Polyphase Downsampling Based Multiple Description Image Coding using Optimal Filtering with Flexible Redundancy Insertion

Çağlar Ateş, Yılmaz Ürgün, Begüm Demir, Oğuzhan Urhan, Sarp Ertürk,
Kocaeli University Laboratory of Image and Signal Processing,
Electronics and Telecommunications Eng., University of Kocaeli,
41040 Kocaeli, Turkey,
e-mail: {caglar.ates, yilmazurgun}@gmail.com, {begum.demir, urhano, sertur}@kocaeli.edu.tr

Abstract— **In this paper, a novel multiple description image coding scheme is proposed to facilitate the transmission of images over unreliable networks. The target is to minimize the received image distortion over error prone channels. The proposed method is based on adding redundancy to multiple descriptions coding structure to increase correlation between descriptions. The novelty of the paper is that optimal reconstruction filter coefficients are obtained, that will be used to combine the multiple descriptions in an optimal way, based on least squares minimization of the reconstruction error. It is shown that the proposed method provides better results compared to existing approaches in the literature.**

Index Terms— image coding, multiple description coding (MDC), polyphase downsampling, post-processing

I. INTRODUCTION

Nowadays wired and especially wireless networks are widespread, and therefore transmitting data packets over error-prone networks is very important. Moreover many multimedia applications such as video conferencing, digital video broadcast require large bandwidth and limited delay time. These constraints may cause important problems in some applications that require real time data transmission [1]. Multiple descriptions coding (MDC) approaches provide a reasonable solution for such mediums. MDC creates self decodable data packets, so called descriptions, to transfer data over unreliable networks [2]. In the case of packet loss some of the descriptions might not reach the decoder. In such cases, the decoder utilizes only the received descriptions to reconstruct the original data as good as possible. If all descriptions are received, a higher quality signal can be reconstructed. However, since certain level of redundancy to is introduced to recover lost data, the overall coding efficiency may be degraded.

The purpose of MDC is to effectively transmit data such as voice, image and video, over error-prone networks. Multiple description scalar quantization is presented in [3] to provide rate distortion bounds. It is proposed to use MDC for voice communication in [4] creating two descriptions using odd and even samples of the speech signal. Then, these descriptions are encoded using two separate differential pulse code modulation (DPCM) encoders. The method in [5], which is referred to as Pairwise Correlating Transform (PCT) performs encoding on the transformed variables.

Typically if one of the two transformed input variables is not received, the original signal is estimated making use of the other received variable. The lapped orthogonal transform is used for multiple description image coding in [6]. Some polyphase down-sampling (PD) based approaches are also presented in the literature. PD based MDC approaches can be grouped into two main categories. The first one introduces the redundancy by employing different quantization levels for each description as described in [7]. The second approach is to insert the redundancy using a zero-padding approach in the DCT (Discrete Cosine Transform) domain, as in [8]. Then, multiple descriptions are generated by simply sub-sampling the oversampled image data. MDC can also be used for 3-D model coding. The method in [9] employs a wavelet subdivision surfaces strategy for 3-D model coding and introduces 3-D multiple description capabilities.

In the literature, it is shown that down-sampling prior to image coding at the encoder, and succeeding up-sampling after decoding at the decoder provides better results compared to conventional image coding at low bit-rates. In [10], an optimal filtering approach is investigated for the minimization of the difference, introduced by decimation and encoding, between the original and decoded image. Hence, it is shown that the performance of standard JPEG coding can be improved in this way. In [11], the optimal filtering concept presented in [10] is applied for MDC, employing a PD framework. Four sub-images are created by simply downsampling the original image. Next, each sub-image is encoded using standard baseline JPEG and using the optimal filtering concept a total of sixteen optimal filters are obtained. The sub-images and related filter coefficients are combined to create equally important and independent multiple descriptions. It is shown that such a scheme can provide better performance with respect to standard schemes.

In this paper it is proposed to combine the zero padding based PD approach in [8] with the optimal filtering based MDC approach in [11] to enable even better performance for multiple description coding. Following this idea we present a MDC approach that is superior compared to the methods presented in [8] and [11].

II. PROPOSED METHOD

This paper proposes to combine the PD based approach presented in [8] with the optimal filtering approach presented in [11]. The method in [8] firstly introduces redundancy simply inserting zeros in the DCT domain and taking the inverse DCT to return to the image domain again. This process accomplishes oversampling of the original image. This process is depicted in Fig. 1 where D and M show the original image size and the number of inserted zeros in both directions, respectively.

