Multiple Description Coding using Whitening Transform

Kwang-Pyo Choi and Keun-Young Lee
Department of Electrical and Computer Engineering
University of Sungkyunkwan, Republic of Korea
Tel. +82-31-290-7193, Fax.: +82-31-290-7180
E-mail: sinawe@ece.skku.ac.kr

Abstract: In the communications systems with diversity, we are commonly faced on needing of new source coding technique, error resilient coding. The error resilient coding addresses the coding algorithm that has the robustness to unreliability of communications channel. In recent years, many error resilient coding techniques were proposed such as data partitioning, resynchronization, error detection, concealment, reference picture selection and multiple description coding (MDC).

Especially, the MDC using correlating transform explicitly adds correlation between two descriptions to enable the estimation of one set from the other. However, in the conventional correlating transform method, there is a critical problem that decoder must know statistics of original image. In this paper, we propose an enhanced method, the MDC using whitening transform that is not necessary additional statistical information to decode image because the DCT coefficients to apply whitening transform to an image have uni-variance statistics.

Our experimental results show that the proposed method achieves a good trade-off between the coding efficiency and the reconstruction quality. In the proposed method, the PSNR of images reconstructed from two descriptions is about 0.7dB higher than conventional method at the 1.0 BPP and from only one description is about 1.8dB higher at the same rate.

1. Introduction

In the multiple description coding (MDC), especially, source coder generates several coded bit-streams (descriptions). We are able to reconstruct image of reasonable quality even though we have only received one description of them. The highlight of designing a multiple description image coding is how to achieve a good trade-off between the coding efficiency and the reconstruction quality from one description.

Like so much of communications technologies, MDC was invented at Bell Laboratories in connection with communicating speech over the telephone network. In the 1981, N.S. Jayant proposed a separation of odd and even samples in a speech coding method for channel splitting [1]. Later, V.A. Vaishampayan developed a theory for designing multiple description scalar quantizers (MDSQ) [2] and Y. Wang et al. proposed multiple description coding using correlating transforms [3].

In the MDC, statistical dependencies between transform coefficients can be useful for improving the estimation of transform coefficients of a lost description. Especially, the MDC using correlating transform explicitly adds correlation between two descriptions to enable the estimation of one set from the other. In the designing of MDC using correlating transform, there are ambiguous problems that decoder must know statistics of original image because the statistical properties vary significantly among different blocks and image. Therefore, in the conventional method, prior to encoding an image, the blocks in each image are classified into several classes so that the transform coefficients of the blocks within same class have similar statistical properties. Within each class, the coefficient variances are calculated and ordered. Finally, each descriptions including classification information, variances of each coefficients and ordering information are sent to channel [3].

In this paper, we propose an enhanced MDC scheme using whitening transform, which is not necessary additional statistical information to decode image because the DCT coefficients to apply whitening transform to an image have uni-variance statistics.

This paper is organized as follows. First, we give a brief description of conventional correlating transform. Next, we show effects of the whitening transform and it's approximation. Finally, we then discuss experiments on rate-PSNR.

2. Correlating Transform

Multiple description coding of image is a technique that used to obtain two or more descriptions of an source image, which are transmitted over different channels to the decoder. These descriptions of the image support two levels of reconstruction quality. When all the descriptions are received, the decoder (central decoder) can reconstruct the image, with high fidelity. When only a subset of the description is received, the decoder (side decoder) can still reconstruct image of acceptable quality. In this paper, we only considered the case of two descriptions as shown at figure 1.

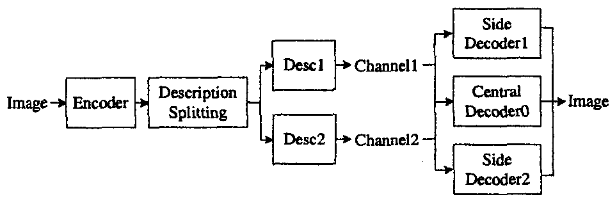

Figure 1. MDC using two descriptions

In the conventional MDC using correlating transform, shown at figure 2, two DCT coefficients of an image block are transformed by using linear orthogonal or bi-orthogonal matrix, and the two transformed coefficients are sent to channels separately.

