Optimal and Generalized Multiple Descriptions Image Coding Transform in the Wavelet Domain

Bahi brahim, El hassane Ibn Elhaj,Driss Aboutajdine

Abstract—In this paper we propose a Multiple Description Image Coding(MDIC) scheme to generate two compressed and balanced rates descriptions in the wavelet domain (Daubechies biorthogonal (9, 7) wavelet) using pairwise correlating transform optimal and application method for Generalized Multiple Description Coding (GMDC) to image coding in the wavelet domain. The GMDC produces statistically correlated streams such that lost streams can be estimated from the received data. Our performance test shown that the proposed method gives more improvement and good quality of the reconstructed image when the wavelet coefficients are normalized by Gaussian Scale Mixture (GSM) model then the Gaussian one ,.

Keywords—Multiple description coding (MDC), gaussian scale mixture (GSM) model, joint source-channel coding, pairwise correlating transform, GMDCT.

I. Introduction

THE main problem encountered in transmitting visual information over heterogeneous packet switched networks is the rapid degradation in the reconstructed image quality due to packet loss [1]. If packet retransmission is not guaranteed as opposed to the TCP protocol [1], then we should think of another appropriate mean to receive meaningful data despite the loss of packets. This problem finds its natural solution in the so-called multiple description framework [1]. Recently, multiple description coding (MDC) has taken considerable attention as a method of communication over unreliable packet switched networks. MDC is a technique which can be considered a joint source-channel coding (JSC) code for erasure channels [4]. It can efficiently combat packet loss without any retransmission, thus not only guarantee the real-time communication, but also relieve the network congestion. In the MDC scheme, several representations of the source, called descriptions, are generated. The descriptions are designed in such a way that the quality of the received signal degrades gracefully with the increase in the number of descriptions that are lost. Also, the descriptions are designed so that the quality of the reconstructed image is dependent only on the number of received descriptions and not on which descriptions are actually received.

II. Multiple description coding

In this section, we describe two of the most popular multiple description image coding techniques are multiple description scalar quantization (MDSQ) [11] and pairwise

B.Bahi and D. Aboutajdine is with the Department of physics, GSCM-LRIT Laboratory, Faculty of sciences,University of Mohamed V,Avenue Ibn Battouta B.P.1014 RP, Rabat, Maroc
E. Ibn Elhaj Institut National des Postes et Télécommunications Av. Allal Al Fassi-Madinat Al Irfane-Rabat-Instituts Maroc

correlating transform (PCT) [3]. A relatively new approach is frame based multiple description (MD) image coding [14]. We briefly review these techniques before describing our proposed scheme. The fundamental idea behind PCT based multiple description coding is the introduction of controlled redundancy in the de-correlated source using a correlating transform. The correlation thus introduced is used to estimate lost coefficients from those received error-free. A popular DCT domain MD image coding scheme is presented in [3]. It assumes that DCT coefficients are uncorrelated and Gaussian distributed. Pairs of DCT coefficients are correlated using a rotation matrix and the resultant coefficients are treated as two descriptions. These descriptions are entropy coded and packetized before transmission over an erasure channel. The essence of MDSQ is the simultaneous generation of two quantization indices (instead of one) for every input. The two indices are generated such that either index provides an acceptable reconstruction level and a much higher quality reconstruction results from the use of both indices. In wavelet based MDSQ image coding, each sub-band is MD scalar quantized to generate two descriptions followed by entropy coding. One of the first MDC coder was designed by Vaishampayan [11] in which multiple description scalar quantizers were used in an extension of the old JPEG coder. Other methods for the design of MDC coders use correlation-inducing transforms. Wang et al. [14], [13] , [12] and [5] proposed applying a pairwise correlating transforms to introduce dependencies between two descriptions. Goyal et al. further generalized Wang's work to any number of descriptions, and coined the term generalized multiple description coding (GMDC) [4] and [10]. GMDC was later applied to image coding with correlating transforms [10]and [12]. Servetto has designed and implemented error-resilient data compression algorithms based on the use of wavelets and MD scalar quantizers [8]. Pereira [6] and Sumohana [9] have studied MDC techniques based on wavelet transform, but not considering correlating transforms in wavelet domain.

