. However, the quantization of the prediction reduces the prediction accuracy, and the staggered quantizers become asymmetric around 0, which lose some coding efficiency, especially at low rates. In addition, it is well known that deadzone quantizer has better rate-distortion (R-D) performance than uniform quantizer, but TRPCSQ only uses shifted uniform quantizers to ensure uniform offsets.

A three-layer MDC (TLMDC) scheme is developed in [23], which generalizes MDLTPC 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 usually the dominant error scenario. It is shown that TLMDC can have better performance than TRPCSQ.

In this paper, we propose two improved MDC schemes for MDLTPC, TRPCSQ and TLMDC. As in TRPCSQ [22], both new methods use two-rate predictive coding. Sequential prediction is also used as in TLMDC [23]. In the first method, instead of simply averaging the low-rate reconstructions from different descriptions, an improved reconstruction is obtained by finding the intersection of all received bins, which have random offsets due to predictions. This is different from the uniform offsets in TRPCSQ. Another difference with TRPCSQ is that deadzone quantizers can be used in this method. Moreover, different from [21]–[23], the reconstructions of the high-rate coded subsets are also refined by the refined low-rate reconstruction. The second method is similar to TRPCSQ in the sense that it uses near-uniform offsets among different low-rate quantizers, which also requires the quantization of the predictions. However, different from TRPCSQ, the near-uniform offsets are achieved by employing unequal deadzones in different quantizers, which avoids the asymmetric quantizer problem in TRPCSQ, and has better R-D performance. We call the two methods multiple description coding with randomly offset quantizers (MDROQ) and uniformly offset quantizers (MDUOQ), respectively.

Although staggered quantizers and unequal deadzone quantizers have been used in various MD schemes, their theoretical and image coding performances have not been systematically studied and compared to each other, especially for  $M > 2$ . For example, in [24], 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 [25], 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 [26], both unequal quantizations and unequal deadzones are suggested for MD video coding, but no theoretical analysis is given.

In this paper, by generalizing the random quantization theory recently developed in [27], we obtain the closed-form expressions of the theoretical performances of MDROQ for any  $M$ . Our generalization also unifies all the results in [27]. We also obtain a lower bound for the distortion of MDUOQ. We then apply the two methods to lapped transform-based MD image coding. The closed-form expressions enable the optimization of the pre/post-filters in the lapped transform. An iterative algorithm is developed to facilitate the optimization. Theoretical analyses and image coding results show that the proposed schemes achieve better performance than [21]–[23].

## II. MULTIPLE DESCRIPTION CODING WITH PREDICTION-INDUCED RANDOMLY OFFSET QUANTIZERS

In this section, we describe the MDC framework using prediction-induced randomly offset quantizers (MDROQ), and derive the closed-form expression of its expected distortion.

### A. System Description

In the proposed MDROQ method, to get  $M$  descriptions, the input samples are first partitioned into  $M$  subsets, which

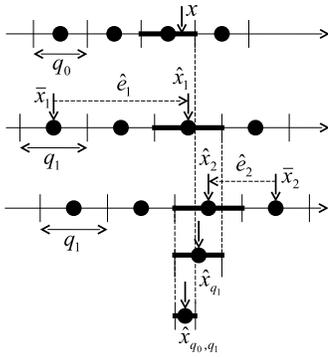


Fig. 1. A three-description coding example of joint de-quantization from prediction-induced random quantizers.

can be done at sample or block level, depending on the application. For one-dimensional (1-D) signals, the  $i$ -th subset includes the  $(nM + i)$ -th samples or blocks for all possible  $n$ , where  $i = 0, \dots, M - 1$ . Two-dimensional (2-D) signals can be partitioned into  $M$  subsets of 2-D blocks using some 2-D patterns, as shown in Fig. 5 in Sec. IV-D. In practices, block level partition is more useful, together with block transforms such as Discrete Cosine Transform (DCT). We can also partition the source at sample level, and apply DPCM to encode the data. This can be used to study the asymptotic performance of block coding, since the performance of block coding approaches that of the DPCM as the increase of the block size. The general formulas derived in this section are applicable to both sample-based and block-based methods. Later on, we will give detailed expressions for these cases.

In the encoder, the  $i$ -th description encodes the  $i$ -th subset using a uniform or deadzone scalar quantizer with a 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 by a uniform or deadzone scalar quantizer with a larger quantization stepsize  $q_1 > q_0$ . This is simpler, more flexible and has better R-D performance than the shifted uniform quantizers in TRPCSQ [22], especially when deadzone quantizers are used.

A more important difference of the proposed scheme from [21]–[23] is in the decoder. The key is to jointly reconstruct each sample from all received descriptions, based on the intersection of all received quantization bins.

A three-description example with uniform quantizers is shown in Fig. 1. We first look at the case where only low-rate coded versions of a sample  $x$  are available. Let  $x$  be quantized with stepsize  $q_0$  in Description 0, and  $\bar{x}_i$  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  $x$ . In the  $i$ -th description,  $e_i$  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:

$$\hat{x}_i = \bar{x}_i + \hat{e}_i. \quad (1)$$

As a result, the quantization of  $e_i$  induces a quantization partition for  $x$  with the same stepsize  $q_1$ , but the partition is shifted from that of  $e_i$  by  $\bar{x}_i$ , as shown in Fig. 1.

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. For example, at high rates, the middle point of the intersection can be used.

In Fig. 1,  $\hat{x}_{q_1}$  is the refined reconstruction from the two low-rate quantizers. Since the intersection is smaller than  $q_1$ , the refined reconstruction has less error.

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, as shown by the point  $\hat{x}_{q_0, q_1}$  in Fig. 1. 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.

### B. General Formula of the Expected Distortion

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

$$D = \sum_{k=0}^M p_k D_k, \quad (2)$$

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 mean squared error (MSE). 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  $k$  ( $k > 0$ ) 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

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

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

1) *Expression of  $D_{1,k}$* : Finding the expression of  $D_{1,k}$  is generally quite challenging, due to the random length of the intersection of the  $k$  low-rate quantization bins that a signal belongs to, as shown in Fig. 1. However, when all quantizers are uniform quantizers and when the bit rate is high ( $q_1$  is very small), a simple closed-form expression can be obtained.

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 in [27], 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 [27]. However, as pointed out in [27], when the thresholds are uniformly and independently chosen from  $[0, q_1)$ , all the intervals have the same distribution. Therefore using order statistics theory and with the assumption of high rate coding, the expected distortion is found to be (Eq. (4) in [27])

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

where

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

Clearly, the more quantizers are available, the smaller the distortion will be.

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 [27], the rate formula in Eq. (2) of [27] is not applicable here, as it is the rate of encoding the index of a single scalar quantizer with random thresholds, whereas  $R_1$  in our system is the average rate of several uniform scalar quantizers.

Let the rate and entropy of each residual subset be  $R_{1,i}$  and  $h_{1,i}$ ,  $i = 1, \dots, M - 1$ . Assuming the rate is high and entropy coding is applied to encode the quantized coefficients, their relationship with  $q_1$  is [28]

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

where we assume all the data are Gaussian, and  $\sigma_{1,i}^2$  is the variance of the prediction residual of the  $i$ -th subset. When block coding is used, its expression will be derived in Eq. (32) in Sec. IV.  $R_1$  is the average of all  $R_{1,i}$ 's. That is,

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

We can then represent  $q_1$  by

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

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

Therefore  $D_{1,k}$  in (4) becomes

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

2) *Expression of  $D_{0,k}$* : We next find the expression of  $D_{0,k}$  in (3) when uniform quantizers are used in both low-rate and high-rate quantizers. In this case, the problem is to find the MSE of source subsets with one high-rate and  $k - 1$  low-rate quantizers, and the quantization partitions of the  $k - 1$  low-rate quantizers have random offsets with respect to that of the high-rate quantizer. Note that when  $k = 1$ , the distortion is simply given by the high-rate quantizer, so we focus on  $k > 1$ .

This is more general than the special case of random quantization studied in Sec. IV of [27], where only two uniform quantizers with different stepsizes are considered. However, the proof in Eq. (7)–(8) of [27] can be generalized to study the more general problem here.

We can focus on finding the MSE of the joint reconstruction within one high-rate bin, *e.g.*,  $[0, q_0]$ , in which  $x$  is uniformly distributed. Since the random offsets of the low-rate quantizers are in  $[0, q_1)$  and  $q_1 > q_0$ , the number of random thresholds in  $[0, q_0]$  is between 0 and  $k - 1$ . This is different from the low-rate joint reconstruction in finding  $D_{1,k}$ , where there are always  $k - 1$  thresholds in each bin. Our goal is to find the size of the refined bin into which the source  $X$  falls. Since all offsets are random, all refined bin lengths are identically distributed; hence we only need to work with the first refined bin with left boundary at 0 [27]. We denote its length by  $V$ .

