Spectral Entropy-Based Bit Allocation

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 scheme to allocate bits amongst the chosen coefficients. We show that the proposed scheme can outperform the classical method under certain conditions. We then design quantization matrices (QMs) based on the proposed bit allocation method and 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).**

I. Introduction

Transform coding paradigms are very popular today in audio, image and video compression schemes because of their energy compaction properties. In applications 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 $X(t)$. The spectral entropy is the quantity in the exponent and is given by $h(S) = \log(Q)$. 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 [3] examined Campbell's results and observed that a compression scheme based on spectral entropy was not apparent.

Entropy-based indicators have been employed in adaptive coding applications independent of Campbell's work. In a two-dimensional discrete cosine transform (DCT) based coding scheme, Mester and Franke [4] used two factors to classify data blocks and adopted different coding strategies for the different classes. For two-dimensional transform coefficients $C(i,j)$, they defined the activity measure $A = \sum_{i,j}|C(i,j)|$ which reflects the total energy of the data block. The spectral entropy, $E = -\sum_{i,j} a(i,j)\log a(i,j)$ where $a(i,j) = |C(i,j)|/A$, was defined as a measure of the energy compaction achieved by the transform. These two measures were used 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. Coifman and Wickerhauser [5] applied a similar approach to select the best basis for signal representation. For transform coefficients x_n, they defined the theoretical dimension of the signal as $d = \exp\left(-\sum_n p_n \log p_n\right)$, where $p_n = |x_n|^2/\|x\|^2$. They interpreted d as a measure of the number of coefficients to be coded in a wavelet transform, and chose the wavelet packet basis with minimum d as optimal, since it produces the minimum number of coefficients. This method was extended to other transforms for speech processing in [6].

Around the same time, Gibson, Stanners, and McClellan [7] investigated the properties of Campbell's coefficient rate for autoregressive (AR) processes and speech signals. In speech experiments, the signal energy and spectral entropy were shown to be different indicators of activity, similar to the results of Mester and Franke [4] for transform coding of images. McClellan and Gibson [8] utilized spectral entropy for voiced/unvoiced decisions with applications to speech coding and showed that for the computation of spectral entropy in two subbands, the voice activity decision could be made robust to additive noise. In [8], McClellan and Gibson used more sophisticated variable rate indicators based on spectral entropy calculations to develop a variable rate tree coder operating at 4 to 5 kbits/s. These spectral entropy based rate indicators were then combined with a code-excited linear predictive (CELP) coder in [9] to achieve good quality speech at average rates in the 2 kbits/s range.

Almost thirty years after Campbell's work, Yang and Gib-

This research has been supported by NSF Grant Nos. CCF-0728646 and CCF-0917230 and by the UC Discovery Grant Program and Applied Signal Technology, Inc.

son [1], [10]–[13] 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. Another interesting outcome of Yang and Gibson's research was the interpretation of the Campbell bandwidth $W_c = Q/2$ as the minimum average bandwidth for encoding the process across all possible distortion levels and its relationship to the well-known Fourier [11] and Shannon bandwidth [13].

Kokes and Gibson [14] 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 biorthogonal transform (NMLBT) [15], [16]. The more precise frequency selectivity was shown to improve the performance of a wideband speech coder for both speech and audio signals.

The previous research discussed above uses Campbell's coefficient rate and spectral entropy as a basis for efficiently sampling frequency coefficients. Yang and Gibson [13] 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 [17] 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, no efforts have been made towards developing a spectral entropy-based coding scheme. Hence in this work, we build on the spectral entropy-based coefficient selection mechanism and derive a scheme that can be used to allocate bits amongst the chosen significant coefficients. We show that it is possible for the spectral entropy-based bit allocation method to out-perform the classical bit allocation scheme.

