Multimed Tools Appl
DOI 10.1007/s11042-012-1093-1

Optimizing the deadzone width to improve the polyphase-based multiple description coding

Chunyu Lin · Tammam Tillo · Jimin Xiao · Yao Zhao

Abstract The polyphase-based mechanism is the basis of many multiple descriptions coding schemes. Its main drawback is the inefficient exploitation of the inserted redundancy, especially when the redundancy is large. In this paper we propose a novel approach that uses mid-tread quantizers with tunable deadzone, in order to efficiently exploit the inserted redundancy. In particular, the deadzone width is selected based on the statistical distribution of the data, and the approximated level of redundancy to be inserted. The proposed approach is tailored for those codecs that use mid-tread quantizers with tunable step-size and deadzone width. This is particularly interesting given that the majority of codecs use this topology of quantization, and they rarely allow changing it. Moreover, the proposed scheme

C. Lin (✉) · Y. Zhao
Institute of Information Science, Beijing Jiaotong University, Beijing, China, 100044
e-mail: yuailian@gmail.com

Y. Zhao
e-mail: yzhao@bjtu.edu.cn

C. Lin · Y. Zhao
Beijing Key Laboratory of Advanced Information Science and Network, Beijing, China, 100044

T. Tillo · J. Xiao
Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China, 215123

T. Tillo
e-mail: tammam.tillo@xjtlu.edu.cn

J. Xiao
Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, UK
e-mail: jimin.xiao@liverpool.ac.uk

Y. Zhao
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China, 100044

Published online: 03 May 2012

can be extended to the case of more than two descriptions. Finally, it is worth reporting that the results of the proposed approach outperform that of state-of-the-art schemes. In fact with the same side performance, the central quality can achieve up to 1 dB gain.

Keywords Image coding · Multiple description coding · Scalar quantization

1 Introduction

Multiple description coding (MDC) is an effective scheme for error resilient transmission of image and video [3]. In MDC, one source is encoded into two or more bitstreams (descriptions), which are mutually refinable and can be decoded independently. In general, the reconstructed quality is proportional to the number of received descriptions. When the network is reliable, and all the descriptions are received, the best quality is obtained, which is usually referred to as the central performance. On the other hand, the delivered quality when only one description is received is lower than the central one. In general, it is still acceptable and referred to as side performance. To achieve this resiliency, some redundancy should be inserted in the descriptions, which is useful for mitigating packet losses but is unfavorable to the central performance when no packets are lost. Hence, the main task of MDC is to design an effective scheme to tune the tradeoff between central and side performance according to the network status.

Jiang et al. propose a polyphase transform and selective quantization (PTSQ) scheme in which different polyphase samples are coded with different quantization steps [4]. With high flexibility and simplicity, the polyphase-based MDC scheme is currently receiving more and more attention for practical applications. For the two-description case, this method divides the source into two sub-streams. One description is composed of a coarse version of the first sub-stream and a fine version of the second sub-stream; whereas, the other description is formed in a complementary manner. Therefore, the redundancy is explicitly controlled by the bit rate of the coarse sub-stream, and the central performance is determined by the bit rate of the fine sub-stream. Such schemes are flexible and adaptable to other conventional single description coding (SDC) schemes. PTSQ [4] and the Lagrangian rate-allocation-based method [10] are two typical schemes of this kind. The main disadvantage of these schemes is that, when both descriptions are received, the coarse sub-stream version is disregarded, consequently it will not provide any contributions to the central performance. Therefore, Tillo et al. propose a modified scheme to exploit the coarse information, so as to improve the fine information [9]. This improvement is achieved by using different combinations of mid-tread and mid-rise quantizers. However, the usage of quantizers with different structure represents the main limitation of Tillo's method [9]. That means, when only mid-tread quantizers are used, which is the most frequently used quantizer structure, the improvement can only be obtained for even ratios between the coarse and fine quantization step. In addition, the effect of the coefficients' statistical distribution (i.e. the probability density function (*pdf*) of the source) has not been considered in Tillo's method, as we will see later, this has a noticeable impact on the performance.