Let  $a_i \in [0, q_1)$  be the random offset of the  $i$ -th low-rate quantizer with stepsize  $q_1$ . Since  $q_1 > q_0$ , it is possible that the range  $[0, q_0]$  is completely included in a bin of the low-rate quantizer. Therefore the length  $V$  is given by

$$V = \begin{cases} a_{(1)}, & \text{if } a_{(1)} \in [0, q_0], \\ q_0, & \text{if } a_{(1)} \in (q_0, q_1), \end{cases} \quad (10)$$

where  $a_{(1)}$  is the minimum value of all  $a_i$ 's, or their first order statistic, whose pdf is [27]:

$$f_{a_{(1)}}(v) = \frac{1}{q_1^{k-1}} (k-1)(q_1 - v)^{k-2}, \quad 0 \leq v \leq q_1. \quad (11)$$

To find the distribution of  $V$ , note that the probability that  $V = q_0$  is  $\int_0^{q_1} f_{a_{(1)}}(v) dv = (1 - q_0/q_1)^{k-1}$ , *i.e.*, all  $a_i$ 's are greater than  $q_0$ . Define  $r \triangleq q_0/q_1$ , the pdf of  $V$  is thus

$$f_V(v) = f_{a_{(1)}}(v) + (1-r)^{k-1} \delta(v - q_0). \quad (12)$$

Let  $U$  be the size of the intersection into which  $X$  falls, whose pdf is the product of those of  $X$  and  $V$  [27]. Since  $X$  is uniformly distributed, we have

$$f_U(u) = \frac{uf_V(u)}{\int_0^{q_0} v f_V(v) dv}. \quad (13)$$

It can be shown that

$$\int_0^{q_0} v f_V(v) dv = \frac{q_1 (1 - (1-r)^{k-1})}{k} \triangleq C_1 \quad (14)$$

and the MSE of the joint reconstruction is

$$\begin{aligned} D_{0,k} &= \frac{E[U^2]}{12} = \frac{1}{12C_1} \int_0^{q_0} u^3 f_V(u) du \\ &= \frac{1}{12C_1} \left( q_0^3 (1-r)^{k-1} + \int_0^{q_0} u^3 f_{a_{(1)}}(u) du \right) = \frac{q_0^2}{12} S_{0,k}, \end{aligned} \quad (15)$$

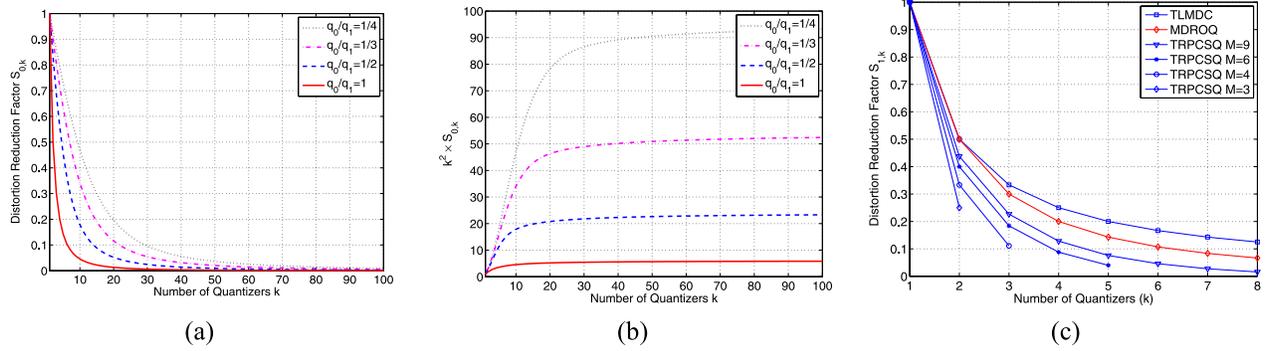


Fig. 2. (a)  $S_{0,k}$  in MDROQ with different  $k$  and  $q_0/q_1$ . (b)  $k^2 S_{0,k}$  in MDROQ with different  $k$  and  $q_0/q_1$ . (c)  $S_{1,k}$  of different reconstruction methods.

where

$$S_{0,k} = \frac{1}{1 - (1-r)^k} \left( \frac{6}{r^2(k+1)(k+2)} - (1-r)^k C_2 \right), \quad (16)$$

$$C_2 = \frac{6}{(k+1)(k+2)} \left( \frac{1-r}{r} \right)^2 + \frac{6}{k+1} \left( \frac{1-r}{r} \right) + 3. \quad (17)$$

It can be verified that  $S_{0,1} = 1$ . Also, when  $k = 2$ ,  $D_{0,k}$  reduces to  $\frac{q_0^2}{12} \times \frac{q_1 - 3/4q_0}{q_1 - 1/2q_0}$ , which is Eq. (8) of [27] for the two-quantizer case. Moreover, when  $r = 1$  ( $q_1 = q_0$ ),  $S_{0,k}$  reduces to  $S_{1,k}$  in Eq. (5). Therefore this is a more general result that subsumes all cases in [27].

Fig. 2 (a) plots  $S_{0,k}$  for different  $k$  and  $q_0/q_1$ , which shows that  $S_{0,k}$  decreases with  $k$ , but the decaying rate drops as  $q_1$  grows with respect to  $q_0$ . To further understand the impact of  $k$  and the ratio  $r$  (i.e.,  $q_0/q_1$ ), we compare with the joint reconstruction from  $k$  uniformly offset quantizers, all with the same stepsize  $q_0$ . The corresponding MSE is simply  $q_0^2/(12k^2)$ . Therefore the ratio  $k^2 S_{0,k}$  represents the penalty factor of using random offsets and larger stepsize  $q_1$  in  $k-1$  quantizers. It is easy to see that  $k^2 S_{0,k}$  approaches  $6/r^2$  as  $k \rightarrow \infty$ , as shown in Fig. 2 (b). In the special case of  $r = 1$ ,  $k^2 S_{0,k} \rightarrow 6$ , which was observed in [27].

We can rewrite  $S_{0,k}$  and  $D_{0,k}$  in terms of  $R_0$  and  $R_1$  using Eq. (8) and  $q_0 = \sqrt{2\pi e \sigma_0} 2^{-R_0}$ , where  $\sigma_0^2$  is the entropy power of the signal in the high-rate coding. This leads to

$$D_{0,k} = \frac{2\pi e}{12} S_{0,k} \sigma_0^2 2^{-2R_0}. \quad (18)$$

When block coding is used, the expression of  $\sigma_0^2$  will be derived in Eq. (29) in Sec. IV.

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

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

$$\triangleq \frac{2\pi e}{12} \bar{S}_0 \sigma_0^2 2^{-2R_0} + \frac{2\pi e}{12} \bar{S}_1 \bar{\sigma}_1^2 2^{-2R_1} + p_0 D_0 \quad (19)$$

where  $D_0$  is the variance of the input signal. It should be noted that the factor  $\bar{S}_0$  depends on  $R_0$  and  $R_1$ . This poses some

difficulties for the optimization of the system. An iterative algorithm will be developed in Sec. IV to resolve this problem.

### III. MULTIPLE DESCRIPTION CODING WITH UNEQUAL-DEADZONE-INDUCED UNIFORMLY OFFSET QUANTIZERS

In this section, we develop another MDC scheme using deadzone-based near-uniformly offset quantizers, motivated by the theoretical advantage of uniformly offset quantizers over randomly offset quantizers.

#### A. Comparison of Low-Rate Joint Reconstruction Methods

The proposed scheme improves the TRPCSQ scheme in [22], where the low-rate quantization bins of the same subset in different descriptions are obtained by uniformly shifting the bins of a uniform quantizer. It is shown in [22] that after joint reconstruction from  $k$  ( $k \geq 2$ ) descriptions, the distortion of the low-rate-coded samples is also in the form of (4), but the distortion reduction factor  $S_{1,k}$  also depends on  $M$ :

$$S_{1,k} = \frac{1}{(M-1)^2 \binom{M-1}{k}} \sum_{l=1}^{M-k} \binom{M-2-l}{k-2} l^3. \quad (20)$$

Fig. 2 (c) compares the factor  $S_{1,k}$  for the low-rate coded subsets in MDROQ and TRPCSQ, which shows that the uniformly offset quantizer in TRPCSQ outperforms the random quantizer in MDROQ. For the same  $k$ , the gap between the two joint reconstruction methods reduces as the increase of  $M$ , whereas for the same  $M$ , the gap increases with  $k$ . Note that Fig. 2 (c) is only for  $S_{1,k}$ , and does not consider the rates of different quantizers. The overall R-D performance of TRPCSQ is actually worse than TLMDC and MDROQ.

