Spectral Entropy-Based Quantization Matrices for H.264/AVC Video Coding

Malavika Bhaskaranand and Jerry D. Gibson
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA - 93106
Email: {malavika, gibson}@ece.ucsb.edu

***Abstract*—In transform-based compression schemes, the task of choosing, quantizing, and coding the coefficients that best represent a signal is of prime importance. As a step in this direction, Yang and Gibson [1] have designed a coefficient selection scheme based on Campbell's coefficient rate and spectral entropy [2]. Building on the spectral entropy-based coefficient selection mechanism, we develop a method to allocate bits amongst the chosen coefficients that can outperform the classical method under certain conditions. We then design quantization matrices (QMs) based on the proposed bit allocation scheme. Results show that the newly designed QMs perform better than the default QMs for H.264/AVC encoding in terms of both peak signal to noise ratio (PSNR) and structural similarity (SSIM). The proposed method entails delay but is not computationally intensive.**

I. Introduction

The Fidelity Range Extensions (FRExt) of the H.264/AVC standard allow the use of quantization matrices (QMs) that can be updated at frame level. Although default QMs are specified in the standard, the encoder can specify a customized QM for each transform block size and separately for intra and inter prediction, for use in inverse-quantization scaling by the decoder [3]. Several methods have been proposed for the design of QMs for image and video coding, most common of them being psycho-visual model based techniques [4], [5] and rate-distortion (RD) optimization techniques [6]–[8].

In this work, we propose a QM design algorithm using Campbell's concepts of spectral entropy and coefficient rate [2]. Since spectral entropy methods entail delay, this scheme can be used to customize QMs on a per-frame basis in contrast to macroblock adaptive QM schemes proposed for H.264 video encoders [9]–[11]. First, we build on Yang and Gibson's spectral entropy-based coefficient selection scheme [1] and derive a scheme that can be used to allocate bits amongst the chosen significant coefficients. Then we design QMs based on the proposed bit allocation scheme and demonstrate that the newly designed QMs outperform the default QMs in a H.264/AVC encoder. The performance evaluation is done using PSNR, the most commonly used video quality metric and structural similarity (SSIM), an objective metric for assessing perceptual video quality [12]. The proposed QM design method involves delay but is not computationally intensive. Hence, it can be used in applications such as video streaming and entertainment-quality encoding where low-latency encoding is not a necessity.

This paper is organized as follows. Section II briefly discusses previous research on spectral entropy including Yang and Gibson's coefficient selection mechanism, develops the spectral entropy-based bit allocation method, and suggests a way to employ the proposed bit allocation scheme for QM design. Section III provides implementation details and discusses experimental results that show the improved performance of the newly designed QMs. It also briefly outlines scope for future improvements. Finally, Section IV summarizes the work.

This research has been supported by the UC Discovery Grant Program and Applied Signal Technology, Inc.

II. Spectral Entropy-based Quantization Matrix Design

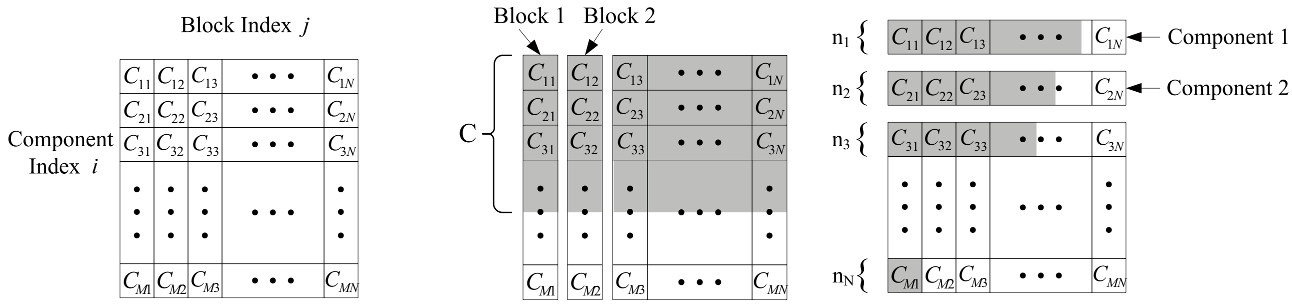
A. Previous research on Spectral Entropy