The main contribution of this paper is the proposal of a new quantization-based MD scheme that could be easily applied for various images codecs. The main features of this approach are: the deadzone of the quantizer is optimized based on the statistics of the data and the approximated level of redundancy to be inserted. Secondly, the information carried by the coarse quantized values will be employed in any case to improve the performance when more than one description is received. Finally, the framework for more than two descriptions is proposed.

2 The proposed scheme

Firstly, we will define the parameters that will be used to describe the proposed algorithm. For a two-description scheme, denote the coarse and fine quantization steps as q_c and q_f respectively. With this notation, the central performance of the polyphase-based MDC is determined by q_f, whereas, the redundancy is controlled by q_c. When more than one description is received, the coarse quantized information will be discarded in most of polyphase-based schemes [4, 10]. Hence, Tillo et al. proposes to exploit the coarse quantizer to refine the fine quantizer for the central decoder [9]. However, it can only be applied to some certain case, such as the ratio between the two quantization step is even for the mid-tread quantizers.

Figure 1 shows an illustrative example of the coarse and fine quantization, where the top line and the middle line correspond to the fine and coarse quantizers, respectively. Let us refer to the i-th quantizer by $\mathbf{Q}_i$, with $q_i \leq q_j$ if $i < j$. On the top line, $(a_{i-1}, a_i]$ is the interval in which the values are reconstructed with x_i. In a uniform scalar deadzone quantizer, every interval has the same length (called quantization step), except for the central interval which is denoted as the deadzone of the quantizer. When both the coarse and fine information are received, the signal can be reconstructed with a fused quantizer, as shown in the middle line of Fig. 1.

There are two main practical problems with the approach proposed by Tillo et al. [9]. The first, is that achieving optimal performance requires to choose a set of two quantizers from all the four combinations of mid-tread and mid-rise quantizers. This choice has been shown to be dependent on the ratio between the coarse and fine quantization steps. The second, and more important issue, is that the effect of the

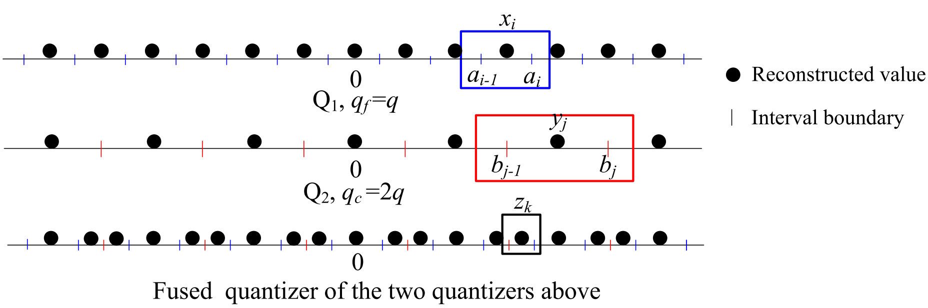

Fig. 1 The refined quantizer for two mid-tread quantizers. The *vertical lines* and *round dots* denote the interval boundary and reconstructed level for the quantizers

data statistics has not been taking into account to select the quantizer type. Although this simplifies the selection process, it degrades the performance.

Hence, we propose to use two mid-tread quantizers with optimized deadzones to mitigate these drawbacks. The width of the deadzone will be tuned so as to control the structure of the fused, i.e., merged quantizer. This consequently will improve the performance of the central decoder. In order to select the best fused quantizer structure, which minimizes the reconstructed distortion, the statistical distribution (or equivalently the *pdf* of the data) of the coefficients to be quantized will be taken into account. This is justified by the fact that it is better to have smaller steps of the fused quantizer, where values of data are more likely to occur, rather than having them for the range of values which are less probable. For this reason the central distortion will be expressed as following

$$D_0 = f(q_f, d_f, q_c, d_c, pdf) \tag{1}$$

where d_f and d_c represent the deadzone width of the fine and coarse quantizer, respectively. These two parameters will affect the rate and distortion of the two side descriptions, however, they have major impact on the central decoder performance. Because any changes to the deadzones' width would affect many intervals (i.e., the refined intervals) of the fused quantizer, whereas, for the side quantizers, it will mainly affect the deadzone interval alone. The above analysis becomes more accurate at higher coding rate, and smoother *pdf* function of the quantized data. For this reason, to simplify the problem at hand, in the following, we will consider the high rate quantization scenario. Furthermore, we will address the scenario where the coarse deadzone is a parameter to be selected by the user. Then the fine quantizer's deadzone becomes the parameter that will be tuned and automatically optimized by the following proposed approach. In this scenario, we can assume that the effect of the fine deadzone on the rate and distortion of side description is negligible in comparison to that caused to the central decoder, this means that the side distortion can be expressed in mathematical form as