This paper is organized as follows. Section II briefly explains Yang and Gibson's coefficient selection mechanism for the sake of completeness. Section III develops the spectral entropy-based bit allocation method. Section IV compares the proposed method with the classical bit allocation method. Mathematical results used in this section, but not directly related to the work, are relegated to Appendix A. Section V discusses the application of the proposed scheme to design quantization matrices (QMs) for H.264/AVC video coding and presents results that show improvement over the default QMs. Finally, Section VI summarizes the work presented.

II. SPECTRAL ENTROPY-BASED COEFFICIENT SELECTION

In this section, we briefly go over the mathematical basis and a description of the spectral entropy-based coefficient selection proposed by Yang and Gibson [13]. 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 $\sigma^2 = \sum_{i=1}^{M} \lambda_i$. Assuming that $S(f)$, the PSD of $X(t)$ is normalized (integrates to 1), it can be shown that $\sum_{i=1}^{M} \lambda_i = T$.

Let $x_1(t_1), x_2(t_2), \ldots, x_N(t_N)$ be N independent sample functions of $X(t)$ where each sample function can be expressed as

$$x_i(t_j) = \sum_{i=1}^{M} c_{i_j} \phi_i(t_j), \qquad 0 \le t_j \le T,\ j = 1, 2, \ldots, N. \tag{3}$$

Therefore the product of these N independent sample functions can be written as

$$\begin{aligned} y(t_1, t_2, \ldots, t_N) &= x_1(t_1) x_2(t_2) \cdots x_N(t_N) \\ &= \sum_{k=1}^{M^N} c^{(k)} \phi^{(k)}(t_1, t_2, \ldots, t_N), \end{aligned} \tag{4}$$

where $c^{(k)}$ is the product of c_{i_j}'s and $\phi^{(k)}(t_1, t_2, \ldots, t_N)$ are the corresponding products of $\phi_i(t_j)$'s. Assume that the $c^{(k)}$'s are ordered in decreasing order of their variances in the sum in (4).

Campbell [2] approximated $y(t_1, t_2, \ldots, t_N)$ by choosing $\mu (< M^N)$ $c^{(k)}$'s with the largest variances such that the average energy loss of the approximation is small. The approximation to $y(t_1, t_2, \ldots, t_N)$ is

$$y_\mu(t_1, t_2, \ldots, t_N) = \sum_{k=1}^{\mu} c^{(k)} \phi^{(k)}(t_1, t_2, \ldots, t_N), \tag{5}$$

Since all the sample functions are independent, we have $C^{(k)} = C_{i_1} C_{i_2} \cdots C_{i_N}$ and

$$\begin{aligned} \mathrm{E}[(C^{(k)})^2] &= \mathrm{E}[(C_{i_1})^2] \mathrm{E}[(C_{i_2})^2] \cdots \mathrm{E}[(C_{i_N})^2] \\ &= \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_N}. \end{aligned} \tag{6}$$

Therefore, the total energy of the product is

$$\begin{aligned} \mathrm{E}[y(t_1, \ldots, t_N)^2] &= \sum_{k=1}^{M^N} \mathrm{E}[(C^{(k)})^2] \\ &= \sum_{i_1, i_2, \ldots i_N = 1}^{M} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_N} = \left(\sum_{i=1}^{M} \lambda_i \right)^N \\ &= \sum_{\{n_i : \sum_{i=1}^{M} n_i = N\}} \left[\frac{N!}{n_1! \cdots n_M!} \right] \lambda_1^{n_1} \cdots \lambda_M^{n_M}. \end{aligned} \tag{7}$$

In (7), the quantity in square brackets is the count of repetitions of the energy term that follows it. As $N \to \infty$, the largest term in the sum in (7) grows much faster than the others and dominates the total energy of $y(t_1, t_2, \ldots, t_N)$. Hence finding the largest term of $\frac{N!}{n_1! \cdots n_M!} \lambda_1^{n_1} \cdots \lambda_M^{n_M}$ subject to $\sum_{i=1}^{M} n_i = N$ would give us μ in (5) [13].

