FREQUENCY-BASED RECONSTRUCTION OF MULTI-DESCRIPTION CODED JPEG2000 IMAGES

Hang Yu and Benjamin W. Wah

Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
{yuhang, wah}@manip.crhc.uiuc.edu

ABSTRACT

Multiple description coding (MDC), together with reconstruction, is a popular technique for concealing packets losses in a best-effort network like the Internet. In our previous study, we have developed a sample-based MDC scheme for subband images and an optimal reconstruction-based subband transform (ORB-ST) at senders that takes into account the reconstruction process at receivers. Due to the partitioning of a description into segments before coding the segments by the ORB-ST coder, our previous scheme has severe degradation in quality because, among other reasons, the coder cannot exploit redundancies across multiple segments in a description, especially at low bit rates. To address this issue, we propose in this paper to divide a description into subbands in the frequency domain, instead of segmenting a description in the sample domain. Based on the observation that the corresponding subbands across two descriptions after inverse transformation are correlated, we present a scheme that reconstructs a lost subband using the corresponding subband received in the other description. We show that our scheme is more effective than reconstruction in the full band; that it has similar quality and lower delays in low-loss scenarios and slightly degraded quality in high-loss scenarios, when compared to MDC, unsegmented, and JPEG2000-coded images sent by TCP; and that it has much better quality and similar delays when compared to MDC, segmented, and ORB-ST-coded images sent by UDP.

1. INTRODUCTION

Quality and delay are two important considerations in real-time transmissions of still images in the Internet. In general, users are willing to tolerate limited quality degradations in Web downloads of images, especially when the downloading delay is long. However, existing image coding algorithms are very sensitive to losses, and the quality of compressed images may be poor, even in low- or moderate-loss scenarios. As an illustration, Figure 1a shows *lena*, with 30.97 dB in quality, compressed at 0.125 bpp by JPEG2000 and transmitted in 4.01 seconds by TCP at 12 midnight on Dec. 25, 2002, in a round-trip route between UIUC and Thailand$_2$ (`www.kmitnb.ac.th`). In contrast, Figure 1b shows the same image packed in eight packets and transmitted by UDP in 0.71 seconds with 20.51 dB quality when two of the eight packets were lost. Obviously, the delay by TCP transmissions is too long, and the quality by UDP transmissions is not acceptable.

Existing loss-concealment schemes can be classified into receiver-based, sender-based and sender receiver-based.

Receiver-based techniques involve heuristic and computationally expensive post processing algorithms at receivers that do not always lead to high quality.

Sender-based techniques involve some preprocessing of an image before it is sent. Two common techniques are layered coding (LC) and multiple description coding (MDC). LC decomposes an image into descriptions of different importance, and descriptions of lower importance are discarded by routers when congestion happens. Such a strategy requires the support of priority transfers and extra bandwidth for unequal error protection. MDC, on the other hand, exploits redundancies in source data in order to create multiple descriptions of equal importance that can be sent without priority support. The redundancies exploited can be inherent in the data stream or introduced artificially. Examples of artificially introduced redundancies include correlating transforms [5], redundancies added during quantization [2], and FEC codes for error correction [1]. In MDC, a receiver reconstructs a lost description from descriptions received and produces images of acceptable quality.

Finally, *sender receiver-based techniques* involve senders and receivers that optimize jointly the encod-

Research supported by the Motorola Center for Communication, University of Illinois, Urbana.

Proc. IEEE Int'l Conf. on Computer Networks and Mobile Computing, 2003.

a) SDC by TCP (30.97 dB and 4.01 sec.) b) SDC by UDP (20.51 dB and 0.71 sec.) c) Proposed MDC by UDP (25.21 dB and 0.71 sec.)

Figure 1. Quality-delay trade-offs in round-trip transmissions of *lena* compressed at 0.125 bpp by JPEG2000 between UIUC and Thailand$_2$ (`www.kmitnb.ac.th`). In UDP transmissions, two out of the eight packets were lost.

Table 1. PSNR Quality in dB when sample-domain segmentation is applied to *lena* compressed at five bit rates

Segment Size	Bits Per Pixel (bpp)				
	2	1	0.5	0.25	0.125
No segmentation	43.91	40.07	37.16	34.08	30.97
256×256 segments	43.51	39.34	36.02	32.76	29.41
128×128 segments	42.55	37.93	33.92	29.42	24.14
64×64 segments	40.01	33.85	27.01	–	–

ing/reconstruction process. They either assume a channel model or require feedbacks from receivers, both of which are impractical when network losses are dynamic.

In our previous study on loss concealments of MDC subband-coded images sent by UDP, we have developed an optimal reconstruction-based subband transform (ORB-ST) [3, 4]. In this scheme, a sender interleaves an image into multiple descriptions in the sample domain, decomposes each description into segments in such a way that a coded segment fits in a UDP packet, compresses each segment using a subband coder, and sends each coded segment in a UDP packet to a receiver. By ignoring quantization errors, by assuming a subset of the descriptions were lost consistently, and by using linear interpolations to reconstruct loss pixels in the sample domain at receivers, ORB-ST was derived to minimize the total error of a reconstructed image at receivers. Note that the decomposition of a description into segments is needed in order to prevent errors propagating from one lost segment to another segment received.

