

Visual Communication via Trellis Coding and Transmission Energy Allocation

Fady I. Alajaji, *Member, IEEE*, Saud A. Al-Semari, *Member, IEEE*, and Philippe Burlina

Abstract—An unequal error protection approach for the reliable communication of visual information over additive white Gaussian noise channels is proposed and studied. This method relies on a bandwidth-efficient coded modulation scheme that employs selective channel coding and transmission energy allocation in conjunction with sequence maximum *a posteriori* soft-decision detection. Experimental results indicate that this scheme exhibits graceful performance degradation as the channel conditions deteriorate and provides substantial objective and subjective improvements over uncoded and equal-error protection systems. Coding gains of up to 4 dB in E_b/N_0 are achieved.

Index Terms—Image/video communication, MAP detection, transmission energy allocation, trellis-coded modulation, unequal error protection.

I. INTRODUCTION

THE SOURCE and channel coding functions of a communication system are usually designed independently of one another. This is justified by Shannon's separation principle [13], which indicates that no performance loss is suffered if the two functions are thus partitioned. However, Shannon's theorem is an asymptotic result that permits unlimited delay and complexity; given a constraint on complexity/delay, joint source-channel coding may outperform separately designed pairs (e.g., [1]–[8], [10]–[12], [14], [18]).

In this work, we investigate the robustness of a bandwidth-efficient joint source-channel coding system for the communication of visual information (compressed gray-level images or video intraframes) in the presence of channel noise. An unequal error protection (UEP) scheme is proposed for transmitting discrete cosine transform (DCT)-compressed images over an additive white Gaussian noise (AWGN) used in conjunction with coherent M -ary phase-shift keying (PSK) modulation. More specifically, it consists of a trellis-coded modulation (TCM) scheme that uses selective transmission energy allocation (TEA) to the DCT coefficients. It also employs

Paper approved by J. Huber, the Editor for Coding and Coded Modulation of the IEEE Communications Society. Manuscript received May 25, 1998; revised January 8, 1999 and April 12, 1999. The work of F. I. Alajaji was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant OGP0183645. This paper was presented in part at the IEEE International Symposium on Information Theory, Ulm, Germany, June 1997.

F. I. Alajaji is with the Department of Mathematics and Statistics and the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada K7L 3N6.

S. A. Al-Semari is with the Electrical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.

P. Burlina is with the Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 USA.

Publisher Item Identifier S 0090-6778(99)08917-5.

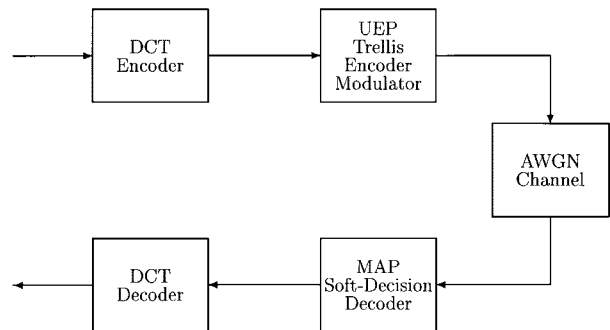


Fig. 1. Block diagram of the proposed UEP trellis coding system.

a sequence maximum *a posteriori* (MAP) detection scheme that exploits both the channel soft decision information and the statistical image characteristics. Comparisons of our scheme with uncoded systems that do not exploit the image residual redundancy as well as with equal error protection systems for identical overall throughputs (in pixels per signaling period) indicate substantial objective and subjective gains.

In previous related work, Fuldseth and Ramstad employed the concept of power allocation and multilevel modulation in the design of a channel-optimized vector quantization system for video transmission [4], [5]. In [6] and [7], Gadkari and Rose developed optimal energy allocation strategies for the transmission of compressed data over hard-decision binary phase-shift keying (BPSK) modulated AWGN channels. MAP source/channel decoding has recently received increased attention. It has been successfully applied in the contexts of image coding over memoryless channels (e.g., [18], [11], [12]) as well as channels with memory (e.g., [2]) and speech communication (e.g., [1], [3]).

II. UEP TRELLIS CODING SCHEME

The block diagram of the proposed system is shown in Fig. 1. We first briefly describe the image coding scheme.

A. Image Compression Model

Our source coding scheme retains all but the run-length and entropy coding¹ component of standard image and video codes, such as JPEG or MPEG [15]. The image is compressed as follows. It is subdivided in 8 by 8 blocks, and for each

¹A system that uses variable length coding (VLC) will break down in the presence of high channel noise (which is the case in our study). Our approach is to compensate for the higher source coding rate that results from not utilizing VLC by exploiting the residual redundancy of the quantized DCT parameters in the decoding scheme.