Using the approximation $\log N! = N \log N - N$ for large N and Lagrangian optimization, Yang and Gibson showed that for the largest term in (7)

$$n_i = \frac{\lambda_i}{\sigma^2} N, \qquad i = 1, 2, \ldots, M, \tag{8}$$

that is, the number of λ_i occurrences in $\lambda_1^{n_1} \lambda_2^{n_2} \cdots \lambda_M^{n_M}$ is proportional to λ_i [13].

Therefore for large N, μ, the number of high energy coefficients in $y(t_1, t_2, \ldots, t_N)$ is

$$\begin{aligned} \mu &\approx \frac{N!}{n_1! n_2! \cdots n_M!} \\ &= \exp\left[-N \sum_{i=1}^{M} \frac{n_i}{N} \log \frac{n_i}{N} \right] \\ &= e^{NH(S)} \end{aligned} \tag{9}$$

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.

This alternative derivation of Campbell's result provides insights into coder design and implies a method of selecting the coefficients that best represent a signal [13]. Equation (8) 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 variance 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 coefficients 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. In classical transform based coding, coefficient bit allocation is accomplished on a block-by-block basis as illustrated in Fig. 1(b). That is, given a particular block (fixed j), a fixed number of bits is allocated across the M coefficients according to their relative energies. 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 significant values of that component in the sequence should be determined based on its energy. In other words, 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 achieves better signal-to-noise ratio (SNR) (with no rate control) [10] and subjective quality [12].

III. Spectral Entropy-based 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 the components are ordered based on their energies *i.e.* $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$. 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 [1], [13] dictates that the number of coefficients n_i coded in each component be proportional to the variance λ_i of that component *i.e.* $n_i = p_i L$, where $p_i = \frac{\lambda_i}{\sigma^2}$ and $\sum_{i=1}^{M} \lambda_i = \sigma^2$.

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

$$\begin{aligned} d_i^{(S)} &= n_i \times \mathrm{E}(\text{quantization error}) + \\ &\qquad (N - n_i) \times \mathrm{E}(\text{energy of discarded coefficients}) \\ &= n_i \times h_i \lambda_i 2^{-2b_i^{(S)}} + (N - n_i) \times \lambda_i \end{aligned} \tag{10}$$

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 $C_i/\sqrt{\lambda_i}$ [18].

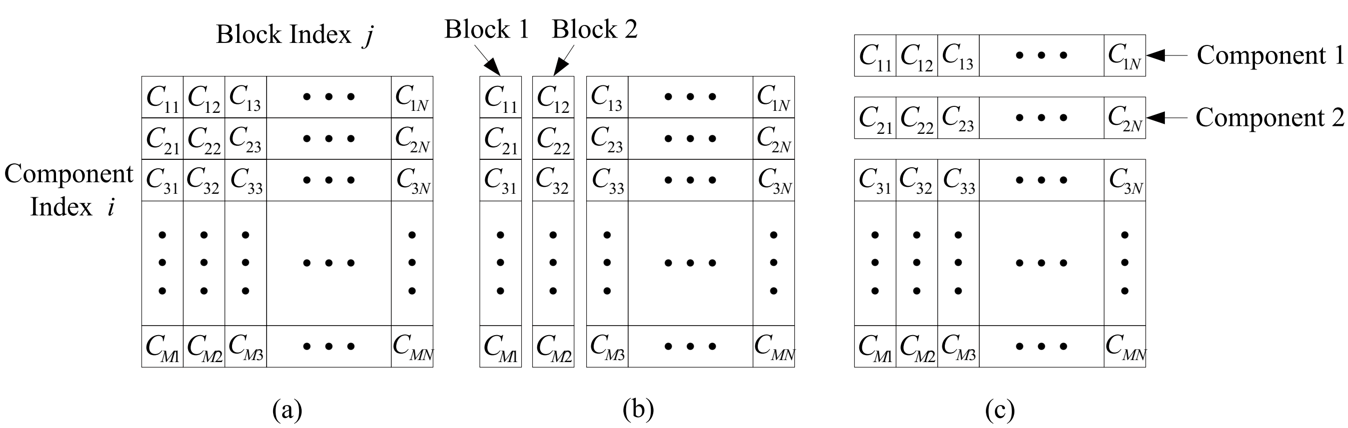


Fig. 1. Encoding transform coefficients. (a) Coefficients of N blocks, (b) Encoding block by block, (c) Encoding component by component.

