Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **6**, November **2018**

JND based Multi-description Stereo Image Coding

Lili Meng[12], Guina Sun[1], Jia Zhang[12], Yanyan Tan[12], Chunjing Wang[1], Jingxiu Zong[1]

[1]Department of Information Science and Engineering, Shandong Normal University
[2]Institute of Data Science and Technology, Shandong Normal University
No.88 Wenhua East Road, Jinan, 250014, China
mengll_83@hotmail.com

Received July, 2018; Revised August, 2018

Abstract. *In this paper, two multi-description stereo image coding schemes based on JND (just noticeable difference) are proposed. JND is based on the minimum distortion characteristics that the human eye can detect. JND model can be divided into JND model of pixel domain and JND model of DCT domain, and the latter is applied in this paper. Firstly, this paper calculates the JND model of DCT domain according to the visual characteristics of human eyes. Then, it is applied to multi-description stereo image coding system. And JND-based MDROQ and JND-based MDLTPC are used in the stereo image coding scheme respectively. In both schemes, in order to improve the coding efficiency, the DCT values that less than JND threshold are set to zero, while others remain unchanged. Experimental results show that the JND based multi-description stereo image coding schemes can improve coding performance compared with other methods.*

Keywords: Just noticeable difference, multi-description stereo image coding, visual characteristics.

1. **Introduction.** Stereo images are able to give a 3D impression to the viewer. However, the data of 3D image is huge, and the internet and wireless communications networks are unreliable transmission channel. On the one hand, because of the internet has a strong heterogeneity, the transmission of the internet can not ensure the reliability of the data reception. Which means that the packet will be loss and transmission delay. On the other hand, wireless channel has high bit error rate, multipath interference, fading and other properties. These properties determine that the wireless transmission environment is worse than the internet, which may result in the loss of the whole transmission image or the loss of the whole video frame, leading to the complete failure of the transmission channel. Therefore, it is significance to design efficient compression and robust coding schemes for stereo images and video. And a variety of multi-description coding (MDC) methods are proposed.

The appearance of multi-description coding is intended to solve the needs of practical problems, and then the researchers make corresponding theoretical analysis and innovation, and finally returned to practical application. The MDC idea is first proposed by Bell LABS in the 1970s [1]. Later, in order to improve the reliability of speech transmission, Miller and Boyle propose channel segmentation technology for the first time [2]. Vaishampayan et al. proposes the first practical MDC scheme: Multiple Description Scalar Quantization (MDSQ) [3] and applies it to the entropy constrained coding scheme [4]. It requires complex index allocation. Since then, MDC research has shifted from theory to practical research, and many practical MDC models have emerged, for example, MDC based on sampling [5], MDC based on quantification [6][7], MDC based on correlation transformation [8][9], and MDC based on unequal protection [10][11]. [12] proposes a two-stage Modified Multiple Description Scalar Quantization (MMDSQ), which generates the first layer of each Description by setting two staggered quantifiers. MMDSQ does not require the design and implementation of complex index assignment scheme, and its simplicity makes it suitable for image coding. Source separation is another way to generate multiple descriptions, which first appeared in [2]. In [13], the transformation encoding method is proposed, and each description includes a subset of high code rate and a subset of low code rate. In [14], a Rate Distortion method is applied to MDC based on JPEG 2000, i.e., Rate-based Multiple Description Coding (RD-MDC).

A Prediction Compensated Multiple Description Coding scheme (PCMDC) is proposed in [15], where M=2. PCMDC does not encode another subset directly at a lower rate, predictive redundancy of another subset is encoded. In [16] a M channel MDC scheme is proposed, according to Two-Rate Predictive Coding and Staggered Quantization (TR-PCSQ). In [17], a three-layer multi-description coding (TLMDC) scheme is proposed, and PCMDC is extended to $M > 2$ through sequence prediction. In [18], according to MDLTPC, TRPCSQ and TLMDC, an improved MDC is proposed, called Multiple Description Coding with Randomly Offset Quantizers (MDROQ). In [18], another scheme is also proposed, that is Multiple Description Coding with Uniformly Offset Quantizers (MDUOQ), which adopts the uniform quantization.