To gain more insights, we consider the two special cases of  $k = 2$  and  $k = M - 1$ . When  $k = 2$ ,  $S_{1,k}$  in Eq. (5) becomes  $1/2$ , and (20) reduces to  $\frac{M-2}{2(M-1)}$ , which approaches to  $1/2$  as  $M$  increases. Therefore the uniformly staggered quantizer would have similar performance to the random quantizer. This is because for large  $M$ , the offset between neighboring uniformly staggered quantizers is very small, thus the offset between any two quantizers is essentially random.

When  $k = M - 1$ ,  $S_{1,k}$  for TRPCSQ in (20) becomes  $1/k^2$ . Therefore the combination of the  $k$  quantizers reduces to a uniform quantizer with stepsize of  $q_1/k$ . In this case,

as pointed out in [27], the random quantizer is worse than the uniform quantizer by a factor of  $6k^2/((k+1)(k+2))$ , which approaches to 6 for large  $k$ .

Fig. 2 (c) also includes the scaling factor  $S_{1,k}$  in TLMDC [23], where the simple average of all lower-rate coded reconstructions of the subset from the received descriptions is used as the final reconstruction. Note that the theoretical derivation in [23] is more general as it allows different quantization stepsizes for different subset residuals (see Eq. (6) in [23]), but to reduce the complexity, the experimental results in [23] are obtained with the same stepsize for all residuals, as in TRPCSQ and this paper. To get  $S_{1,k}$  for TLMDC in this case, suppose  $\hat{x}_i = x + n_i$  ( $i = 1, \dots, k$ ) are the  $k$  low-rate reconstructions of  $x$ , where  $n_i$ 's are reconstruction noises. Their average is  $\hat{x} = x + \frac{1}{k} \sum_i n_i$ . Since the reconstructions in different descriptions involve different predictions,  $n_i$ 's can be assumed uncorrelated. Therefore the noise variance of the averaged reconstruction is  $\frac{q_1^2}{12k}$ . This is a special case of Eq. (14) in [23] when the same quantization stepsize  $q_1$  is used for all subset residuals. Compared to (4), we get

$$S_{1,k} = 1/k, \quad \forall M. \quad (21)$$

As shown in the figure, except for  $k = 2$ , the averaging method in TLMDC has worse performance than the random quantization in MDROQ, and the gap increases with  $k$ .

The comparison above shows that uniformly offset quantizers are theoretically superior to randomly offset quantizers. Therefore it is necessary to investigate how to improve the performance of the MDROQ using uniformly offset quantizers.

### B. Description of the MDUOQ Scheme

In TRPCSQ [22], 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 stepsize  $q_1$ . Therefore during the reconstruction, after shifting the shifted quantizers according to the reconstructed prediction, the quantization bin boundaries remain the same. Thus the uniform offsets among different quantizers are always maintained. The problem 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 MDUOQ.

In the  $i$ -th description, 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 subtracted from subset  $j$  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,  $l = \text{mod}(j-i-1, M)$ ,  $\text{mod}(a, b) = a - b\lfloor a/b \rfloor$ ,

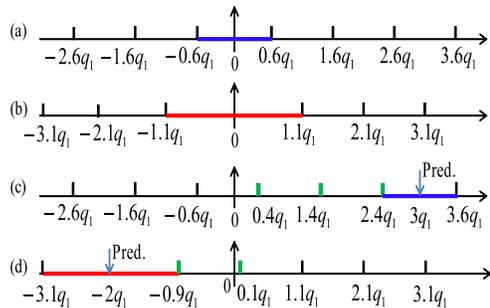


Fig. 3. Examples of uniformly offset quantizers with unequal deadzone and  $\delta = 0.6$  and  $M = 3$ . Quantizers (c) and (d) are shifted from (a) and (b).

and the floor operator rounds towards  $-\infty$ . As a result, across all the  $M$  descriptions, each subset is predictively coded by  $M-1$  low-rate quantizers with deadzones of  $2(\delta + \frac{n}{M-1})q_1$ ,  $n = 0, \dots, M-2$ , respectively; hence there is a uniform offset of  $q_1/(M-1)$  among these quantizers.

An important difference between MDUOQ and TRPCSQ is that the uniform offsets among low-rate quantizers are not always preserved in the MDUOQ, due to predictive coding and deadzone quantizers, *i.e.*, some quantization bin boundaries will be changed after adding the reconstructed prediction.

An example is given in Fig. 3 for  $M = 3$  and  $\delta = 0.6$ , where quantizers (a) and (b) are the original quantizers with unequal-deadzone-induced uniform offsets. The partitions in (c) and (d) are obtained from (a) and (b) by shifting to  $3q_1$  and  $-2q_1$  respectively, due to the quantized prediction. As shown in the figure, the partitions in (c) and (d) have three and two different bin boundaries from (a) and (b) respectively, leading to some nonuniformly offset bins.

In general, if the partition of a quantizer with deadzone  $2\delta q_1$  ( $\delta < 1$ ) is shifted to the right by the prediction  $nq_1$  ( $n > 0$ ), it can be seen from Fig. 3 that only the  $n$  quantization bin boundaries within  $[(1-\delta)q_1, (n-\delta)q_1]$  can have different values from those of the original quantizer, but all other bin boundaries remain the same. Similarly, if the prediction is  $-nq_1$  ( $n > 0$ ), only the  $n$  boundaries in  $[-(n-\delta)q_1, -(1-\delta)q_1]$  can be affected.

A special case can be noticed from Fig. 3. When  $M = 3$ , if we choose  $\delta = i/2$  for any nonzero integer  $i$ , the bin boundaries of both quantizers will not change after the shift. Unfortunately for  $M > 3$ , no  $\delta$  can maintain the boundaries of all quantizers after the shift. The reason is that in order for all shifted quantization partitions to yield the same bin boundaries as the original quantizers, the deadzone size  $2(\delta + \frac{l}{M-1})q_1$  should be multiple of  $q_1$  for all possible  $l$ . This is true only when  $M = 3$  and  $\delta = i/2$ .

The problem of using unequal deadzones to achieve uniform offsets is somewhat related to the requantization problem studied in [29] for application in H.264 SP-frames, where the data are first quantized by a finer quantizer. The reconstruction is then quantized by a coarser quantizer. It is shown in [29] that the requantization error can be minimized if we choose the quantization stepsizes and deadzones of the two quantizers such that the boundaries of the second quantizer perfectly align with that of the first one.

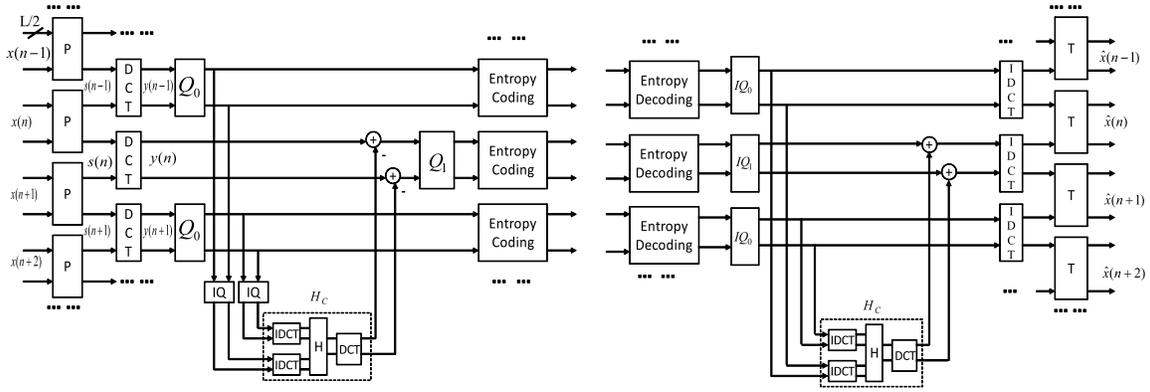


Fig. 4. Encoder and decoder of one description of block-based MDROQ for a 1-D signal with  $M = 2$ .

Our problem is different from [29] since we need uniform offsets. It is more challenging as we have  $M - 1$  quantizers, and our partitions are not static, *i.e.*, each quantizer's partition can be dynamically shifted by prediction.

Finding the exact closed-form expression of  $D_{1,k}$  in MDUOQ is more difficult than in MDROQ, because different quantizers have different deadzones, and each can be shifted by its own prediction. Therefore the joint reconstruction behaviors in different quantization bins could be different, and the simplification in the proof of the MDROQ cannot be applied here. The problem is further complicated by the lack of closed-form R-D formula for deadzone quantizers.