In transform coding-based compression schemes where the bandwidth is limited, it is not possible to transmit all transform coefficients and hence some coefficients need to be discarded. Therefore it is important to choose or sample the transform coefficients that best represent a signal and code them with high fidelity. In 1960, Campbell [2] first examined the problem of sampling a random process with non-rectangular power spectral densities (PSD) and demonstrated that the products of a large number of sample functions of such a random process require average sampling rates less than the Nyquist rate. Using a version of the asymptotic equipartition property (AEP), he proved that a Karhunen-Loéve (K-L) expansion of the product of N sample functions of a stationary random process $X(t)$ could be separated into two sets: one with average power very close to that of the product and the other with very low average power. Asymptotically in the number of sample functions N and the support interval T of $X(t)$, he showed that the average number of terms in the high energy set approached the coefficient rate Q defined as

$$Q = \exp\left[-\int S(f)\log S(f)\,\mathrm{d}f\right], \tag{1}$$

where $S(f)$ is the normalized PSD of $X(t)$. Since the quantity in the exponent $h(S) = \log(Q)$ resembles the differential entropy of $S(f)$, it was termed the spectral entropy by Gibson *et al.* [13]. Campbell presented this as a sampling result but did not explore the application of the spectral entropy and coefficient rate for data compression. Additionally, Abramson [14] examined Campbell's results and observed that a compression scheme based on spectral entropy was not apparent.

Almost thirty years after Campbell's work, Yang and Gibson [1], [15]–[18] examined Campbell's coefficient rate and theoretically derived a new mechanism for selecting the significant coefficients *i.e.* those that best represent a signal. They proved that the number of significant coefficients in each component should be proportional to the energy/variance of that component. Kokes and Gibson [19] applied Yang and Gibson's spectral entropy-based coefficient selection to wideband speech coding and showed results that were perceptually better than those of conventional speech coders. More importantly, they developed a band combining strategy based on spectral entropy to formulate an adaptive nonuniform modulated lapped bi-orthogonal transform (NMLBT) [20], [21]. The more precise frequency selectivity was shown to improve the performance of a wideband speech coder for both speech and audio signals. Kikkawa and Yoshida [22] examined the relation between the equivalent rectangular bandwidth,



(a) Coefficients of N blocks (b) Classical method (c) Spectral entropy method

Fig. 1. Encoding N blocks of data each with M transform coefficients for $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$.

Blackman-Tukey's bandwidth, and Campbell's bandwidth of a random process and developed a unified representation of these using the Rényi entropy of the spectrum.

Recently, Roy and Vetterli [23] have defined the effective rank of a matrix as $e^{H(p)}$ where $H(p)$ is the entropy of its singular value distribution. On lines similar to those followed for coefficient rate in [2], [18], they show that the effective rank of a matrix represents the average number of significant dimensions in its range and thus interpret it as the "effective dimension" of the matrix. Their work provides a theoretical justification to the heuristic entropy-based algorithm proposed by Coifman and Wickerhauser [24] for selecting the best basis for signal representation by minimizing the theoretical dimension $d = \exp\left(-\sum_n p_n \log p_n\right)$, where $p_n = |x_n|^2 / \sum_n |x_n|^2$ and x_n are the transform coefficients.

Most of the previous research discussed above uses Campbell's coefficient rate and spectral entropy as a basis for efficiently sampling frequency coefficients. Yang and Gibson [18] have shown that the Campbell bandwidth is the minimum average bandwidth for encoding the process across all possible distortion levels. In addition, Jung and Gibson [25] have obtained an expression for coefficient rate using the logarithm of the ratio of rate distortion function slopes of the given source and a uniform source, where the logarithm is averaged over large distortions. These results indicate a relationship between coefficient rate and the rate-distortion function of a source. However, to the best of our knowledge, this is the first time a spectral entropy-based coding scheme has been developed.