In this paper, we propose an error-resilient image communications scheme that exploits the properties of the GSM model and uses MDC techniques to achieve error resilience. We first derive the redundancy rate-distortion function for the pair-wise correlated GSM source. Finally, we simulate an image communication system that uses an optimal pairwise correlating transform to form multiple descriptions and we simulate an image communication system that uses the transform is implemented as parallel 2-by-2 transforms to form four descriptions and uses an erasure channel.

III. THE GAUSSIAN SCALE MIXTURE MODEL

In this section, we describe the model used to represent the statistics of natural images in the wavelet domain. It has been shown in [7]. That the statistics of wavelet coefficients fit the GSM model very accurately. A random vector Y is a GSM if it satisfies the relation $Y = z.U$ where z is a scalar random variable with $z \geq 0$ and $U \sim N(0; Q)$ is a Gaussian random vector and z, U are independent that normalized wavelet coefficients $U(Y/z)$ are jointly Gaussian. Where, z corresponds to local standard deviation of wavelet coefficients that scales the Gaussian random vector U. In [7] show log histograms of raw$(f_{Y/z}(Y/z))$ and normalized $(f_{Y/z}(Y/z))$ wavelet coefficients respectively. It is clear that normalizing the coefficients makes them closer to a Gaussian distribution.

IV. MD CODING WITH CORRELATING TRANSFORMS

In [7] description the redundancy rate-distortion function for the pair-wise correlated GSM source based orthogonal transform. In this paper using, optimal transforms that give balanced rates [10], [4] the redundancy rate-distortion function for the pair-wise correlated GSM source. We assume a two channel scenario with channel failure probabilities p_1 and p_2 respectively. Since Y is a GSM random variable, the pdf of the components of U i.e., u_1 and u_2 can be expressed as $f_{u_1/z}(u_1/z) \sim N(0; z^2\sigma_{u_1}^2)$ where z is the scalar random variable (which is assumed to be known). We derive the expression for redundancy ρ at a given two channel distortion D_0 needed to achieve a one-channel distortion D_1. A correlating transform T adds redundancy between transform coefficients and give balanced rates is applied to the uncorrelated source pair $U = [u_1; u_2]^T$ to generate $V = [v_1; v_2]^T$ i.e., $V = TU$. We evaluate

$$Cov(V) = E[VV^T]$$

as

$$Cov(V) = TE[UU^T]T^T \qquad (1)$$

where,

$$E[UU^T] = \begin{pmatrix} z^2\sigma_{u_1}^2 & 0 \\ 0 & z^2\sigma_{u_2}^2 \end{pmatrix}$$

Optimal Transforms that Give Balanced Rates:

$$T = \begin{pmatrix} a & 1/(2a) \\ -a & 1/(2a) \end{pmatrix}$$

Substituting T in (1)

$$Cov(V) = z^2.W$$

where,

$$W = \begin{pmatrix} a^2\sigma_{u_1}^2 + \frac{\sigma_{u_2}^2}{4a^2} & -a^2\sigma_{u_1}^2 + \frac{\sigma_{u_2}^2}{4a^2} \\ -a^2\sigma_{u_1}^2 + \frac{\sigma_{u_2}^2}{4a^2} & a^2\sigma_{u_1}^2 + \frac{\sigma_{u_2}^2}{4a^2} \end{pmatrix}$$

We now compute the redundancy (or excess rate) required to encode $V = [v_1; v_2]^T$ with respect to the rate required to encode $U = [u_1; u_2]^T$ at two-channel distortion D_0. Let $R_{u_1,u_2/z}$ denote the encode rate for U, and $R_{v_1; v_2/z}$ denote

the encode rate for V. Since U, V are conditionally Gaussian, we can use the rate-distortion function for Gaussian variables and arrive at

$$R_{u_1,u_2/z} = \frac{1}{2}\log\frac{z^4\sigma_{u_1}^2\sigma_{u_2}^2}{D_0} + K \qquad (2)$$

$$R_{v_1,v_2/z} = \frac{1}{2}\log\frac{z^4\sigma_{v_1}^2\sigma_{v_2}^2}{D_0} + K \qquad (3)$$

where, $\sigma_{v_1} = \sigma_{v_2}$

$$R_{v_1,v_2/z} = \frac{1}{2}\log\frac{z^4\sigma_{v_1}^4}{D_0} + K \qquad (4)$$

$$\rho = R_{v_1,v_2/z} - R_{u_1,u_2/z} = \frac{1}{2}\log\frac{z^4\sigma_{v_1}^4}{z^4\sigma_{u_1}^2\sigma_{u_2}^2} \qquad (5)$$

