3 WADDY, D.S., LU, P., CHEN, L., and BAO, X.: 'Fast state of polarization changes in aerial fiber under different climatic conditions', *IEEE Photonics Technol. Lett.*, 2001, **13**, (9), pp. 1035–1037

4 CAPONI, R., RIPOSATI, B., and SCHIANO, M.: 'Wdm design issues with highly correlated pmd spectra of buried optical cables'. In *OFC Technical Digest*, Anaheim, California, USA, 2002(Optical Fiber Conference) pp. 453–455

Source/channel coding of still images using lapped transforms and block classification

A.L.A. da Cunha, W.A. Finamore and E.A.B. da Silva

A novel scheme for joint source/channel coding of still images is proposed. By using efficient lapped transforms, channel-optimised robust quantisers and classification methods it is shown that significant improvements over traditional source/channel coding of images can be obtained while keeping the complexity low.

Introduction: Joint source/channel coding is an efficient approach for coding and transmitting images over noisy channels. It is envisioned that better performance with less complexity and reduced delay can be achieved by coding the source and channel jointly rather than separately. Yet, for images transmitted over binary symmetric channels (BSC), one of the best known results is obtained with a *tandem* framework, that is, compression of large size blocks followed by efficient forward error correction [1]. Such schemes are however subject to an inherent probability of incomplete decoding. They are, in addition, designed for a fixed channel and perform poorly under channel mismatch conditions. To date no joint source/channel scheme with good performance has been found and it remains unclear whether it is possible to obtain similar performance to traditional techniques with the joint source/channel approach.

Classical examples of joint source/channel coding schemes are the channel-optimised quantisers in which the encoder dispenses with the use of channel codes and protection is attained by trading-off empty encoding regions in $\mathbb{R}^n$ and quantisation performance. One can seek channel optimisation with both the Lloyd-Max type quantizers [2] or with Trellis-coded quantisers [3]. For Gaussian memoryless sources these quantisers can attain better MSE performance, for moderate block length, than their tandem counterparts [2] albeit severely perturbed by annoying impulsive noise artifacts.

Several schemes have been proposed to tackle the problem of transmitting compressed images over noisy channels [1,4–7]. In [6] a scheme that uses subband transforms achieves good performance by scrambling the DFT phase of the subband coefficients in such a manner that the Laplacian-like coefficients are transformed into near-Gaussian coefficients prior to channel-optimised scalar quantisation (COSQ). Note that the source distribution reshaping increases quantisation performance since, as pointed out in [6], COSQ, for generalised Gaussian distributions, performs better for larger shape parameters. Also phase scrambling spreads the impulsive noise thus drastically reducing the perceptual effect. An extension of [6] to channel-optimised trellis-coded quantisers provided further performance improvement at the expense of the extra complexity brought by the TCQ [7].

Following the same approach of [6] we propose some modifications seeking to improve the overall performance of joint source channel coding methods. To begin with we replaced wavelet decomposition by the lapped transform (LT). The block-based nature of LT allows for the use of efficient block classification strategies coupled with a steepest descent bit allocation method. It was in fact shown that significant gains can be obtained over the traditional scheme in [6] without increasing complexity.

Algorithm: Fig. 1 shows the proposed scheme building blocks. The image is first segmented into 8 × 8 blocks, each to be lapped transformed. Let $G_i(k,\ell)$ $0 \le k \le 7$, $0 \le \ell \le 7$ be the set of coefficients corresponding to the ith image block. The ith set of 64 coefficients so obtained are classified according to their block classification gain g_i (AC energy squared) expressed as

$$g_i^2 = \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} G_i^2(k, l) - G_i^2(0, 0). \tag{1}$$

Classification uses the equal mean-normalised standard deviation (EMNSD) criterion presented in [8]. Guided by the indication in [5], that two classes provide a good classification gain versus overhead tradeoff, we have chosen to use two classes only in this work.

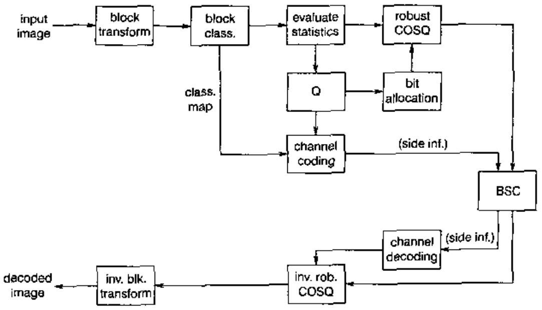

Fig. 1 *Proposed scheme building blocks*

After classification the 64 variances (one for each subband in each class) are computed. These are quantised with 16 bits each. Together with the classification map and the bit allocation matrices (one for each class), these constitute the data that is transmitted as side information. Since they represent just a small amount of information, a simple code can protect it, just as in [6]. After normalisation, the subbands are robust quantised. The robust quantiser works by first scrambling the subband coefficients and then scalar quantising each sequence of coefficients. The scrambling is accomplished by adding a reference pseudo-noise sequence to the phase of the DFT of the subband coefficients [6].