B. Coefficient selection [18]

Consider a zero-mean stationary continuous-time random process $X(t)$. Using the K-L expansion in the time interval $[0, T]$, the process can be decomposed as

$$X(t) = \sum_{i=1}^{M} C_i \phi_i(t), \tag{2}$$

where $\phi_i(t)$'s are normalized eigenfunctions and C_i's are uncorrelated random variables with $\mathbf{E}[C_i] = 0$ and $\mathbf{E}[C_i^2] = \lambda_i$. Hence, the random process can be represented by a random vector $\{C_1, C_2, \ldots, C_M\}$ and the total average energy of the process is $\sigma^2 = \sum_{i=1}^{M} \lambda_i$.

Let $x_1(t_1), x_2(t_2), \ldots, x_N(t_N)$ be N independent sample functions of $X(t)$. Consider the product of these N independent sample functions

$$y(t_1, t_2, \ldots, t_N) = x_1(t_1) x_2(t_2) \cdots x_N(t_N). \tag{3}$$

Yang and Gibson [18] showed that for large N, the number of occurrences of λ_i in the high energy terms of the energy of $y(t_1, t_2, \ldots, t_N)$ is proportional to λ_i and given by

$$n_i = \frac{\lambda_i}{\sigma^2} N, \qquad i = 1, 2, \ldots, M, \tag{4}$$

and the number of high energy coefficients μ in the product is

$$\mu = \exp\left[-N \sum_{i=1}^{M} \frac{n_i}{N} \log \frac{n_i}{N}\right] = e^{NH(S)} \tag{5}$$

where $H(S) = -\sum_{i=1}^{M} \frac{\lambda_i}{\sigma^2} \log \frac{\lambda_i}{\sigma^2}$ is the spectral entropy in discrete form.

Equation (4) suggests that in a sequence of N samples of a particular coefficient, the number of coefficient samples that should be coded is proportional to the energy of the coefficient. The basic approach to source compression implied by the spectral entropy results is illustrated in Fig. 1. Fig. 1(a) shows M transform components for N blocks of source data, denoted C_{ij}, $i = 1, \ldots, M, j = 1, \ldots, N$, where i is the component index and j the block index. Without loss of generality, we can assume that the components are ordered based on their energies *i.e.* $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$. In classical transform based coding, coefficient bit allocation is accomplished on a block-by-block basis as illustrated in Fig. 1(b). Given a particular block (fixed j), a fixed number of bits is allocated across the M coefficients based on their energies and only coefficients of the C components with the highest energies are coded. However, the spectral entropy approach implies that each component should be considered as a separate sequence, $\{C_{ij},\ j = 1, \ldots, N\}$, as shown in Fig. 1(c), and the number of significant coefficients n_i of that component in the sequence should be determined based on its energy. Therefore, a coefficient is more likely to be coded if it has high energy. In contrast to the classical method, this coefficient selection mechanism entails delay and has been shown to achieve better SNR (with no rate control) [15] and subjective quality [17].

C. Bit Allocation

Consider the decomposition in (2). As before, let the M components $\{C_i,\ i = 1, 2, \ldots, M\}$ be independent with $\mathrm{E}[C_i] = 0$ and $\mathrm{E}[C_i^2] = \lambda_i$. Without loss of generality, we can assume that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$. Let there be N sampling functions (blocks/frames) each with M such components. Out of the total $M \times N$ coefficients, let L coefficients be coded. Then the spectral entropy-based coefficient selection derived in the previous sub-section dictates that the number of coefficients n_i coded in each component be proportional to the energy λ_i of that component *i.e.* $n_i = p_i L$, where $p_i = \frac{\lambda_i}{\sigma^2}$ and $\sigma^2 = \sum_{i=1}^{M} \lambda_i$.

If $b_i^{(S)}$ is the average number of bits spent to code a coefficient of component i, the total number of bits spent is $B = \sum_{i=1}^{M} n_i b_i^{(S)}$. The coding distortion is generated by two sources: quantization and discarding coefficients. Hence the expected value of the distortion of the ith component can be written as

$$\begin{aligned} d_i^{(S)} &= n_i \times \text{E(quantization error)} + \\ &\quad (N - n_i) \times \text{E(energy of discarded coefficients)} \\ &= n_i \times h_i \lambda_i 2^{-2b_i^{(S)}} + (N - n_i) \times \lambda_i \end{aligned} \tag{6}$$

In this equation, the quantization error is computed assuming that the overload distortion is negligible and the high-resolution approximation holds and h_i is a constant determined by the distribution of the normalized random variable $C_i/\sqrt{\lambda_i}$ [26].