Although ORB-ST has led to some improved quality as compared to that of the original subband transform (ST) in concealing packet losses, we have found significant quality degradations due to segmentation, especially at low bit rates. Table 1 demonstrates the effect of segmentation by showing the quality when the 512-by-512 *lena* is decomposed into, respectively, four 256-by-256 subimages, sixteen 128-by-128 subimages, and sixty-four 64-by-64 subimages, each compressed using JPEG2000, decompressed, and reassembled. Here, we do not use the tiling option in JPEG2000 but compress each subimage individually. The results for 64-by-64 subimages coded at 0.25 bpp and 0.125 bpp are not shown because each subimage has a size of 128 and 64 bytes, respectively, which are smaller than the basic JPEG2000 image header size.

The results clearly show that segmentation degrades coding efficiency in JPEG2000. This is not surprising since the coder is only able to remove redundancies within each subimage, whereas redundancies between subimages remain. In addition, segmentation leads to new image boundaries and more severe finite-window effects in FIR filtering. Further, the uniform allocation of a fixed bit budget across all subimages is suboptimal. These factors all contribute to performance degradations in segmentation.

In this paper, we present a new loss-concealment algorithm for MDC JPEG2000 images that divides a description into subbands in the frequency domain, instead of segmenting it in the sample domain, and that reconstructs a lost subband using the corresponding subband received in the other description. Since our scheme does not create new segment boundaries by segmenting correlated samples in a description before coding them, it leads to better coding efficiency. Further, since it uses UDP for transmission, it has smaller delays when compared to those of images sent by TCP.

2. INTERNET DELAY AND LOSS BEHAVIOR

We summarize in this section the delay and loss behavior of TCP and UDP transmissions [4]. In our experiments, we sent 2,000 packets of 512 bytes each by TCP and by UDP, at a rate of ten packets per second, to a remote server's echo port, and collected the sending time of each packet and the time it was bounced back.

Figure 2a plots the round-trip delays to three destinations that are typical of low-, medium- and high-loss connections. The results show that TCP delays can be one to two orders

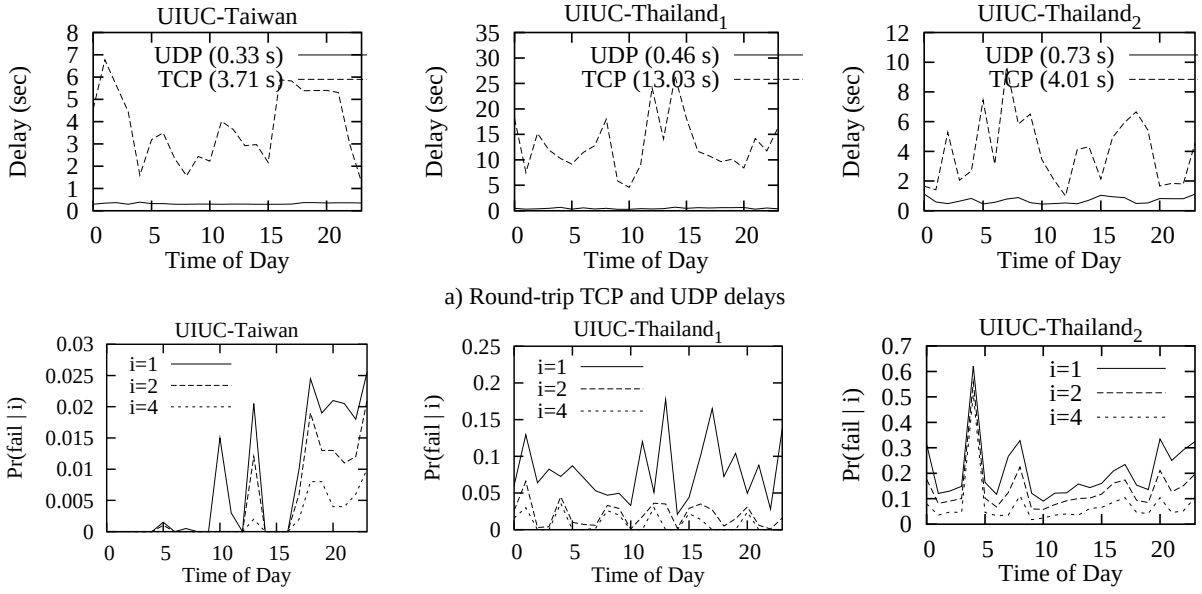


Figure 2. Round-trip TCP and UDP delays and UDP loss rates to three international sites collected in Dec 2002 (Taiwan: `pager.mit.com.tw`; Thailand$_1$: `www.iced.moe.go.th`; Thailand$_2$: `www.kmitnb.ac.th`).

longer than UDP delays.

Figure 2b plots the UDP loss rates, with and without loss concealments. Since packet losses in the Internet tend to be bursty, interleaving can be used to convert bursty losses to random ones. Using an interleaving factor of i, i blocks are put into one interleaving set, and two adjacent bits in a block are distributed to two adjacent packets. Consequently, all bursty losses of length smaller than i packets can be reconstructed by interpolations. For a burst length between i and $(2i-2)$, there is a chance that a bursty loss falls between two interleaving sets, and some packets in each set are received and are used to reconstruct the lost packets. Only when a bursty loss is longer than $(2i-2)$ would there be unrecoverable losses.