Hence, the problem of bit allocation is to find $b_i^{(S)}$ for $i = 1, 2, \ldots, M$ so as to minimize $D^{(S)} = \sum_{i=1}^{M} d_i^{(S)}$ 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

$$b_i^{(S)} = \frac{B}{L} + \frac{1}{2}\log_2\left(\frac{h_i\lambda_i}{\prod_{i=1}^{M}(h_i\lambda_i)^{\frac{\lambda_i}{\sigma^2}}}\right). \tag{11}$$

This is similar to the result of classical bit allocation [18] except that the geometric mean of h_i's $\prod_{i=1}^{M} h_i^{\frac{1}{M}}$ has been replaced by $\prod_{i=1}^{M} h_i^{\frac{\lambda_i}{\sigma^2}}$ and the geometric mean of λ_i's $\prod_{i=1}^{M} \lambda_i^{\frac{1}{M}}$ has been replaced by $\prod_{i=1}^{M} \lambda_i^{\frac{\lambda_i}{\sigma^2}}$. The corresponding total distortion is

$$D^{(S)} = L2^{-\frac{2B}{L}}\prod_{i=1}^{M}(h_i\lambda_i)^{\frac{\lambda_i}{\sigma^2}} + N\sigma^2 - \sum_{i=1}^{M} n_i\lambda_i. \tag{12}$$

A. Bit-allocation scheme

Given N blocks with M components each (a total of MN coefficients) and a bit budget of B, the problem is to choose and code L coefficients such that the overall distortion is minimized. Summarizing the spectral entropy-based coding method, the steps involved in choosing significant coefficients and allocating bits to them are as follows.

1) Compute the empirical variance $\widehat{\lambda}_i$ of the ith component using N samples.
2) Compute the number of coded coefficients n_i in the ith component using $n_i = \frac{\widehat{\lambda}_i}{\widehat{\sigma}^2}L$ where $\widehat{\sigma}^2 = \sum_{i=1}^{M}\widehat{\lambda}_i$.
3) Treating the empirical variance as the true variance, compute $b_i^{(S)}$, the bits allocated to each coefficient of the ith component using (11). Alternatively, a practical non-negative integer constrained allocation method such as the one proposed in [19] can be used with the cost function as defined in (10).
4) Of the N coefficients of the ith component, choose n_i coefficients with the largest magnitudes and code each of them using $b_i^{(S)}$ bits.

However, for this scheme, we first need to determine L, the number of coefficients that need to be coded. L can be computed by starting with an estimate for the average bits per coded coefficient and then refining it around the initial estimate based on a rate-distortion cost. Alternatively, L can be computed as the total bit budget divided by the average bits per coded coefficient, where the average bits per coded coefficient is estimated as a function of the distribution of energy amongst the transform components and the function is determined using offline training.

IV. Comparison of spectral entropy-based and classical bit allocation schemes

In a classical bit allocation scheme, where B bits are to be allocated among the M components of N blocks, each block is coded independently. Hence, B/N bits are allocated to the M coefficients within a block. If C coefficients are coded out of the M coefficients in each block, the total number of coefficients that are coded becomes CN. Assuming that the coded components are those with the largest variances, the first C coefficients in each block are coded. The bits allocated to each coefficient of the ith component can be written as

$$b_i^{(C)} = \begin{cases} \frac{B}{CN} + \frac{1}{2}\log_2\left(\frac{h_i\lambda_i}{\prod_{i=1}^{C}(h_i\lambda_i)^{\frac{1}{C}}}\right), & \text{if } 1 \le i \le C \\ 0, & \text{if } C+1 \le i \le M. \end{cases} \tag{13}$$