In the MDC schemes mentioned above, the characteristics of the Human Vision System (HVS) are not fully considered. The human eye is the final receiver of the image, so it is necessary to optimize the encoding algorithm with human visual features. The method of human visual features can be used to optimize the image coding algorithm to make the coding more consistent with human subjective experience. Therefore, a new multi-description stereo image coding method based on human visual characteristics is proposed in this paper, that is JND based multi-description stereo image coding. In this paper, two MDC systems based on JND are exploited, one is JND based MDLTPC, the other is JND based MDROQ.

The outline of the paper is as follows. Sec 2 overviews the mathematical model of DCT domain JND. Sec 3 introduces the schemes of JND based multi-description stereo image coding. Experimental results are given in Sec 4, and Sec 5 summarizes the paper.

2. **The Model of Just Noticeable Difference.**

In this paper, DCT-based JND model is exploited [19]. And, the corresponding mathematical expression is as follows:

$$JND(r,c,i,j) = JND_{Basic}(r,c,i,j) \times JND_{lum}(r,c,i,j) \times JND_{contrast}(r,c,i,j). \tag{1}$$

where r and c are the index of the block in image, i and j are the index of DCT coefficients $(i,j = 1:8)$. $JND_{Basic}(r,c,i,j)$ is spatial contrast sensitivity function (CSF). $JND_{lum}(r,c,i,j)$ is brightness adaptive weighting factor. $JND_{contrast}(r,c,i,j)$ is contrast masking weighting factor.

$JND_{Basic}(r,c,i,j)$ represents the sensitivity of HVS to the visual signal, which is a spatial contrast sensitivity function(CSF), and whose mathematical expression is as follows:

$$JND_{Basic}(r,c,i,j) = \frac{s}{\phi_i \phi_j} \cdot \frac{exp(c\omega_{ij})/(a+b)\omega_{ij}}{r+(1+r)\cdot cos^2(\varphi_{i,j})} \tag{2}$$

where s represents a collection effect, and its empirical value is 0.25, ϕ_i and ϕ_j are the normalization factor of DCT, ω_{ij} is the corresponding spatial frequency of the DCT subband coefficients at the (i,j) position, $r+(1+r)\cdot cos^2(\varphi_{i,j})$ represents the tilt efficiency of the human eye (r=0.6), and $\varphi_{i,j}$ represents the direction angle of the corresponding DCT component. Besides, the three parameters a, b and c are set to 1.33, 0.11 and 0.18 [20].

$JND_{lum}(r,c,i,j)$ represents the brightness adaptive weighting factor, which is used to measure the weight of the perceived error in a stable background, and only depends on the characteristics of the local image. The mathematical expression of $JND_{lum}(r,c,i,j)$ is as follows:

$$JND_{lum}(r,c,i,j)=\begin{cases}(60-\overline{I})/150+1 & \overline{I}\leq 60\\ 1 & 60<\overline{I}<170\\ (\overline{I}-170)/425+1 & \overline{I}\geq 170\end{cases} \tag{3}$$

where $\overline{I}$ denotes the average brightness.

$JND_{contrast}(r,c,i,j)$ is a contrast masking weighting factor, which is usually associated with the perceived degree of a signal. In the calculation, the image is processed by the Canny edge detection firstly. Then, the image blocks are divided into three categories: smooth area, edge area and texture area. Different regions have different weights. The weighted factor of smooth region and the edge region is ψ=1. For the texture region, if the coefficient index satisfies the condition $(i^2+j^2)\leq 16$, ψ=2.25, otherwise ψ=1.25. Considering the masking effect of the adjacent sub-bands, the contrast masking weighting factor can be obtained by following (4).

$$JND_{contrast}(r,c,i,j)=$$
$$\begin{cases}\psi, & for(i^2+j^2)\leq 16\ in\ Plan\ and\ Edge\ Block\\ \psi\cdot min(4,max(1,(\frac{C(r,c,i,j)}{JND_{Basic}(r,c,i,j)}\times JND_{lum}(r,c))^{0.36})), & \\ & Others\end{cases} \tag{4}$$

In this paper, JND is used to preprocess the DCT coefficients. First, DCT coefficients are extracted from an image. Then, the DCT coefficients are compared with the corresponding transform domain JND values of the same position. The DCT coefficient will be set to zero, if it is less than the JND value at the corresponding position.