Decoding is simply the reverse process: after the side information is decoded, the coefficients obtained from the SQ indices are phase-descrambled and de-normalised. The reconstructed image is thus obtained after the LT synthesis stage.

Bit allocation is crucial to ensure good performance. We have used an algorithm similar to that presented in [9], based on a steepest descent method. The algorithm attempts to find a set of rates $\mathcal{R} = \{r_1, r_2, \ldots, r_N\}$ such that distortion is minimised while still maintaining the overall bit rate $\bar{r}$. With $(r_i, d_i(r_i))$ denoting the rate-distortion pair associated with the ith string of coefficients, the algorithm can be summarised as follows:

```
1. Set $k = \bar{r}N$; Set $r_i = 0$, $i = 1, \ldots, N$
2. Set $k = k - 1$; find $i_k$ satisfying $\Delta_{i_k}(r_{i_k}) = \max_{i=\{1,\ldots,N\}} d_i(r_i) - d_i(r_i + 1)$
3. Set $r_{i_k} = r_{i_k} + 1$. If $k = 0$ stop; else go to step 2.
```

$d_i(r_i)$ is the distortion produced by the transmission of the quantised coefficients (assumed Gaussian) over the BSC. The algorithm is fast and yields near-optimal allocation.

Results: The proposed scheme has be analysed by simulating the transmission of Lena and Goldhill 512 × 512 pixel images over a BSC. Several of the LTs discussed in [10] have been investigated—the results presented were obtained with the 8 × 16 generalised lapped bi-orthogonal transform (GBLT), selected for its good reconstruction performance. All the results presented are averages taken over 10 simulation runs. Fig. 2 shows the performance of the proposed scheme, at a rate of 1 bpp, for several channel cross probabilities. To highlight the effectiveness of variant I of the novel scheme (i.e. the one with classification) its performance is displayed against the results obtained for variation II (without classification) as well as that of the A-RQ scheme of [6] and of the robust CO-TCQ (Lena image only) of [7]. As can be seen, a performance better than that in [6], and equivalent to that of the CO-TCQ method of [7], has been achieved with the lapped transforms together with the bit allocation strategy (variant I scheme). If in addition, classification under the EMNSD criterion is incorporated (variant II), gains up to 2.2 dB can be obtained. In fact, the curves show that the proposed method outperforms those in [6] and [7] by a large margin—behaviour observed also at lower rates.

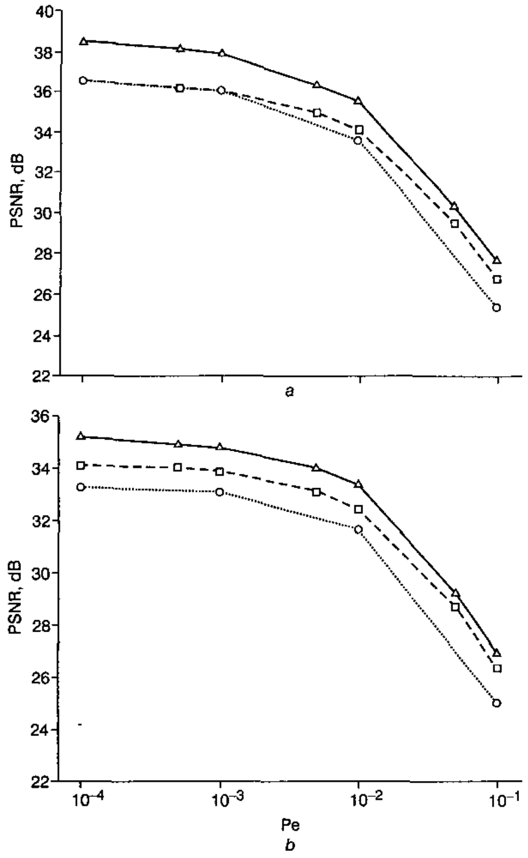

Fig. 2 *Proposed scheme performance for various BER at 1 bpp*

—△— proposed
--□-- proposed - no class.
···○··· A-RQ
a Lena
b Goldhill

Conclusion: We have proposed a novel scheme for joint source/channel coding of still images over BSCs. We have shown that with simple classification and bit allocation strategies, significant improvements can be achieved, thus narrowing the gap between tandem schemes such as [1] and purely joint source/channel coding approaches. In [4] a natural extension of [7] using a classification based on coefficient significance is presented. For moderate bit error probability their results are similar to the results in this Letter. However, the method in [4] still suffers from bad bit allocation when the channel is very noisy, thus producing worse results than our scheme (by around 1.6 dB for $Pe = 0.1$).