Hence the total distortion for all the MN coefficients is the sum of the quantization error and the error from discarding coefficients and is given by

$$D^{(C)} = CN2^{-\frac{2B}{CN}}\prod_{i=1}^{C}(h_i\lambda_i)^{\frac{1}{C}} + N\sum_{i=C+1}^{M}\lambda_i. \tag{14}$$

Given a bit budget B, we compare the distortions of the two bit allocation schemes, assuming that the same number of coefficients are coded in both cases *i.e* $L = CN$. For easier analysis, we also assume that all the M components have the same normalized distribution *i.e.* $h_i = h$ for $i = 1, 2, \ldots, M$. Then, with $p_i = \frac{\lambda_i}{\sigma^2}$, the difference in the distortions per coded coefficient is given by

$$\frac{D^{(C)} - D^{(S)}}{CN} = 2^{-\frac{2B}{CN}}h\left[\prod_{i=1}^{C}\lambda_i^{\frac{1}{C}} - \prod_{i=1}^{M}\lambda_i^{p_i}\right] + \left[\sum_{i=1}^{M}p_i\lambda_i - \frac{\sum_{i=1}^{C}\lambda_i}{C}\right]. \tag{15}$$

Using (A.21) and (A.22),

$$\frac{D^{(C)} - D^{(S)}}{CN} \ge 2^{-\frac{2B}{CN}}h\underbrace{\left[\prod_{i=1}^{M}\lambda_i^{\frac{1}{M}} - \prod_{i=1}^{M}\lambda_i^{p_i}\right]}_{\le 0} + \underbrace{\left[\frac{\sum_{i=1}^{M}\lambda_i}{M} - \frac{\sum_{i=1}^{C}\lambda_i}{C}\right]}_{\le 0}. \tag{16}$$

Similarly, we have

$$\frac{D^{(C)} - D^{(S)}}{CN} \le 2^{-\frac{2B}{CN}}h\underbrace{\left[\prod_{i=1}^{C}\lambda_i^{\frac{1}{C}} - \prod_{i=1}^{M}\lambda_i^{\frac{1}{M}}\right]}_{\ge 0} + \underbrace{\left[\sum_{i=1}^{M}p_i\lambda_i - \frac{\sum_{i=1}^{M}\lambda_i}{M}\right]}_{\ge 0}. \tag{17}$$

From Equations (16) and (17), we observe that $D^{(C)} - D^{(S)}$ is lower bounded by a non-positive number and upper bounded by a non-negative number. Thus in certain cases $D^{(C)} \ge D^{(S)}$ *i.e.* the spectral entropy-based bit allocation scheme outperforms the classical bit allocation scheme. However, a closed form expression for the range of C under which $D^{(C)} \ge D^{(S)}$ is not evident from (15). Additionally, when $\lambda_i = \sigma^2/M$ for all $1 \leq i \leq M$ and $C = M$, the two bit allocation schemes become identical and $D^{(C)} = D^{(S)}$.

V. Application: Quantization Matrix Design for H.264/AVC Video Coding

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 [20].

In order to explore the practical implications of the proposed bit allocation scheme, we employ it to design QMs for the H.264 encoder. For each P frame, the residual transform coefficients of all the inter luminance blocks are buffered and used to design the luma inter 4×4 QM. The coefficients selected using the spectral entropy-based coefficient selection [13] are quantized using the designed QMs and finally entropy coded. We compare the coding performance with the proposed QMs with that of the default QMs based on peak SNR (PSNR) and structural similarity (SSIM) [21], a perceptual video quality metric.

