ITJ

INFORMATION TECHNOLOGY JOURNAL

An Efficient MDC Framework Based on DCT and SPIHT

[1]Lin-Lin Tang and [2]Zhe-Ming Lu
[1]Department of Electronic and Information Engineering, Harbin Institute of Technology,
Shenzhen Graduate School, Shenzhen, China
[2]School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China

Abstract: Multiple Description Coding (MDC) is one of the promising methods for robust transmission over non-prioritized and unpredictable networks. Based on the Discrete Cosine Transform (DCT) and the Set Partition in Hierarchical Trees (SPIHT) compression method, this study proposes a new MDC framework. We make full use of the energy concentration of DCT and the similarity among the blocks composed of reordered DCT coefficients to apply the SPIHT algorithm to the transform-domain images composed of reordered DCT blocks. The purpose of using the reordered coefficients is to realize the energy redistribution. Redundancy is introduced by the full and partial encoding method which means the three descriptions, each using different bit rates to encode the information from three different orientations, i.e., vertical, horizontal and diagonal directions. For transmission we adopt three channels, each containing the hybrid information from three different directions. Experimental results demonstrate that present technique is effective and practical.

Key words: Multiple description coding, discrete cosine transform, energy redistribution, set partition in hierarchical trees algorithm, full and partial encoding method

INTRODUCTION

With the rapid development of wireless communication techniques, more and more attention is being paid to the transmission robustness. Multiple Description Coding (MDC), which was first introduced in the Shannon Conference in 1979, has already been widely used in the network (Goyal, 2001). The main idea of MDC is to introduce some redundancy among different descriptions. Thus, if one of descriptions is lost, we can use the received others to estimate the original information and obtain an acceptable result. If all descriptions are perfectly received, we can obtain a fairly good reconstruction of the original information but with lower efficiency due to more redundancy. Based on the different ways to introduce redundancy, MDC techniques can be classified into many categories. The most common methods are the quantization-based and transform-based MDC schemes. The Multiple Description Scalar Quantization (MDSQ) coding method (Vaishampayan, 1993) and the Multiple Description Vector Quantization (MDVQ) coding method (Servertto *et al.*, 1999) belong to the quantization-based framework. Generally speaking, more efficient algorithms need more complex quantizer design processes. Therefore, more and more researchers have turned to the transform-based frameworks (Goyal and Kova-Cevic, 1998, 2001).

As one of the most useful compression tools for images, wavelet also plays an important role in MDC. This study proposes a novel MDC framework based on the SPIHT algorithm (Said and Pearlman, 1996) which is based on the wavelet multi-resolution and the traditional DCT transform. Firstly, we apply the traditional DCT to the original image. After reshaping the DCT coefficients into a series of similar subimages, we divide the whole information into three parts, representing three different orientations, i.e., vertical, horizontal and diagonal directions. Then the SPIHT algorithm is applied to different parts with different bit rates. The three-channel framework is chosen here. Every channel contains the hybrid information from three orientations with different bit rates. Thus, if only one of three channels is received, we can obtain an acceptable reconstruction. If only two channels are received, we can reconstruct the image by choosing among two channels, for each direction, the information with the higher bit rate, from two channels. If all channels are received, we can generate a fairly good reconstruction of the original image by choosing among three channels, for each direction, the information with the highest bit rate. Finally, simulation results show the efficiency of present method.

As early as in 1999, Miguel *et al.* (1999) introduced a simple and efficient SPIHT based MDC scheme. In their studies, the redundancy lies in repeated zerotrees. Extra

Corresponding Author: Lin-Lin Tang, Room 1307, Building L, Harbin Institute of Technology, Shenzhen Graduate School, Shenzhen, 518055, People's Republic of China Tel: 008615920555956, 008613868074895
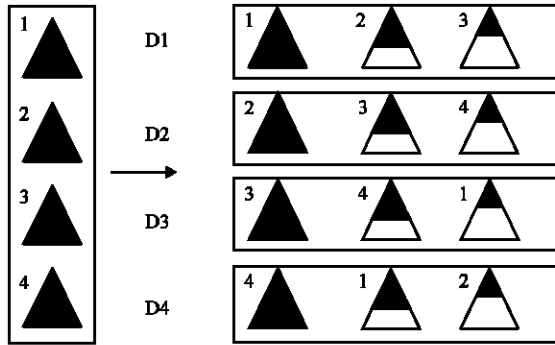
Fig. 1: Four original trees are sent in four descriptions (D1-D4) with two sets of redundant trees (the last two columns). This figure gives the information distribution in four different descriptions based on four trees. The redundancy is introduced by the same information coded at different bit rates. The different ratios of black part in triangles represent the different coding bit rates