$$D_1 \approx f(q_f, q_c, d_c, pdf) \tag{2}$$

It is worth pointing out that the high rate quantization theory is widely used in source coding to derive closed-form solutions. In general, the optimal system designed according to this theory still has good performance in low-rate scenarios, due to the fact that the high-rate quantization theory is generally accurate enough at bit rate down to 2 bits per sample [6].

3 Analysis of the proposed scheme

For the non-improved (traditional) polyphase-based MDC, the introduced redundancy is determined by q_c. In fact, in this traditional approach, the sub-stream quantized with q_c is completely dropped when both descriptions are received. Whereas, for the proposed approach q_c indicatively determines the amount of introduced redundancy, since the exact amount depends also on the width of the deadzone.

Here, we are considering a simplified case where the only parameter to be determined is the fine deadzone. However, the nonlinearity of the expected distortion of the fused quantizer makes it quite difficult, if not impossible, to find a closed form solution to this problem. For this reason, in the following, we devise a practical procedure to address this problem. Firstly, a set of candidates values for d_f is examined, the set is $\mathbf{d}_f = [0, 2q_f]$, with $2q_f$ being its maximum value. For each value, the corresponding joint quantizer is determined, and by taking into account the *pdf* of the source, the expected central distortion is evaluated. At the end, the fine deadzone candidate that leads to the minimum reconstruction error will be selected and used with the other parameters to quantize the source so as to generate the descriptions. The process of getting d_f can be described as

$$d_f = \underset{d_f}{\operatorname{argmin}} D_0(q_f, q_c, d_f, d_c, pdf) \tag{3}$$

It is worth noticing, that the proposed approach does not require quantizing the source to determine the optimal fine deadzone (this is shown in the dashed box in Fig. 2), since it is based on a semi-analytical approach, thus the complexity of the above procedure is low.

Figure 2 shows the flowchart of the proposed scheme. Firstly, the quantization parameter of q_f and q_c are initialized according to the redundancy to be inserted. For example, q_c will be close to q_f when more redundancy are required to protect the data. Hence, the redundancy is easy to be controlled by tuning the ratio between q_c and q_f. The d_c is set by the users, and in the simulations, it is made equal to the quantization step, i.e., $d_c = q_c$. After the initialization, the statistical distribution of the data is estimated. Based on the estimated *pdf* and the initialized parameters, the optimized d_f is searched to minimize function (3). It should be noted that d_c could also be changed to get probably different minimized value for (3). However, with two parameters the optimization task would become more complex. In the meantime, the original signal is polyphase subsampled into sub-signals, e.g. temporal subsampling. Finally, the subsampled signals are quantized with the above parameters and formed into descriptions. Take two-description for an example, Fig. 3 gives the encoding process with the obtained quantization parameters. Since the side decoder is based on a simple dequantization, only the central decoder is given in the Fig. 3, which uses refined quantizer to decode the two descriptions.

3.1 The proposed four-description scheme

The proposed scheme can be generalized to more than two-description case. For Tillo's method, it is more difficult to generate N-description with respect to the proposed approach [9]. This difficulty is because it requires a combinations of mid-tread and mid-rise quantizers, whereas, the proposed scheme only needs to have mid-tread quantizers here. As an example, Fig. 4 presents a four-description case. If the four quantizers have been designed to have no overlap, any two quantizers can be fused to form a refined one. This can be seen in the bottom line in Fig. 4, where interval bounds that have different colors are not superposed. Hence, the more the descriptions are received, the finer the fused quantizer can be obtained.