Hence, the problem of bit allocation is to find $b_i^{(S)}$ for $i = 1, 2, \ldots, M$ so as to minimize the overall distortion $D^{(S)} = \sum_{i=1}^{M} d_i^{(S)}$ subject to the constraint that $\sum_{i=1}^{M} n_i b_i^{(S)} = B$. Using Lagrangian optimization methods, the number of bits $b_i^{(S)}$ allocated to each of the n_i coded coefficients of component i can be shown to be

$$b_i^{(S)} = \frac{B}{L} + \frac{1}{2}\log_2\left(\frac{\lambda_i h_i}{\prod_{i=1}^{M}(\lambda_i h_i)^{\frac{\lambda_i}{\sigma^2}}}\right). \tag{7}$$

This is similar to the result of classical bit allocation [26] except that the geometric mean of $(\lambda_i h_i)$'s $\prod_{i=1}^{M}(\lambda_i h_i)^{\frac{1}{M}}$ has been replaced by $\prod_{i=1}^{M}(\lambda_i h_i)^{p_i}$, the generalized geometric mean with weights $p_i = \frac{\lambda_i}{\sigma^2}$. The corresponding total distortion is

$$D^{(S)} = L 2^{-\frac{2B}{L}} \prod_{i=1}^{M}(\lambda_i h_i)^{\frac{\lambda_i}{\sigma^2}} + N\sigma^2 - \sum_{i=1}^{M} n_i \lambda_i. \tag{8}$$

The proposed bit allocation method examines the input transform coefficients and chooses to code only those that are significant for retaining signal fidelity. Therefore, it is "realization-adaptive" in the sense that it adapts to the actual coefficient values that need to be coded. In contrast, the classical bit allocation method relies entirely on the energies of the transform components and hence is designed for a class of inputs, all having the same component energies but different coefficient values. Ortega and Ramchandran [27] have noted that "input-by-input" approaches that adapt to the source data being compressed are likely to be superior to "one size fits all" approaches that are designed to perform well on average for a class of inputs. In line with their observation, it can be shown that the proposed bit allocation scheme outperforms (gives lower distortion than) the classical bit allocation scheme for a given bit budget B and number of coded coefficients L under certain conditions [28].

For the case when all the components C_i have the same normalized distribution, $h_i = h, i = 1, \ldots, M$ and (7) can be rewritten as

$$\begin{aligned} b_i^{(S)} &= \frac{B}{L} + \frac{H^{(2)}(S)}{2} + \frac{1}{2}\log_2\frac{\lambda_i}{\sigma^2} \\ &= \frac{B}{L} + \underbrace{\frac{1}{2}\log_2\frac{Q}{M}}_{②} + \underbrace{\frac{1}{2}\log_2\frac{\lambda_i}{\sum_{i=1}^{M}\lambda_i/M}}_{③} \end{aligned} \tag{9}$$

where $H^{(2)}(S)$ is the spectral entropy expressed in bits and $Q = e^{H(S)}$ is the coefficient rate. In (9), the first term is an average bit rate, the second term ② depends on the source, and the last term ③ depends on the current component i being coded. In the corresponding expression for classical bit allocation, Q is replaced by M in ② and the arithmetic mean of (λ_i)'s in the denominator of ③ is replaced by the geometric mean.'

In a two-dimensional discrete cosine transform (DCT) based coding scheme, Mester and Franke [29] used the spectral entropy (corresponding to ②) and the energy (corresponding to ③) of the transform coefficients to classify data blocks and adopted different coding strategies for the different classes. They used these two measures to estimate the amount of tolerable errors and sensitivity to quantization and/or truncation and thus develop an adaptation scheme for a threshold coding system. The spectral entropy and energy were treated as orthogonal entities in their work, where as (9) provides a way to combine these two measures and the average bit rate for bit allocation.