$$\rho = \log\frac{\sigma_{v_1}^2}{\sigma_{u_1}\sigma_{u_2}} \qquad (6)$$

Where K is a constant that accounts for entropy coding. We observe that since we are conditioning on z, z needs to be sent to the decoder without error for successful reconstruction. The rate needed to encode z is lower bounded by its entropy $h(z)$ [7]. Therefore the total excess rate needed is bounded by

$$\rho_{tot} \geq h(z) + \rho \qquad (7)$$

$$= h(z) + \log\frac{\sigma_{v_1}^2}{\sigma_{u_1}\sigma_{u_2}} \qquad (8)$$

$$= \log[T(\gamma)] + \gamma - (\gamma - 1)\psi(\gamma) + \log\frac{\sigma_{v_1}^2}{\sigma_{u_1}\sigma_{u_2}} \qquad (9)$$

Where $\psi(\gamma)$ is the digamma or psi function and $T(\gamma)$ is the Gamma function [7]. The single channel distortion D_1 is defined as the average single channel distortion per random variable [14], [7]. Express the single channel distortion D1 in terms of the excess rate in order to obtain the redundancy rate-distortion bound [10]. Assuming $p_1 = p_2 = 1/2$ neglecting effects of quantization and simplifying, we get,

$$D_1(\rho_{tot}) = D_2(\rho_{tot})$$

$$= E[z^2](\frac{1}{2}\sigma_{u_2}^2 + \frac{\sigma_{u_1}^2 - \sigma_{u_2}^2}{4.2^{2\rho}(2^{2\rho} + \sqrt{2^{4\rho} - 1})})$$

$$= E[z^2](\frac{\sigma_{u_1}^2 + \sigma_{u_2}^2}{4} - \frac{\sigma_{u_1}^2 - \sigma_{u_2}^2}{4}\sqrt{1 - 2^{-4\rho}}) \qquad (10)$$

where $\rho = (\rho_{tot} - h(z))$

The MMSE optimal linear estimates $\hat{U}^{(1)}$ and $\hat{U}^{(2)}$ from v_1 and v_2 respectively neglecting quantization noise are given by:

$$\hat{U}^{(1)} = \frac{2a}{4a^4\sigma_{u_1}^2 + \sigma_{u_2}^2}\begin{pmatrix} 2a^2\sigma_{u_1}^2 \\ \sigma_{u_2}^2 \end{pmatrix}v_1 \qquad (11)$$
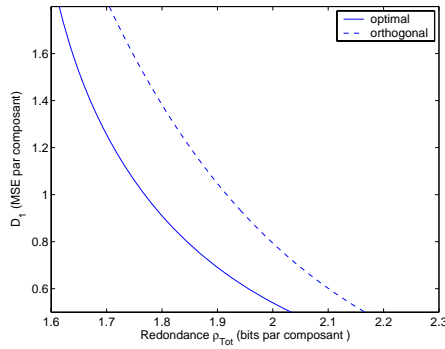
Fig. 1.   Theoretical redundancy rate-distortion plot

Fig. 2.   PSNR results for reconstruction from description 1 with GSM Model and Gaussian Assumption

and

$$\hat{U}^{(2)} = \frac{2a}{4a^4\sigma_{u_1}^2 + \sigma_{u_2}^2} \begin{pmatrix} -2a^2\sigma_{u_1}^2 \\ \sigma_{u_2}^2 \end{pmatrix} v_2 \qquad (12)$$

Figure.1 shows the plot of the redundancy rate-distortion function of the proposed multiple description coding system is better than the case redundancy rate-distortion function based orthogonal transform [7]. The figure is plotted for, $u_1 \sim N(0;1)$ and $u_2 \sim N(0;0.16)$. We see that the theoretical curve.

V.   SYSTEM IMPLEMENTATION

The algorithm proposed in [7] based on transform orthogonal and in [2] the algorithm based on optimal transforms that Give Balanced Rates basic premise of the proposed algorithm is the fact that normalizing wavelet coefficients by appropriate scale factors makes them Gaussian . Furthermore, it is assumed pairs of wavelet coefficients are uncorrelated and jointly Gaussian [7].