Just as the two-description case, we should consider the *pdf* of the source to increase gains. If the ratios between the steps of the four quantizers are given as well as the deadzone width of the finest quantizer, there will be then three deadzone widths that require to be optimized, which is not a trivial work. To solve this problem, we devise a heuristic approach, which takes advantages of these facts: a) the quantizer $\mathbf{Q}_1$ has the largest contribution to the whole performance; b) the smaller the ratio between the quantization steps of two quantizers is, the larger the gain that

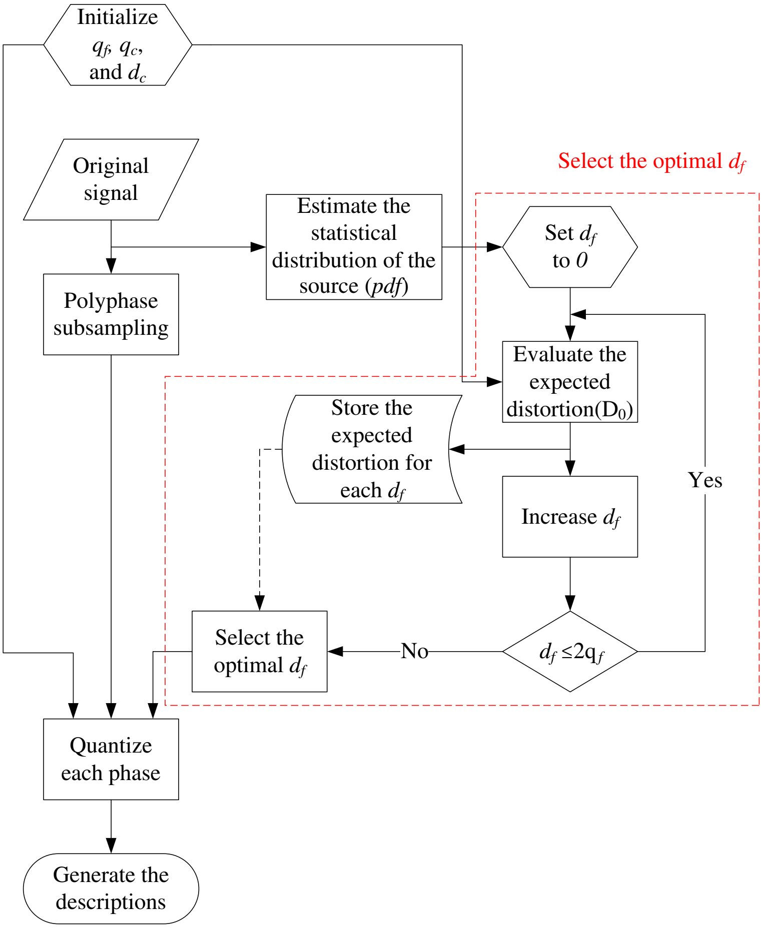

Fig. 2 The flowchart of the proposed scheme

can be obtained by the refinement process. Therefore, we will firstly get the optimal deadzone width of $\mathbf{Q}_2$ so as to refine $\mathbf{Q}_1$. Then we determine the optimal deadzone width of $\mathbf{Q}_3$ and then $\mathbf{Q}_4$. This process can be generalized to produce any number of descriptions.