copies of each wavelet coefficient tree are placed in other descriptions. It means that the information from one tree is also partially located in other descriptions as shown in Fig. 1. Thus, if two descriptions, e.g., D2 and D3, are lost, they can be partially recovered from their copies in the received descriptions. To incorporate the effect of expected description loss in their method, they used the following equation:

$$D = \sum_{i=1}^{M} \sum_{j=0}^{N-1} p_{i,j} D_{i,j} \qquad (1)$$

where, $p_{i,j}$ is the probability of using the jth copy of a tree i and $D_{i,j}$ is its distortion. M is the number of trees and N is the number of bit-rate choices to encode one tree.

PROPOSED SCHEME

Decomposition and reordering: As one of the most useful image compression techniques, the Discrete Cosine Transform (DCT) has better energy concentration than wavelet transform and has been widely used in various compression standards. We firstly make full use of this characteristic to perform the block DCT with size 8×8 on the original image, obtaining 8×8 decomposition coefficients. For the 512×512 Lena image, we can get a transform-domain image composed of 64×64 DCT blocks. By using the following formula, we can reshape the coefficients into a new transform-domain image with 64 sub-images.

$$D\big(((m-1)\times(R/8)+i),((n-1)\times(C/8)+j)\big)$$
$$= B\big(((i-1)\times 8+m),((j-1)\times 8+n)\big) \qquad (2)$$

where, D represents the reshaped transform-domain image and B is the original transform-domain image, with (i, j) being the block coordinates in the whole 64×64 blocks and (m, n) being the element coordinates in each 8×8 block. Here, R is the number of rows in the transformed image and C is the number of columns in the transformed image.

In this study, R = C = 64. After this operation, we reshape the original image into 64 subimages and the energy concentration is perfectly similar to the wavelet decomposition.

Present MDC scheme: After above operations, we can further divide the reshaped image into three parts with different orientations, i.e., vertical, horizontal and diagonal directions. Then the SPIHT algorithm is applied to them with three different bit rates to generate the whole code for the main information and partially codes for the other two redundant information similar to the former work proposed by Miguel *et al.* (1999), as shown in Fig. 2b. The basic three-channel MDC framework is used in this study as shown in Fig. 2a. Each side-information considers the hybrid information from all orientations with different bit rates as shown in Fig. 2b. Here, we don't pay more attention to the optimization of the bit rates and the packetization algorithm. For description 1, we use the whole information of the first direction as the main information and use one-half information of the second direction and one-third information of the third direction as the redundant information. The similar scheme is adopted for description 2 and 3.

As shown in Fig. 2b, if only one of three channels is received, e.g., the first description, then the full horizontal information, one-half diagonal information and one-third vertical information, can be used to reconstruct the original image. If any two channels are received, e.g., the first two channels, then we can use the full horizontal information, the full diagonal information and one-half vertical information to obtain a better reconstruction. When all three channels are received, the full information from three orientations can be used to generate a fairly good reconstruction. For the fixed three channels, we don't care about the probability of using the different copies as shown in Eq. 1.

To evaluate the results, we use the Mean Squared Error (MSE) to denote the distortion between the reconstructed image and the original image as follows:
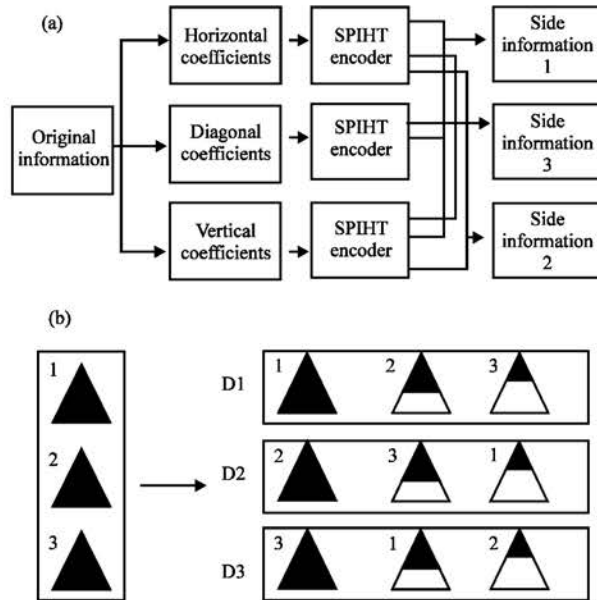
Fig. 2: Basic framework used in the proposed algorithm, (a) the three-channel MDC system used in present algorithm and (b) the constitution of three descriptions (D1-D3), each with three-direction information. Figure a shows the three-channel frame used in present algorithm. Figure b shows the specific composition of three channels used in proposed frame