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. Due to space restrictions in this paper, we limit our results to 2 video sequences at 176×144 (QCIF) resolutions: "container" (still camera on a slow moving scene) and "mobile" (complex motion with camera panning and zooming; high spatial and color detail). Fig. 2 plots the distortion of the luma component versus the average bits per frame for the encoder using the default QMs (reference method) and newly designed QMs (proposed method). Curves are provided using both PSNR and SSIM as distortion metrics. It can be seen that the proposed QMs perform better than the default QMs in terms of both PSNR and SSIM. A PSNR improvement of up to 1dB can be observed accompanied by a slight improvement in perceptual quality as indicated by SSIM. Fig. 3 provides the 39th frame of the QCIF "mobile" sequence encoded at QP = 35 using both the reference and proposed methods. It is evident that the frame compressed using the proposed method retains more details as seen in the inset around the numbers 21-23.

VI. Conclusions

This work derives and develops a bit allocation scheme based on the concepts of coefficient rate and spectral entropy. It has been shown that the proposed scheme can outperform the classical bit allocation method under certain conditions. An application of the proposed scheme to design QMs for a H.264 video encoder on a per-frame basis is discussed and shown to achieve better compression performance than the default QMs of H.264.

Appendix

A. Relation between various arithmetic and geometric means

The geometric mean(GM)-arithmetic mean(AM) inequality states that for positive real numbers $x_1, x_2, \ldots, x_n$

$$\prod_{i=1}^{n} x_i^{\frac{1}{n}} \leq \frac{\sum_{i=1}^{n} x_i}{n} \tag{A.18}$$

with equality when $x_i = x$ for all $i = 1, 2, \ldots, n$. The generalization of the GM-AM inequality states that for any $w_1, w_2, \ldots, w_n$ such that $\sum_{i=1}^{n} w_i = 1$

$$\prod_{i=1}^{n} x_i^{w_i} \leq \sum_{i=1}^{n} w_i x_i. \tag{A.19}$$

Hence, $\prod_{i=1}^{M} \lambda_i^{\frac{1}{M}} \leq \frac{\sum_{i=1}^{M} \lambda_i}{M}$ and $\prod_{i=1}^{M} \lambda_i^{p_i} \leq \sum_{i=1}^{M} p_i \lambda_i$, with equality in both cases iff $\lambda_i = \sigma^2/M$ for all $i = 1, 2, \ldots, M$. Recollect that $p_i = \lambda_i/\sigma^2$.

Additionally, using the Jensen's inequality on the $\log$ function that is convex $\cap$ (concave), we have

$$\begin{aligned} \log \prod_{i=1}^{M} \lambda_i^{p_i} &= \sum_{i=1}^{M} p_i \log p_i + \log \sigma^2 \\ &\geq (\sum_{i=1}^{M} p_i) \log \frac{\sum_{i=1}^{M} p_i}{\sum_{i=1}^{M} 1} + \log \sigma^2 \\ &= \log \frac{\sigma^2}{M} \end{aligned} \tag{A.20}$$

Hence $\prod_{i=1}^{M} \lambda_i^{\frac{\lambda_i}{\sigma^2}} \geq \sigma^2/M$ with equality iff $\lambda_i = \sigma^2/M$ for all $i = 1, 2, \ldots, M$.

Therefore,

$$\prod_{i=1}^{M} \lambda_i^{\frac{1}{M}} \leq \frac{\sum_{i=1}^{M} \lambda_i}{M} \leq \prod_{i=1}^{M} \lambda_i^{p_i} \leq \sum_{i=1}^{M} p_i \lambda_i, \tag{A.21}$$

with equality at all points iff $\lambda_i = \sigma^2/M$ for all $i = 1, 2, \ldots, M$.

Also, since $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$, it is evident that

$$\frac{(\sum_{i=1}^{C} \lambda_i)}{C} \geq \frac{\sum_{i=1}^{M} \lambda_i}{M} \text{ and } \prod_{i=1}^{C} \lambda_i^{\frac{1}{C}} \geq \prod_{i=1}^{M} \lambda_i^{\frac{1}{M}}. \tag{A.22}$$

References

[1] W. Yang and J. Gibson, "Coefficient rate in transform coding," in *Proceedings, 35th Annual Allerton Conf. on Communication, Control, and Computing*, Sept. 29-Oct. 1 1997, pp. 128–137.