$Pr(fail|i)$, the probability of unrecoverable losses conditioned on interleaving factor i, can be computed by first grouping packets into interleaving sets of i packets each, and by checking the fraction of interleaving sets in a transmission trace with all i packets lost. Figure 2b show that an interleaving factor of two can reduce the level of unrecoverable losses to well below 5% in the UIUC-Taiwan and UIUC-Thailand$_1$ connections. However, an interleaving factor of four is needed to conceal most of the packet losses in the UIUC-Thailand$_2$ connection.

Interleaving, on the other hand, reduces the coding efficiency because it distributes correlated information to different packets, leaving behind information of lower correlation. Hence, wee need to find the minimum interleaving factor in order to keep the amount of unrecoverable bursty losses to below a maximum threshold.

3. FREQUENCY-BASED LOSS CONCEALMENTS

To overcome performance degradations caused by the segmentation of a description in the sample domain, we describe in this section an MDC algorithm that divides a description into subbands in the frequency domain and that reconstructs a lost subband by exploiting the correlation of the corresponding subband received in the other description. Since data from different subbands in a description is usually orthogonal and independent of each other, there is little redundancy between subbands and, consequently, little degradation in compression efficiency.

3.1. Correlation of subbands across descriptions

To prevent error propagation, we use JPEG2000 to segment the source data stream in the frequency domain into independently decodable units called *subbands*. By performing quantization and entropy coding on a subband, or a fraction thereof, and by adding some side information, the resulting *codeblock* can be decoded independently.

In order to reconstruct lost information in one description from that in another description, the information in different descriptions must be correlated. To evaluate the correlation of the corresponding subbands in different descriptions, we tested thirteen 512×512 8-bit gray-scale images divided into two groups. Group 1 includes *lena*, *zelda*, *goldhill*, *barbara*, *boat*, and *peppers*, which are about people, objects and sceneries. Group 2 includes *cloth*, *thumb*, *trick*, *smoke*, *pines*, *teeth*, and *grape*, which all contain textures.

Table 2. Correlation coefficient ρ between the corresponding inverse-transformed subbands of the two descriptions of *lena* and *teeth* and the average distortion per pixel (d_0^2) of three reconstruction methods in each subband.

Subband	Image *lena* in Group 1				Image *teeth* in Group 2			
	ρ	Average Distortion Per Pixel			ρ	Average Distortion Per Pixel		
		Duplication	Padding-0	Interpolation		Duplication	Padding-0	Interpolation
Unfiltered	0.972	64.60	1151.86	22.20	0.993	46.79	3609.21	18.36
LH1	0.369	31.44	25.07	20.54	0.367	19.75	15.43	13.28
HL1	0.370	2.12	1.68	1.96	0.528	3.16	3.39	2.97
HH1	0.119	3.44	1.96	2.08	0.174	4.30	2.61	2.67
LH2	0.820	12.76	42.83	14.64	0.827	8.77	25.52	9.37
HL2	0.851	1.06	2.97	1.11	0.901	1.57	7.97	1.59
HH2	0.674	2.43	3.73	2.41	0.728	2.37	4.37	2.36
LH3	0.954	6.50	72.78	8.43	0.956	3.55	40.92	4.67
HL3	0.954	0.41	4.50	0.46	0.984	0.61	19.34	0.65
HH3	0.914	1.07	6.28	1.38	0.935	1.32	10.06	1.65
LH4	0.992	2.41	160.93	1.65	0.989	1.29	61.83	1.77
HL4	0.981	0.21	5.79	0.23	0.996	0.27	38.59	0.31
HH4	0.977	0.47	10.80	0.63	0.984	0.59	18.79	0.81
LL4	0.999	2.07	811.08	4.01	0.999	0.87	3356.90	0.71

We first interleaved each image by two-way column interleaving and evaluated the correlation between the corresponding pixels in the two descriptions. We then coded each description using JPEG2000 in four levels and thirteen subbands, given that the default number of levels in JPEG2000 is five and that it is no longer the best since interleaving has been performed. The coefficients in a subband of a description were then inverse transformed into the sample domain, after setting other subband coefficients to zero. The resulting subimage is a filtered version of the original description.

Table 2 shows the correlation of the pixels in the corresponding subimages in the two descriptions for one image in Group 1 and another in Group 2. It shows that the correlations of the two subimages in lower subbands are high, and that their correlations in higher subbands are low. This is not surprising because lower subbands contain the basic image structure, whereas higher subbands contain details.

Table 2 also shows d_0^2, the average distortion per pixel, of three reconstruction methods, assuming one description is lost consistently. Note that $PSNR = 10\ \log_{10}(255^2/d_0^2)$ is only used to measure completely decoded images, whereas d_0^2 is used to evaluate partially filtered images. We have evaluated three reconstruction methods: duplicating a lost image by the corresponding filtered image in the description received, padding it by zeroes, and reconstruction by linear interpolation of the received filtered image. The results show that padding by zeroes may even perform better than duplication when the subimage to be reconstructed has low correlation to the one received. Note that the total distortion introduced in padding by zeroes in a subband is actually the total image energy in that subband.