© IEE 2002
17 April 2002
Electronics Letters Online No: 20020759
DOI: 10.1049/el:20020759

A.L.A. da Cunha and W.A. Finamore (*Center for Telecommunication Studies, Pontifical Catholic University of Rio, Rua M. S. Vicente 255, Rio de Janeiro, Brazil*)

E-mail: weiler@cetuc.puc-rio.br

E.A.B. da Silva (*Universidade Federal do Rio de Janeiro, PEE/COPPE/DEL/EE, Rio de Janeiro, RJ - Brazil*)

References

1 SHERWOOD, P.G., and ZEGER, K.: 'Progressive image coding on noisy channels'. Proc. IEEE Data Compress. Conf., March 1997, pp. 72–81
2 FARVARDIN, N.: 'A study of vector quantization for noisy channels', *IEEE Tran Inf. Theory*, July 1990, **36**, (4), pp. 799–809
3 WANG, MIN, and FISCHER, T.R.: 'Trellis-coded quantization designed for noisy channels', *IEEE Trans. Inf. Theory*, November 1994, **40**, (5), pp. 1792–1802
4 ABOUSLEMAN, G.P.: 'Adaptive image coding with robust channel-optimized trellis-coded quantization'. Proceedings, 29th Applied Imagery Pattern Recognition Workshop, 2000, 1999, pp. 109–114
5 CAI, JIANFEI, and WEN CHEN, CHANG: 'Uniform threshold tcq with block classification for image transmission over noisy channel', *IEEE Trans. Circuits Syst. Video Technol.*, January 2001, **11**, (1), pp. 105–110
6 CHEN, QING, and FISCHER, T.R.: 'Image coding using robust quantization for noisy digital transmission', *IEEE Trans. Image Process.*, April 1998, **7**, (4), pp. 496–505
7 LAM, TUYET-TRANG, ABOUSLEMAN, G.P., and LINA, L.J.: 'Image coding with robust channel-optimized trellis-coded quantization', *IEEE JSAC*, 2000, **18**, (6), pp. 940–951
8 JOSHI, R., JAFARKHANI, H., KASNER, J., FISCHER, T., FARVARDIN, N., MARCELLIN, M., and BAMBERGER, R.: 'Comparison of different methods of classification in subband coding of images', *IEEE Trans. Image Process*, 1997, **6**, (11), pp. 1473–1486
9 VAISHAMPAYAN, V.A., and FARVARDIN, N.: 'Optimal block cosine transform image coding for noisy channels', *IEEE Trans. Commun.*, March 1990, **38**, (3), pp. 327–336
10 DE QUEIROZ, R.L., and TRAN, T.D.: 'Lapped transforms for image compression' *in* RAO, K.R., YIP, P. (Eds.): The Handbook on Transforms and Data Compression' (CRC Press, Oct. 2000), Chap. 5

MPEG-4 video object-based rate allocation with variable temporal rates

Jeong-Woo Lee, A. Vetro, Yao Wang and Yo-Sung Ho

In object-based coding, bit allocation is performed at the object level and temporal rates of different objects may vary. The proposed algorithm deals with these two issues when coding multiple video objects (MVOs). The proposed algorithm is able to successfully achieve the target bit rate, effectively code arbitrarily shaped MVOs with different temporal rates, and maintain a stable buffer level.

Introduction: The problem that we consider in this Letter is the object-based rate-distortion (R-D) encoding of multiple video objects (MVOs) for MPEG-4 video coding. In a recent paper by Vetro *et al.* [1], models that estimate the distortion for coded as well as non-coded frames have been proposed. The frame-based optimisation process determines if it is better to code more frames with lower quality or fewer frames with higher quality. However, this algorithm is not directly applicable to the coding of MVOs. In [2], an algorithm that considered the trade-off in spatial and temporal quality for coding MVOs was presented; however, it did not address the possibility of coding the objects with varying temporal rates, i.e. every object in the current time is either coded or all objects are skipped. In this Letter we propose a framework that supports object-based coding with different temporal rates, aiming to improve the coding efficiency of an object-based coder.

Object-based coding with variable frameskip: When each video object (VO) has a different frameskip, we define the constraints on the rate and buffer. Let M denote the set of VOs, and L denote the complete set of time indices. $M(l)$ denotes the set of coded VO at time index l $(t = t_l)$. The constraint on the rate can then be written as

$$\sum_{l\in L}\sum_{j\in M(l)} R_j(t = t_l) \le R_{budget} \qquad (1)$$

where $R_j(t = t_l)$ is the rates used for jth VO at time index l. (1) essentially says that the total rates for the coded VOs at all time instants within the specified time interval must be less than the calculated bit rate budget over that time interval.

To ensure that buffer overflow and underflow are avoided at every coded time, we have a set of buffer constraints which are given by