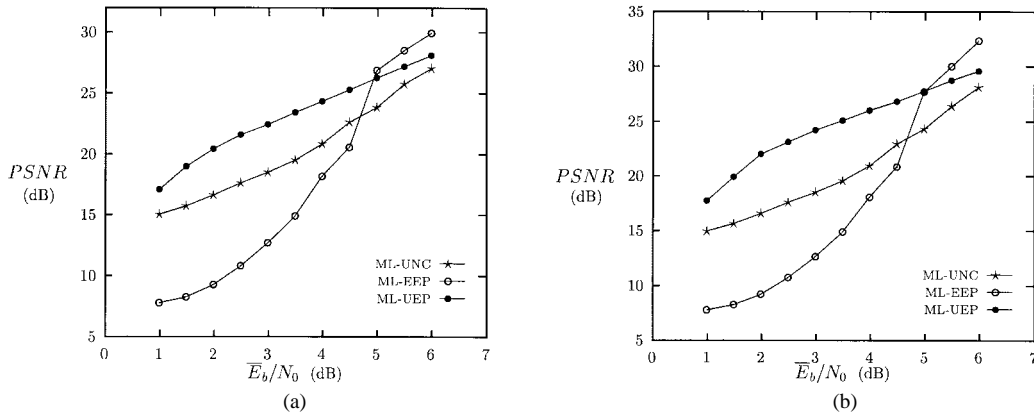


Fig. 2. ML-UNC, ML-EEP, and ML-UEP systems over AWGN channel with throughput R ; PSNR versus \bar{E}_b/N_0 of decoded Lena. Results are averaged over 30 experiments. (a) $R = 2.21$ pixels/T; (b) $R = 1.33$ pixels/T.

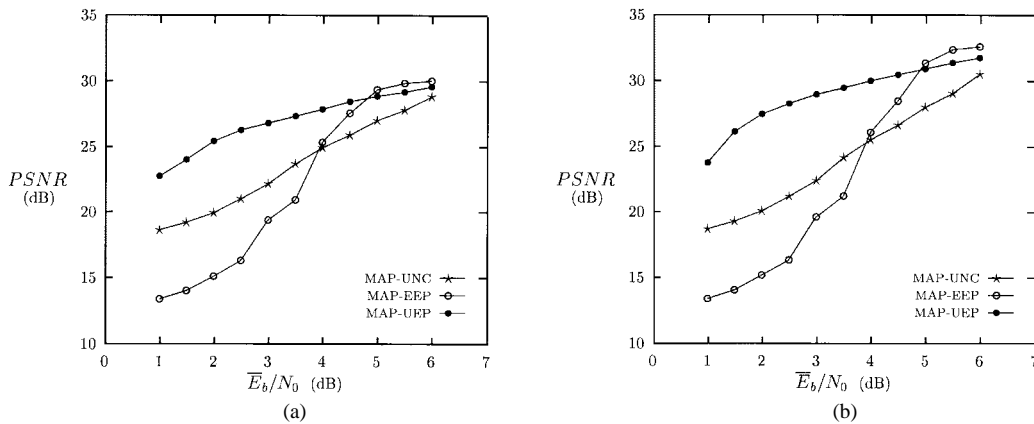


Fig. 3. MAP-UNC, MAP-EEP, and MAP-UEP systems over AWGN channel with throughput R ; PSNR versus \bar{E}_b/N_0 of decoded Lena. Results are averaged over 30 experiments. (a) $R = 2.21$ pixels/T; (b) $R = 1.33$ pixels/T.

of these blocks the DCT is computed. The resulting 64 DCT coefficients are uniformly quantized using a default quantization matrix proposed in [15] and derived from the human visual system sensitivity. We use a (fixed) zonal coding [15] bit allocation technique by keeping the first L zig-zag scanned DCT coefficients. The retained coefficients are then converted to a binary stream using a folded binary code (FBC) representation [9]. We order the FBC data for transmission over the channel in order to exploit the *interblock* redundancy of the coefficients: for each FBC codebit of each DCT coefficient, we form a bit stream by zigzag-scanning the same FBC codebit position across all the image blocks. Each such FBC bit stream is then modeled as an independently, identically distributed (i.i.d.)² nonuniform binary source $\{u_i\}$.