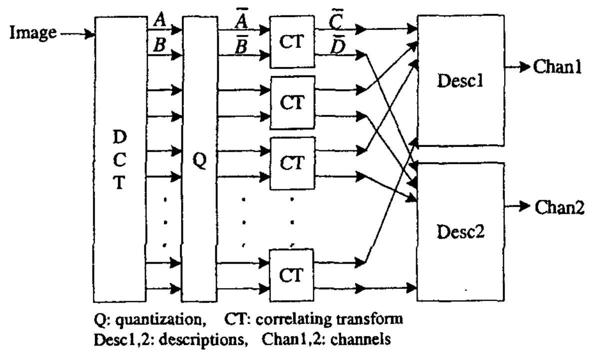

Figure 2. MDC using correlating transform

First, two DCT coefficients A and B are quantized with Q through equation (2). And then, the quantized value $\bar{A}$ and $\bar{B}$ are transformed into $\bar{C}$ and $\bar{D}$ by correlating transform $\boldsymbol{T}$ such as equation (1), (3) and (4):

$$\boldsymbol{T} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{1}$$

$$\bar{A} = \left[\frac{A}{Q}\right], \quad \bar{B} = \left[\frac{B}{Q}\right], \tag{2}$$

$$W = \bar{B} + \left[\frac{1+c}{d}\bar{A}\right], \tag{3}$$

$$\bar{D} = [dW] - \bar{A}, \quad \bar{C} = W - \left[\frac{1-b}{d}\bar{D}\right], \tag{4}$$

where [] implies rounding off operation. After entropy coding the transformed values $\bar{C}$ and $\bar{D}$, the two descriptions are sent to channel 1 and channel 2 respectively.

If decoder have received both descriptions, we can easily reconstruct image through following equations:

$$W = \bar{C} + \left[\frac{1-b}{d}\bar{D}\right], \tag{5}$$

$$\bar{A} = [dW], \quad \bar{B} = W - \left[\frac{1+c}{d}\bar{A}\right], \tag{6}$$

$$\hat{A} = \bar{A}Q, \quad \hat{B} = \bar{B}Q. \tag{7}$$

When decoder have only received one description, we can estimate lost description by linear estimation such as following equations and reconstruct image through above equations:

$$\tilde{D}(\bar{C}) = \frac{ac\sigma_{\bar{A}}^2 + bd\sigma_{\bar{B}}^2}{a^2\sigma_{\bar{A}}^2 + b^2\sigma_{\bar{B}}^2}\bar{C}, \tag{8}$$

$$\tilde{C}(\bar{D}) = \frac{ac\sigma_{\bar{A}}^2 + bd\sigma_{\bar{B}}^2}{c^2\sigma_{\bar{A}}^2 + d^2\sigma_{\bar{B}}^2}\bar{D}, \tag{9}$$

where, $\sigma_{\bar{A}}^2$ and $\sigma_{\bar{B}}^2$ are variances of $\bar{A}$ and $\bar{B}$ respectively.

3. Whitening Transform

3.1 Effects of Whitening Transform

In the conventional MDC using correlating transform, decoder must have knowledge of statistics of original image for reconstructing image properly as shown at equation (8) and (9). However, if we apply whitening transform to the conventional method, encoder does not need to send statistics to decoder anymore.

Given colored random variable x, we can diagonalize variance of x such as following equation (10)

$$\boldsymbol{C}_x = \boldsymbol{E}\boldsymbol{\Lambda}\boldsymbol{E}^H, \tag{10}$$

where $\boldsymbol{C}_x$ is variance of x, $\boldsymbol{E}^H$ is orthogonal linear transform and $\boldsymbol{\Lambda}$ is eigenvalue matrix. By following equation (11), we can transform colored random signal x to white random signal y,

$$y = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{E}^H x. \tag{11}$$

According to equation (12), when the whitening transform, $\boldsymbol{\Lambda}^{-1/2}\boldsymbol{E}^H$, is applied to colored random signal x, the y has uni-variance statistics,

$$\begin{aligned} \boldsymbol{C}_y &= (\boldsymbol{\Lambda}^{-1/2}\boldsymbol{E}^H)\boldsymbol{C}_x(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{E}^H)^H \\ &= \boldsymbol{\Lambda}^{-1/2}\boldsymbol{E}^H\boldsymbol{E}\boldsymbol{\Lambda}\boldsymbol{E}^H\boldsymbol{E}\boldsymbol{\Lambda}^{-1/2} = \boldsymbol{I}. \end{aligned} \tag{12}$$

Therefore, if we perform correlating transform not on x but on y such as figure 3, we don't need to send statistics of original image to decoder and result in improved coding efficiency and reconstruction quality. For example, let

$$\underbrace{\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}}_{z} = \boldsymbol{T} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}}_{y}, \tag{13}$$

where a, b, c and d are real numbers and y_1 and y_2 are independent Gaussian random variables with variances σ_1^2 and σ_2^2. In this case, $\sigma_1^2 = \sigma_2^2 = 1$ because y_1 and y_2 has uni-variance.