Nevertheless, since many bin boundaries still have uniform offsets, the offsets among the low-rate quantizers in MDUOQ still share some similarity with the uniform offset in TRPCSQ. Therefore, the low-rate joint reconstruction MSE of MDUOQ can be lower-bounded by that of TRPCSQ, *i.e.*, the  $D_{1,k}$  of MDUOQ can be lower-bounded by (4), with  $S_{1,k}$  in (5) replaced by that of TRPCSQ in (20). After this, all the derivations in Sec. II-B can be reused to derive a lower bound for the expected distortion  $D_k$  of MDUOQ.

Note that for  $M = 2$ , there is only one low-rate quantizer, thus the quantization of the prediction is not necessary, and MDUOQ would reduce to MDROQ.

#### IV. OPTIMIZATION FOR BLOCK-BASED MD CODING

The block transform we use is the time-domain lapped transform (TDLT) [30], [31], which employs a prefilter at block boundaries before the DCT and a postfilter after the inverse DCT, thereby providing improved coding efficiency and reduced blocking artifact. The TDLT has been adopted by the JPEG XR standard [32], which is a low-cost alternative to JPEG 2000 with competitive performance.

An attractive feature of TDLT is that its pre/postfilters can be optimized for different applications. In MDC, they can be designed to control the redundancy. In this section, we formulate the optimization of the pre/postfilters and the corresponding prediction filters for MDROQ/MDUOQ by first deriving the expressions of  $\sigma_0^2$  and  $\bar{\sigma}_1^2$  in (19), and then developing an iterative algorithm to find the optimized filters.

To apply the block-based MDROQ/MDUOQ to MD image coding, the TDLT and the corresponding prediction filters are

applied to each two-dimensional (2-D) block of an image in a separable way, *i.e.*, row by row and column by column, as will be discussed in Sec. IV-D. The separable TDLT-based framework has achieved the state-of-the-art MD image coding performance [21]–[23].

#### A. Block Diagram and DCT-Domain Wiener Filter

Fig. 4 illustrates the block diagrams of the encoder and decoder of one description in the TDLT-based MDROQ for an 1-D signal. The block size is  $L$ , where each line represents half block, *i.e.*,  $L/2$  samples. For simplicity, only the codec for  $M = 2$  is shown, and the joint-dequantization with more than one description is not shown. The framework of MDUOQ is similar to Fig. 4, except for the quantization of the prediction and the use of low-rate quantizers with different deadzones.

We first introduce some notations of the TDLT. An  $L \times L$  prefilter  $\mathbf{P}$  is applied at the boundary of two blocks. The  $L$ -point DCT  $\mathbf{C}$  is then applied to each block. As a result, the basis functions of the TDLT cover two blocks, and the bases of neighboring blocks are overlapped by one block. At the decoder, the inverse DCT and the postfilter  $\mathbf{T}$  at block boundaries are applied.  $\mathbf{P}$  and  $\mathbf{T}$  have the following structures to yield linear-phase filters [30]:  $\mathbf{P} = \mathbf{W} \text{diag}\{\mathbf{I}, \mathbf{V}\} \mathbf{W}$ ,  $\mathbf{T} = \mathbf{P}^{-1} = \mathbf{W} \text{diag}\{\mathbf{I}, \mathbf{V}^{-1}\} \mathbf{W}$ , where  $\text{diag}\{\mathbf{A}, \mathbf{B}\}$  denotes a block diagonal matrix with matrices  $\mathbf{A}$  and  $\mathbf{B}$  on the diagonal, and zeros elsewhere,  $\mathbf{I}$  is the  $\frac{L}{2} \times \frac{L}{2}$  identity matrix,  $\mathbf{V}$  is an  $\frac{L}{2} \times \frac{L}{2}$  invertible matrix that can be optimized for different purposes, and  $\mathbf{W}$  is the butterfly matrix defined by

$$\mathbf{W} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I} & \mathbf{J} \\ \mathbf{J} & -\mathbf{I} \end{bmatrix}, \quad (22)$$

where  $\mathbf{J}$  is the  $\frac{L}{2} \times \frac{L}{2}$  counter-identity matrix.

Let  $\mathbf{P} = [\mathbf{P}_0^T \ \mathbf{P}_1^T]^T$  where  $\mathbf{P}_0$  and  $\mathbf{P}_1$  contain the first and the last  $L/2$  rows of  $\mathbf{P}$ , respectively, and the superscript  $T$  denotes matrix transpose. Define  $\mathbf{P}_{12} = \text{diag}\{\mathbf{P}_1, \mathbf{P}_0\}$ . The  $L \times 2L$  forward transform can be written as

$$\mathbf{F} = \mathbf{C}\mathbf{P}_{12}. \quad (23)$$

Similarly, to obtain the inverse transform, let  $\mathbf{T} = [\mathbf{T}_0 \ \mathbf{T}_1]$ , where  $\mathbf{T}_0$  and  $\mathbf{T}_1$  are the first and the last  $L/2$  columns of  $\mathbf{T}$ ,

respectively. Define  $\mathbf{T}_{21} = \text{diag}\{\mathbf{T}_1, \mathbf{T}_0\}$ . The  $2L \times L$  inverse transform is thus

$$\mathbf{G} = \mathbf{T}_{21}\mathbf{C}^T. \quad (24)$$

As in [21], [23], since a block transform is used, to get the subsets required by the proposed MDC schemes, we partition the input at block level, after applying the TDLT. In each description, one subset of the output blocks is encoded at a high rate, and all other subsets are sequentially predicted from the nearest neighbors from both sides using Wiener filter. For example, in Fig. 4,  $\mathbf{y}(n)$  is a low-rated coded block, which will be predicted by the reconstructed  $\mathbf{y}(n-1)$  and  $\mathbf{y}(n+1)$ . When there are more than two descriptions, multiple Wiener filters could be required. For example, for  $M = 3$ , two filters will be needed. One of them uses the reconstructed  $\mathbf{y}(n-1)$  and  $\mathbf{y}(n+2)$  to predict  $\mathbf{y}(n)$ , and another uses the reconstructed  $\mathbf{y}(n)$  and  $\mathbf{y}(n+2)$  to predict  $\mathbf{y}(n+1)$ .

In [21], [23], [33], the prediction and the calculation of the prediction residual were obtained in the spatial domain before the DCT, where it was shown that it is necessary to normalize the Wiener filter matrix such that each row has a unit sum. This will produce a constant output when the input is constant.

In this paper, to facilitate the joint dequantization of random quantizers, the Wiener filtering and the calculation of the residual should be performed after the DCT and before the quantization, as shown in Fig. 4. However, we still need to impose the normalization constraint discussed above. This can be done by transforming the DCT-domain prediction back to the spatial domain, applying the normalization, and transforming back to the DCT domain.

The relationship between the spatial-domain and DCT-domain Wiener filters can also be understood from Fig. 4. If we reuse the spatial-domain Wiener filter  $\mathbf{H}$ , we can apply inverse DCT to the reconstructed  $\hat{\mathbf{y}}(n-1)$  and  $\hat{\mathbf{y}}(n+1)$ , then apply the normalized spatial-domain Wiener filter  $\mathbf{H}$  to them, and finally apply the DCT transform to the prediction to convert it into the DCT domain. The concatenation of these steps is the DCT-domain Wiener filter  $\mathbf{H}_c$ , which is related to the spatial-domain Wiener filter by

$$\mathbf{H}_c = \mathbf{C}\mathbf{H}\text{diag}\{\mathbf{C}^T, \mathbf{C}^T\}. \quad (25)$$

Once the optimal filter  $\mathbf{H}_c$  is found, it can be implemented directly in the DCT domain, which is faster than going through the inverse DCT,  $\mathbf{H}$ , and DCT as in (25).

The DCT-domain Wiener filter described here is more general than the transform-domain prediction in [22], where only the DC coefficient is predicted. The prediction of other coefficients does not have good performance in [22], because the prediction is also quantized to ensure uniformly staggered quantizers. In the current method, the prediction not only improves the coding efficiency, but also enables the random quantizer-based joint reconstruction.

It should be noted that when more than one description is received, the simple inverse quantizers shown in Fig. 4 should be replaced by the joint reconstruction methods in MDROQ and MDUOQ to get refined results.