B. TCM and Channel MAP Decoding

The FBC bitstream is arranged in a sequence of binary k -tuples $\{u_1, u_2, \dots\}$. At time i , each binary k -tuple u_i is an input to a trellis encoder which outputs a $(k + 1)$ -tuple of

²Markov chain modeling can also be used [1], [2]; this however increases the number of source statistics. Note that we assume that the statistics of the i.i.d. source are available at the decoder. This can be achieved by providing them to the decoder *a priori* via training [1] or transmitting them as overhead information via a forward error correction code. It has also been found empirically that the DCT coefficients statistics are relatively stable over various coded natural images [18].

binary bits c_i . The encoder output c_i is mapped to a $2^{(k+1)}$ -ary two-dimensional (2-D) channel symbol x_i (complex number). Note that for the uncoded system with the same throughput as the coded system, each u_i is directly mapped to a 2^k -ary channel symbol x_i . The sequence $\mathbf{x}_n = \{x_1, \dots, x_n\}$ is then transmitted over the AWGN channel

$$y_i = x_i + z_i, \quad i = 1, 2, \dots, n \tag{1}$$

where y_i is the received 2-D channel symbol, and z_i is a 2-D noise tuple with independent components. Each noise component is a zero-mean Gaussian random variable with variance $N_0/2$.

We use the MAP decision rule at the decoder, which is optimal in the sense of minimizing the sequence error probability and can be implemented via a modified version of the Viterbi algorithm [1]. The sequence MAP decoding rule is to choose $\hat{\mathbf{x}}_n$, which maximizes

$$f(\mathbf{y}_n | \hat{\mathbf{x}}_n) \Pr(\hat{\mathbf{x}}_n) \tag{2}$$

or equivalently

$$\ln[f(\mathbf{y}_n | \hat{\mathbf{x}}_n) \Pr(\hat{\mathbf{x}}_n)] \tag{3}$$

where $f(\cdot)$ is the channel conditional density function. For the AWGN channel, this consists of choosing $\hat{\mathbf{x}}_n$ which

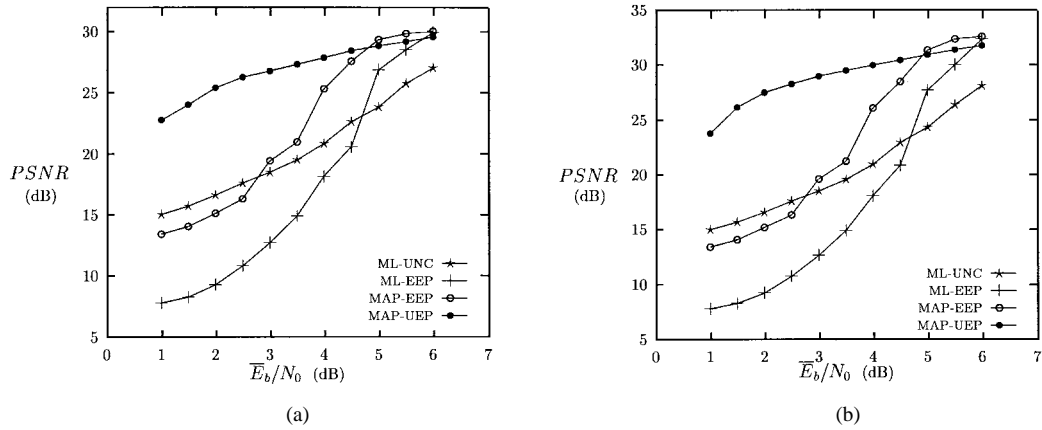


Fig. 4. Comparison of MAP-UEP with MAP-EEP, ML-EEP, and ML-UNC over AWGN channel with throughput R ; PSNR versus \bar{E}_b/N_0 of decoded Lena. Results are averaged over 30 experiments. (a) $R = 2.21$ pixels/T; (b) $R = 1.33$ pixels/T.