3. **System Description.** In this section, we will introduce the two $2-description$ MDC schemes based on JND for stereo image. Firstly, the stereo image are divided into two subsets. The Fig. 1 shows that how to get the two subsets from the image. For one description, one subset $S0$ is encoded with small stepsize. Another subset $S1$ is predicted by the reconstructed $S0$,and the prediction residuals are quantized by large stepsize. For another description, one subset $S1$ is encoded with small stepsize. Another subset $S0$ is predicted by the reconstructed $S1$,and the prediction residuals are quantized by large stepsize. The detailed processing is given in following subsection 3.1 and subsection 3.2.

3.1. **JND based MDLTPC Scheme for Stereo image.** In this section, the JND based MDLTPC scheme for stereo image will be introduced. The MDLTPC scheme is given in [15]. The JND preprocessing will be used to decrease the redundancy. The processing of one description is as follows:

At the encoder, the stereo image is transformed by the time-domain lapped transform (TDLT) [21] firstly. Secondly, the coefficients of TDLT are divided into two subsets ($S0$, $S1$) using the method of Fig. 1. Then, the subsets of $S0$ are processed by DCT, the JND model is introduced to make the DCT coefficients be sparse. The DCT coefficients are compared with the DCT-based JND thresholds (1). We can neglect the data that is less than the JND thresholds, while others are retained. Which will improve the coding efficiency under the premise of ensuring the visual quality. Then, the quantization with small stepsize and entropy coding are used to process the preprocessed DCT coefficients.

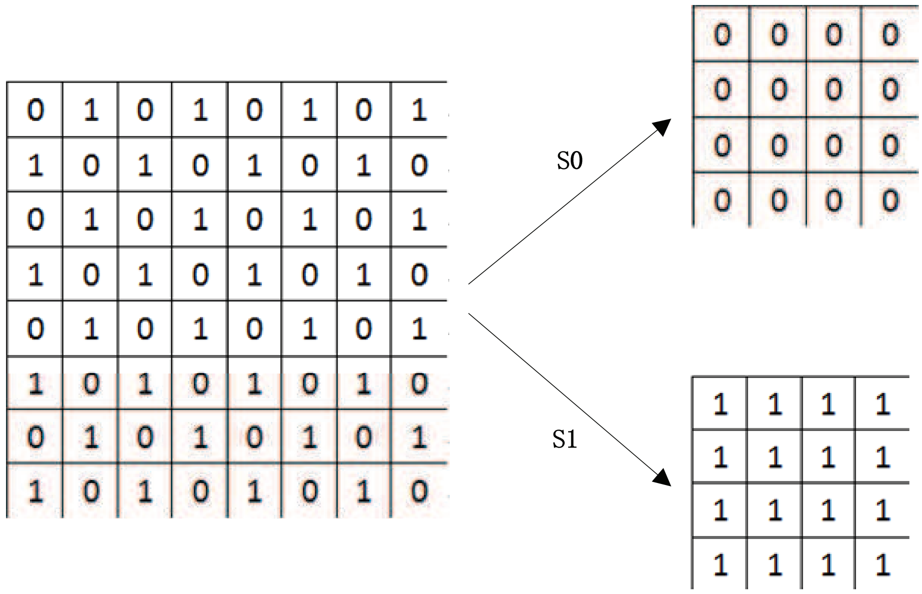

Figure 1. The block subset definition in $2 - description$ image coding.

We can get the bit streams $R0$. The subsets $S1$ are predicted by the neighboring reconstructed $S0$, which can be gotten at the encoder. The prediction residuals are processed by DCT, quantization with large stepsieze and entropy coding. The bit streams of $R1$ of $S1$ can be received.

At the decoder, the bit streams $R0$ are processed by entropy decoding, inverse quantization with small stepsize and IDCT. The reconstructed subsets $\hat{S}0$ can be gotten. Another bit streams $R1$ are processed by entropy decoding, inverse quantization with large stepsize and IDCT. Then, the reconstruction subsets $\hat{S}1$ can be received by the prediction using the reconstructed $\hat{S}0$. Finally, the inverse TDLT is applied to the reconstructed subsets $\hat{S}0$ and $\hat{S}1$, the reconstructed stereo image is gotten by this description.

In the reconstructed stereo image of this description, the reconstructed subsets $S0$ are refined, and the subsets $S1$ are reconstructed coarsely. If we get the reconstructed stereo image by another description, the reconstructed subsets $S1$ are refined, and the subsets $S0$ are reconstructed coarsely. If we receive both of the two descriptions, all the reconstructed subsets $S0$ and $S1$ are refined.