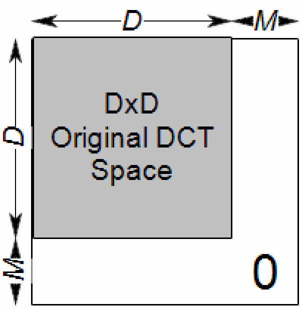

Fig. 1. Redundancy insertion by zero-padding in the DCT domain [8].

After oversampling, i.e. zero insertion, a larger image of size $(D+M)\times(D+M)$ is obtained. The oversampling ratio (OSR) can be defined as the ratio of original image with respect to the zero padded image. This stage can be regarded as a pre-processing step as demonstrated in Fig. 2. Next, multiple descriptions are generated by simply sub-sampling the oversampled input image. Typically two or four descriptions are generated for MDC. It is possible to simply use odd and even samples of the input image as the descriptions. Subsequently, these images are encoded using JPEG and the resultant compressed descriptions are dispatched over the channel. The decoder uses the received description to reconstruct the original image as good as possible. If some of the descriptions are lost, received descriptions are utilized for the reconstruction. Fig. 3. shows all steps of this process in detail.

In this paper, results are demonstrated for four sub-images obtained as follows:

$$\begin{aligned} X_1(i,j) &= X(2i,2j+1) \\ X_2(i,j) &= X(2i+1,2j+1) \\ X_3(i,j) &= X(2i+1,2j) \\ X_4(i,j) &= X(2i,2j) \end{aligned} \qquad (1)$$

where X represents the original image, while $X_n, n=1,2,3,4$ show the sub-images. After creating the sub-images, each sub-image can be encoded using standard baseline JPEG at the desired quality level for MDC, creating the resultant $\tilde{X}_n$ image. At this stage we employ the optimal filtering approach presented in [11].

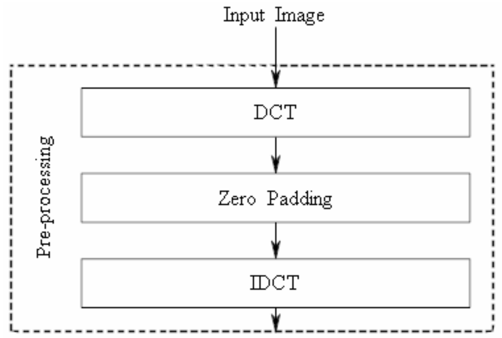

Fig. 2. Pre-processing block: it consists of DCT, zero-padding and IDCT steps [8].

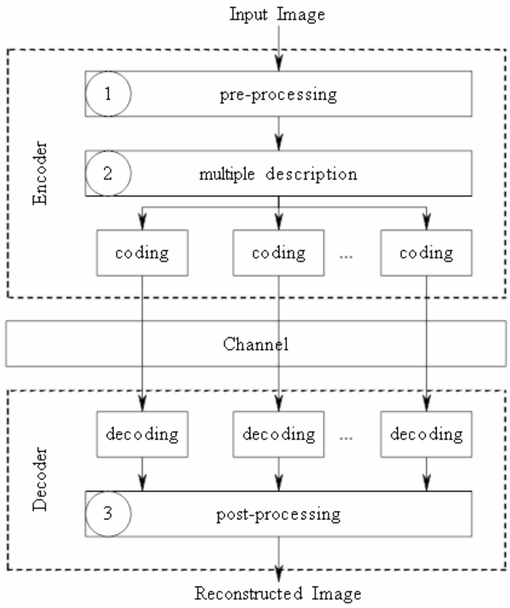

Fig. 3. Block representation of oversampling based MDC approach. (1): preprocessing, (2): multiple description, and (3) post-processing [8].

We use only the first sub-sampled image X_1 to demonstrate the idea of the optimal filtering concept, which is basically the same for all other multiple descriptions as well. The idea is to create estimates of all sub-images, i.e. $\hat{X}_1, \hat{X}_2, \hat{X}_3$, and $\hat{X}_4$, using the $\tilde{X}_1$ and related filters as

$$\begin{aligned} \hat{X}_1 &= \tilde{X}_1 * \mathbf{G}_{1,1} \\ \hat{X}_2 &= \tilde{X}_1 * \mathbf{G}_{1,2} \\ \hat{X}_3 &= \tilde{X}_1 * \mathbf{G}_{1,3} \\ \hat{X}_4 &= \tilde{X}_1 * \mathbf{G}_{1,4} \end{aligned} \qquad (2)$$

where * represent the convolution operation in the image domain. The optimal filters, i.e. $G_{1,n}$, have a dimension of $l \times l$ and are obtained in least squares sense as in (3).

$$\min_{g_{1,n}} \left\| X_n - \hat{X}_1 \right\|_2^2 = \min_{g_{1,n}} \left\| X_n - \tilde{X}_1 * G_{1,n} \right\|_2^2 \qquad (3)$$

The Iterative Preconditioned Conjugate Gradients method (IPCG) [12] is used to solve (3). IPCG is capable of solving large systems fast. The reader is referred to [12, 13] for the details of IPCG.