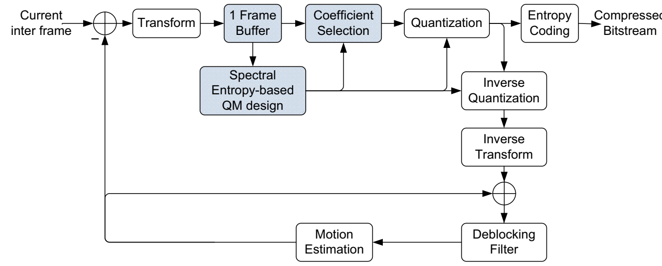


Fig. 2. Block diagram of the H.264/AVC encoder using the proposed QM design method

D. Quantization Matrix Design

Since the number of bits $b_i^{(S)}$ allocated for the ith component C_i is indicative of the number of steps in the uniform quantizer of C_i, entropy constrained uniform scalar quantizers optimal for the component distribution can be used. Considering C_i to be uniformly distributed within its range, the ith element of the QM can be designed as follows,

$$\mathbf{QM}(i) = \text{range}(C_i)/(2^{b_i^{(S)}}), \quad i = 1, 2, \ldots, M. \tag{10}$$

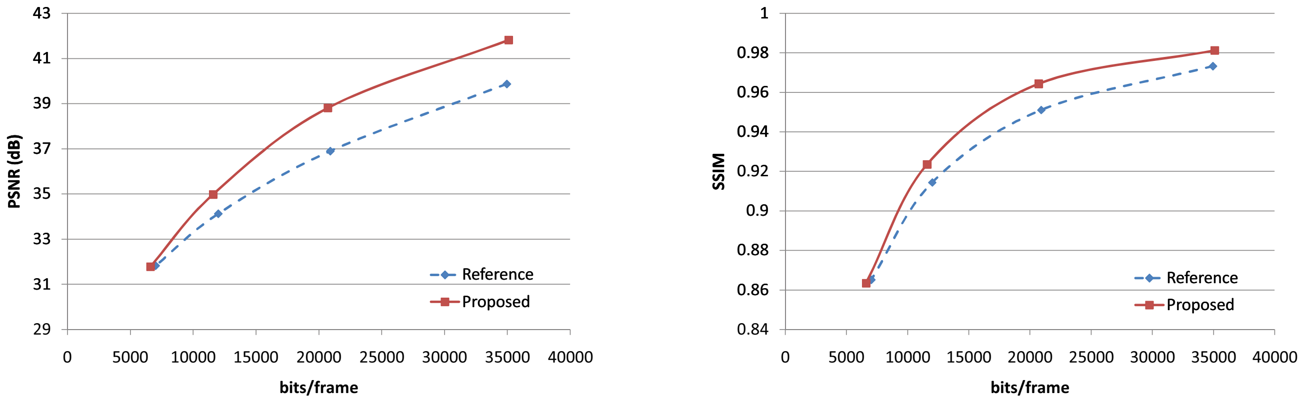
Alternatively, quantizers can be designed to be optimal for Laplacian distributed random variables [30], [31] which are commonly used to model DCT coefficients of image and video.

III. Experiments and Results

We have employed the proposed bit allocation scheme to design inter 4×4 QM for the luminance component in the H.264/AVC encoder. The overall block diagram of the modified inter-frame encoder is given in Fig. 2 with the modifications indicated by shading. For each inter frame, the residual transform coefficients of all the inter 4×4 luminance blocks are first buffered *i.e.* $M = 16$ and $N = \frac{\text{width} \times \text{height}}{16}$. The energies of the transform components are estimated as the empirical variances of the buffered coefficients. These are used to choose the significant coefficients and design the luma inter 4×4 QM as described in Section II. The coefficients selected using the spectral entropy-based coefficient selection scheme are quantized using the designed QMs and finally entropy coded.