## B. Expressions Required by the Expected Distortion

To find the expression of  $\sigma_0^2$  in (19) in the TDLT framework, let  $\mathbf{y}(k)$  be a high-rate coded DCT block with quantization stepsize  $q_0$ . Let  $\mathbf{q}_y(k)$  be the quantization error for  $\mathbf{y}(k)$ . After the inverse TDLT, the reconstruction error becomes  $\mathbf{G}\mathbf{q}_y(k)$ , where  $\mathbf{G}$  is the inverse TDLT in (24). As usual, we assume the quantization noises of different subbands are uncorrelated, so the average reconstruction error per sample is

$$\frac{1}{L} \sum_{j=0}^{L-1} \|\mathbf{g}_j\|^2 \sigma_{q_y(j)}^2, \quad (26)$$

where  $\sigma_{q_y(j)}^2$  is the variance of the quantization noise of the  $j$ -th entry of  $\mathbf{y}(k)$ , and  $\mathbf{g}_j$  is the  $j$ -th column of  $\mathbf{G}$ . At high rates,  $\sigma_{q_y(j)}^2$  can be written as

$$\sigma_{q_y(j)}^2 = \frac{2\pi e}{12} \sigma_{y(j)}^2 2^{-2R_{0,j}}, \quad (27)$$

where  $R_{0,j}$  is the bits allocated to the  $j$ -th entry of the block, and  $\sigma_{y(j)}^2$  is the variance of the  $j$ -th entry of  $\mathbf{y}(k)$ , which is the  $j$ -th diagonal element of the autocorrelation matrix  $\mathbf{R}_{yy}$ . The latter can be obtained from the TDLT-transformed input,

$$\mathbf{R}_{yy} = \mathbf{F}\mathbf{R}_{x_2x_2}\mathbf{F}^T, \quad (28)$$

where  $\mathbf{F}$  is the forward lapped transform in (23), and  $\mathbf{R}_{x_2x_2}$  is the autocorrelation function of two input blocks. In this paper, we assume the input is a first order Gaussian-Markov source with unit variance and correlation coefficient  $\rho = 0.95$ , so the  $(i, j)$ -th entry of  $\mathbf{R}_{x_2x_2}$  is  $\rho^{|i-j|}$ .

Minimizing the distortion in (26) subject to the bit rate constraint of  $\frac{1}{L} \sum_{j=0}^{L-1} R_{0,j} = R_0$ , the minimal value for (26) can be found to be

$$\frac{2\pi e}{12} \left( \prod_{j=0}^{L-1} \|\mathbf{g}_j\|^2 \sigma_{y(j)}^2 \right)^{\frac{1}{L}} 2^{-2R_0} \triangleq \frac{2\pi e}{12} \sigma_0^2 2^{-2R_0}. \quad (29)$$

The  $\sigma_0^2$  obtained above can be used in (19) to find the expected distortion.

Next, we derive  $\sigma_{1,i}^2$  in (8) for  $i = 1, \dots, M-1$ . For a low-rate coded DCT blocks with average rate of  $R_{1,i}$ , its average reconstruction error after the inverse transform is

$$\frac{1}{L} \sum_{j=0}^{L-1} \|\mathbf{g}_j\|^2 \sigma_{q_e(i,j)}^2, \quad (30)$$

where  $\sigma_{q_e(i,j)}^2$  is the quantization noise variance of the  $j$ -th entry of the block with average rate of  $R_{1,i}$ :

$$\sigma_{q_e(i,j)}^2 = \frac{2\pi e}{12} \sigma_{e_i(j)}^2 2^{-2R_{1,i,j}}, \quad (31)$$

where  $\sigma_{e_i(j)}^2$  is the variance of the  $j$ -th residual in the block with average rate of  $R_{1,i}$ .  $R_{1,i,j}$  is the bits allocated to the  $j$ -th coefficient of the block.

Minimizing (30) with the constraint of  $\frac{1}{L} \sum_{j=0}^{L-1} R_{1,i,j} = R_{1,i}$ , the minimal distortion is found to be

$$\frac{2\pi e}{12} \left( \prod_{j=0}^{L-1} \|\mathbf{g}_j\|^2 \sigma_{e_i(j)}^2 \right)^{\frac{1}{L}} 2^{-2R_{1,i}} \triangleq \frac{2\pi e}{12} \sigma_{1,i}^2 2^{-2R_{1,i}}. \quad (32)$$

Plugging this into (8) and (19) yields the expression of the expected distortion in the TDLT framework.

### C. An Iterative Optimization Algorithm

Our objective is to find the optimal TDLT prefilter that minimizes the distortion  $D$  in (19), 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 need to find the optimal bit allocation and the corresponding minimal distortion. The optimal pre/post-filter can then be found by minimizing the distortion.

For simple optimization problems such as those in [21]–[23], a closed-form solution can be found by the Lagrangian multiplier method. However, in this paper it is difficult to apply this method, because  $S_{0,k}$  in (15) also depends on  $R_0$  and  $R_1$ .

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  $\lambda$ .

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 (19) can be easily minimized by the Lagrangian multiplier method, and the optimal bit allocation is given by

$$\begin{aligned} R_0 &= \min \left( MR, R + \frac{M-1}{2M} \log_2 \frac{(M-1)\bar{S}_0\sigma_0^2}{\bar{S}_1\bar{\sigma}_1^2} \right), \\ R_1 &= \max \left( 0, R - \frac{1}{2M} \log_2 \frac{(M-1)\bar{S}_0\sigma_0^2}{\bar{S}_1\bar{\sigma}_1^2} \right). \end{aligned} \quad (33)$$

We then use the  $R_0$  and  $R_1$  above to calculate the  $S_{0,k}$  in (15), which can then be used to update the bit allocation in (33). Each time the bit allocation is updated, the distortion in (19) is re-calculated, and the iteration terminates when the distortion change is less than a threshold.

This iteration method does not need to select  $\lambda$ . The bit rate constraint is strictly met by (33). 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.

Numerical optimization results show that as long as the clipping operators in the bit allocation (33) are not in effect, a stable optimal TDLT prefilter can be obtained independent of  $R$  and  $p$ . Since practical image coding usually operates at low rates, experimental results show that slightly better MD image coding results can be obtained by designing the TDLT using a total rate of  $MR = 1 \sim 1.5$  bpp and  $p = 0.1 \sim 0.16$ . The corresponding coding gain of TDLT is about 9.53 dB. In this case, the clippings in (33) are active, which essentially increase the weighting parameter of the low-rate term in the expected distortion in (19).

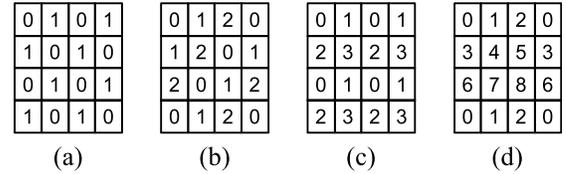


Fig. 5. Block subset definitions in MD image coding. (a)  $M = 2$ ; (b)  $M = 3$ ; (c)  $M = 4$ ; (d)  $M = 9$ .

### D. Application in MD Image Coding

To apply the block-based MDROQ/MDUOQ to MD image coding, we partition the image blocks into  $M$  subsets in a periodic pattern. Some examples are shown in Fig. 5. The patterns are designed to simplify the implementation and increase the distance of blocks in the same subset, so that when one subset is lost, there will be more available neighboring blocks to recover the missing ones. Similar methods have been used in other MDC schemes such as [34].

In the  $i$ -th description, the  $i$ -th subset is coded at high rate, and all other subsets are sequentially predicted and the residuals are coded at low rate. The TDLT and Wiener filter are applied to each block in a separable way, *i.e.*, row by row and column by column, according to Fig. 4.

There are a couple of differences when applying the 1-D optimized TDLT and Wiener filter to 2-D images. First, in sequential prediction, sometimes a block can be predicted from both horizontal and vertical neighboring blocks. For example, in Fig. 5 (b), in the first description, Subset 1 can be predicted from reconstructed Subset 0, in both directions. After that, Subset 2 can be predicted from reconstructed Subsets 0 and 1 in both directions. In this case, the average of the two predictions can be used as the final prediction to improve the performance, as shown in Eq. (21). This method was also used in [21], [23]. Second, due to the 2-D partition of block subsets, in many cases, the horizontal and vertical distances of blocks within the same subset are closer than those in 1-D signals with the same  $M$ , as shown in Fig. 5 (c-d). Therefore filters optimized based on the 1-D signal model for a smaller  $M$  can be reused for some larger  $M$ . For example, the filters obtained from 1-D model with  $M = 2$  and  $M = 3$  can be reused for image coding with  $M = 4$  and  $M = 9$  respectively.