3.2. **JND based MDROQ Scheme for Stereo image.** In this section, the JND based MDROQ scheme for stereo image will be introduced. The MDROQ scheme is given in [18]. The JND preprocessing will be used to decrease the redundancy. The processing of one description is as follows:

At the encoder, firstly, the stereo image is transformed by the time-domain lapped transform (TDLT) [21]. Secondly, the coefficients of TDLT are divided into two subsets ($S0$, $S1$) using the method of Fig. 1. Then, the subsets of $S0$ are processed by DCT, the JND model is introduced to make the DCT coefficients be sparse. The DCT coefficients are compared with the DCT-based JND thresholds (1). We can neglect the data that is less than the JND thresholds, while others are retained. Which will improve the coding efficiency under the premise of ensuring the visual quality. Then, the quantization with small stepsize and entropy coding are used to process the preprocessed DCT coefficients. We can get the bit streams $R0$ of $S0$. The subsets $S1$ are processed by DCT and JND preprocessing, then, the coefficients are predicted by the neighboring reconstructed DCT coefficients of $S0$, which can be gotten at the encoder. The prediction residuals are processed by quantization with large stepsieze and entropy coding. The bit streams of $R1$ of $S1$ can be received.

At the decoder, the bit streams $R0$ are processed by entropy decoding, inverse quantization with small stepsize and IDCT. We get the reconstructed subsets $\hat{S0}$. Another bit streams $R1$ are processed by entropy decoding, inverse quantization with large stepsize, prediction with wiener filter [22]and IDCT. The reconstruction subsets $\hat{S1}$ can be received. Finally, the inverse TDLT is applied to the reconstructed subsets $\hat{S0}$ and $\hat{S1}$, the reconstructed stereo image is gotten by this description..

In the reconstructed stereo image of this description, the reconstructed subsets $S0$ are refined, and the subsets $S1$ are reconstructed coarsely. If we get the reconstructed stereo image by another description, the reconstructed subsets $S1$ are refined, and the subsets $S0$ are reconstructed coarsely. If we receive both the descriptions, all the reconstructed subsets $S0$ and $S1$ are refined.

4. **Experimental Results.** In this section, we evaluate the performance of the proposed scheme. The peak signal-to-perceptual ration (PSPNR) in [23] is used as the criterion for evaluating the quality of the decoding image. The expression of the PSPNR is as follows:

$$\begin{aligned} PSPNR &= 10log_{10} \\ &\times \frac{255 \times 255}{\frac{1}{WH}\sum_{x=1}^{W}\sum_{y=1}^{H}|I(x,y)-\hat{I}(x,y)|^2\delta(x,y)} \end{aligned} \tag{5}$$

where

$$\delta(x,y) = \begin{cases} 1, if|I(x,y)-\hat{I}(x,y)| \geq Pixel\_JND(r,c,i,j) \\ 0, otherwise \end{cases} \tag{6}$$

$I(x,y)$ and $\hat{I}(x,y)$ represent the original value and the reconstructed value of the pixel located at (x,y) respectively. W and H denote the width and height of the image.

We compare the performances of proposed JND-based MDLTPC and JND-based MDROQ with MDLTPC [15] and MDROQ[18].The stereo images *rabbit_L*.000, *rabbit_R*.000, *soccer_L*.000 and *soccer_R*.000 are used to test the experimental results. The resolution of the test images is 720×480.

Fig. 2 and Fig. 3 report the expected PSPNR and central PSPNR for *rabbit* and *soccer* respectively, by adjusting the values of $q0$ and $q1$ while maintaining the same total bit rates. It can be seen that for the same central PSPNR, the expected PSPNR of JND-based MDLTPC and JND-based MDROQ are better than the MDLTPC and MDROQ.

Fig. 4 and Fig. 5 compare the relationships between the side PSPNR D_i and central PSPNR D_M of JND-based MDLTPC, JND-based MDROQ, MDLTPC and MDROQ for *rabbit* and *soccer*. It can be seen that the JND-based MDLTPC and JND-based MDROQ outperform MDLTPC and MDROQ respectively.

It can be seen from all the experimental results that the stereo multiple description schemes based on JND have better performance. The reason is that the preprocessing of JND can neglect some DCT coefficients, which don't impact the human vision.