The results show that a lost subband can be reconstructed from the corresponding subband of another description. Although the results show that linear interpolation is the best for low-frequency subbands and that padding by zeroes performs better in high-frequency subbands, duplication is preferable in practice for two reasons. First, duplication can be done in the frequency domain without an inverse wavelet transformation because duplication in the frequency domain is the same as duplication after the inverse transform. In contrast, the computational overhead of 100-200 ms for an inverse transform of a lost codeblock may be too high in real-time transmissions. Second, since interpolation is not a strictly linear process, the distortion introduced in each subband will not be orthogonal and does not add up to the overall distortion. As a result, minimizing the distortion in each subband does not guarantee the minimum overall distortion. In contrast, minimizing the distortion introduced by a linear process like duplication in each subband will be the same as minimizing the overall distortion.

3.2. Optimal Linear Reconstruction

To design a good reconstruction method, we first describe the optimal linear reconstruction of one random variable from another. Since the method may require image-specific information to be transmitted, we propose a more practical approach that uses parameters obtained by analyzing a class of benchmark images. Our experimental results demonstrate that the latter approach incurs relatively small errors as compared to the former approach.

Given two zero-mean random variables X and Y with the same variance σ_0^2, finding the optimal linear reconstruction for X from Y is equal to finding a that minimizes $E(X - aY)^2$. The solution, obtained by setting its derivative to 0, is $a = E(XY)/E(Y^2) = \rho_{XY}$, the correlation

coefficient between X and Y. When X and Y are highly correlated, ρ_{XY} is close to one and duplication ($a = 1$) is very effective. However, duplication does not work well when ρ_{XY} is not close to one. As ρ in Table 2 varies from one subband to another, we conclude that subband-specific reconstruction methods should be developed, instead of using a common reconstruction method for the full band.

Without loss of generality, consider two frequency bands: band 1 for the low band, and band 2 for the high band. (It is straightforward to generalize the result to the multi-band case.) Assume two variables X_0^e and X_0^o that represent the even and odd descriptions of an image, where $\mathbf{X}_0^e = \mathbf{X}_1^e + \mathbf{X}_2^e$ and $\mathbf{X}_0^o = \mathbf{X}_1^o + \mathbf{X}_2^o$. Further, assume X_i^e and X_i^o, $i = 0, 1, 2$, with zero mean and the same variance to represent, respectively, the filtered versions in band i; that is, $E((\mathbf{X}_i^o)) = E((\mathbf{X}_i^e)) = 0$, $E((\mathbf{X}_i^o)^2) = E((\mathbf{X}_i^e)^2) = \sigma_i^2$. The zero-mean assumption is justified in JPEG2000 because the coder first subtracts 127 from each pixel in an 8-bit image in order to remove its DC component. We have also tested statistically the assumption that the corresponding subbands of the two descriptions contain the same amount of energy. The analysis shows that the ratio of $E((\mathbf{X}_i^e)^2)$ to $E((\mathbf{X}_i^o)^2)$ falls into the interval $[1.01 \pm 0.05]$ with 99% confidence. Finally, we assume, for $i = 0, 1, 2$:

$$E(\mathbf{X}_i^o \mathbf{X}_i^e) = s_i^2, \quad (1)$$

$$E(\mathbf{X}_i^o \mathbf{X}_j^o) = E(\mathbf{X}_i^e \mathbf{X}_j^e) = E(\mathbf{X}_i^o \mathbf{X}_j^e) = 0, i \neq j \quad (2)$$

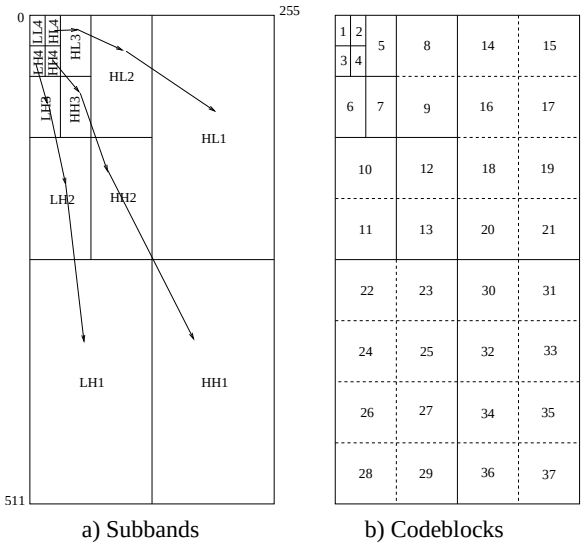
$$\rho_{X_i^o X_i^e} = \rho_i = s_i^2/\sigma_i^2. \quad (3)$$