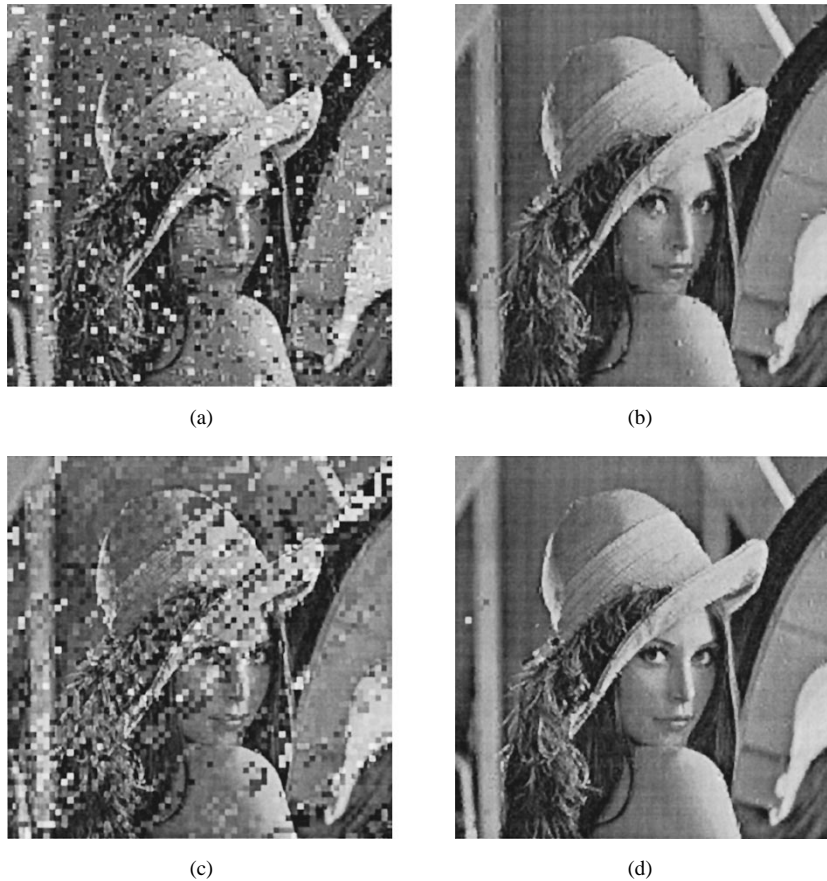


Fig. 5. Decoded Lena. (a) ML-UNC: PSNR = 16.51 dB. (b) MAP-UEP: PSNR = 25.96 dB. (c) MAP-EEP: PSNR = 16 dB. (d) MAP-UEP: PSNR = 28.25 dB. For (a) and (b): $\bar{E}_b/N_0 = 2.0$ dB and Throughput $R = 2.21$ pixels/T. For (c) and (d): $\bar{E}_b/N_0 = 2.5$ dB and Throughput $R = 1.33$ pixels/T.

maximizes

$$\begin{aligned} \ln[f(\mathbf{y}_n | \hat{\mathbf{x}}_n) \Pr(\hat{\mathbf{x}}_n)] &= \ln \prod_{i=1}^n \left[\frac{1}{\pi N_0} e^{-\frac{\|y_i - \hat{x}_i\|^2}{N_0}} \Pr(\hat{u}_i) \right] \\ &= n \ln \left(\frac{1}{\pi N_0} \right) - \frac{1}{N_0} \sum_{i=1}^n \|y_i - \hat{x}_i\|^2 \\ &\quad + \sum_{i=1}^n \ln \Pr(\hat{u}_i). \end{aligned} \quad (4)$$

Thus, the MAP decoding rule reduces to choosing $\hat{\mathbf{x}}_n$, which minimizes

$$\sum_{i=1}^n [\|y_i - \hat{x}_i\|^2 - N_0 \ln \Pr(\hat{u}_i)] \quad (5)$$

where $\|\cdot\|$ denotes the Euclidean norm, and \hat{u}_i is the information binary k -tuple corresponding to \hat{x}_i .

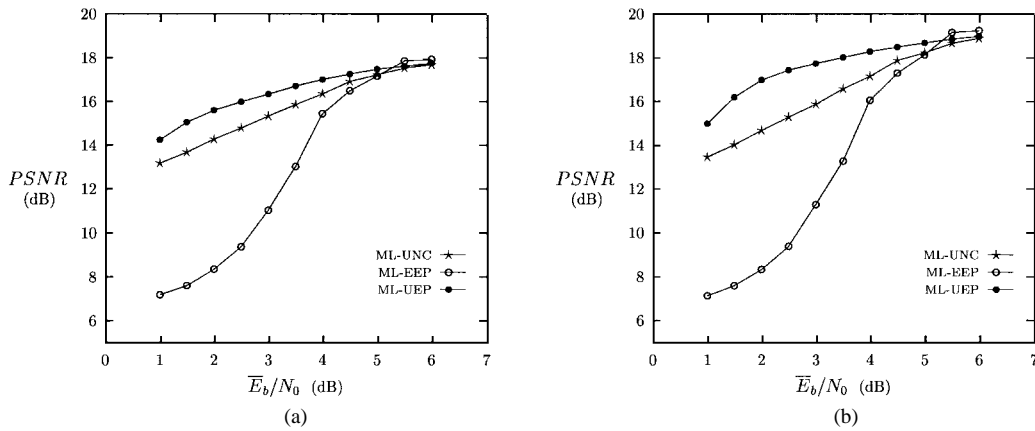


Fig. 6. ML-UNC, ML-EEP, and ML-UEP systems over AWGN channel with throughput R ; PSNR versus \bar{E}_b/N_0 of decoded Pentagon. Results are averaged over 30 experiments. (a) $R = 2.21$ pixels/T; (b) $R = 1.33$ pixels/T.

