FRAME BASED MULTIPLE DESCRIPTION IMAGE CODING IN THE WAVELET DOMAIN

Sumohana S. Channappayya, Joonsoo Lee, Robert W. Heath Jr., Alan C. Bovik

The University of Texas at Austin,
Department of Electrical and Computer Engineering,
1 University Station C0803 Austin, TX 78712-0240 USA
{sumohana, jslee509, rheath, bovik}@ece.utexas.edu

ABSTRACT

Multiple description codes generated by quantized frame expansions have been shown to perform well on erasure channels when compared to traditional channel codes. In this paper we propose a multiple description image coding scheme in the wavelet domain using quantized frame expansions. We form zerotrees from wavelet coefficients and apply a tight frame operator to the zerotrees. We then group appropriate expansions to form packets and evaluate the performance of the scheme over an erasure channel. We compare the performance of the proposed scheme with a conventional channel coding scheme.

1. INTRODUCTION

Error resilient image coding has received a lot of attention in the recent past. This trend can be attributed to advances made in capture devices that has in turn led to a phenomenal increase in the amount of digital image content generated and shared over packet based networks. Packet based networks generally tend to be lossy due to several factors such as buffer overflows, transitions from a high capacity network to one with lower capacity, etc. Since compressed image data is particularly sensitive to bit errors and packet errors, it is important to design error-resilient communication schemes that mitigate the effects of bit errors and packet losses.

Several techniques such as forward error correction, selective retransmission, and multiple description image coding (MDIC) have been proposed to solve this problem. Of these techniques, MDIC schemes perform particularly well over lossy packet networks. MDIC schemes can be broadly classified into multiple description scalar quantization based, pairwise correlating transform based, and frame based. Frame based MDIC was shown to be comparable to traditional channel coding schemes by Goyal [1].

Image analysis and processing using wavelet transforms has become extremely popular. Tasks like image compression, de-noising, communication, and security have greatly benefited from the use of multiscale representations such as wavelet transform. For example, state-of-the-art image compression algorithms such as embedded zerotree wavelet (EZW) based compression, set partitioning in hierarchical trees (SPIHT), and JPEG2000 operate in the wavelet domain. Therefore, we consider it natural to explore frame based MDIC in the wavelet domain.

In this paper, we propose an error-resilient image communication scheme that exploits the erasure resilience properties of quantized frame expansions [2]. There is a previous work done to design this scheme, but it was reported in the discrete cosine transform (DCT) domain. The novelty of our approach is that we operate in the wavelet domain and thus we deal with the challenges presented by the wavelet domain. We first design a packetization scheme that uses zerotrees as a natural unit to form packets. We then design a frame operator that is optimal in the mean squared error sense and compare its performance with conventional channel coding. We also use a pre-designed frame operator and study its error resilience capabilities.

2. FRAMES

A set of vectors $\Phi = \{\varphi_i\}_{i\in I}$ in a Hilbert space $\mathbb{H}$ form a frame if for every $x \in \mathbb{H}, (\ x \neq 0)$,

$$0 \ < \ A||x||^2 \ \leq \ \sum_{i\in I} |\langle x, \varphi_i\rangle|^2 \ \leq \ B||x||^2, \qquad (1)$$

where I is the index set and the constants $A,\ B$ are called frame bounds where $A > 0, A \leq B$ [3]. A frame is *tight* if $A = B$ and *uniform* if $||\varphi_i|| = 1$. We deal exclusively with *uniform tight frames* in the N-dimensional real Hilbert space $\mathbb{R}^N$. Frame expansions offer a number of useful properties such as stable reconstruction and resilience to erasures of expansions.

In this paper, we are interested in the robustness of frame expansions to quantization noise and erasures. We briefly discuss the class of frames that exhibit these properties. The robustness of frame expansions to quantization noise can be attributed to the stable reconstruction property of frames (resulting from the restrictions placed on frame bounds A and B). It has been shown in [4] that a class of frames called *harmonic tight frames* provide resilience to erasures. If $F = \{\varphi_i\}_{i=1}^{M}$ is a harmonic tight frame where $\varphi_i \in \mathbb{R}^N$, then any subset of F containing at least N of the component frame vectors φ_i also forms a frame. Harmonic frames are generated using powers of the M^{th} roots of unity [4]. We use real harmonic tight frames in this paper.