5. **Conclusion.** Two JND based multi-description stereo image coding schemes are proposed in this paper. JND model of DCT domain is used. We calculates the JND model of DCT domain according to the visual characteristics of human eyes firstly. Then, the JND model is applied to the multi-description stereo image coding system. JND-based

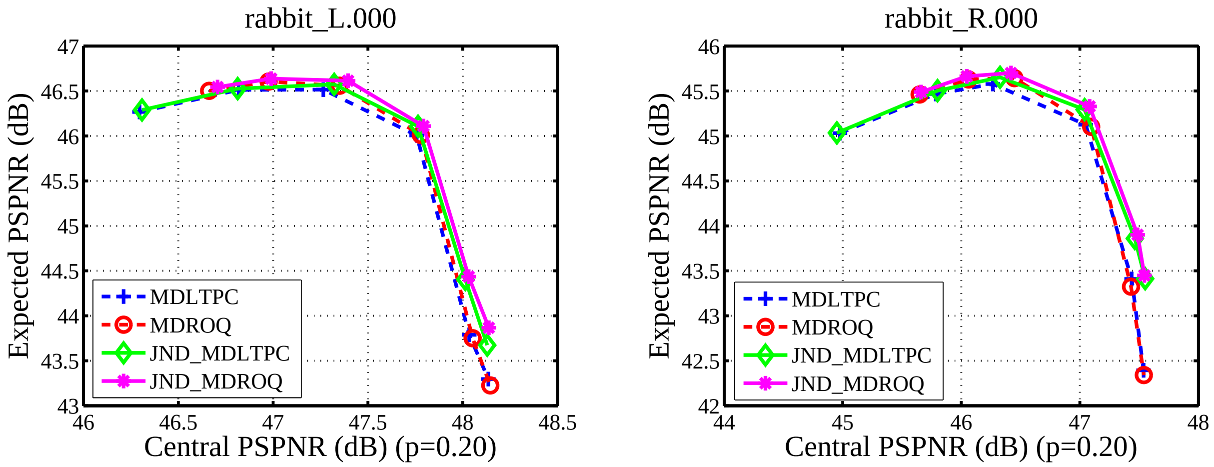

FIGURE 2. The expected PSPNR vs. central PSPNR of various methods for Rabbit.

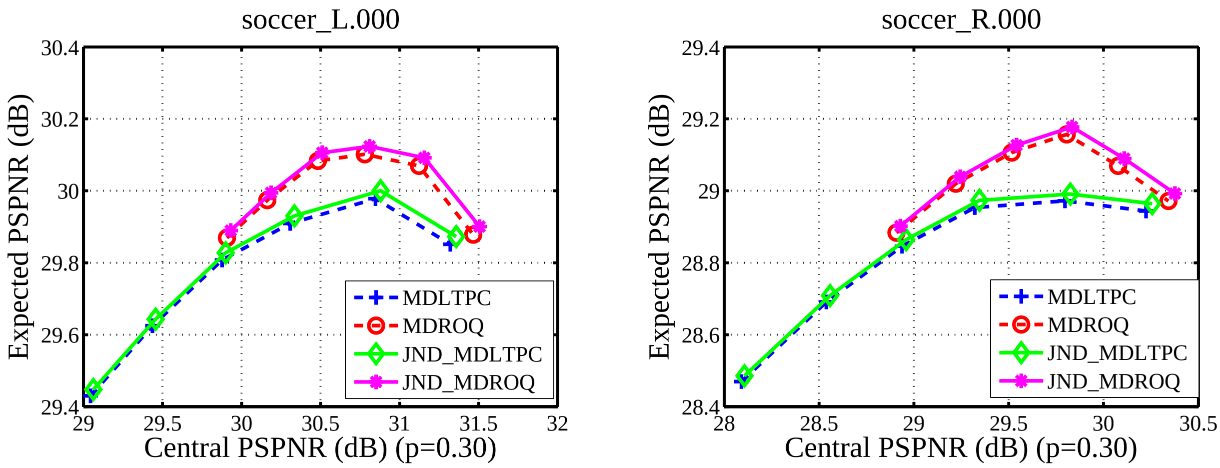

FIGURE 3. The expected PSPNR vs. central PSPNR of various methods for Soccer.