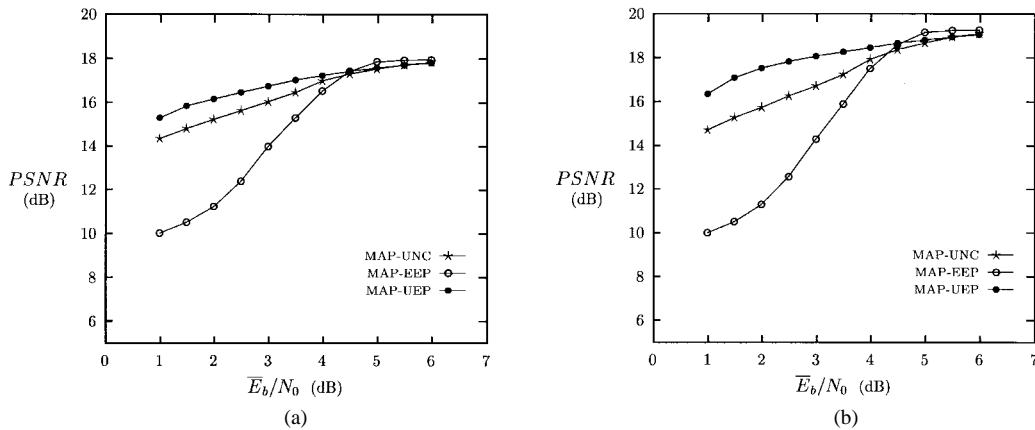


Fig. 7. MAP-UNC, MAP-EEP, and MAP-UEP systems over AWGN channel with throughput R ; PSNR versus \bar{E}_b/N_0 of decoded Pentagon. Results are averaged over 30 experiments. (a) $R = 2.21$ pixels/T; (b) $R = 1.33$ pixels/T.

C. UEP via TEA and TCM

The TCM code used consists of Ungerboeck's rate $2/3$ (i.e., $k = 2$), 16-state trellis code (with $d_{\text{free}}^2 = 5.172$, where d_{free} is the code's minimum Euclidean free distance) used with 8-PSK modulation [16]. It has a throughput of two information bits per signaling period. Other TCM codes can similarly be used.

The concept of transmitting signals with different levels of energy was previously examined in [6], [7], and [17]. We herein propose a UEP system that employs two signal constellations (one quadrature PSK (QPSK) and one 8-PSK). The scheme, which offers different levels of protection to the DCT bitstream, is described as follows.

1) *MAP-UEP*: For each block, the first N (among the L) DCT coefficients are 8-PSK trellis-coded with signal energy $E'_s > 1$ over the AWGN channel in conjunction with MAP decoding; the other $L - N$ DCT coefficients are processed *uncoded* but using MAP decoding over the QPSK-modulated channel with signal energy $E''_s < 1$ so that the average signal energy $\bar{E}_s = 1$. In other words, UEP is used by applying channel coding and assigning selective energy allocation: the *important* DCT bits are encoded and are provided with additional transmission energy, while the rest of the DCT bits are sent uncoded with less transmission energy (but the overall average energy per signal is equal to 1).

The relationship between E'_s , E''_s and \bar{E}_s is described by

$$\begin{aligned} \bar{E}_s &= \frac{\text{total transmitted signal energy per image}}{\text{total number of signals per image}} \\ &= \frac{l_N E'_s + (B - l_N) E''_s}{\frac{B}{k}} \\ &= \frac{1}{B} [l_N E'_s + (B - l_N) E''_s] \end{aligned} \quad (6)$$

where B is the total number of bits required to describe all the image DCT coefficients [$B = L * (\text{number of blocks per image})$], and l_N is the total number of bits describing the first N coefficients of each block. Therefore, to maintain an overall average energy per signal equal to 1, we choose E'_s and E''_s so that the above equation is satisfied with $\bar{E}_s = 1$.

We systematically examine the effect of UEP (via TCM coding and TEA) and of MAP decoding on the system performance by comparing the above scheme with reference systems of comparable overall complexity and identical throughput. The reference systems are as follows:

- *ML-UEP*: As in MAP-UEP but uses maximum-likelihood (ML) decoding.