Suppose X_0^o is lost and needs to be reconstructed from X_0^e. Let e_1 (*resp.* e_2) be the error of reconstruction in the full band (*resp.* separate subbands). We have:

$$\begin{aligned} e_1 &= E((\mathbf{X}_0^o - \rho_0 \mathbf{X}_0^e)^2) \\ &= E((\mathbf{X}_0^o)^2 - 2\rho_0 \mathbf{X}_0^o \mathbf{X}_0^e + \rho_0^2 (\mathbf{X}_0^e)^2) \\ &= \sigma_0^2 - 2(s_0^2/\sigma_0^2)s_0^2 + \sigma_0^2 (s_0^2/\sigma_0^2)^2 \\ &= (\sigma_0^2 - s_0^4/\sigma_0^2) \quad (4) \\ e_2 &= \sum_{i=1,2} (\sigma_i^2 - s_i^4/\sigma_i^2) \quad (5) \end{aligned}$$

$$\begin{aligned} e_1 - e_2 &= (\sigma_0^2 - s_0^4/\sigma_0^2) - \sum_{i=1,2} (\sigma_i^2 - s_i^4/\sigma_i^2) \\ &= s_1^4/\sigma_1^2 + s_2^4/\sigma_2^2 - s_0^4/\sigma_0^2 \\ &= s_1^4/\sigma_1^2 + s_2^4/\sigma_2^2 - (s_1^2 + s_2^2)^2/(\sigma_1^2 + \sigma_2^2) \\ &= (s_1^2\sigma_2^2 - s_2^2\sigma_1^2)^2/\sigma_1^2\sigma_2^2(\sigma_1^2 + \sigma_2^2) \geq 0. \end{aligned}$$

The above derivation shows that the total error of reconstruction in separate subbands can lead to smaller errors than reconstruction in the full band. Further, knowing the correlation of the subbands in the two descriptions can lead to better reconstruction errors than duplication.



a) Subbands b) Codeblocks

Figure 3. The left figure shows a four-level decomposed frequency-band structure of a 512×256 image. The right figure shows the structure when subbands are further divided into codeblocks

Although a linear reconstruction of individual subbands at receivers can outperform that of full band, it requires the correlations of all corresponding subbands to be sent to receivers. An alternative is for senders to analyze some benchmark images ahead of time and transmit a single set of parameters to receivers for reconstructing missing subbands.

Rewrite the correlation in subband i as follows:

$$\begin{aligned} \rho_i &= \frac{E(\mathbf{X}_i^o \mathbf{X}_i^e)}{\sqrt{E(\mathbf{X}_i^{o2}) E(\mathbf{X}_i^{e2})}} \\ &= \frac{E(\mathbf{X}_i^{e2}) + E(\mathbf{X}_i^{o2}) - d_i^2}{2\sigma_i^2} \\ &= 1 - 0.5 d_i^2/\sigma_i^2. \quad (6) \end{aligned}$$

Here $d_i^2 = E((\mathbf{X}_i^o - \mathbf{X}_i^e)^2)$ is the energy of the difference between these two variables in subband i. It can also be interpreted as the distortion in using duplication to reconstruct one variable from another.

According to (6), σ_i^2 and d_i^2 must be known in order to reconstruct a missing subband. Parameter σ_i^2 can be estimated from the subband received. However, in order to estimate d_i^2, it is necessary to know the model of energy distribution for d_0^2, where d_0 is expected to be random like.

To model the distribution of d_0^2, we show in Figure 3a a four-level decomposed frequency-band structure of a 512×256 image, whose LH (also HL and HH) subbands in the four levels are connected by arrows into a group. If the difference between a pair of corresponding subimages in the two descriptions is perfect white noise, then d_0^2 is expected to be uniformly distributed at every frequency level. That is, the ratio of the energy of two subbands just depends on the

Table 3. Results of k_i found by linear regression on each image, each group of images, and all the images together. In each case, the R^2 measure of linear regression is shown.

Image	HL Band		LH Band		HH Band		Image	HL Band		LH Band		HH Band	
	k_1	R_1^2	k_2	R_2^2	k_3	R_3^2		k_1	R_1^2	k_2	R_2^2	k_3	R_3^2
barbara	3.76	0.97	3.86	0.98	4.61	0.85	*cloth*	4.13	0.80	1.31	0.52	0.97	0.23
boat	3.95	0.99	2.28	0.99	2.54	0.88	*grape*	2.49	0.98	1.97	0.99	1.56	0.80
goldhill	2.88	0.99	2.80	0.99	2.70	0.99	*pines*	3.95	0.97	2.46	0.95	2.50	0.90
lena	2.36	0.99	2.16	0.99	2.04	0.99	*smoke*	2.54	0.99	2.31	0.99	1.97	0.99
peppers	1.82	0.98	2.57	0.96	2.08	0.99	*teeth*	3.79	0.95	2.53	0.87	2.03	0.54
zelda	2.07	0.98	2.27	0.98	1.92	0.82	*thumb*	3.59	0.68	1.75	0.19	1.40	0.02
Group 1	2.69	0.79	2.60	0.80	2.53	0.68	*trick*	3.79	0.99	1.42	0.86	2.05	0.96
Groups 1 & 2	3.05	0.80	2.20	0.68	2.05	0.40	**Group 2**	3.40	0.82	1.91	0.55	1.71	0.22

Table 4. Performance gain in dB of our four proposed reconstruction methods as compared to duplication, both under no quantization loss and with quantization. G_u is the gain when unified k_i's are used across all images; G_t, the gain when k_i's depend on image type; G_i, the gain when image-dependent parameters are used; and G_0, the gain when actual ρ_i's are used.