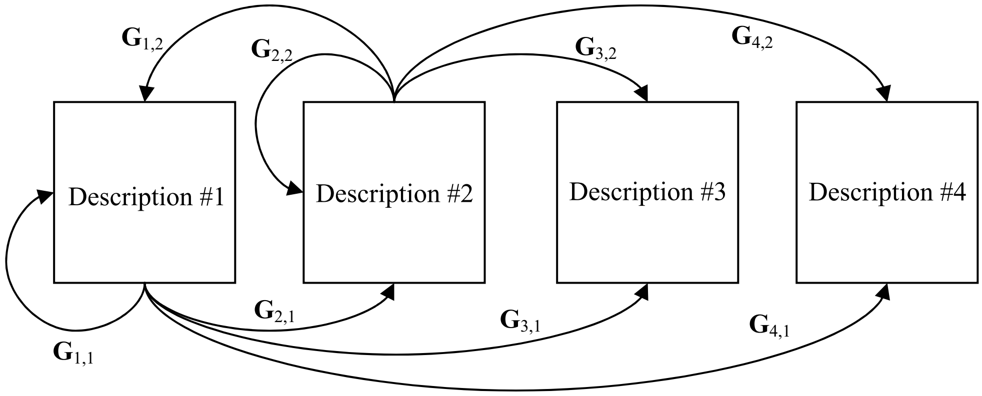

Fig. 4. Reconstruction of the lost descriptions using received descriptions.

The method presented in [11] utilizes a total of sixteen optimal filters to effectively reconstruct the image at the receiver. The lost descriptions are created using the received description and optimal filtering. This situation is demonstrated in Fig. 4 when the third and fourth descriptions are lost. In this case the lost descriptions are estimated from the first and second descriptions and the average of these estimations is used as the final estimation. On the other hand, for the received descriptions only the related, $G_{n,n}$ is used to create the descriptions at the decoder. Finally all created descriptions are combined to reconstruct the image at the decoder by simply performing the inverse of the operation given in (1).

III. EXPERIMENTAL RESULTS

We use several test images to show the effectiveness of the proposed approach with respect to the methods we combined. Rate distortion plots are generally used to assess the performance of MDC methods. We used the Peak Signal to Noise Ratio as distortion measure. Fig 5. shows rate distortion results for different OSR and different description reception cases for the Lena image. The first column in Fig 5. shows the results when only one description is received. Following columns show the two, three, and four description reception cases respectively.