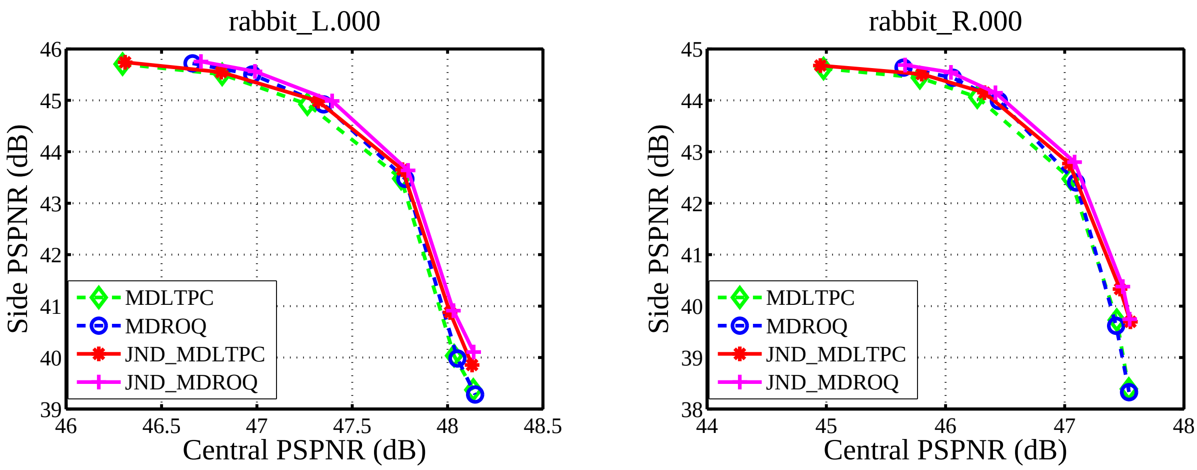

FIGURE 4. The side PSPNR vs. central PSPNR of various methods for Rabbit and total rate of 0.2 bpp

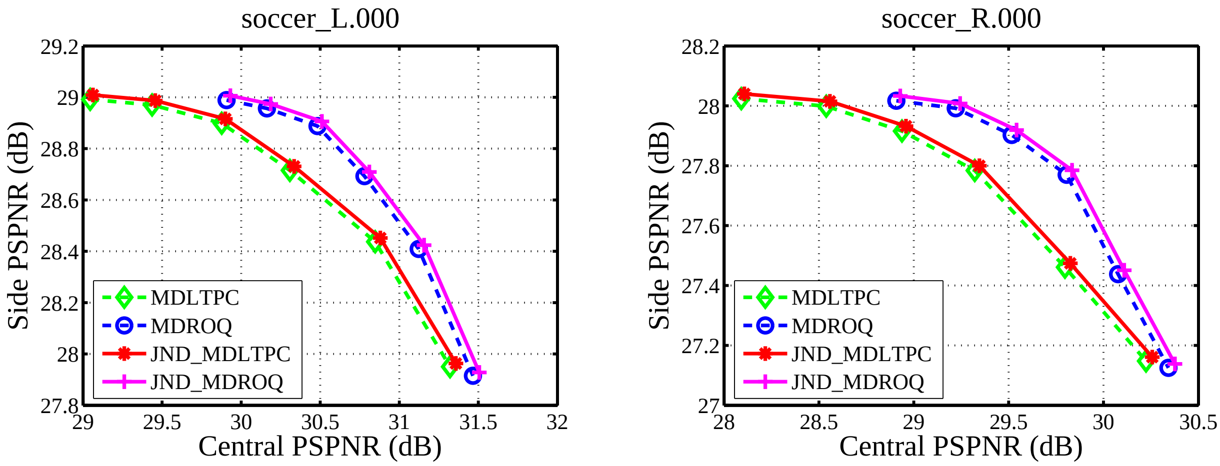

FIGURE 5. The side PSPNR vs. central PSPNR of various methods for Soccer and total rate of 0.5 bpp

MDROQ and JND-based MDLTPC are exploited to compress the stereo image. The experimental results show that the proposed scheme achieve better performance than other methods.

Acknowledgment. The work is partially supported by the National Natural Science Foundation of China (No. 61402268, 61572298, 61601269, 61602285, 61601268, U1736122) and the Natural Science Foundation of Shandong China (No. ZR2016FB12).