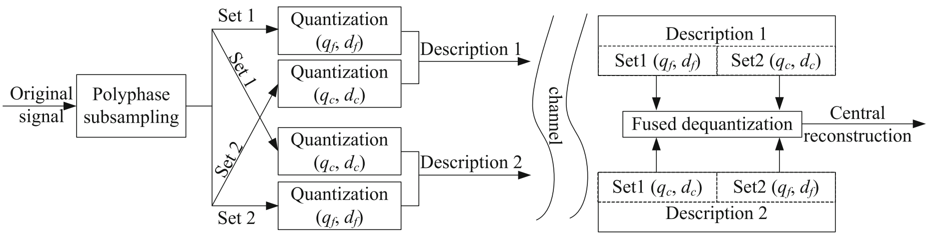

Fig. 3 The two-description case with the fused dequantization

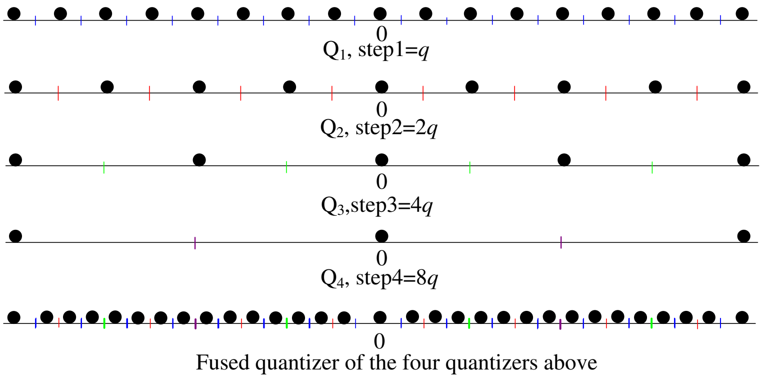

Fig. 4 The refined quantization case for the four quantizers. The *vertical lines* and *round dots* denote the interval boundary and reconstructed level for the quantizers

3.2 The proposed scheme for image coding

In this section, we apply the proposed scheme to a simple wavelet based image encoder. The proposed Quadtree classification and Scalar Quantization scheme (QSQ) is implemented, and it is mainly a modified version of the Quadtree classification and Trellis Coded Quantization scheme (QTCQ) proposed by Banister et al. [2]. The proposed QSQ scheme will classify the wavelet coefficients into some classes by a threshold determined with its quantization step. The classified coefficients are quantized subsequently. More details about QTCQ can be found from Banister's paper [2]. In the proposed MD image coding scheme, the coefficients of the wavelet transformed image are polyphase subsampled by odd and even quadtree fashion into two parts. After subsampling, each phase will undergo the procedure that has been described in Section 2 to get the optimal set of parameters. With these parameters, the QSQ scheme is employed to encode the coefficients. Finally, the description is generated by composing the two parts, i.e. the phase obtained with fine quantization step and the second phase generated by the coarse quantization step, which is similar to Fig 3. Whereas, the other description is formed in a complimentary way. When both of the two descriptions are received, the coarse information can be used to improve the fine information because the deadzone width of the quantizers have been designed to enhance the performance of the refined central quantizer. Most importantly, the proposed approach will improve the performance for whatever amount of inserted redundancy, in contrast with the approach proposed by Tillo et al. [9].

For the four-description scheme, the wavelet coefficients are subsampled into four sets by quadtree fashion and encoded by four different quantization steps, following the QSQ approach. Each description will be composed of four sets encoded with four different quantizers steps. Meanwhile, each set is quantized differently into the four descriptions. Thus, when more than two descriptions are received, a fused quantizer for each set can be employed to assure a better performance. Moreover, any additionally received description will be used to further improve the refined quantizer.

4 Experimental results

A set of simulations have been carried out on vectors having 10^5 samples of a zero mean and unit variance memoryless Gaussian random process. The total rate is 5 bps, and the ratio between the two quantization steps (i.e., q_f/q_c) varies from 0.10 to 1. The step is selected as 0.1 in the range of $\mathbf{d}_f = [0, 2q_f]$, which is good for the tradeoff of complexity and performance. Figure 5 shows the central performance versus side performance in terms of SNR for the proposed approach and that from Tillo et al. [9]. It is obtained by tuning q_f to meet the total bit rate and using (3) to get the optimized performance. The reported results show that the proposed algorithm outperforms Tillo's method in general [9], especially when the ratio between the quantization step of side quantizer and central quantizer is odd. The maximum gain is up to 2 dB. In the simulation, the estimation of *pdf* is obtained using Raykar's non-parametric histogram method [7] which, despite its simplicity, yields a reasonable estimate of the original *pdf*. More details about estimation process can be found Raykar's paper [7].

A second set of simulations have been carried out to verify the effectiveness of the proposed approach for image coding.

Figure 6 shows the results of Lena image with dimension 512×512 pixels in terms of central versus side PSNR. For comparison, we also give the results of various state-of-the-art schemes. This includes the wavelet based MDSQ approach proposed by Servetto et al. [8], PTSQ with SPIHT engine proposed by Jiang et al. [4], JPEG2000-based rate-distortion MDC (RDMDC) approach proposed by Tillo et al. [10] and feature-oriented MDC (FOMDC) proposed by Liu et al. [5]. All the reported methods use efficient image compression framework with wavelet transform in their coding schemes, such as SPITH and JPEG2000. Hence, the results are very competitive. In our scheme, wavelet transform has been applied with 6 levels of decomposition, and the Daubechies 9/7 filters are adopted. The bit rate is set to be 0.5 bpp per description. The different side/central tradeoff points are obtained with different ratios between coarse and fine quantization step. The ratios are also denoted in Fig. 6. The results of the proposed scheme with no refinement (i.e. the