Image	Parameters				No Quantization				0.5 bpp				0.25 bpp				0.125 bpp			
	N_1	N_2	N_3	N_e	G_u	G_t	G_i	G_0	G_u	G_t	G_i	G_0	G_u	G_t	G_i	G_0	G_u	G_t	G_i	G_0
barbara	13	3	20	311	2.28	2.39	2.57	2.81	2.02	2.02	2.15	2.35	1.56	1.61	1.70	1.78	0.77	0.78	0.78	0.84
boat	3	2	8	102	1.44	1.32	1.42	1.46	1.09	1.00	1.08	1.11	0.65	0.67	0.62	0.68	0.28	0.28	0.28	0.29
goldhill	5	2	9	68	1.02	1.14	1.13	1.15	0.50	0.50	0.49	0.59	0.28	0.27	0.27	0.32	0.17	0.17	0.17	0.18
lena	10	2	11	65	0.90	0.98	0.98	0.98	0.64	0.69	0.70	0.70	0.37	0.42	0.44	0.46	0.16	0.17	0.17	0.18
peppers	15	2	16	79	0.86	0.96	1.09	1.15	0.64	0.64	0.66	0.76	0.45	0.44	0.39	0.52	0.25	0.25	0.25	0.28
zelda	7	2	11	28	1.13	1.25	1.25	1.28	0.70	0.76	0.77	0.78	0.45	0.49	0.53	0.55	0.20	0.20	0.20	0.22
cloth	3	4	12	212	0.26	0.21	0.09	0.49	0.17	0.08	0.00	0.38	0.06	0.06	0.06	0.12	0.06	0.06	0.06	0.11
grape	5	3	8	67	0.68	0.60	0.73	0.75	0.41	0.36	0.45	0.52	0.19	0.14	0.25	0.32	0.06	0.06	0.07	0.08
pines	4	2	5	166	1.06	1.05	1.05	1.08	0.48	0.42	0.41	0.50	0.23	0.23	0.23	0.27	0.07	0.07	0.07	0.07
smoke	3	2	5	47	0.92	0.71	1.02	1.02	0.19	0.19	0.19	0.26	0.06	0.06	0.06	0.06	0.03	0.03	0.03	0.03
teeth	3	3	7	127	0.73	0.69	0.67	0.77	0.48	0.44	0.43	0.51	0.14	0.14	0.13	0.22	0.04	0.04	0.04	0.04
thumb	15	3	14	196	0.05	-0.03	-0.03	0.49	0.48	0.11	0.11	0.39	0.14	0.16	0.16	0.25	0.05	0.06	0.06	0.10
trick	1	1	2	1724	1.48	1.61	1.63	1.65	0.79	0.76	0.76	0.83	0.22	0.22	0.22	0.23	0.06	0.06	0.07	0.08

area of the two subbands. In Figure 3a, this means $d_{LH1}^2 = 4d_{LH2}^2 = 16d_{LH3}^2 = 64d_{LH4}^2$ (similarly for the HH and LH subbands). This argument leads us to build an exponential model for the distribution of d_0^2.

Assuming $d_{LL4}^2 = b$, $d_{LH4}^2 = c_1 b$, $d_{HL4}^2 = c_2 b$, $d_{HH4}^2 = c_3 b$, we can model the remaining parameters as:

$$d_{LHi}^2 = c_1 k_1^{(4-i)}; \;\; d_{HLi}^2 = c_2 k_2^{(4-i)}; \;\; d_{HHi}^2 = c_3 k_3^{(4-i)}. \quad (7)$$

Here, c_1, c_2, c_3, b are image-dependent, and k_1, k_2, k_3 are parameters that can be generated from a group of images.

After taking the logarithm of both side of (7), we use linear regression to find k_i, first for each image, then for each group of images, and finally for all the images together. Table 3 shows that the k_i's found are mostly between two and four, thereby verifying our hypothesis that the energy of a subband depends on its area. We also show the R^2 measure to test the linearity of the data, where R^2 of one means that the data is perfectly linear, and R^2 of zero means no linearity in the data. The results show that, for most subband groups and most images, the linearity assumption holds. Only for *cloth* and *thumb*, which contain almost pure texture with very fine structures, the assumption does not hold. This is not surprising since pure texture has periodic patterns and the difference signal becomes periodic as well.

Based on the k_i's, we can derive d_0^2 as follows:

$$d_0^2 = b + \sum_{i=1}^{3} c_i b(1 + k_i + k_i^2 + k_i^3). \quad (8)$$

By transmitting the c_i's, d_0^2, and k_i's to receivers, a receiver can derive the d_i's using (7) and ρ_i using (6) and apply the reconstruction method outlined earlier to reconstruct a missing subband. Note that if image-independent or group-specific k_i's are used, they can be sent to receivers ahead of time. On the other hand, if image-specific k_i's were used, they must be sent with each image to receivers.

In the thirteen images evaluated, since c_2 is always the smallest and smaller than one, we transform the three c_i's into integers as follows: $N_1 = \lfloor c_1/c_2 + 0.5 \rfloor$, $N_2 = \lfloor 1/c_2 + 0.5 \rfloor$, and $N_3 = \lfloor c_3/c_2 + 0.5 \rfloor$. The total energy is also rounded to an integer $N_e = \lfloor d_0^2 + 0.5 \rfloor$. As shown in Table 4, N_1 and N_2 can be encoded in four bits, N_3 in six bits, and N_e in ten bits. Therefore, the overhead is as small as three bytes, an acceptable overhead for a 512×512 image that may require over 4 KB in a 1:64 compression ratio.