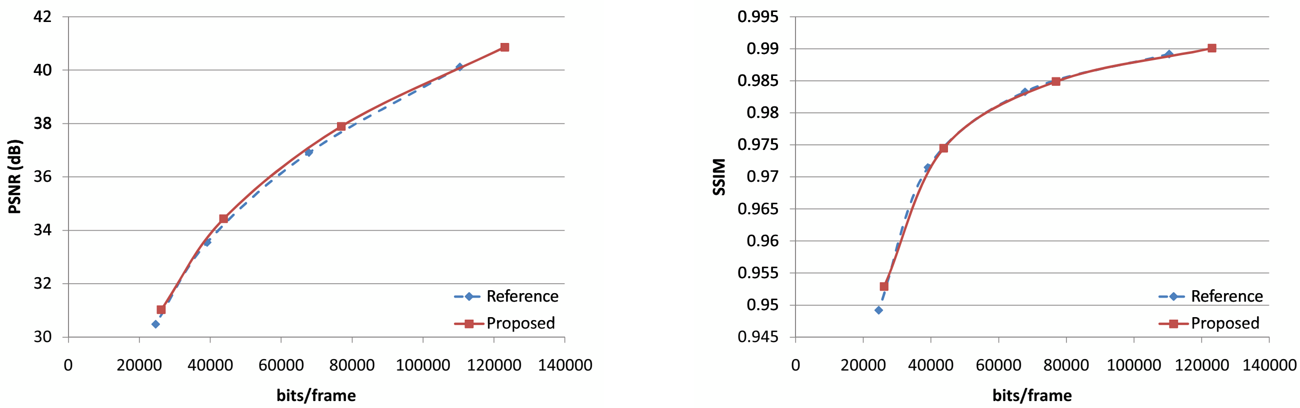
The discrete cosine transform (DCT) coefficients of luminance inter-prediction residue have been shown to be Laplacian distributed [32], [33]. Since the integer transform used in the H.264 standard and the DCT have similar coding gains for prediction residuals [34], we assume that all components of the H.264 integer transformed data have the same normalized distribution *i.e.* $h_i = h$ for $i = 1, 2, \ldots, 16$ and can be approximated by a Laplacian distribution.

In our experiments, the high profile of the JM17.0 encoder was used with IPPP... group of pictures (GOP) structure and an intra-period of 15. The RD curves were obtained by encoding the test video sequences at four quantization parameters (QP): 20, 25, 30, 35. Fig. 3 and 4 plot the distortion of the luma component versus the average bits per frame for the encoder using the default QMs (reference



(a) PSNR vs. bits/frame (b) SSIM vs. bits/frame

Fig. 3. Comparison of performance for CIF “silent” sequence.



(a) PSNR vs. bits/frame (b) SSIM vs. bits/frame

Fig. 4. Comparison of performance for CIF “stefan” sequence.

method) and newly designed QMs (proposed method) for 2 video sequences at 352×288 (CIF) resolution: “silent” (still camera, fast but restricted subject motion and relatively low background detail) and “stefan” (fast content motion with camera pan and high amount of background detail). Curves are provided using both PSNR and SSIM as distortion metrics. It can be seen that the proposed QMs perform as well as or better than the default QMs in terms of both PSNR and SSIM. The PSNR improvement observed is without any loss in perceptual quality as indicated by SSIM.

We have also evaluated the performance improvement for various video sequences in terms of BD-rate [35]. BD-rate for a compression method with respect to a reference method indicates the average difference in bit rate of the two methods at constant PSNR. It can be viewed as a metric that captures the bit rate vs. PSNR curve in a single number. Given multiple bit rate-PSNR points $\{(R_{i_j}, D_{i_j}), j = 1, 2, \ldots\}$ for the two methods $i = 1, 2$, the BD-rate is computed in the following steps [35].

1) Fit a third order polynomial to each of the PSNR vs. $\log_{10}$(bit rate) curves for methods 1 and 2. In other words, find optimal a_i, b_i, c_i and d_i such that

$$\log_{10}(R_{i_j}) \approx a_i + b_i D_{i_j} + c_i D_{i_j}^2 + d_i D_{i_j}^3,\ i = 1, 2$$

2) Evaluate the area under each curve as I_1 and I_2 between the limits L_{min} and L_{max} given by

$$L_{min} = \max(\min\{D_{1_j}\}, \min\{D_{2_j}\})$$
$$L_{max} = \min(\max\{D_{1_j}\}, \max\{D_{2_j}\})$$

3) Considering method 1 as the reference, compute the average difference in the bit rate as

$$\text{BD-rate} = 10^{\left[\frac{(I_2 - I_1)}{L_{max} - L_{min}}\right]} - 1. \quad (11)$$

The BD-rate values presented in Table I are for H.264 compression using the newly designed QMs compared to the reference H.264 compression with default QMs. Therefore negative values of BD-rate imply bit rate savings at constant quality for the proposed method with respect to the reference method. It can be seen that the bit rate savings vary from 0.5% to 21.3% with an average savings of about 10.4%.