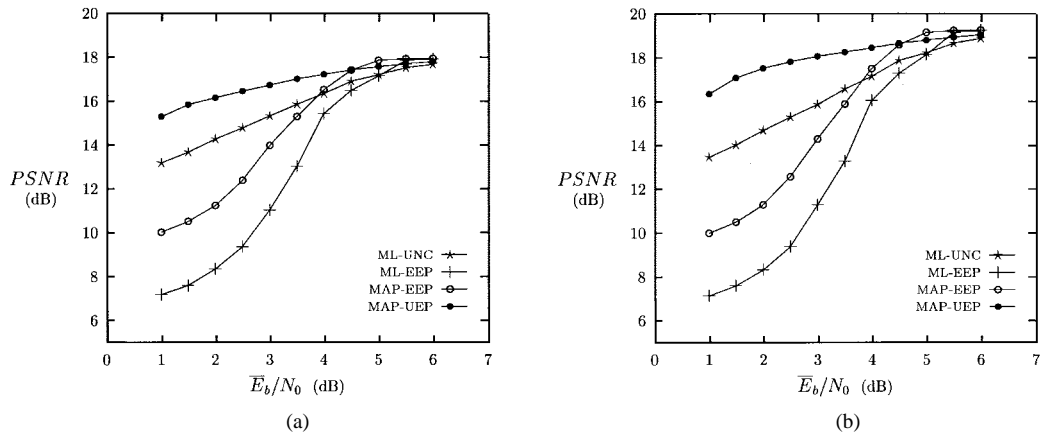


Fig. 8. Comparison of MAP-UEP with MAP-EEP, ML-EEP, and ML-UNC over AWGN channel with throughput R ; PSNR versus \bar{E}_b/N_0 of decoded Pentagon. Results are averaged over 30 experiments. (a) $R = 2.21$ pixels/T; (b) $R = 1.33$ pixels/T.

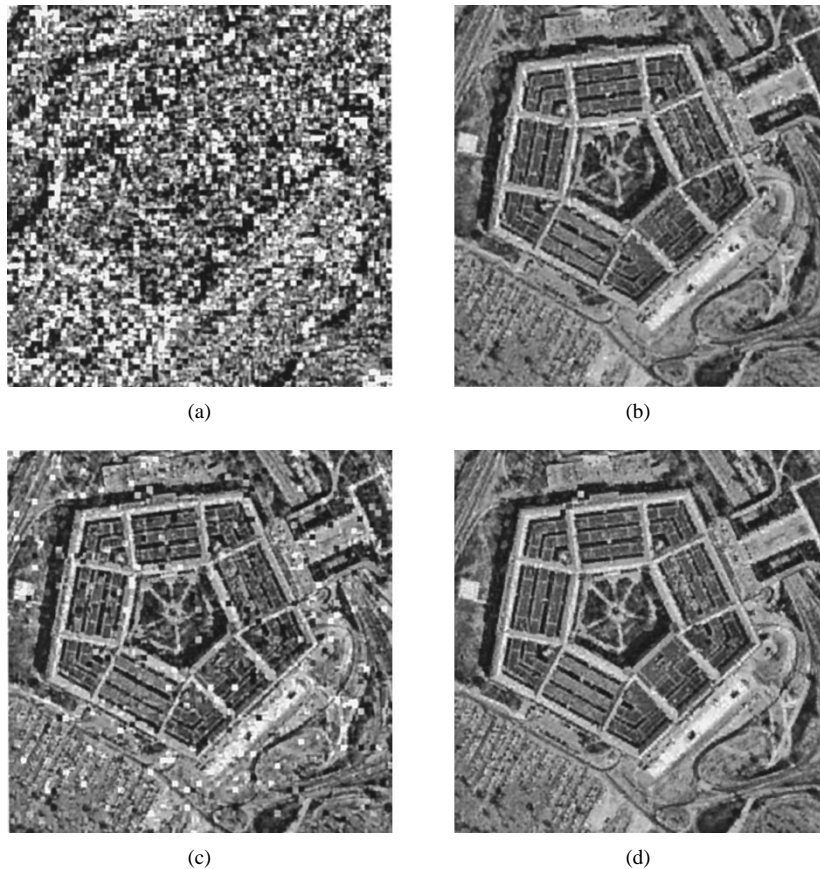


Fig. 9. Decoded Pentagon. (a) ML-EEP: PSNR = 8.43 dB. (b) MAP-UEP: PSNR = 16.26 dB. (c) ML-UNC: PSNR = 14.75 dB. (d) MAP-UEP: PSNR = 17.50 dB. For (a) and (b): $\bar{E}_b/N_0 = 2.0$ dB and Throughput $R = 2.21$ pixels/T. For (c) and (d): $\bar{E}_b/N_0 = 2.0$ dB and Throughput $R = 1.33$ pixels/T.

- ML-UNC: Uncoded system using ML decoding over AWGN with QPSK modulation with signal energy $E_s = 1$.
- MAP-UNC: Uncoded system using MAP decoding over AWGN used in conjunction with QPSK modulation with signal energy $E_s = 1$.
- ML-EEP: This is a typical equal-error protection (EEP) system. All L DCT coefficients are TCM-coded using ML decoding over AWGN used with 8-PSK modulation with signal energy $E_s = 1$.
- MAP-EEP: As in ML-EEP but uses MAP decoding.