REFERENCES

[1] L. Ozarow, On a source-coding problem with two channels and three receivers, *Bell System Technical Journal*. vol. 59, no. 10, pp. 1909-1921, 2013.
[2] N. S. Jayant, Subsampling of a DPCM Speech Channel to Provide Two "Self-Contained" Half-Rate Channels. *Bell System Technical Journal*, vol. 60, no. 4, pp. 501-509, 2014.
[3] V.A. Vaishampayan, Design of multiple description scalar quantizers. *IEEE Transactions on Information Theory*. vol. 39, no. 3, pp. 821-834, 1993.
[4] V.A. Vaishampayan and J. Domaszewicz, Design of entropy-constrained multiple-description scalar quantizers. *IEEE Transactions on Information Theory*. vol. 40, no. 1, pp. 245-250, 1994.
[5] Y. Wang, A.R. Reibman, and S. Lin, Multiple Description Coding for Video Delivery. *Proceedings of the IEEE*. vol. 93, no. 1, pp. 57-70, 2005.
[6] H. Bai, Y. Zhao, and C. Zhu, Optimized multiple description image coding using lattice vector quantization, *Ksii Transactions on Internet & Information Systems*. vol. 9, no. 3, pp. 1140-1154, 2005.
[7] H. Jafarkhani and V. Tarokh, Multiple description trellis coded quantization. *IEEE Transactions on Communications*. vol. 47, no. 6, pp. 799-803, 1999.
[8] J.C. Batllo and V.A. Vaishampayan, Asymptotic performance of multiple description transform codes. *IEEE Transactions on Information Theory*. 43, no. 2, pp. 703-707, 2002.
[9] Y. Wang, et al., An improvement to multiple description transform coding. *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2843-2854, 2002.
[10] A. E. Mohr, E. A. Riskin, and R. E. Ladner. Generalized multiple description coding through unequal loss protection. *International Conference on Image Processing, 1999. ICIP 99. Proceedings*. 1999.
[11] M.R. Ardestani, A. Aminlou, and A.A.B. Shirazi. Unequal loss-protected multiple description coding of scalable source streams using a progressive approach. *IEEE International Conference on Image Processing*. 2009.
[12] C. Tian and S.S. Hemami, A new class of multiple description scalar quantizer and its application to image coding. *IEEE Signal Processing Letters*, vol. 12, no. 4, pp. 329-332, 2005.
[13] W. Jiang and A. Ortega, Multiple description coding via polyphase transform and selective quantization. *Proceedings of SPIE-The International Society for Optical Engineering*, 1998. 3653.

[14] E. Baccaglini, T. Tillo, and G. Olmo, A Flexible R-D-Based Multiple Description Scheme for JPEG 2000. *IEEE Signal Processing Letters*, 2007, vol.14, no. 3, pp. 197-200.

[15] G. Sun, et al., Multiple description coding with prediction compensation. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Societ*y, vol. 18, no. 5, pp. 1037-47, 2009.

[16] U. Samarawickrama, J. Liang, and C. Tian, M-Channel Multiple Description Coding With Two-Rate Coding and Staggered Quantization: *IEEE Press.* , pp. 933-944, 2010.

[17] U. Samarawickrama, J. Liang, and C. Tian. A three-layer scheme for M-channel multiple description image coding. *in IEEE International Conference on Image Processing.* 2011.

[18] L. Meng, et al., Multiple Description Coding With Randomly and Uniformly Offset Quantizers. *IEEE Transactions on Image Processing*, 2014, vol. 23, no. 2, pp. 582-595.

[19] Y. Jia, W. Lin, A. Kassim. Estimating just-noticeable distortion for video. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, vol. 16, no. 7, pp. 820-829.

[20] Z Wei, K. N. Ngan. Spatio-temporal just noticeable distortion profile for grey scale image/video in DCT domain. *IEEE Transactions on Circuits and Systems for Video Technology*, 2009, 19, no. 3, pp. 337-346.

[21] T. Tran, j. Liang, C. Tu. Lapped transform via time-domain pre-and post-filtering. *IEEE Transactions on Signal Processing*, 2003, vol. 51, no. 6, pp. 1557-1571.

[22] J. Liang, C. Tu, L. Gan, et al. Wiener filter-based error resilient time-domain lapped transform. *IEEE transactions on image processing*, vol. 16, no. 2, pp. 491-502, 2007.

[23] Y Gao, X Xiu, J Liang, et al. Fast synthesized and predicted just noticeable distortion maps for perceptual multiview video coding. *Journal of Visual Communication and Image Representation* , vol. 24, no. 6, pp. 700-707, 2013,.