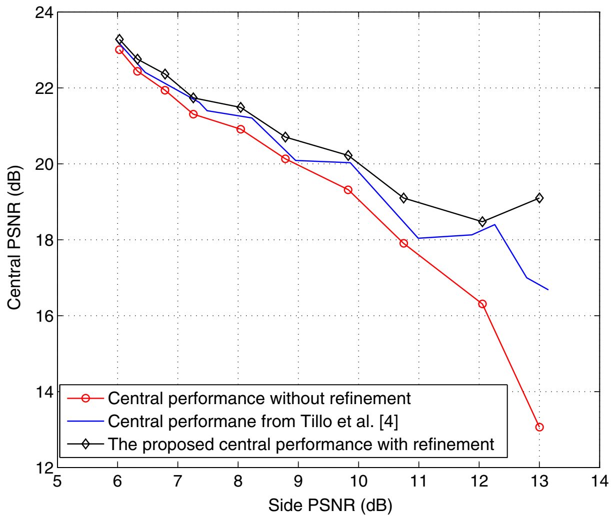

Fig. 5 Central versus side performance for a Gaussian i.i.d. source

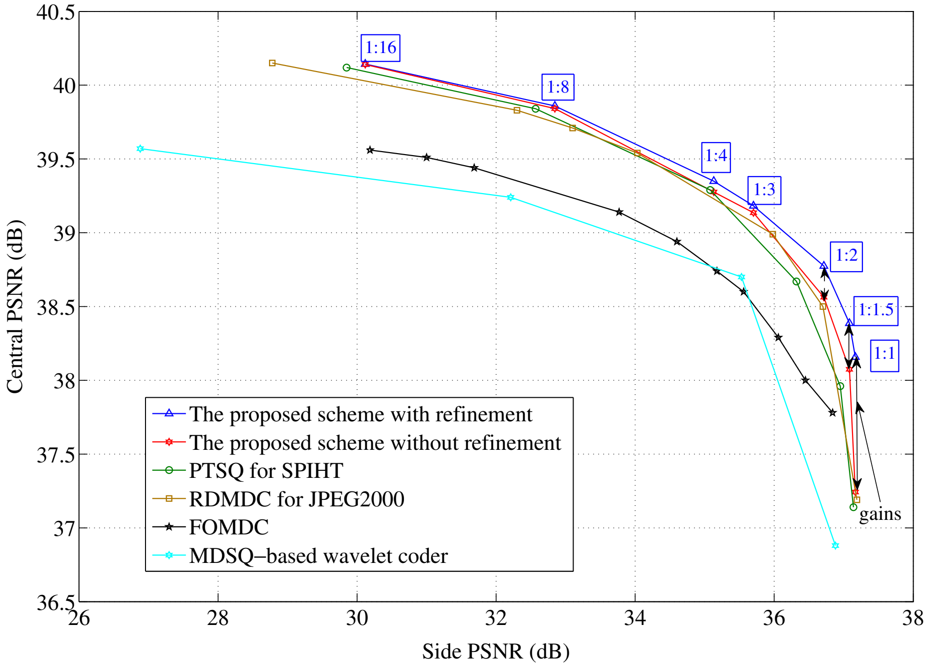

Fig. 6 Central versus side PSNR for image Lena, encoded at 0.5 bpp per description; the ratios of different quantization steps are labeled in the figure and the gains over no refinement case are indicated

coarse and fine quantizers are not fused) is also presented. These, along with the results of the fused quantizers have identical side PSNR, while the central PSNR gap between the two approaches, shows the obtained gain by the proposed fused quantizer. The gain depends on the inserted redundancy and it is up to 1 dB when the