Implement the system is described below. The image is sub-band decomposed to three levels using the Daubechies biorthogonal (9,7) wavelet and decomposition applied on standard test image Zelda of size 512×512,shown in figure 2. The low-low sub-band is assumed to be transmitted without error. In order to normalize the wavelet coefficients, non-overlapping blocks of size 4x4 are formed and the covariance matrix Q and scaling factor estimate z are determined from these blocks. Once z is estimated, the coefficients are normalized by dividing them by z. The pairwise correlating transform is then applied to the normalized coefficients. The resulting correlated coef-ficients are transmitted over a pair of erasure channels with erasure probability $p_1 = p_2 = 1/2$. Since z has to be transmitted without error, it is sent over both channels. At the receiver, lost normalized coefficients are estimated using the MMSE estimator. The wavelet coefficient estimates are found by multiplying the normalized coefficient estimates with z.

VI.   THE GENERALIZED MULTIPLE DESCRIPTION IMAGE CODING

A block of n independent, Zero-mean variables with different variances are transformed to a block of transform coefficie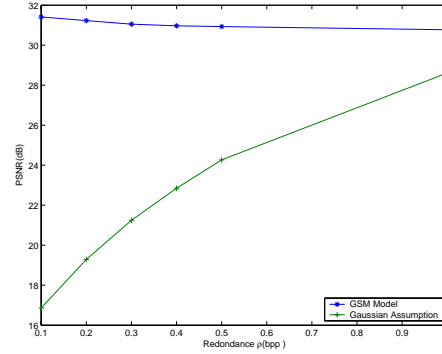nts in order to create a known statistical correlation between transform co- efficients. The transform coefficients from one block are distributed to different packets so in the case of a packet loss, the lost coefficients can be estimated from the received coefficients. The redundancy comes from the relative inefficiency of scalar entropy coding on correlated variables. This method is a generalization of the technique proposed in [13], [5] for two channels. The coding of a source vector x proceeds as follows:

1. x is quantized with a uniform scalar quantizer with step size $\triangle$: $x_{q_i} = [x_i]_\triangle$ where $[.]_\triangle$ rounding to the nearest Multiple of $\triangle$.

2. The vector $x_q = [x_{q_1}, x_{q_2}, ..., x_{q_n}]^T$ is transformed with an invertible, discrete transform $\hat{T} :\triangle Z^n \rightarrow \triangle Z^n, y = \hat{T}(x_q)$.

3. The components of y are independently entropy coded. The discrete transform is related to a continuous transform T through "lifting" Starting with a linear transform T with determinant one, the first step in deriving a discrete version is to factor into a product of upper and lower triangular matrices with unit diagonals $T = T_1 T_2...T_k$. The discrete version of the transform is then given by

$$\hat{T}(x_q) = [\hat{T_1}[\hat{T_2}...[\hat{T_k}x_q]_\triangle]_\triangle]_\triangle \qquad (13)$$

The lifting structure ensures that the inverse of $\hat{T}$ can be implemented by reversing the calculations in (1):

$$\hat{T}^{-1}(y) = [\hat{T_k}^{-1}...[\hat{T_2}^{-1}[\hat{T_1}^{-1}y]_\triangle]_\triangle]_\triangle \qquad (14)$$

When all the components of y are received, the reconstruction process is to (exactly) invert the transform to get the distortion is precisely the quantization error from Step 1.

If some components of y are lost, they are estimated from the received components using the statistical correlation introduced by the transform $\hat{T}$.

Recall that the variances of the components of x are $\sigma_1^2, \sigma_2^2, ..., \sigma_N^2$ and denote the correlation matrix of x by $R_x = diag(\sigma_1^2, \sigma_2^2, ..., \sigma_N^2)$. With fine quantization, the correlation matrix of y is $R_y = TR_xT^T$. By renumbering the variables if necessary, assume that $y_1, y_2, ..., y_{N-l}$ are received and $y_{N-l+1}, ..., y_N$ are lost. Partition y into "received" and "not received" portions as $y = (\tilde{y_r}, \tilde{y_{nr}})$ where $\tilde{y_r} = (y_1, y_2, ..., y_{N-l})$ and $\tilde{y_{nr}} = (y_{N-l+1}, ..., y_N)$. The