The improved performance of the proposed method can be attributed to two reasons. First, the quantization matrix adapts to the relative energies of the transform coefficients in the current frame. Second, the spectral entropy-based coefficient selection mechanism exploits latency to examine all the coefficients in the current frame and chooses those that are more important for retaining fidelity. Thus

TABLE I
BD-RATE FOR VARIOUS TEST VIDEO SEQUENCES

Video Sequence	Resolution	BD-Rate(%)
akiyo	176×144	$-\ 9.003$
container	176×144	-21.294
foreman	176×144	$-\ 0.558$
mobile	176×144	-11.125
mother-daughter	176×144	$-\ 5.003$
news	176×144	-12.086
silent	176×144	-18.822
silent	352×288	-21.145
stefan	352×288	$-\ 3.834$
tempete	352×288	$-\ 1.327$

it adapts to the actual coefficient values that need to be quantized. To achieve further perceptual quality improvement, perceptual weighting of the coefficients can be incorporated into the QM design by using a weighted distortion metric. Additionally, the proposed QM design method can be used for customizing H.264 QMs for other transform sizes and chrominance components.

IV. CONCLUSIONS

In this work, we derive and develop a quantization matrix (QM) design scheme based on Campbell's concepts of coefficient rate and spectral entropy. The proposed method exploits latency to examine all the coefficients in a given frame of data, determines the energies of the transform components, chooses the significant coefficients, and designs the QM based on the statistics of the frame. Thus, it adapts not only to the changing frame statistics but also to the actual values of the coefficients in the frame. We also show that the QMs thus designed on a per-frame basis can outperform the default QMs in the H.264/AVC encoder and give an average bit rate savings of roughly 10.4%. Although the proposed method for QM design involves one frame-time delay, it is not computationally intensive. Hence, it can be used in video encoding applications where latency can be tolerated.

REFERENCES

[1] W. Yang and J. Gibson, "Coefficient rate in transform coding," in *Proc., 35th Allerton Conf. on Communication, Control, and Computing*, Sep. 1997, pp. 128–137.

[2] L. L. Campbell, "Minimum coefficient rate for stationary random processes," *Information and Control*, vol. 3, no. 4, pp. 360 – 371, 1960.

[3] G. J. Sullivan, P. N. Topiwala, and A. Luthra, "The H.264/AVC advanced video coding standard: overview and introduction to the fidelity range extensions," vol. 5558, no. 1. SPIE, 2004, pp. 454–474.

[4] A. Ahumada and H. A. Peterson, "Luminance-model-based DCT quantization for color image compression," vol. 1666, no. 1. SPIE, 1992, pp. 365–374.

[5] A. B. Watson, "DCT quantization matrices visually optimized for individual images," vol. 1913, no. 1. SPIE, 1993, pp. 202–216.

[6] D. Monro and B. Sherlock, "Optimum DCT quantization," in *Data Compression Conf. (DCC)*, 1993, pp. 188 –194.

[7] S. Wu and A. Gersho, "Rate-constrained picture-adaptive quantization for JPEG baseline coders," in *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, vol. 5, Apr. 1993, pp. 389–392.

[8] V. Ratnakar and M. Livny, "An efficient algorithm for optimizing DCT quantization," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 267 –270, Feb. 2000.

[9] J. Hu and J. Gibson, "Intra-mode indexed nonuniform quantization parameter matrices in AVC/H.264," in *Conf. Record of the 39th Asilomar Conf. on Signals, Systems and Computers*, Oct. 2005, pp. 746 –750.

[10] J. Chen, J. Zheng, and Y. He, "Macroblock-level adaptive frequency weighting for perceptual video coding," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 775 –781, May 2007.