Table 1 Central and side PSNR for image Lena at 0.5 bpp per description

Compared schemes	Side PSNR (dB)						
	Central PSNR (dB)						
Proposed scheme with refinement	37.17	37.08	36.71	35.7	35.13	32.84	30.12
	38.16	38.39	38.77	39.18	39.35	39.86	40.14
Proposed scheme without refinement	37.17	37.08	36.71	35.7	35.13	32.84	30.12
	37.17	38.07	38.56	39.14	39.28	39.84	40.14
PTSQ for SPIHT [4]	37.14	36.95	36.32	35.08	32.57	29.85	23.18
	37.14	37.96	38.67	39.29	39.84	40.12	40.3
MDSQ-based wavelet coder [8]			36.88	35.53	32.21	26.88	23.06
			36.88	38.7	39.24	39.57	39.71
RDMDC for JPEG2000[10]	37.19	36.7	35.98	34.03	33.1	32.3	28.78
	37.19	38.5	38.99	39.54	39.71	39.83	40.15
FOMDC [5]	36.84	36.45	35.56	35.18	33.77	31.68	30.18
	37.78	38	38.6	38.74	39.14	39.44	39.56

The side and central performance are shown in the first and second row for each method, respectively

Table 2 The results comparison of four-description MDC scheme

Compared schemes	PSNR with different number of received descriptions (dB)			
	One	Two	Three	Four
Baccaglini et al. [1] (0.5 bpp)	34.59	36.19	38.11	41.53
Proposed scheme without refinement(0.5 bpp)	34.55	37.94	39.79	41.00
Proposed scheme with refinement (0.5 bpp)	34.55	38.09	40.02	41.35
Baccaglini et al. [1] (1.0 bpp)	37.77	39.33	41.78	47.68
Proposed scheme without refinement (1.0 bpp)	37.71	41.56	44.1	46.13
Proposed scheme with refinement (1.0 bpp)	37.71	41.75	44.38	46.68

ratio between the two quantization step is 1:1. As the ratio between the quantization steps becomes close to 1:16, there would be no obvious gain because the bit rate for the coarse quantization only accounts for 5% of the total rate. Since some of the results are very competitive and overlapped, detailed results are provided in Table 1, where the side and central performance are shown in the first and second row for each method, respectively. It can be noticed that the proposed scheme achieves the best performance at all the points. Moreover, it can be seen that the improvement increases as the ratio between the two quantizers steps approaches one, i.e. when the redundancy increases, which is particularly important since the main drawback of the polyphase-based approaches is their inefficient exploitation of the inserted redundancy at high range. It should be noted the proposed scheme needs some extra computation cost for the estimation of the *pdf* for the source.

The four-description coding results are shown in Table 2 for the Lena image of dimension 512×512. The target bit rate is 0.5 bpp and 1.0 bpp per description. For comparison, the performance without refinement is provided for the proposed scheme. Moreover, we report the performance curve of the algorithm proposed by Baccaglini et al. in [1]; this algorithm exploits the data rate-distortion characteristics to generate multiple descriptions for JPEG 2000, which is also based on polyphase subsampling. The proposed side performance (only one description) and that of Baccaglini's method are tuned to be close, so that we can have a good comparison for the more than one description received case. It can be seen that the gain is about 0.2–0.5 dB and it increases as more descriptions are received, compared with that without refinement. In addition, the gain increases with the bit rate because the refinement model is much more accurate at high bit rate. The proposed scheme is inferior to that of Baccaglini's method about 0.2–1 dB when all of the four descriptions are received. Nevertheless, the proposed scheme outperforms Baccaglini's method by 1.9–2.6 dB when two or three descriptions are received. This proves the effectiveness of the proposed scheme.

5 Conclusion

This paper proposes a novel approach to improve the performance of those schemes that use polyphase-based mechanism to generate the descriptions. By tuning the width of the deadzones, the scheme can exploit the coarse version of the data to improve the fine one. The tuning process is carried out by taking into account statistical characteristics of the source data. The proposed scheme overcomes the inefficient exploitation of the redundancy, which is a typical problem with the polyphase-based approaches. Most importantly, the proposed scheme can be easily generalized to more than two descriptions case. Finally, the experimental results on Gaussian and image data demonstrate the effectiveness of the proposed scheme. However, the good performance is achieved with some extra computation cost that is required to estimate the statistical distribution of the source.

Acknowledgements This work was supported by 973 Program, under Grant 2011CB302204, the National Science Foundation of China for Distinguished Young Scholars, under Grant 61025013, Sino-Singapore JRP, under Grant 2010DFA11010, National Natural Science Foundation of China, under Grants 60776794, 60972085.