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(O(i,j) - R(i,j) \right)^2 \qquad (3)$$

where, O(i, j) and R(i, j) represent the original image and the reconstructed image, respectively and M×N is the image size. Based on the MSE value, we can easily get the corresponding PSNR value.

EXPERIMENTAL RESULTS

We adopt the gray-level Lena image of size 512×512 as the test image. All experiments are performed with a 6 scale biorthogonal wavelet transform based on a 4/4-tap filter bank. Table 1 shows the simulation results. Here, one channel means that only one-channel information is received, two channels means that two-channel information is received and three channels means that all channels are received. The bit rate here is calculated after averaging, since there are $C_3^1 = 3$ and $C_3^2 = 3$ sub-cases for the one channel and two channels cases, respectively.

Figure 3 shows the comparisons of the reconstructed image quality between two cases, i.e., without redundancy and with 35% redundancy. For each case, there are two
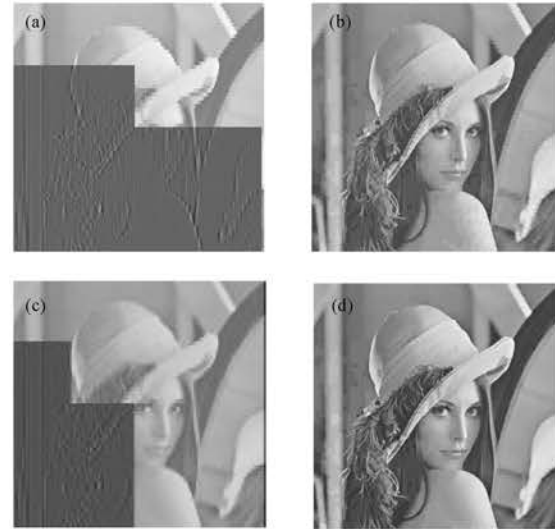

Fig. 3: Comparisons of PSNR values between the case without redundancy (left column) and the case with 35% redundancy (right column), (a) and (b) are for the one channel sub-case. PSNR = 8.33 and 29.68 dB (c) and (d) are for the two channels sub-case. PSNR = 12.76 and 32.17 dB. All results are obtained at the bit rate of 1.215 bpp. These four figures are present experimental results to show the efficiency of present new algorithm. They are achieved through MATLAB program

Table 1: Simulation results for the 512×512 Lena image

	Bit rate (bpp)				
Cases	0.55	1.10	1.65	2.20	2.75
One channel (dB)	26.50	30.34	32.70	34.46	35.73
Two channels (dB)	28.45	32.82	35.13	36.96	37.86
Three channels (dB)	30.42	34.59	36.97	38.70	40.20

These experimental results are achieved by us through the MATLAB program and the experimental image is the 512×512 Lena image

sub-cases, i.e., one channel and two channels sub-cases. Here, all results are obtained at the same bit rate of 1.215 bpp. For the two channels case, we assume that description 1 and 2 are received. The great improvement in image quality shows the importance of the added redundancy.

The comparisons of PSNR values among the non-redundancy method, the MD-SPIHT method (Miguel *et al.*, 1999) and present proposed method are shown in Table 2. Here, all results are obtained at the bit rate of 1.215 bpp and we use the same scheme as adopted in the MD-SPIHT method to packet the information into three descriptions in order to give a relatively fair comparison between the MD-SPIHT method and our proposed method. From Table 2, we can see that when the

Table 2: Comparisons of PSNR values among three methods

| | Methods | | |
Cases	Non-redundancy	Present method	MD-SPIHT
One channel (dB)	8.33	29.68	28.55
Two channels (dB)	12.76	32.17	31.84
Three channels (dB)	38.77	36.96	38.28