III. EXPERIMENTAL RESULTS

Simulations were performed using the above schemes for the transmission of the images Lena and Pentagon (both of size 512×512) over the AWGN channel. For illustrative purposes, we present results for two different source compression rates $R_s = 0.906$ bits per pixel (bpp) and $R_s = 1.5$ bpp. These result in overall throughputs of $R = 2.21$ pixels/T and $R = 1.33$ pixels/T, where T is the signaling period (note that the throughput is given by the relation $R = k/R_s$ pixels/T).

For the $R = 2.21$ pixels/T system, $L = 10$ DCT coefficients are zonally retained [15] with FBC bit allocations $\{8, 7, 7, 6, 6, 6, 4, 5, 5, 4\}$; the MAP-UEP parameters applied are $N = 1$, $E'_s = 2$, and $E''_s = 0.84$. For the $R = 1.33$ pixels/T system, $L = 21$ DCT coefficients are zonally retained with FBC bit allocations $\{8, 7, 7, 6, 6, 6, 5, 5, 5, 5, 4, 4, 4, 4, 2, 3, 3, 3, 3, 2\}$; the MAP-UEP parameters utilized are $N = 3$, $E'_s = 2$, and $E''_s = 0.70$. The bit allocations were chosen according to the zonal coding bit allocation technique described in [15], and the values of E'_s and E''_s were chosen via a simple numerical search subject to the constraint that (6) is satisfied. The bit allocation and energy values were obtained based on Lena; and the same values were applied to the transmission of Pentagon.

Note that when only source compression is performed and the channel is noiseless, the peak signal-to-noise ratio (PSNR) is 30.48 dB (when $R = 2.21$ p/T) and 33.75 dB (when $R = 1.33$ p/T) for the image Lena. For the image Pentagon, the PSNR is 18.45 dB (when $R = 2.21$ p/T) and 20.92 dB (when $R = 1.33$ p/T). Since Pentagon has a higher energy content in its high AC coefficients than Lena, its performance under our zonal coding compression scheme, which favors the low frequency components, is inferior to that of Lena. Higher PSNR's can be obtained if we improve the compression scheme, such as using a wavelet-based system with optimal quantization.

Simulation results for the images Lena and Pentagon are shown in Figs. 2–5 and Figs. 6–9, respectively. The PSNR of the decoded images is displayed for different values of \bar{E}_b/N_0 , where \bar{E}_b is the average transmission energy per information bit. In the figures, we also show the graphical representation of the ML-UNC, MAP-EEP, and MAP-UEP systems. It can clearly be remarked from all the figures that the MAP-UEP scheme substantially outperforms all the other schemes; it also offers a very graceful degradation as the channel conditions deteriorate. In particular, MAP-UEP achieves coding gains over ML-UNC of up to 4 dB in \bar{E}_b/N_0 (at a PSNR of 27 dB) for Lena, and up to 2.5 dB (at a PSNR of 17 dB) for Pentagon. Inspection of the figures also demonstrates a significant advantage of MAP decoding over ML decoding for all coding systems, particularly when the channel is very noisy.

Furthermore, for low to medium values of \bar{E}_b/N_0 , the MAP-UEP scheme yields considerable gains over the MAP-EEP scheme. For example, for Lena, coding gains of up to 12 dB in PSNR and up to 2.5 dB in \bar{E}_b/N_0 (for a PSNR of 27 dB) are obtained. Then at high values of \bar{E}_b/N_0 , the MAP-EEP scheme provides a better performance (due to the asymptotic coding gain obtained by channel coding all the DCT coefficients). This suggests that when the system is operating at low to medium channel signal-to-noise ratios (SNR), it is recommended to use the MAP-UEP scheme; at high channel SNR's, the system can then switch to the MAP-EEP scheme. Examples illustrating the subjective performance improvements of MAP-UEP over ML-UNC and MAP-EEP are respectively presented in Fig. 5 for Lena and Fig. 9 for Pentagon.

IV. CONCLUSION

We introduced a bandwidth efficient UEP-MAP trellis coding scheme for the reliable communication of compressed images over very noisy AWGN channels used with PSK modulation. The scheme employs selective channel coding and TEA of the modulation constellation; it also exploits the residual image redundancy at the decoder. Substantial coding gains are achieved over uncoded systems and EEP systems. Our results reveal that the joint use of MAP decoding and UEP can substantially enhance the error resilience of a visual communication system, particularly during severe channel conditions, without sacrificing bandwidth efficiency. We expect that similar performance improvements would extend for the case of subband or wavelet-encoded images.