Table 5. A typical example of packing codeblocks into eight packets divided in two descriptions, each packet containing 536 bytes maximum. There are 37 codeblocks for a 512-by-256 image coded at 0.125 bpp with four-level decomposition and 64-by-64 subimages. Only codeblocks with non-zero bit allocation are shown.

Des.	Pac.	Codeblock (Size in Bytes)
0	0	Header (209), 0 (292), Unused (35)
	1	1 (222), 6 (88), 7 (151), 5 (54), Unused (21)
	2	4 (388), 17 (26), 18 (91), Unused (31)
	3	8 (290), 19 (67), 20 (26), 3 (77), 2 (55), Unused (21)
1	0	Header (209), 0 (287), 16 (32), Unused (8)
	1	1 (248), 6 (87), 7 (151), 5 (50), Unused (0)
	2	4 (381), 17 (75), 2 (57), Unused (23)
	3	8 (287), 18 (66), 15 (32), 19 (28), 3 (74) Unused (49)

4. EXPERIMENTAL RESULTS

Table 4 shows the results on testing our proposed reconstruction algorithms on the thirteen images in Table 3 with no quantization and with quantization under three bit rates. In both cases, our algorithms have obvious performance gains over simple duplication. However, the gain is less significant when the compression ratio is high because more high frequency parts will be removed by quantization, and our reconstruction algorithms resemble simple duplication on the low-frequency parts, whose correlations are close to one.

Before experiments on Internet transmissions can be carried out, we need to pack codeblocks into UDP packets, subject to the MTU limit of 536 bytes. Based on the codeblocks generated in a subband by JPEG2000, the encoder further decomposes a codeblock into smaller ones if it cannot fit in a packet. We further need to pack the header with the codeblock of the lowest frequency because the loss of either information would render it impossible to decode the image, and placing them together would minimize the probability of losing either. The other codeblocks are then packed in an arbitrary fashion as long as each fits in a UDP packet. Table 5 shows the packing of the 37 codeblocks of *lena* in two descriptions into eight packets.

Since reconstructions in a real implementation are done at the codeblock level instead of at the subband level and our algorithm uses the d_i^2's in the subband level, we propose to apportion the d_i^2's in a subband uniformly to its codeblocks. For instance, we divide the d_i^2 in subband LH1, HL1 and HH1 in Figure 3b by 8 in order to approximate the distortion of each codeblock in 14 to 37. We have tested the validity of this heuristic statistically and have found that the ratio of the actual distortion in one codeblock to that predicted falls in $[1.05 \pm 0.04]$ with 95% confidence.

Next, we present the results of our proposed algorithm based on losses captured in round-trip traffic traces between UIUC and three destinations: Taiwan, Thailand$_1$, and Thailand$_2$. We have used two-way MDC for the UIUC-Taiwan and UIUC-Thailand$_1$ connections due to their low loss rates, but have used four-way MDC for the UIUC-Thailand$_2$ connection. We implement four-way interleaving by recursive two-way interleaving and maintain the four parameters N_1, N_2, N_3, and N_e for each two-way interleaving. As a result, three sets of such parameters are sent to receivers in order for them to carry out reconstructions.

Table 6 shows the quality of the three reconstruction methods averaged over a 24-hour period at three bit rates for *lena* and *smoke*. (The results on other images are similar and are not shown due to space limitations.) Figure 4 further shows the quality of the three methods, each averaged over forty runs at the beginning of each hour. We compare the results against the reference results obtained by sending the eight packets in two-way (*resp.* four-way) MDC by TCP and no segmentation. The results show that our proposed reconstruction method can achieve almost the same quality as that of the reference results in low-loss scenarios and much better quality than that of segmented MDC images coded by JPEG2000 with ORB-ST. Further, in contrast to the high delays incurred by TCP, our proposed method incurs the same delays as that of UDP.

Figure 1 in Section 1 illustrates *lena* obtained by our proposed loss-concealment scheme and transmitted in 0.71 second. Figure 5 further shows the quality of *smoke* compressed at 0.25 bpp in a trace-driven experiment on the UIUC-Thailand$_1$ connection. The image has little subjective difference with that transmitted by TCP using SDC but incurs almost thirty times less delay.

5. REFERENCES

[1] E. A. R. Alexander E. Mohr and R. E. Ladner. Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction. *IEEE Journal on Selected Areas in Communications*, 18(6):819–828, June 2000.

[2] S. D. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt. Multiple-description wavelet based image coding. In *Proc. IEEE Int'l Conf. on image processing*, pages 659–663, 1998.