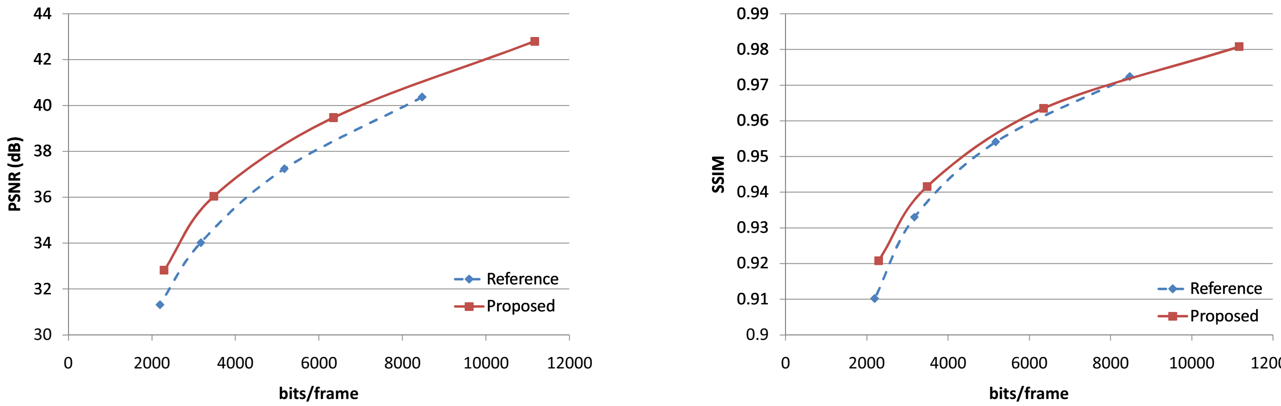
[2] L. L. Campbell, "Minimum coefficient rate for stationary random processes," *Information and Control*, vol. 3, no. 4, pp. 360 – 371, 1960.

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

[4] R. Mester and U. Franke, "Spectral entropy-activity classification in adaptive transform coding," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 5, pp. 913–917, Jun 1992.

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

[6] E. Wesfreid and M. Wickerhauser, "Adapted local trigonometric transforms and speech processing," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3596–3600, Dec 1993.



(a) (b)

Fig. 2. Comparison of performance for QCIF "container" sequence. (a) PSNR vs. bits/frame (b) SSIM vs. bits/frame

(a) (b) (c)

Fig. 3. Frame 39 of QCIF "mobile" sequence encoded at QP = 35 with inset showing an enlarged portion. (a) Original (b) Encoded using reference method (c) Encoded using proposed method

[7] J. Gibson, S. Stanners, and S. McClellan, "Spectral entropy and coefficient rate for speech coding," in *Conference Record of Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 2, Nov 1993, pp. 925–929.

[8] S. McClellan and J. Gibson, "Spectral entropy: an alternative indicator for rate allocation?" in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. i, Apr 1994, pp. I/201–I/204.

[9] ——, "Variable-rate CELP based on subband flatness," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 2, pp. 120–130, Mar 1997.

[10] W. Yang and J. Gibson, "Coefficient rate and significance maps in transform coding," in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, vol. 2, Nov 1997, pp. 1373–1377.

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

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

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

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

[15] M. Kokes, J. Gibson, and G. Schuller, "A wideband speech codec based on nonlinear approximation," in *Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers*, vol. 2, 2001, pp. 1573–1577.

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

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

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

[19] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 36, no. 9, pp. 1445 –1453, sep 1988.

[20] 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," A. G. Tescher, Ed., vol. 5558, no. 1. SPIE, 2004, pp. 454–474. [Online]. Available: http://link.aip.org/link/?PSI/5558/454/1

[21] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600 –612, april 2004.