These useful properties are being exploited in several practical applications such as space time coding for multiple input multiple output systems, multiple description coding, quantizer design, etc. Frame expansions have been used as joint source channel codes for communication over erasure channels [5]. Such joint source channel codes provide a more graceful degradation in distortion as the channel deteriorates when compared to traditional block codes. We discuss the application of frame expansions to multiple description image coding (in the wavelet domain) in the following sections.

3. FRAME BASED MULTIPLE DESCRIPTION CODING OF IMAGES

Two of the most popular multiple description image coding techniques are multiple description scalar quantization (MDSQ) [6] and pairwise correlating transform (PCT) [7]. A relatively new approach is frame based MD image coding [1]. 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 [7]. 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.

A frame based approach to MD image coding (in the DCT domain) was introduced in [1]. A uniform tight frame operator designed using the procedure outlined in [5] is used. This frame operator is applied to DCT coefficients and the resulting coefficients are encoded using a JPEG encoder. The encoded bitstream is then transmitted over an erasure channel. The performance of this system is shown to be better than that of a traditional channel coding scheme when the erasure rate is high. The fundamental difference between frame based error protection and channel coding is the position of the quantizer. The frame operator is applied before quantization while channel coding is done after quantization.

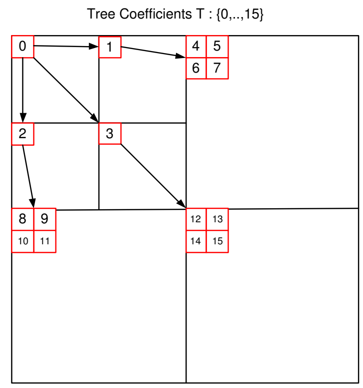

Fig. 1. Zerotree formation in the proposed MD image coding scheme

4. PROPOSED SCHEME

The frame based MDIC scheme proposed in [1] works with discrete cosine transform (DCT) coefficients. In this section, we propose a frame based MDIC scheme in the wavelet domain. The algorithm proposed in [1] applies a frame operator to vectors formed by groups of DCT coefficients and encodes the resultant frame expansions. An important point to note is the absence of an explicit packetization scheme in [1]. We propose a zerotree based packetization scheme for frame expansions of wavelet coefficients. Zerotrees lend themselves well to packetization as demonstrated in [8]. The zerotree structure used in our scheme is shown in Fig. 1. Zerotrees are formed for every coefficient in the low-low (LL) sub-band. Each LL coefficient has three children in high-low (HL), low-high (LH), and high-high (HH) sub-bands at the same decomposition level. Each child in the HL, LH, and HH sub-band has four children each in the HL, LH, and HH sub-band at the next decomposition level.

The packetization scheme is shown in Fig. 2. Once the zerotrees are formed, we apply a tight frame operator of size $M \times N$ to the zerotrees. To do so, we form vectors of length $N \times 1$ from the zerotree coefficients. An $M \times 1$ vector results from after frame expansion. We repeat this procedure for all the vectors that result from a zerotree and for all zerotrees in a given image. We collect all the inner

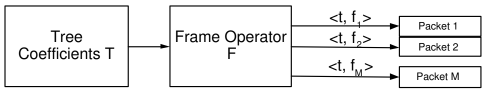

Fig. 2. Packetization procedure in the proposed MD image coding scheme

product elements corresponding to a given frame vector in the same packet. We then quantize and entropy code these packets before transmission.