Minimum MSE (MMSE) estimate of x given $\tilde{y_r}$ is $E[x/\tilde{y_r}]$, which has a simple closed form because x is a jointly Gaussian vector. Using the linearity of the expectation operator gives the following sequence of calculations:

$$\hat{x} = E[x/\tilde{y_r}] = E[T^{-1}Tx/\tilde{y_r}] = T^{-1}E[Tx/\tilde{y_r}]$$

$$= T^{-1}E[\begin{pmatrix} \tilde{y_r} \\ \tilde{y_{nr}} \end{pmatrix}|\tilde{y_r}] = T^{-1}\begin{pmatrix} \tilde{y_r} \\ E[\tilde{y_{nr}}/\tilde{y_r}] \end{pmatrix} \quad (15)$$

If the correlation matrix of y is partitioned compatibly with the partitioning of y as

$$R_y = \begin{pmatrix} R_1 & B \\ B^T & R_2 \end{pmatrix} \quad (16)$$

then $\tilde{y_{nr}}/\tilde{y_r}$ is a Gaussian random variable with mean $B^T R^{-1}\tilde{y_r}$ and correlation matrix $R_2 - B^T R^{-1} B^T$ Thus and the reconstruction is

$$\hat{x} = T^{-1}\begin{pmatrix} I \\ B^T R^{-1} \end{pmatrix}\tilde{y_r} \quad (17)$$

The estimate $\hat{x}$ is then generated by inverting T. The optimal design of the transform $\hat{T}$ for Gaussian sources, where arbitrary (unequal, dependent) packet loss probabilities are allowed, is discussed in [12]. Here we consider the simpler case where packet losses are i.i.d. and the transform is implemented as parallel and/or cascade combinations of 2-by-2 transforms. It is shown in [12] that for coding a two-tuple source over two channels, where each is equally like to fail, it is sufficient to consider transforms of the form

$$\hat{T} = \begin{pmatrix} \alpha & 1/(2\alpha) \\ -\alpha & 1/(2\alpha) \end{pmatrix} \quad (18)$$

We use this as a building block to form larger transforms. The transform used in this paper is parallel 2x2 transforms(pairing)

$$T = \begin{pmatrix} T_\alpha & 0 \\ 0 & T_\beta \end{pmatrix} \quad (19)$$

If T_α used to transform variables with variances σ_1^2 and σ_2^2 and T_β is used to transform variables with variances σ_3^2 and σ_4^2 then the equal-slope condition implies that we should have

$$\beta^4 = \frac{\gamma(16\alpha^8\alpha_1^4 - \alpha_2^4\sqrt{\gamma^2(16\alpha^8\alpha_1^4 - \alpha_2^4)^2 + 64\alpha^8\sigma_3^4\sigma_4^4})}{32\alpha^4\sigma_3^4}$$

where

$$\gamma = \frac{\sigma_3^2\sigma_4^2(\sigma_3^2 - \sigma_4^2)}{\sigma_1^2\sigma_2^2(\sigma_1^2 - \sigma_2^2)} \quad (20)$$

VII. Image communication system

The generalized multiple descriptions image coding scheme proposed in [3] works with discrete cosine transform (DCT) coefficients. In this section, we propose a Generalized Multiple Descriptions Image Coding scheme in the wavelet domain. This to demonstrate the efficacy of the correlating transform method for image coding, we consider the case of coding for
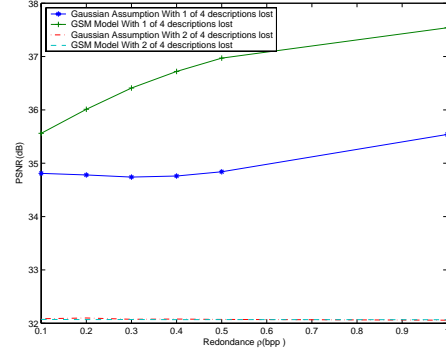

Fig. 3. PSNR results for reconstruction results with 1 of 4 descriptions lost and reconstruction results with 2 of 4 descriptions lost with GSM Model and Gaussian Assumption

four channels. This method is designed to operate on source vectors with uncorrelated components. We (approximately) obtain such a condition by forming vectors from coefficients wavelets. A straightforward application proceeds in the following steps:

The image is sub-band decomposed to three levels using the Daubechies biorthogonal (9,7) wavelet and decomposition applied on standard test images Zelda, Lena of size 512×512, shown in figure 3 of image Zelda and Figure 2 of image Lena . The low-low subband is assumed to be transmitted without error. In order to normalize the wavelet coefficients, non-overlapping blocks of size 4x4 are formed and the covariance matrix Q and scaling factor estimate z are determined from these blocks. Once z is estimated, the coefficients are normalized by dividing them by z. Correlating transforms are applied to each 4-tuple is then applied to the normalized coefficients. The resulting correlated coef-ficients are transmitted over a four of erasure channels. Since z has to be transmitted without error, it is sent over four channels. At the receiver, lost normalized coefficients are estimated using the MMSE estimator. The wavelet coefficient estimates are found by multiplying the normalized coefficient estimates with z.

the results of the above experiment on standard test images of size 512x512. We clearly see from the figure 2 and 3 that using the GSM model results in an improvement in the reconstructed image quality when compared to the case where wavelet coefficients are assumed to be Gaussian. GSM model performs better than the case where wavelet coefficients are assumed to be Gaussian. In the figure 3 quality of the reconstructed image depends on the number of descriptions arrived. In the figure 2 is summarized as follows: The correlating transform with GSM model method is better at low redundancies and the correlating transform with wavelet coefficients are assumed to be Gaussian method is better at high redundancies.

VIII. Conclusion

In this paper, we have considered a wavelet transform-based MDCT and GMDTC coder for coding still images. Our performance test shown that the proposed method gives more

improvement and good quality of the reconstructed image when the wavelet coefficients are normalized by Gaussian Scale Mixture (GSM) model.

a

REFERENCES

[1] Goyal, V. Multiple description coding: Compression meets the network. IEEE Signal Processing Magazine, 18(5), September 2001.

[2] Bahi, Brahim, ElHaj, E. I., and Aboutajdine, D. Optimal multiple description image coding transform in the wavelet domain. IEEE International Confrence on Signal Processing and Communications (ICSPC), 2007.

[3] Goyal, Vivek K., Kovacevic, J., Arean, R., and Vetterli, M. Multiple descriptions transform coding of images. Proc. IEEE Intl. Conf. on Image Processing, ICIP, pages 674678, Oct 1998.

[4] Kovacevic, J. and Goyal, V. Optimal multiple description transform coding of gaussian vectors. Proc.IEEE Data Compression Conf., pages 388 397, Mars 1998.

[5] Orchard, M., Y., W., V., V., and A., R. Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms. IEEE international conference on image proceedings, pages 60811, 1997.

[6] Pereira, M., Antonini, M., and Barlaud, M. Multiple description image and video coding for wireless channels. Session special EURASIP: Image Commun Special Issue Recent AdvWireless Video, 10(18):925945, 2003.

[7] S. S. Channappayya, R. W. H. J. and Bovik, A. C. Multiple description image coding using natural scene statistics. in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),Philadelphia, PA, USA., 2005.

[8] Servetto, S. Multiple description wavelet based image coding. IEEE Trans Image Process, (9):813826, 2000.

[9] Sumohana, S. C., Lee, J., Jr., R. W. H., and Bovik, A. C. Frame based multiple description image coding in the wavelet domain. in Proc. of IEEE International Conference on Image Processing (ICIP),Genova, Italy, 2005.

[10] V., K. G. and Jelena, K. Generalized multiple description coding with correlating transforms. IEEE Transactions on Information Theory, 47(6):21992224, Sept 2001.

[11] Vaishampayan, V. A. Design of multiple description scalar quantizers. IEEE Transactions on Information Theory, 39:821834, May 1993.

[12] Wang, Y., M., O., and A., R. Optimal pairwise correlating transforms for multiple description coding. IEEE international conference on image processing ICIP98, Chicago, IL, USA, 1998.

[13] Wang, Y., Orchard, M., and Arean, R. Multiple description image coding for noisy channels by pairing transform coefficients. IEEE first workshop MMSP97, Princeton, NJ, USA, pages 419 24, 1997.

[14] Wang, Y., Orchard, M., Vaishampayan, V., and Reibman, A. Multiple description coding using pairwise correlating transforms. IEEE Transactions on Image Processing, 10(3):351366, Mars 2001.