References

1. Baccaglini E, Tillo T, Olmo G (2007) A flexible R-D-based multiple description scheme for JPEG 2000. IEEE Signal Process Lett 14(3):197–200
2. Banister BA, Fischer TR (2001) Quadtree classification and tcq image coding. IEEE Trans Circuits Syst Video Technol 11(1):3–8
3. Goyal VK (2001) Multiple description coding: compression meets the network. IEEE Signal Process Mag 18(5):74–93
4. Jiang W, Ortega A (1999) Multiple description coding via polyphase transform and selective quantization. In: Proc. SPIE intl. conf. on visual comm. and image proc., pp 998–1008
5. Liu Y, Oraintara S (2007) Feature-oriented multiple description wavelet-based image coding. IEEE Trans Image Process 16(1):121–131
6. Neuhoff DL (1993) The other asymptotic theory of lossy source coding. In: In coding and quantization of DIMACS series in discrete mathematics and theoretical computer science, vol 14, pp 55–65
7. Raykar VC (2002) Probability density function estimation by different methods. ENEE 739Q SPRING 2002 course assignment 1 report, pp 1–8
8. Servetto SD, Ramchandran K, Vaishampayan VA, Nahrstedt K (2000) Multiple description wavelet based image coding. IEEE Trans Image Process 9(5):813–826
9. Tillo T, Gabriella O (2008) Improving the performance of multiple description coding based on scalar quantization. IEEE Signal Process Lett 15:329–332
10. Tillo T, Grangetto M, Olmo G (2007) Multiple description image coding based on Lagrangian rate allocation. IEEE Trans Image Process 16(3):673–683

Chunyu Lin was born in LiaoNing Province, China. He received the B.S. degree from Bohai University and the M.S. degree from Yanshan University in 2002 and 2005, respectively. He obtained his doctor dgree in Beijing Jiaotong University in 2011. From 2009 to 2010, he was a visiting researcher at the ICT group of Delft University of Technology, Netherlands. From 2011 to 2012, He was a postdoc in Gent Unveristy, Belgium. His research interests are in the areas of image/video compression and robust transmission, 2D-to-3D conversion and 3D video coding.

Tammam Tillo was born in Damascus, Syria. He received the degree in Electrical Engineering from the University of Damascus, Syria, in 1994, and the Ph.D. in Electronics and Communication Engineering from Politecnico di Torino, Italy, in 2005. From 1999 to 2002 he was with Souccar for Electronic Industries, Damascus, Syria. In 2004 he was visiting researcher at the EPFL (Lausanne, Switzerland), and from 2005 to 2008, he worked as a Post-Doctoral researcher at the Image Processing Lab of Politecnico di Torino, and for few months he was Invited Research Professor at the Digital Media Lab, SungKyunKwan University, Suwon, S.Korea. From August 2008 he joined Xi'an Jiaotong-Liverpool University, Suzhou, China. His research interests are in the areas of robust transmission, image and video compression, and hyperspectral image compression. He has been member of the technical program committees for several international conferences.

Jimin Xiao was born in Suzhou, China. He received the BS and M.Eng. degrees in telecommunication engineering from Nanjing University of Posts and Telecommunications, China, in 2004 and 2007, respectively. Then he worked as a software engineer in Motorola Nanjing Center, and later as system engineer in Realsil (Realtek) Semiconductor Corp. Currently, he is persuing his PhD degree in the University of Liverpool, UK. His research interests are in the areas of video streaming, image and video compression, and multiview video coding.

Yao Zhao received the B.S. degree from Fuzhou University in 1989 and the M.E. degree from Southeast University in 1992, both from Radio Engineering Department, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University (BJTU) in 1996. He became an associate professor at BJTU in 1998 and became a professor in 2001. From 2001 to 2002, he worked as a senior research fellow in the Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, Netherlands. He received the National Science Fund of China (NSFC) for Distinguished Young Scholars in 2010. He is now the director of the Institute of Information Science, Beijing Jiaotong University. His research interests include image/video coding, digital watermarking, and image/video analysis and understanding. Now he is leading several national research projects from 973 Program, 863 Program, the National Science Foundation of China.