[11] A. Tanizawa, T. Chujoh, and T. Yamakage, "Improvement of adaptive quantization matrix selection," ITU-T SG16Q.6 VCEG-AI19, Jul. 2008. [Online]. Available: http://ftp3.itu.org/av-arch/video-site/0807_Ber/VCEG-AI19.zip

[12] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600 –612, Apr. 2004.

[13] J. Gibson, S. Stanners, and S. McClellan, "Spectral entropy and coefficient rate for speech coding," in *Conf. Record of 27th Asilomar Conf. on Signals, Systems and Computers*, vol. 2, Nov. 1993, pp. 925–929.

[14] N. Abramson, "Information theory and information storage," in *Proc. Symp. System Theory*, Apr. 1965, pp. 207–213.

[15] W. Yang and J. Gibson, "Coefficient rate and significance maps in transform coding," in *Conf. Record of the 31st Asilomar Conf. on Signals, Systems and Computers*, vol. 2, Nov. 1997, pp. 1373–1377.

[16] ——, "The coefficient rate and equivalent bandwidth in source coding," in *Information Theory Workshop*, Jun. 1998, pp. 24–25.

[17] ——, "Distortion analysis of the coding scheme based on coefficient rate," in *IEEE International Symposium on Information Theory*, Aug. 1998, pp. 224–.

[18] W. Yang, J. Gibson, and T. He, "Coefficient rate and lossy source coding," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 381–386, Jan. 2005.

[19] M. Kokes and J. Gibson, "Spectral entropy-based wideband speech coding," in *Conf. Record of the 34th Asilomar Conf. on Signals, Systems and Computers*, vol. 2, 2000, pp. 1464–1468.

[20] M. Kokes, J. Gibson, and G. Schuller, "A wideband speech codec based on nonlinear approximation," in *Conf. Record of 35th Asilomar Conf. on Signals, Systems and Computers*, vol. 2, 2001, pp. 1573–1577.

[21] M. Kokes and J. Gibson, "Frequency selectivity via the SpEnt methodology for wideband speech compression," in *IEEE International Conf. on Acoustics, Speech, and Signal Processing*, vol. 2, 2001, pp. 777–780.

[22] S. Kikkawa and H. Yoshida, "On unification of equivalent bandwidths of a random process," *IEEE Signal Process. Lett.*, vol. 11, no. 8, pp. 670 – 673, Aug. 2004.

[23] O. Roy and M. Vetterli, "The Effective Rank: A Measure of Effective Dimensionality," in *European Signal Processing Conf. (EUSIPCO)*, 2007, pp. 606–610.

[24] R. Coifman and M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.

[25] J. Jung and J. Gibson, "The interpretation of spectral entropy based upon rate distortion functions," in *IEEE International Symposium on Information Theory*, Jul. 2006, pp. 277–281.

[26] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1991.

[27] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23 –50, Nov. 1998.

[28] M. Bhaskaranand and J. D. Gibson, "Spectral entropy-based bit allocation," in *Proc. International Symposium on Information Theory and its Applications ISITA (in press)*, Oct. 2010.

[29] R. Mester and U. Franke, "Spectral entropy-activity classification in adaptive transform coding," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 5, pp. 913–917, Jun. 1992.

[30] N. Farvardin and J. Modestino, "Optimum quantizer performance for a class of non-gaussian memoryless sources," *IEEE Trans. Inf. Theory*, vol. 30, no. 3, pp. 485 – 497, May 1984.

[31] G. Sullivan, "Efficient scalar quantization of exponential and laplacian random variables," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1365 –1374, Sep. 1996.

[32] F. Bellifemine, A. Capellino, A. Chimienti, R. Picco, and R. Ponti, "Statistical analysis of the 2D-DCT coefficients of the differential signal for images," *Signal Process. Image Commun.*, vol. 4, pp. 477–488, Nov. 1992.

[33] S. Smoot and L. Rowe, "Study of DCT coefficient distributions," *Proc. SPIE*, vol. 2657, pp. 403–411, 1996.

[34] H. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 598 – 603, Jul. 2003.

[35] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," ITU-T SC16/Q.6 VCEG-M33, Apr. 2001. [Online]. Available: http://ftp3.itu.org/av-arch/video-site/0104_Aus/VCEG-M33.doc