## V. THEORETICAL ANALYSES AND EXPERIMENTAL RESULTS

### A. Theoretical and Simulation Results for 1-D Data

We first use 1-D data to compare the theoretical and simulation performances of MDROQ. We consider a unit-variance first-order Gauss-Markov source with correlation coefficient of  $\rho = 0.9$ . The source is partitioned into subsets at sample level. Each sample  $x(n)$  is predicted by  $\rho \hat{x}(n-1)$ , so  $\sigma_0^2 = \bar{\sigma}_1^2 = 1 - \rho^2$  in the theoretical calculation, as in [22]. The prediction residual is uniformly quantized with step size of either  $q_0$  or  $q_1$  according to the MDROQ setup. The results are reported in Fig. 6 for  $M = 2$ ,  $M = 3$ , and  $M = 4$  respectively, with  $R = 5$  bits/sample/description, where the entropy of the quantized results is used as the rate in the simulation. It can

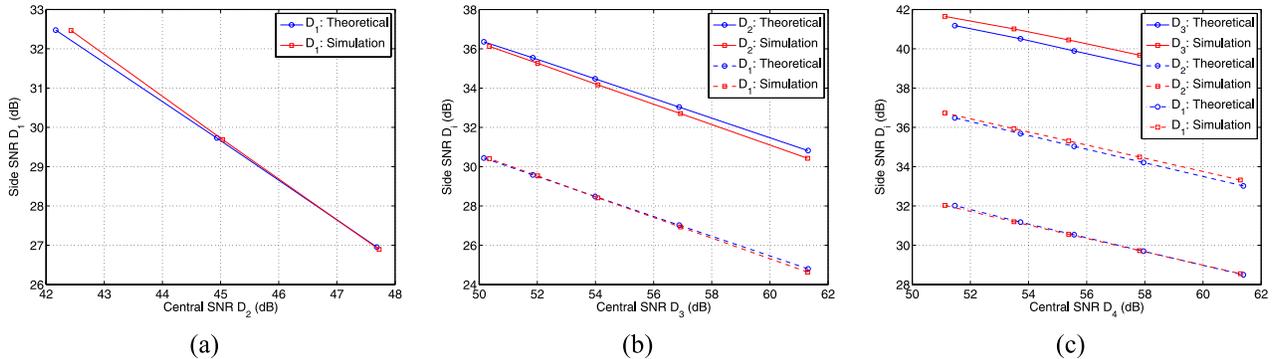


Fig. 6. Comparison between theoretical analysis and simulation result for 1-D data. (a)  $M = 2$ ; (b)  $M = 3$ ; (c)  $M = 4$ .

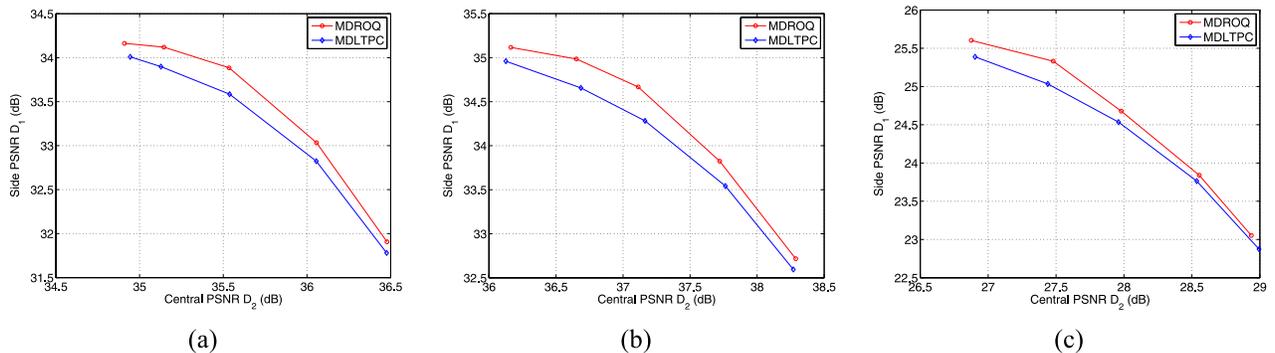


Fig. 7. The side PSNR vs. central PSNR of MDROQ and MDLTPC for  $M = 2$ . (a) Lena. Total rate 0.5 bpp. (b) Boat. total rate 1.0 bpp. (c) Baboon. Total rate 1.0 bpp.

be seen that the theoretical and simulation curves generally agree quite well. The small discrepancy is because the random quantization theory assumes that the offsets are uniformly distributed, which is not exactly the case in MDROQ, as there are still some correlations in different predictions of the same sample in different descriptions.

We also replace  $S_{1,k}$  in MDROQ by that of TRPCSQ in Eq. (20). As discussed before, this can be used to derive a lower bound of the MSE of the MDUOQ scheme. The simulation result of MDUOQ for  $M = 4$  shows that its  $D_i$  is about 2 dB worse than the bound. How to derive a more accurate distortion expression for MDUOQ is a future research topic.

### B. MD Image Coding Results

In this part, we compare the performances of MDROQ and MDUOQ with MDLTPC, TRPCSQ, and TLMDC in MD image coding using testing images of various characteristics. It has been reported in [21]–[23] that these methods have better performances than many other methods such as the PCT/GPCT, MMDSQ, RD-MDC, and MDLVQ.

In MDROQ, deadzone quantizers are used in both high-rate and low-rate coding, with deadzone size of  $1.2q_0$  or  $1.2q_1$ . In MDUOQ, the minimal deadzone factor is also chosen as  $\delta = 0.6$ .

Fig. 7 compares the two-description MDROQ and MDLTPC in [21]. The trade-off between the side PSNR  $D_1$  and the

central PSNR  $D_2$  is reported, by varying the bit allocation or redundancy of the scheme, *i.e.*, adjusting the values of  $q_0$  and  $q_1$  while maintaining the same total rate. 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. Alternatively, for the same central PSNR, the side PSNR of MDROQ achieves up to 0.3 dB gain over MDLTPC.

Fig. 8 to Fig. 10 compare the relationships between the side PSNR  $D_i$  and central PSNR  $D_M$  of MDROQ, MDUOQ and TRPCSQ for different images when  $M = 3, 4$ , and 9, respectively. To avoid too crowded figures, only half of  $D_i$ 's are shown in Fig. 10. It can be seen that the proposed MDROQ and MDUOQ outperform TRPCSQ in almost all cases. Up to 5 dB gain can be obtained when the redundancy is low ( $D_1$  of MDROQ in Fig. 9 (c)), *i.e.*, when  $q_1$  is large, which corresponds to the right corners of the curves. When there is moderate or high redundancy, the gain of  $D_i$  is still up to 0.5 dB. This is due to various limitations in TRPCSQ, including the asymmetric quantizers and the use of uniform quantizer.

Fig. 8 also includes results of the TLMDC in [23], which is better than TRPCSQ. The TLMDC 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 overall performance than TLMDC, due to improved joint dequantization. Note that the third layer can also be added to our new schemes. Therefore it is fairer to compare our schemes with TRPCSQ.

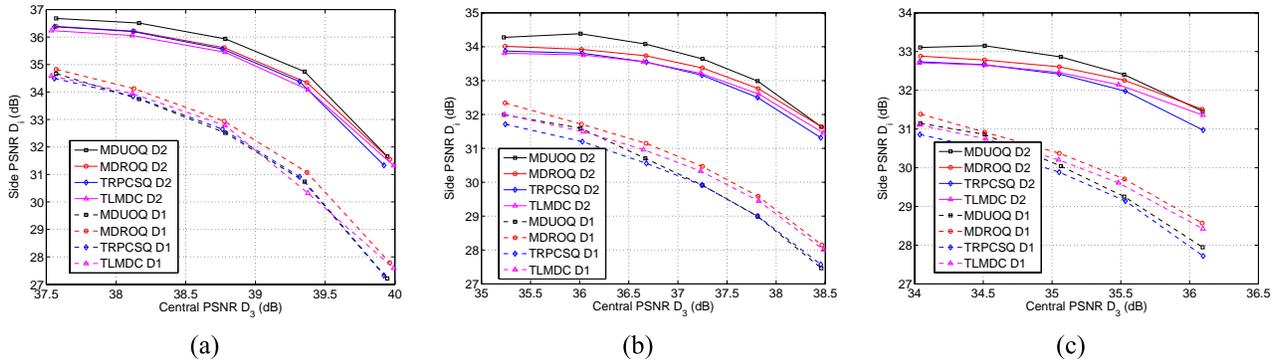


Fig. 8. The side PSNRs vs. central PSNR of various methods for  $M = 3$  and total rate of 1 bpp. (a) Lena. (b) Boat. (c) Goldhill.