These experimental results are also achieved by MATLAB programs. The comparison results show the efficiency of our method. The Non-redundancy case result is achieved based on our MATLAB program and the MD-SPIHT case is emulated by us through MATLAB program
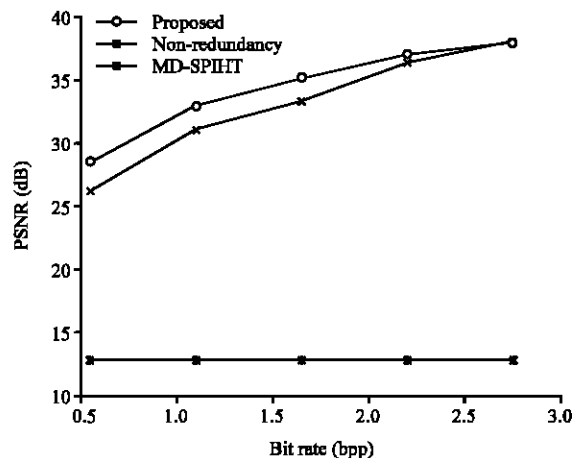

Fig. 4: PSNR values with the increase of bit rates, the above one is ours, the following one is the MD-SPIHT algorithm and the lowest one is for the non-redundancy case. The experimental value is achieved through MATLAB program and the figure is drawn by this tool

number of received channels is less than three, present method gives a better reconstruction of the original image. An even better result will be achieved if we divide the whole information into more descriptions and use more partially-coded redundancy in each description. It is an optimization problem and so is the choice of bit rates.

Another point of view to check the efficiency of present method is to give the reconstruction performance curve with the change of bit rate. We take the two channels case as an example. We compare it with the non-redundancy case and the MD-SPIHT case under the same conditions as shown in Fig. 4. We can see that, for the case without any redundancy, the PSNR value slightly changes with the increase of bit rate. It is almost a horizontal line, where the PSNR value ranges only from 12.74 to 12.77 dB. Obviously, the results are not satisfactory without any redundancy. In contrast, present method performs very well under the same condition, where the PSNR value ranges from 28.45 to 37.86 dB,

which is much larger than the case without any redundancy. From the comparison with the MD-SPIHT case, we can see that present algorithm performs much better when the bit rate is low. Though, our performance grows slower than the MD-SPIHT case as the increase of bit rate, ours is still slightly better than it. Here, the low efficiency of non-redundancy framework results from the whole missing of the information from the third orientation.

DISCUSSION

A novel DCT based MDC framework is proposed in this study. Its main idea is to make full use of the energy concentration of the DCT transform and reshape the DCT coefficients into an image with similar subimages which can make us easily perform the SPIHT algorithm on it. Good performance in the simulation results shows the efficiency of our method. Firstly, the special redundancy introducing method used by us plays an important role in protecting the lost information. Great improvement about 20 dB can be checked both in the one-channel case and the two-channel case in Fig. 3. Secondly, this redundancy introducing method's successful use directly induces the good performance in packet losing reconstruction composition experiment shown in Fig. 4. In fact, the previous findings, the reordered DCT coefficients which can have the similar multiresolution character to the DWT image and the orientation based information division method are the success root of present method. From the special transform domain to the energy redistribution and the orientation based information division, all these methods used in our framework are the difference from other earlier study. To find the optimal bit rates used to encode the redundancy information and to design a more efficient way to divide the whole information are two research topics to be addressed in the future research.

REFERENCES

Goyal, V.K. and J. Kova-Cevic, 1998. Optimal multiple description transform coding of Gaussian vectors. Proceedings of the IEEE Data Compression Conference Snowbird, Mar. 30-Apr. 1, IEEE Computer Society, UT, USA, pp: 388-397.

Goyal, V.K., 2001. Multiple description coding: Compression meets the network. IEEE Signal Process. Maga., 18: 74-94.

Goyal, V.K. and J. Kova-Cevic, 2001. Generalized multiple description coding with correlating transforms. IEEE Trans. Inform. Theor., 47: 2199-2224.

Miguel, A.C., A.E. Mohr and E.A. Riskin, 1999. SPIHT for generalized multiple description coding. Proceedings of the ICIP Kobe, Japan, Oct. 24-28, IEEE, pp: 842-846.

Said, A. and W.A. Pearlman, 1996. A new, fast and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans. Circuits Syst. Video Technol., 6: 243-250.

Servertto, S.D., V.A. Vaishampayan and N.J.A. Sloane, 1999. Multiple description lattice vector quantization. Proceedings of the IEEE Data Compression Conference, Snowbird, Mar. 29-31, IEEE, UT, USA., pp: 13-22.

Vaishampayan, V.A., 1993. Design of multiple description scalar quantizers. IEEE Trans. Inform. Theroy, 39: 812-834.