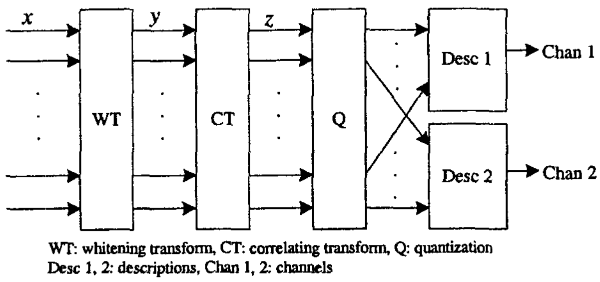

Figure 3. MDC using whitening transform

Now suppose that z is description of y (Quantization is ignored for now). Since z is jointly Gaussian, the conditional expectation of y with given either descriptions alone is a linear function as following equations:

$$\hat{y}^{(1)} = \frac{1}{a^2\sigma_1^2 + b^2\sigma_2^2}\begin{bmatrix} a\sigma_1^2 \\ b\sigma_2^2 \end{bmatrix} z_1 = \frac{1}{a^2+b^2}\begin{bmatrix} a \\ b \end{bmatrix} z_1 \tag{14}$$

$$\hat{y}^{(2)} = \frac{1}{c^2\sigma_1^2 + d^2\sigma_2^2} \begin{bmatrix} c\sigma_1^2 \\ d\sigma_2^2 \end{bmatrix} z_2 = \frac{1}{c^2 + d^2} \begin{bmatrix} c \\ d \end{bmatrix} z_2 \qquad (15)$$

3.2 Approximation

Unfortunately, while we can simplify the designing of correlated transform, we must send the eigenvalue matrix to decoder in order to decode image completely. It is not practical because the eigenvalue matrix of an image is too large to send. However, instead of theoretical approach, we can approximate the whitening transform through prediction technique. In the lossless image compression, we can easily find the approximation method such as linear prediction analysis(LPA)[4], prediction through average with down-sampling[5] and so on. If an image is subtracted by predicted image from original image, the residual image comes to white random signal that has small and same variance all over image. As shown in equation (14) and (15), if residual image has same variance all over image, we can except to send additional statistics from description in the MDC using correlating transform, although encoder sends prediction parameters.

We show the proposed method at figure 4. After

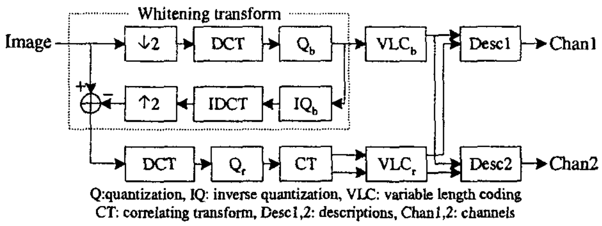

Figure 4. the proposed image coder

down-sampling an original image to quarter image, the quarter image is transformed with DCT, and the DCT coefficients are quantized with coarse non-uniform quantizer(Q_b). And after entropy coding (VLC_b) the quantized coefficients, the coded bit stream of two descriptions are sent to channel 1 and channel 2 respectively. The residual image subtracted the original image from the quarter image is transformed with DCT, and the DCT coefficients are quantized with fine uniform quantizer(Q_r). After applying correlating transform such as equation (13), two descriptions are sent to channels respectively.

4. Experimental Results

In this paper, we have made experiments about rate-PSNR of the conventional correlating transform and the proposed method. In the figures, CT represents correlating transform, WCT the proposed method, and BPP(bit per pixel) total rate of description 1 and description 2. For comparison, we also implemented a sample interleaving (SI) scheme, which split image into even/odd pixel images.

We used the 'lena' image, JPEG standard quantization table in the Q of the figure 2, and Q_b and uniform quantizer in the Q_r of figure 4.

The figure 5 shows comparison of PSNRs with respect to images reconstructed from both description. It can be seen that the proposed method (WCT) achieves a good trade-off between the coding efficiency and the reconstruction quality. The figure 6 shows image reconstructed at about 1.0BPP.