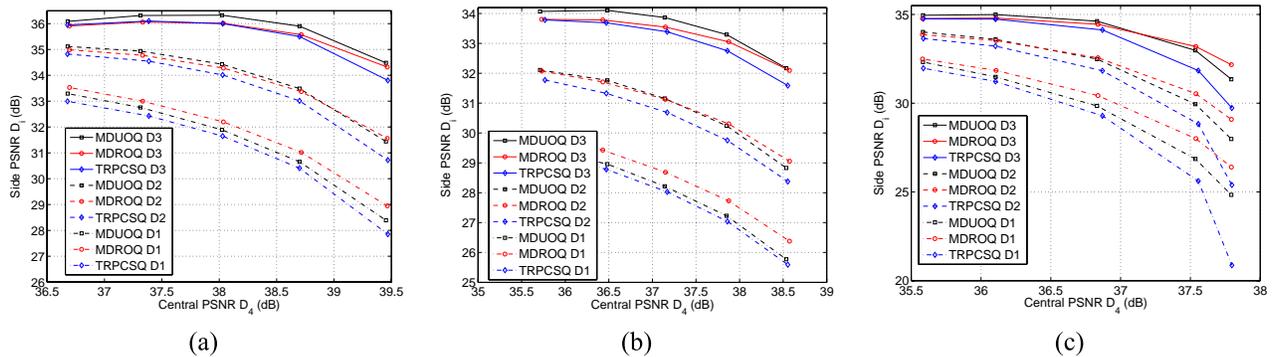


Fig. 9. The side PSNRs vs. central PSNR of various methods for  $M = 4$  and total rate of 1 bpp. (a) Lena. (b) Boat. (c) Pepper.

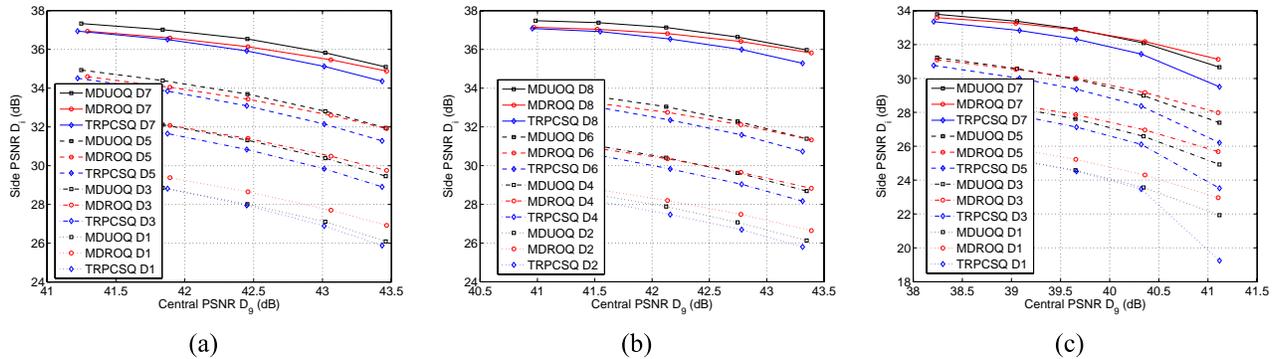


Fig. 10. The side PSNRs vs. central PSNR of various methods for  $M = 9$  and total rate of 2 bpp. (a) Lena. (b) Boat. (c) Couple.

Between MDROQ and MDUOQ, when the number of received descriptions is small, the random quantizer-based MDROQ achieves better performance than MDUOQ. However, when more descriptions are available, the uniform offset-based MDUOQ will generally outperform MDROQ. This indicates that MDUOQ can have better expected distortion  $D$ , defined in Eq. (2), because the description loss probability  $p$  is usually quite small. In this case, the expected distortion is dominated by the cases with more descriptions.

To verify this, Fig. 11 reports the expected PSNRs of the Boat image with various values of bit rates,  $M$  and  $p$ . In each curve, the point with the highest expected PSNR has the optimal bit allocation for the given  $p$ . It is clear from the figure that both MDUOQ and MDROQ have better optimal

expected PSNRs than TRPCSQ and TLMDC, and MDUOQ is better than MDROQ. It also shows that more redundancy (lower central PSNR) should be selected when  $p$  increases.

It can also be observed from these figures that the overall difference between MDUOQ and MDROQ reduces when  $M$  increases, which agrees with Fig. 2 (c). However, as mentioned before, the MDUOQ is not always better than MDROQ in the  $D_i$  vs.  $D_M$  curves, especially when the redundancy is very small ( $q_1$  is large) and when  $M$  is large. One reason for this is that some quantizers in MDUOQ have large deadzones, which are not optimal at low rates. Another reason is that there are more prediction-induced nonuniform offsets in the MDUOQ as  $M$  increases, making the uniformly offset quantizers similar to the randomly offset quantizers. This suggests that the

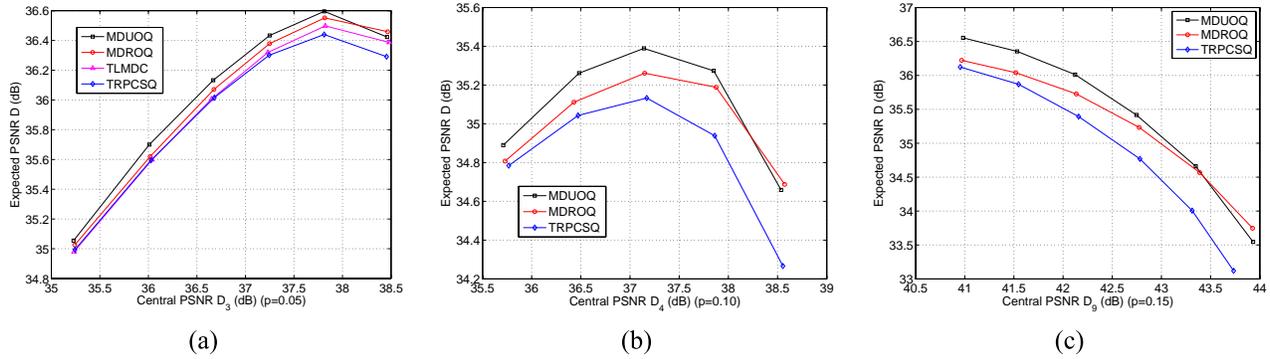


Fig. 11. The expected PSNR vs. central PSNR of the Boat image with various methods. (a) Total rate  $1\text{bpp}$ ,  $M = 3$ ,  $p = 0.05$ . (b)  $1\text{bpp}$ ,  $M = 4$ ,  $p = 0.1$ . (c)  $2\text{bpp}$ ,  $M = 9$ ,  $p = 0.15$ .

predictive coding and unequal deadzone-based MDUOQ still has not fully achieved the full potential of uniformly staggered quantizers as promised by the theoretical curves in Fig. 2 (c); hence it might still be possible to modify the design of MDUOQ to further improve its image coding performance. Another open problem is to derive a more accurate formula than Eq. (20) for the factor  $S_{1,k}$  in the current MDUOQ that takes into account the effect of nonuniform offsets.

## VI. CONCLUSION

Two multiple description coding schemes with randomly and uniformly offset quantizations are proposed, which can be viewed as unified and improved versions of the methods in [21]–[23]. The closed-form expressions of the expected distortions of the proposed schemes are obtained by generalizing the random quantization theory in [27]. The proposed schemes are applied to lapped transform-based MD image coding, and an iterative algorithm is developed to help the optimization. Theoretical analyses and image coding results show that the proposed schemes achieve better performance than other methods in this category.

The proposed schemes can be further improved by adding the third layer in [23], but this is beyond the scope of the paper. As discussed before, how to derive the accurate formula of the distortion of MDUOQ remains as an open problem.

## ACKNOWLEDGEMENT

The authors thank the anonymous reviewers for their suggestions that have significantly enhanced the quality and presentation of this paper.

## REFERENCES

- [1] V. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sep. 2001.
- [2] V. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.
- [3] V. Vaishampayan and J. Batllo, "Asymptotic analysis of multiple description quantizers," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 278–284, Jan. 1998.
- [4] T. Guionnet, C. Guillemot, and S. Pateux, "Embedded multiple description coding for progressive image transmission over unreliable channels," in *Proc. IEEE Conf. Image Process.*, Oct. 2001, pp. 94–97.
- [5] S. M. Satti, N. Deligiannis, A. Munteanu, P. Schelkens, and J. Cornelis, "Symmetric scalable multiple description scalar quantization," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3628–3643, Jul. 2012.
- [6] 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.
- [7] M. Liu and C. Zhu, "Enhancing two-stage multiple description scalar quantization," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 253–256, Apr. 2009.
- [8] T. Y. Berger-Wolf and E. M. Reingold, "Index assignment for multi-channel communication under failure," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2656–2668, Oct. 2002.
- [9] C. Tian and S. S. Hemami, "Sequential design of multiple description scalar quantizers," in *Proc. IEEE Data Compress. Conf.*, Snowbird, UT, USA, Mar. 2004, pp. 32–41.
- [10] J. Ostergaard, J. Jensen, and R. Heusdens, "n-channel entropy-constrained multiple-description lattice vector quantization," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1956–1973, May 2006.
- [11] M. Liu and C. Zhu, "M-description lattice vector quantization: Index assignment and analysis," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2258–2274, Jun. 2009.
- [12] N. Jayant, "Subsampling of a DPCM speech channel to provide two self-contained half-rate channels," *Bell Syst. Tech. J.*, vol. 60, pp. 501–509, Apr. 1981.
- [13] W. Jiang and A. Ortega, "Multiple description coding via polyphase transform and selective quantization," *Proc. SPIE*, vol. 3653, pp. 998–1008, Feb. 1999.
- [14] T. Tillo, M. Grangetto, and G. Olmo, "Multiple description image coding based on Lagrangian rate allocation," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 673–683, Mar. 2007.
- [15] E. Baccaglioni, T. Tillo, and G. Olmo, "A flexible RD-based multiple description scheme for JPEG 2000," *IEEE Signal Process. Lett.*, vol. 14, no. 3, pp. 197–200, Mar. 2007.
- [16] T. Tillo, E. Baccaglioni, and G. Olmo, "A flexible multi-rate allocation scheme for balanced multiple description coding applications," in *Proc. IEEE Workshop Multimedia Signal Process.*, Nov. 2005, pp. 1–4.
- [17] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Process.*, vol. 10, no. 3, pp. 351–366, Mar. 2001.
- [18] Y. Wang, A. Reibman, M. Orchard, and H. Jafarkhani, "An improvement to multiple description transform coding," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2843–2854, Nov. 2002.
- [19] V. Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2199–2224, Sep. 2001.
- [20] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Unequal loss protection: Graceful degradation over packet erasure channels through forward error correction," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 819–828, Jun. 2000.
- [21] G. Sun, U. Samarawickrama, J. Liang, C. Tian, C. Tu, and T. D. Tran, "Multiple description coding with prediction compensation," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1037–1047, May 2009.