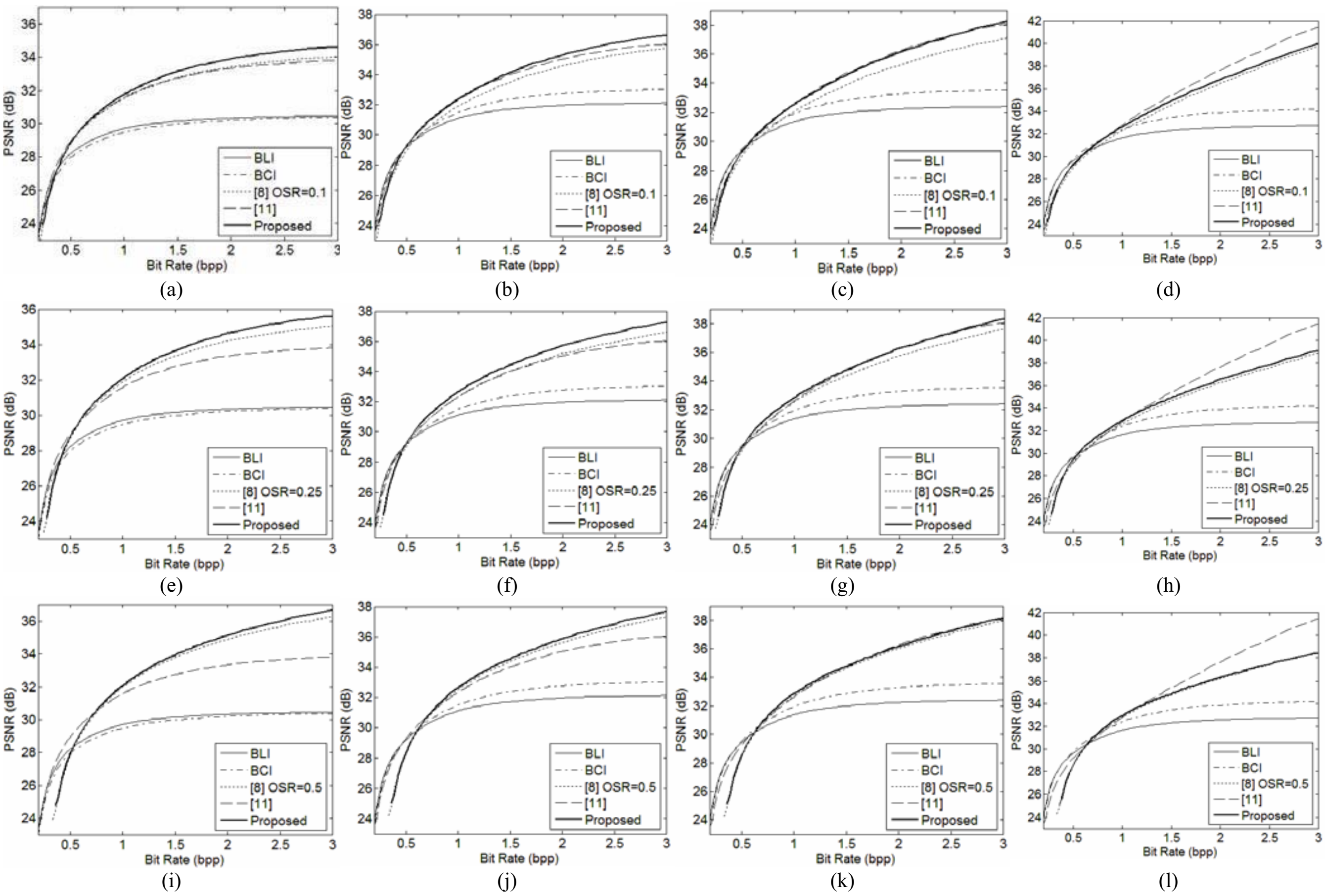

Fig. 5. (a-e-i) One description reception case, (b-f-j) Two description reception case, (c-g-k) Three description reception case, (d-h-l) Four description reception case.

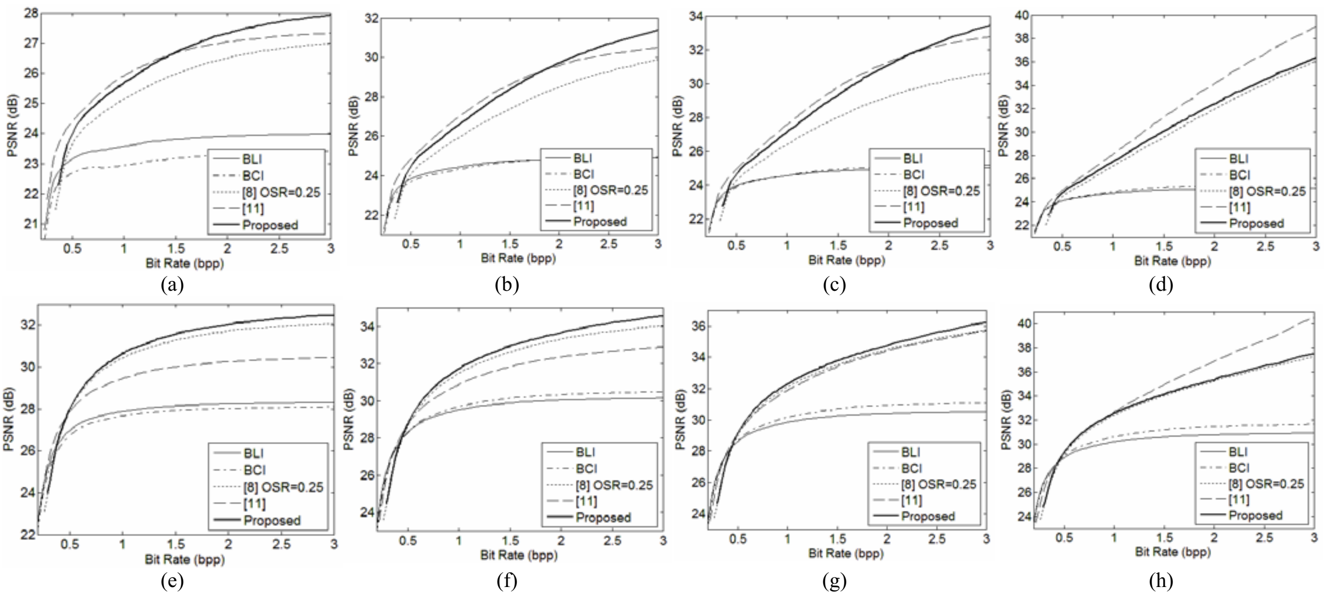

Fig. 6. (a-e) One description reception case, (b-f) Two description reception case, (c-g) Three description reception case, (d-h) Four description reception case.