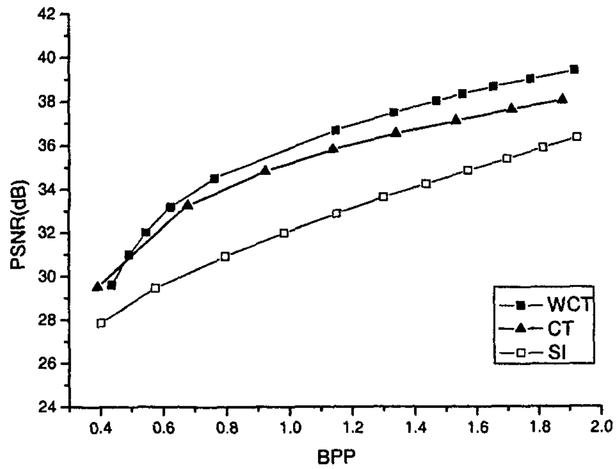

Figure 5. comparison of PSNRs of images reconstructed from both descriptions

Figure 6. images reconstructed from both descriptions: (a) CT: 1.02BPP, 35.27dB (b) WCT: 1.01BPP, 35.95dB

The WCT has a better coding efficiency than CT (the PSNR of images reconstructed from two descriptions in the WCT is about 0.7dB higher than that of CT at the 1.0 BPP), but the reconstruction quality of WCT is worse than CT slightly at lower part of 0.5BPP. The SI shows poor quality because of check-board artifact caused by sample interleaving and DC value mismatch.

When that decoder receives only one description, the proposed method obviously has a better coding efficiency than CT (the PSNR of image reconstructed from only one description in the WCT is about 1.8dB higher than that of CT at the 1.0 BPP)

The figure 7 shows comparison of PSNRs of images reconstructed from one description. Obviously, the

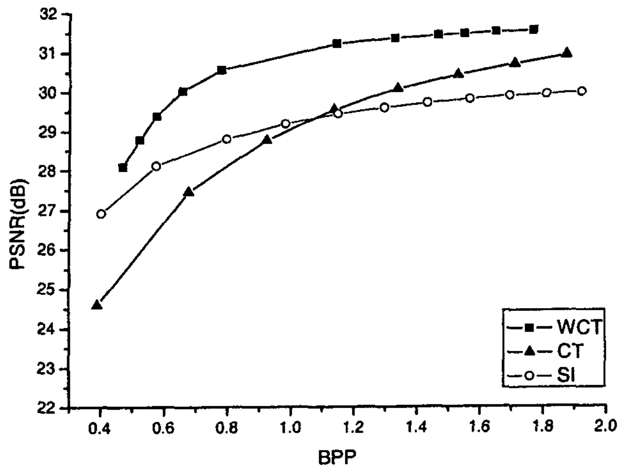

Figure 7. comparison of PSNRs of images reconstructed from one descriptions

Figure 8. images reconstructed from only one description: (a) CT: 1.02BPP, 29.12dB (b) WCT: 1.02BPP, 30.99dB

image reconstructed from one description in the conventional correlating transform is very poor at the low rate case. It has a severe checker-board artifact caused by mismatch of variance in each image block. The mismatch of variance may result in a inaccuracy coefficient estimation and frequently increase high frequency coefficient values. However, the proposed method does not have the artifacts because variance information of each blocks does not need in reconstruction.

5. Conclusions

In this paper we proposed a practical MDC method using correlating transform by the whitening transform that transforms colored random signal to white random signal. After DCT coefficients are transformed by whitening transform, the DCT coefficients have univariance property. Therefore, by using the method, we can remove statistical information necessary for reconstructing an image, and can reduce complexity and ambiguity of the conventional methods, in conclusion, we can achieve a good rate-quality trade off.

To use whitening transform, encoder must send the eigenvalue matrix evaluated at the transform decoder, which causes overhead on bit rate. Therefore, we approximated the whitening transform for practical usage. The approximated algorithm has not uni-variance, but same variance property.

Our experimental results show that the proposed method (WCT) achieves a good trade-off between the coding efficiency and the reconstruction quality. In the proposed method, the PSNR of images reconstructed is about 0.7dB higher than that of conventional method at the 1.0 BPP from both descriptions, and is about 1.8dB higher at the same rate from only one description.

References

[1] N.S. Jayant, "Subsampling of a DPCM speech channel to provide two 'self-contained' half-rate channels," *Bell Syst. Tech. J.*, vol.60, no.4, pp.501-509, April 1981.

[2] V.A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol.39, pp.821-834, May 1993.

[3] Y. Wang et al. "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Proc.*, vol.10, no.3, pp.351-366, March 2001.

[4] Maragos et al. "Two-Dimensional Linear Pre-diction and Its Application to Adaptive Predic-tive Coding of Images," *IEEE Trans. ASSP.*, vol32, no.6, pp.1213-1229, Dec. 1984.

[5] K. Knowlton. "Progressive transmission of grey-scale and binary pictures by simple, effi-cient and lossless encoding schemes," *Proceedings of the IEEE*, vol68, no.7, pp.885-896, 1980