Future work will address the development of an optimal energy allocation strategy for soft-decision decoded TCM constellations; similar strategies for the AWGN channel with uncoded hard-decision BPSK modulation have been reported in [6] and [7]. These schemes will then be applied over fading channels and for systems with higher overall throughputs (e.g., 16-QAM systems and others).

REFERENCES

- [1] F. Alajaji, N. Phamdo, and T. Fuja, "Channel codes that exploit the residual redundancy in celp-encoded speech," *IEEE Trans. Speech Audio Processing*, vol. 4, pp. 325–336, Sept. 1996.
- [2] P. Burlina and F. Alajaji, "An error resilient scheme for image transmission over noisy channels with memory," *IEEE Trans. Image Processing*, vol. 7, pp. 593–600, Apr. 1998.
- [3] T. Fazel and T. Fuja, "Channel-encoded transmission of MELP-compressed speech," in *Proc. Int. Symp. Information Theory*, MIT, Cambridge, MA, Aug. 1998, p. 203.
- [4] A. Fuldseth and T. A. Ramstad, "Combined video coding and multilevel modulation," in *Proc. Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996, pp. 941–944.
- [5] ———, "Channel-optimized subband video coding for channels with a power constraint," in *Proc. Int. Conf. Image Processing*, Chicago, IL, 1997, pp. 428–431.
- [6] S. Gadkari and K. Rose, "Jointly optimized VQ index assignment and transmission energy allocation," in *Proc. Int. Conf. Communications*, Montreal, Quebec, June 1997, pp. 324–328.
- [7] ———, "Transmission energy allocation with low peak-to-average ratio," *IEEE Commun. Lett.*, vol. 1, pp. 166–168, Nov. 1997.
- [8] J. Hagenauer, "Source controlled channel decoding," *IEEE Trans. Commun.*, vol. 43, pp. 2449–2457, Sept. 1995.
- [9] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [10] N. Phamdo, F. Alajaji, and N. Farvardin, "Quantization of memoryless and gauss-markov sources over binary markov channels," *IEEE Trans. Commun.*, vol. 45, pp. 668–675, June 1997.
- [11] K. Sayood and J. C. Borkenhagen, "Use of residual redundancy in the design of joint source/channel coders," *IEEE Trans. Commun.*, vol. 39, pp. 838–846, June 1991.
- [12] K. Sayood, F. Liu, and J. D. Gibson, "A constrained joint source/channel coder design," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 1584–1593, Dec. 1994.
- [13] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, July 1948.
- [14] P. G. Sherwood and K. Zeger, "Error protection for progressive image transmission over memoryless and fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 1555–1559, Dec. 1998.
- [15] A. M. Tekalp, *Digital Video Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [16] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55–67, Jan. 1982.
- [17] L.-F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, pp. 1439–1449, Oct. 1993.
- [18] W. Xu, J. Hagenauer, and J. Hollmann, "Joint source-channel decoding using the residual redundancy in compressed images," in *Proc. Int. Conf. Commun.*, Dallas, TX, June 1996, pp. 142–146.



Fady I. Alajaji (S'90–M'95) was born on May 1, 1966, in Beirut, Lebanon. He received the B.E. degree (with distinction) from the American University of Beirut, Beirut, Lebanon, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, all in electrical engineering, in 1988, 1990, and 1994, respectively. He held a graduate Fellowship from 1990 to 1992 and a postdoctoral appointment in 1994 at the Institute for Systems Research, University of Maryland.

In 1995, he joined the Department of Mathematics and Statistics at Queen's University, Kingston, Ontario, where he is an Assistant Professor of Mathematics and Engineering. Since 1997, he has been also cross-appointed as an Assistant Professor in the Department of Electrical and Computer Engineering at Queen's University. His research interests include information theory, digital communications, error control coding, joint source-channel coding, and data compression.

Philippe Burlina was born on September 1, 1965, in Paris, France. He received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1990 and 1994, respectively.

He is currently a Senior Scientist with ImageCorp, Inc. He is also affiliated with the Institute for Advanced Computer Studies, University of Maryland, College Park. His current research interests include image/video processing, analysis and digital communications.



Saud A. Al-Semari (S'95–M'96) was born on December 25, 1968. He received the B.S. and M.Sc. degrees in electrical engineering from King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1991 and 1992, respectively. He received his Ph.D. degree from the University of Maryland, College Park, in 1995.

Since January 1996, he has been an Assistant Professor of Electrical Engineering at KFUPM, and he is also the Director of the Communications and Computer Research Center. His research interests include topics related to wireless communication systems including error control coding, equalization, diversity, fading, and multiple access.