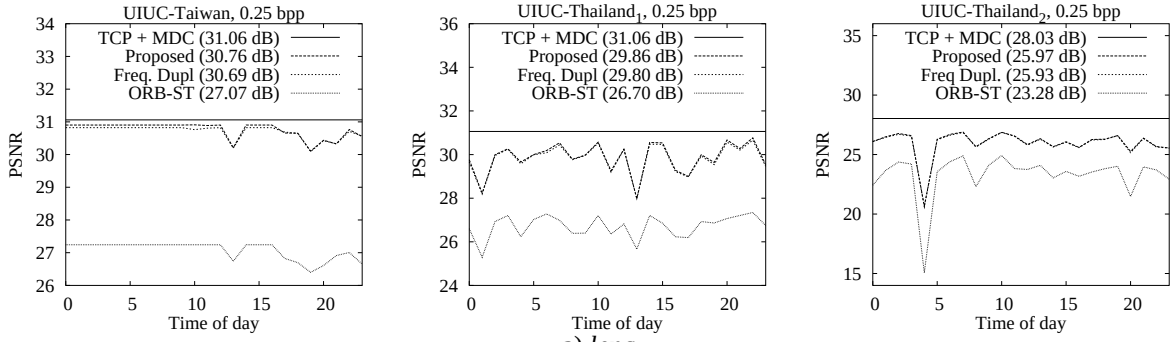
[3] X. Su and B. W. Wah. Reconstruction-based subband image coding for UDP transmissions over the Internet. *J. of VLSI Signal Processing*, pages 29–48, May-June 2003.

[4] B. W. Wah and X. Su. Loss concealments of subband coded images for real-time transmissions in the Internet. In *Proc. Int'l Conf. on Multimedia and Expo*, pages 449–452. IEEE, Aug. 2002.

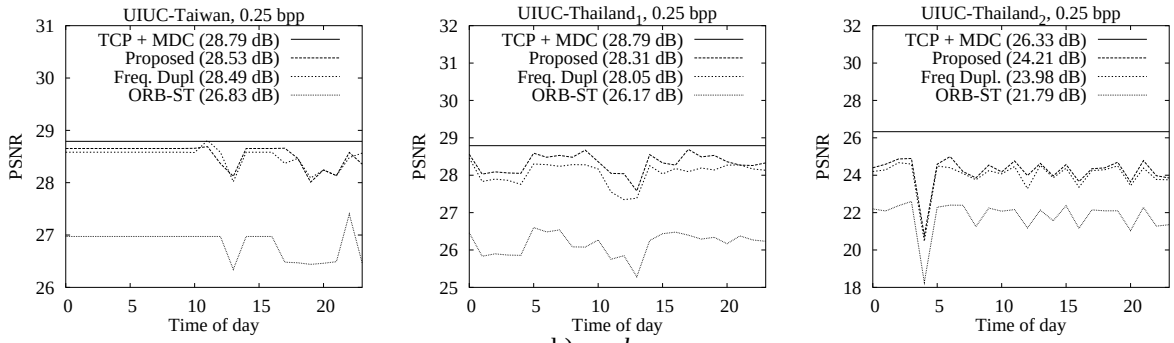
[5] Y. Wang, M. T. Orchard, and A. R. Reibman. Optimal pairwise correlating transforms for multiple description coding. In *Proc. IEEE Int'l Conf. on image processing*, pages 679–683, 1998.

Table 6. Average quality in dB of two reconstructed images by three reconstruction methods (proposed, frequency-based duplication, and interpolations of segmented MDC images coded by JPEG2000 with ORB-ST) when compared to the reference method (MDC with TCP and no segmentation)

Image	bpp	UIUC-Taiwan (2-way MDC)				UIUC-Thailand$_1$ (2-way MDC)				UIUC-Thailand$_2$ (4-way MDC)			
		TCP	Reconstruction Method			TCP	Reconstruction Method			TCP	Reconstruction Method		
		+ MDC	Proposed	Freq Dupl.	ORB-ST	+ MDC	Proposed	Freq Dupl.	ORB-ST	+ MDC	Proposed	Freq Dupl.	ORB-ST
lena	0.5	34.95	**33.42**	33.32	28.86	34.95	**31.79**	31.70	28.27	30.62	**26.32**	26.27	23.96
	0.25	31.06	**30.76**	30.69	27.07	31.06	**29.86**	29.80	26.70	28.03	**25.97**	25.93	23.28
	0.125	27.78	**27.75**	27.73	25.40	27.78	**27.34**	27.31	25.12	25.22	**24.79**	24.76	22.30
smoke	0.5	31.16	**30.04**	29.95	28.23	31.16	**29.12**	29.03	27.63	28.64	**26.75**	26.69	23.07
	0.25	28.79	**28.53**	28.49	26.83	28.79	**28.31**	28.05	26.17	26.33	**24.11**	23.98	21.79
	0.125	26.46	**26.40**	26.25	24.01	26.46	**26.01**	25.95	23.78	24.06	**22.90**	22.81	20.15



a) *lena*



b) *smoke*

Figure 4. Hourly PSNR of the three reconstruction methods as compared to that of MDC with TCP and no segmentation for two benchmark images.

a) SDC by TCP (30.96 dB and 13.03 s) b) SDC by UDP (22.03 dB and 0.46 s) c) Proposed MDC by UDP (28.72 dB and 0.46 s)

Figure 5. Quality-delay trade-offs in round-trip transmissions of *smoke* compressed at 0.25 bpp by JPEG2000 between UIUC and Thailand$_1$ (`www.iced.moe.go.th`). In UDP transmissions, five out of the sixteen packets were lost.