- [22] U. Samarawickrama, J. Liang, and C. Tian, "M-channel multiple description coding with two-rate coding and staggered quantization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 7, pp. 933–944, Jul. 2010.
- [23] U. Samarawickrama, J. Liang, and C. Tian, "A three-layer scheme for M-channel multiple description image coding," *Signal Process.*, vol. 91, no. 10, pp. 2277–2289, Oct. 2011.
- [24] 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, Mar. 2008.
- [25] S. Zhu, S. A. Yeung, B. Zeng, and J. Wu, "Total-variation based picture reconstruction in multiple description image and video coding," *Signal Process., Image Commun.*, vol. 27, no. 2, pp. 126–139, Feb. 2012.
- [26] S. Zhu, S. A. Yeung, and B. Zeng, "A comparative study of multiple description video coding in P2P: Normal MDSQ versus flexible dead-zones," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jun./Jul. 2009, pp. 1118–1121.
- [27] V. K. Goyal, "Scalar quantization with random thresholds," *IEEE Signal Process. Lett.*, vol. 18, no. 9, pp. 525–528, Sep. 2011.
- [28] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Boston, MA, USA: Kluwer, 2002.
- [29] W.-T. Tan and B. Shen, "Methods to improve coding efficiency of SP frames," in *Proc. IEEE Conf. Image Process.*, Oct. 2006, pp. 1361–1364.
- [30] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-filtering," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1557–1571, Jun. 2003.
- [31] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA, USA: Artech House, 1992.
- [32] S. Srinivasan, C. Tu, S. L. Regunathan, and G. J. Sullivan, "HD photo: A new image coding technology for digital photography," *Proc. SPIE*, vol. 6696, pp. 6696A-1–6696A-9, Aug. 2007.
- [33] J. Liang, C. Tu, L. Gan, T. D. Tran, and K.-K. Ma, "Wiener filter-based error resilient time domain lapped transform," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 428–441, Feb. 2007.
- [34] I. V. Bajic and J. W. Woods, "Domain-based multiple description coding of images and video," *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1211–1225, Oct. 2003.



**Lili Meng** received the B.E. degree from Shandong University, Jinan, China, and the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2005 and 2013, respectively. From 2010 to 2011, she was a Visiting Student at Simon Fraser University, Canada. She is currently with the School of Information Science and Engineering, Shandong Normal University, Jinan. Her research interests include image/video coding and 3-D videos.



**Jie Liang** (S'99–M'04–SM'11) received the B.E. and M.E. degrees from Xi'an Jiaotong University, China, the M.E. degree from the National University of Singapore (NUS), and the Ph.D. degree from Johns Hopkins University, Baltimore, MD, USA, in 1992, 1995, 1998, and 2003, respectively. Since May 2004, he has been with the School of Engineering Science, Simon Fraser University, Canada, where he is currently an Associate Professor. In 2012, he visited the University of Erlangen-Nuremberg, Germany, as an Alexander von Humboldt Research

Fellow. From 2003 to 2004, he was with the Video Codec Group, Microsoft Digital Media Division. From 1997 to 1999, he was with Hewlett-Packard Singapore and the Center for Wireless Communications (now part of the Institute for Infocomm Research), NUS.

His research interests include multimedia communications, multirate signal processing, information theory, and wireless communications. He is currently an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, *Signal Processing: Image Communication*, the *EURASIP Journal on Image and Video Processing*, and *Circuits, Systems, and Signal Processing*. He is a member of the IEEE Multimedia Systems and Applications Technical Committee, and is a Professional Engineer in British Columbia.



**Upul Samarawickrama** received the B.Sc. degree from the University of Moratuwa, Sri Lanka, the M.Sc. degree from the University of Manitoba, Canada, and the Ph.D. degree from Simon Fraser University, Canada, in 2002, 2005, and 2011, respectively. He is currently with the Skype Video DSP Team of Microsoft Corporation. His research interests include multiple description coding, joint source-channel coding, GPU computing, and video coding.



**Yao Zhao** (M'06–SM'12) received the B.S. degree from Fuzhou University, Fuzhou, China, in 1989, and the M.E. degree from Southeast University, Nanjing, China, in 1992, both in radio engineering, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University (BJTU), Beijing, China, in 1996.

In 1998, he was an Associate Professor at BJTU, where he became a Professor in 2001. From 2001 to 2002, he was a Senior Research Fellow with the Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, Delft, The Netherlands. He is currently the Director of the Institute of Information Science, BJTU. He is currently leading several national research projects such as 973 Program, 863 Program, and the National Science Foundation of China. His current research interests include image/video coding, digital watermarking and forensics, and video analysis and understanding.

Dr. Zhao serves on the editorial boards of several international journals, including as an Area Editor of *Signal Processing: Image Communication* (Elsevier), and as an Associate Editor of *Circuits, System and Signal Processing* (Springer). He was a recipient of the National Science Foundation of China for Distinguished Young Scholars in 2010.



**Huihui Bai** received the Ph.D. degree in signal and information processing from Beijing Jiaotong University (BJTU), Beijing, China, in 2008. She is currently an Associate Professor with the Institute of Information Science, BJTU. Her research interests include video coding technologies and standards such as HEVC, 3-D video compression, multiple description video coding and distributed video coding. She is leading or participating in several research projects such as 973 Program, 863 Program, National Natural Science Foundation of China, Beijing Natural Science Foundation, and Jiangsu Provincial Natural Science Foundation.



**André Kaup** (M'96–SM'99–F'13) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from RWTH Aachen University, Aachen, Germany, in 1989 and 1995, respectively.

From 1989 to 1995, he was with the Institute for Communication Engineering, RWTH Aachen University, where he was responsible for industrial as well as academic research projects in the area of high resolution printed image compression, object based image analysis and coding, and models for human perception. In 1995, he joined the Networks and

Multimedia Communications Department, Siemens Corporate Technology, Munich, Germany, where he chaired work packages in several European research projects in the areas of very low bit rate video coding, image quality enhancement, and mobile multimedia communications. In 1999, he was appointed the Head of the Mobile Applications and Services Group in the same department with research focusing on multimedia adaptation for heterogeneous communication networks. Since 2001, he has been a Full Professor and the Head of the Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Germany. From 1997 to 2001, he was the Head of the German MPEG delegation, and from 1998

to 2001, he served as an Adjunct Professor with the Technical University of Munich, Germany, and the University of Erlangen-Nuremberg. From 2005 to 2007, he was a Vice Speaker of the DFG Collaborative Research Center 603 Modeling and Analysis of Complex Scenes and Sensor Data. He has authored more than 250 journal and conference papers and has published over 50 patents granted and patent applications. His current research interests include image and video signal processing, coding of images, video, and multiview data, as well as multimedia communication.

Dr. Kaup is a Voting Member of the IEEE Communications Society Multimedia Communications Technical Committee, and a member of the German ITG. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and was a Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. He was elected as a Siemens Inventor of the year 1998 and was a recipient of the 1999 ITG Award. He received the Best Paper Awards at the IEEE International Workshops on Multimedia Signal Processing in 2009, 2011, and 2012, and the International Workshop on Image Analysis for Multimedia Interactive Services in 2010. In 2013, he was a recipient of the Paul Dan Cristea Special Award at the International Conference on Systems, Signals and Image Processing.