We illustrate the scheme with an example. Suppose $F = \{\varphi_i\}_{i=1}^M$, where $\varphi_i \in \mathbb{R}^N$, i.e., F is a matrix of size $M \times N$. Further, if the coefficients of a zerotree form a matrix of size $N \times P$, then the result of the frame expansion is a matrix of size $M \times P$. The elements of each row (of size $1 \times P$) are composed into a packet followed by quantization and entropy coding. The motivation to form packets in this method is two-fold. We form packets from zerotrees to minimize perceptual distortion. If we encoded sub-band by sub-band, erasures in the LL sub-band would result in severe distortions. Using the zerotree approach, we avoid catastrophic errors by distributing important (LL) coefficients across packets. Further, we group the elements of the same row into one packet as these elements correspond to the inner product with the same frame vector (say φ_k). This way, the loss of a packet corresponds to the deletion of that frame vector from the frame operator ($F \backslash \varphi_k$). The source vector is reconstructed using the inverse frame operator corresponding to $F \backslash \varphi_k$. This construction is identical to the construction used in the theoretical analysis provided in [4, 5].

5. IMPLEMENTATION DETAILS

The primary goal of this work is to evaluate the performance of frame based error protection of images (in the wavelet domain) relative to conventional channel coding techniques over erasure channels. We propose to evaluate the performance at two popular coding rates viz., $\frac{1}{2}$ and $\frac{2}{3}$. We use two frame operators that provide the required levels of error protection. We first design a frame operator F of size 4 $\times$ 2 and compare its performance with a block code of size (4, 2). We then use a frame operator G of size 6 $\times$ 4 and compare it with a (6, 4) block code.

In the first case, we focus on the design of a frame operator which minimizes distortion caused by quantization and erasures. To form F, we use Theorem 2.7 in [4]. It states that for the case $M = 4, N = 2$, a tight frame operator can be designed as the union of two orthonormal bases Φ, Ψ in $\mathbb{R}^2$. The design variable is the angle $\alpha(\Phi, \Psi)$ between these two bases that minimizes distortion. The overall average distortion is defined as $D = \sum_{E=0}^{M} p_E MSE_E$, with p_E being the probability of E erasures and MSE_E being the error over E erasures. Since it is shown that the error over more than $(M - N)$ erasures is independent of the frame operator in [9], the overall distortion function used for optimization is

$$\begin{aligned} D(\alpha) = (1-p_e)^4 MSE_0 + \\ (1-p_e)^3 p_e (MSE_1^{\{1\}} + \ldots + MSE_1^{\{4\}}) + \\ (1-p_e)^2 p_e^2 (MSE_2^{\{1,2\}} + \ldots + MSE_2^{\{3,4\}}) \end{aligned} \quad (2)$$

with

$$\mathbf{F} = \begin{bmatrix} cos(0) & sin(0) \\ cos(\pi/2) & sin(\pi/2) \\ cos(\alpha) & sin(\alpha) \\ cos(\alpha + \pi/2) & sin(\alpha + \pi/2) \end{bmatrix}$$

and erasure probability $p_e = 0.2$.

In the second case, we focus on the performance comparison between the frame based scheme and the channel coding scheme since we have larger M in this case. We use the tight frame design mentioned in [4] to form $G = \{\varphi_i\}_{i=1}^M$,

$$\begin{aligned} \varphi_{k+1} = [cos\frac{k\pi}{M}, cos\frac{3k\pi}{M}, \ldots, cos\frac{(N-1)k\pi}{M}, \\ sin\frac{k\pi}{M}, sin\frac{3k\pi}{M}, \ldots, sin\frac{(N-1)k\pi}{M}]^T, \end{aligned}$$

with $M = 6$ and $N = 4$. We compare the performance of both schemes as the number of erasures changes from 0 to $(M - 1)$. We use a uniform quantizer with the same step size in both schemes to make the comparison fair to both schemes.

6. RESULTS

We used the 128$\times$128 8 bpp grayscale Lena image. First, we found the optimal angle $\alpha(\Phi, \Psi) = -3\pi/4$. Thus the optimal frame operator for the $M = 4, N = 2$ case is

$$\mathbf{F} = \begin{bmatrix} cos(0) & sin(0) \\ cos(\pi/2) & sin(\pi/2) \\ cos(-3\pi/4) & sin(-3\pi/4) \\ cos(-1\pi/4) & sin(-1\pi/4) \end{bmatrix}.$$