In the figures, BLI and BCI show bilinear and bicubic interpolation cases. It is clear from Fig. 5 that the proposed combined approach provides mostly the best results. Only, the method presented in [11] provides better results if all descriptions are received at the decoder and the image is encoded at high bit-rates. The main reason for this situation is that the introduced redundancy in the proposed method is not efficient at high bit-rate ranges. When we evaluate the effect of OSR on the performance it is clear that at very low-bit rates higher OSR is not efficient and results in a degradation in performance. Therefore we set the OSR to 0.25 for following experiments since this value balances the performance for both low and high bit-rate ranges. BLI and BCI provide lower performance since these approaches only employ simple interpolation schemes.

Fig. 6 show rate-distortion results for the Barbara and Peppers images, in the first and second rows respectively. As seen from these results in the case of high image detail, i.e. for the Barbara image, the method in [11] provides better results for high bit-rates. This means that oversampling is not very appropriate for images with high details at high bit-rates. On the other hand, in the case of the Peppers image, which contains mostly smooth areas, oversampling works well and the proposed method provides the best results.

IV. CONCLUSIONS

In this paper we present an efficient MDC approach based on a PD based scheme and optimal filtering. The redundancy insertion is carried out using a zero padding based oversampling approach. Experiments show that the proposed approach provides better results with respect to several compared methods in most cases.

ACKNOWLEDGMENTS

This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant EEEAG/107E155.

REFERENCES

[1] Y-C Lee, J. Kim, Y. Altunbasak, and R. M. Mersereau, “Layered Coded vs. Multiple Description Coded Video over Error-Prone Networks”, *Signal Processing: Image Communication*, vol. 18, no. 5, pp. 337–356, May 2003.

[2] V. K. Goyal, “Multiple Description Coding: Compression Meets the Network”, *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74-93, September 2001.

[3] V. A. Vaishampayan, “Design of Multiple Description Scalar Quantizers”, *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 821-834, May 1993.

[4] A. Ingle and V. Vaishampayan, “DPCM System Design for Diversity Systems with Applications to Packetized Speech”, *IEEE Trans. on Speech and Audio Processing*, vol. 3, no. 1, pp. 48-58, January 1995.

[5] Y. Wang, M. Orchard, and A. Reibman, “Multiple Description Coding Using Pairwise Correlating Transforms”, *IEEE Trans. Image Process.*, vol. 10, no. 3, pp. 351–366, March 2001.

[6] D-M. Chung, and Y. Wang, ”Multiple Description Image Coding Using Signal Decomposition and Reconstruction Based on Lapped Orthogonal Transform” *IEEE Trans. Circuits Sys. Video Tech.*, vol. 9, no. 6, pp. 895–908, September 1999.

[7] W. Jiang and A. Ortega, “Multiple Description Coding via Polyphase Transform and Selective Quantization”, *in Proc. SPIE VCIP*, pp. 998-1008, December 1999.

[8] N. Franchi, M. Fumagalli, R. Lancini, “Flexible Redundancy Insertion in a Polyphase Down Sampling Multiple Description Image Coding”, *IEEE International Conference on Multimedia and Expo*, Lausanne, Switzerland, August 2002.

[9] B. Demir, S. Ertürk, “Flexible Multiple Description 3-D Model Coding Using Wavelet Subdivision Surfaces”, *IEEE Signal Processing and Communications Applications Conference*, 2006.

[10] Y. Tsaig, M. Elad, P. Milanfar, GH. Golub, “Variable Projection for Near-Optimal Filtering in Low Bit-Rate Block Coders”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 1, pp. 154 -160, January 2005.

[11] Y. Yapıcı, B. Demir, S. Ertürk, O. Urhan, “Downsampling Based Multiple Description Coding with Optimal Reconstruction Filters”, *IEEE Signal Processing and Communications Applications Conference*, 2007.

[12] M.R. Hestenes and E. Stiefel, “Methods of Conjugate Gradients for Solving Linear Systems”, *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409-436, December 1952.

[13] S. Romero and L.F. Romero, “An Optimized Preconditioned Conjugate Gradient Algorithm”, *Technical Report No: UMA-DAC-02/11*, University of Malaga, September 2002. [Online]: www.ac.uma.es/invdes/inves/reports/2002/UMA-DAC-02-11.ps