We performed the optimization of (2) using Matlab's optimization toolbox and verified the result using exhaustive search (Fig. 3). Next, the performance results of the frame based MDIC scheme and the channel coding scheme for the $M = 6, N = 4$ case are shown in Fig. 4. From the figure, it is apparent that the frame based scheme is better than the channel coding scheme. Also, the performance of the frame based scheme degrades gracefully while the channel coding scheme shows an abrupt degradation after it passes its error

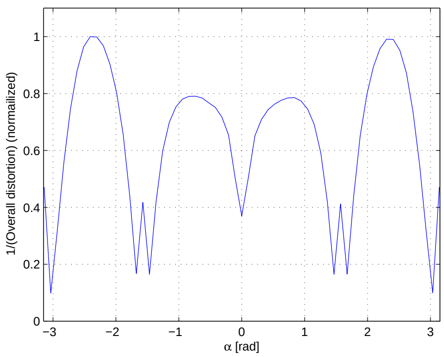

Fig. 3. Normalized overall distortion with $\alpha \in [-\pi, \pi]$

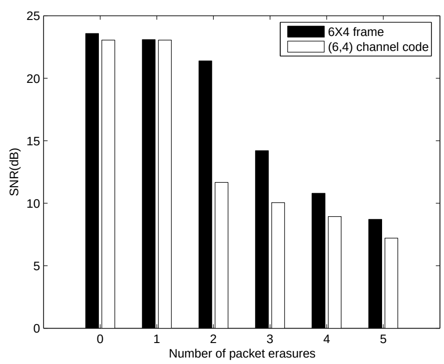

Fig. 4. PSNR plots of the frame based scheme and the channel coding scheme with both BPP around 3.05

correcting capability, which is 1 erasure. (we used Reed-Solomon code whose error correcting capability is $\frac{M-N}{2}$) We can conclude that the frame based MDIC performs better than the traditional channel coding with the same quantization noise and the same erasures. Thus we successfully demonstrated the first-ever frame based MDIC scheme in the wavelet domain.

7. CONCLUSION

We have shown that multiple description codes for images (in the wavelet domain) generated from quantized frame expansions perform well when compared to traditional channel codes. We have demonstrated a zerotree based packetization scheme designed to take advantage of the robustness of quantized frame expansions to erasures. Also, an optimal tight frame operator was designed to minimize mean squared error resulting from erasures. We plan to find a scheme to design an optimal tight frame operator in every case. In this paper, we assumed wavelet coefficients to be Gaussian distributed. We plan to repeat the analysis using a better model for the source such as the Gaussian scale mixture (GSM) model and more appropriate perceptual distortion metrics.

8. REFERENCES

[1] Vivek K. Goyal, *Beyond Traditional Transform Coding*, Ph.D. Thesis, University Of California, Berkeley, 1998.

[2] Ingrid Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[3] Ingrid Daubechies, “The wavelet transform, time-frequency localization and signal analysis,” *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, Sept. 1990.

[4] Vivek K. Goyal, Jelena Kovačević, and Jonathan A. Kelner, “Quantized frame expansions with erasures,” *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 203–233, 2001.

[5] Vivek K. Goyal, Jelena Kovačević, and Martin Vetterli, “Quantized frame expansions as source-channel codes for erasure channels,” in *IEEE Data Compression Conference, 1999. DCC ’99.*, 1999, pp. 326–335.

[6] Sergio D. Servetto, Kannan Ramchandran, Vinay A. Vaishampayan, and Klara Nahrstedt, “Multiple Description Wavelet Based Image Coding,” *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 813 – 826, May 2000.

[7] Yao Wang, Michael T. Orchard, Vinay Vaishampayan, and Amy R. Reibman, “Multiple description coding using pairwise correlating transforms,” *IEEE Transactions on Image Processing*, vol. 10, no. 3, pp. 351–366, Mar. 2001.

[8] Jon K. Rogers and Pamela C. Cosman, “Robust Wavelet Zerotree Image Compression with Fixed-length Packetization,” in *IEEE Data Compression Conference*, 1998, pp. 418–427.

[9] Aurélie C. Lozano, Jelena Kovačević, and Mike Andrews, “Quantized frame expansions in a wireless environment,” in *IEEE Data Compression Conference, 2002. DCC ’02.*, 2002, pp. 232–241.