Analytical Complexity Modeling of Wavelet-based Video Coders

Brian Foo, Yiannis Andreopoulos, and Mihaela van der Schaar

University of California Los Angeles (UCLA), Dept. of Electrical Engineering (EE), Los Angeles, CA, 90095
Tel: +1-310-825-5843, Fax: +1-310-206-4685, Email: {bkungfoo , yandreop, mihaela}@ee.ucla.edu

ABSTRACT

Analytical modeling for video coders can be used in a variety of scenarios where information concerning rate, distortion or complexity is essential for driving system or network interactions with media algorithms. While rate and distortion modeling have been covered extensively in previous works, complexity is not well addressed because it is highly algorithm dependent and hence difficult to model. Based on a stochastic modeling framework for the transform coefficients, we present a novel complexity analysis for state-of-the-art wavelet video coding methods by explicitly modeling several aspects found in operational coders, i.e. embedded quantization and quadtree decompositions of block significance maps. The proposed modeling derives for the first time analytical estimates of the expected number of operations (complexity) for a broad class of wavelet video coders based on stochastic source models, coding algorithm and system parameters.

Index Terms: Multimedia Systems, Resource Modeling, Statistical Modeling of Media Decoding Complexity

I. INTRODUCTION

Recent compression algorithms such as the H.264/AVC standard and wavelet video compression [1] achieve breakthroughs in terms of rate-distortion (R-D) performance at the expense of a significant increase in complexity (in terms of CPU time, or energy dissipation) compared to older coding schemes such as MPEG-1 or MPEG-2. Hence, accurate models that encapsulate the source, algorithm *and system characteristics* are very important for benchmarking existing video coders and facilitating the design of future video coders.

Two methods have been used to determine the complexity characteristics of operational video coders. The first is an empirical approach, where analytical formulations are fitted to experimental data to derive an operational model for a particular class of video sequences, instantiation of a compression algorithm, and a fixed architecture [2]. While this modeling approach is simple, fine-granular adaptation of algorithm or system parameters is not possible, since one can not predict the expected complexity for a different input video source or compression algorithm configuration. This led to current state-of-the-art multimedia compression algorithms and standards providing only very coarse levels (profiles) of complexity [2] and hence quality, thereby neglecting the vast resource diversity and heterogeneity of devices and systems.

The second approach is a theoretical approach, where a stochastic model is used for each pixel or transform coefficient. While several works have modeled complexity using operational source statistics and offline or online training to estimate (learn) the algorithm and system parameters [4], to the best of the authors' knowledge, scarcely any work has addressed the *information-theoretic modeling of complexity* in function of *stochastic source models* and *practical algorithm characteristics*.

In this paper, we follow the second approach to predict complexity in terms of the number of certain operations performed (e.g. the number of symbols read from the bitstream), thereby complementing prior work on information-theoretic R-D modeling [3]. We focus mainly on the quantization and coding process of intra and error frames and present for the first time a stochastic framework for complexity prediction of entropy decoding and the inverse spatial transform in a broad class of wavelet video coders based on easily-obtained source, algorithm and system parameters.

The paper is organized as follows. Section II introduces the coder, a model for wavelet coefficients, and some important nomenclature. Based on these models, in Section III we derive probability estimates for a variety of coding/decoding operations that will be used to determine the complexity (Section IV) for decoding a video sequence. Section V displays theoretical and experimental complexity-quality results that validate the proposed models. Section 0 concludes the paper.

II. CODER STRUCTURE AND WAVELET COEFFICIENT MODELING

A. Coder Structure

Recent state-of-the-art scalable video coding schemes are based on motion compensated temporal filtering (MCTF) [1]. During MCTF, the original video frames are filtered temporally in the direction of motion, prior to performing the spatial transformation and coding. Video frames are filtered into L (low-frequency) and H (high-frequency) frames [1]. The process is recursively applied to subsequently-produced L frames to form a total of T_{MCTF} temporal levels. The derived L and H temporal frames are spatially decomposed in a hierarchy of spatio-temporal subbands.

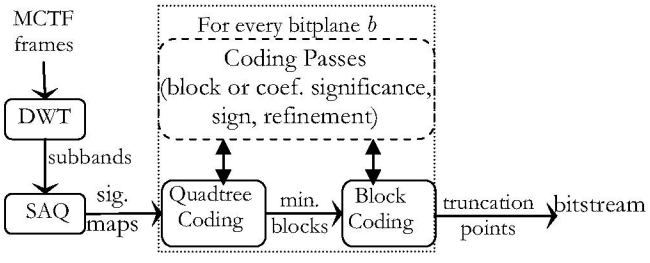

Figure 1. Block diagram of intra-band coding process of state-of-the-art wavelet-based coders encompassing quadtree and block coding of the significance maps.

In all state-of-the-art wavelet coders [5], the coding process exploits intra-band dependencies following a block-partitioning process within each transform subband. A generalized form of this partitioning for every bitplane b is outlined in Figure 1. As indicated there, several coding passes that identify coefficient significance or refine wavelet coefficients with respect to the current SAQ threshold are performed within quadtree coding or within block coding.

B. Wavelet Coefficient Modeling of Spatio-temporal Subbands

Low-frequency wavelet coefficients are typically modeled using independent Gaussian random variables after subtracting the mean value [3] [5]. High frequency wavelet coefficients are often modeled by decorrelated, but non-independent random variables X using a doubly-stochastic process, i.e. a Gaussian distribution parameterized by Θ, which follows a marginal distribution of a Laplacian random variable with variance σ^2 [3] [6]:

$$\Theta \sim p(\theta) = 1/\sigma^2 \cdot \exp\left(-\theta/\sigma^2\right) \quad (1)$$

$$p(x) = \int p(x \mid \theta)p(\theta)d\theta = 1/\sqrt{2}\sigma \cdot \exp\left(-\mid x \mid \sqrt{2}/\sigma\right) \quad (2)$$

The results of Table 1 show an instantiation of this model fitted to a real video sequence. We conclude that the coarser spatio-temporal high-frequency subbands exhibit significant variance and the correlation of the subband statistics (parameter θ) varies significantly as well. Consequently, contrary to the notion that only the low-frequency subbands are essential for complexity prediction, high-frequency subbands contain a significant portion of complexity for a variety of quantization thresholds, and hence an accurate model is important for predicting the overall complexity.

Temporal (T)-Spatial (S) level	Subband variance σ^2, [variance of θ]		
	LH	*HL*	*HH*, *LL* (if exists
2T-2S	6.59, [37.5]	5.55, [22.8]	4.46, [25.7]
4T-4S	39.69, [241]	33.23, [496]	{25.98,[144]}, {53.12, 690]}

Table 1. Examples of subband variances as well as the variance of the correlation θ (for a block of 4×4) formed across the spatio-temporal MCTF decomposition of sequence Foreman.

We now introduce some important nomenclature. Denote the minimum decoded bitplane threshold level as $T_{B_{\min}} = 2^{B_{\min}}$. In addition, we define the following parameters for all bitplanes:

$$\upsilon_b = T_b/\sigma\,,\ \rho_b = e^{-\sqrt{2}\upsilon_b} \quad (3)$$

where υ_b describes the ratio of the threshold of bitplane b to the variance of each wavelet coefficient and ρ_b is the probability of significance of a wavelet coefficient under a certain υ_b under the model of (2). In this paper we analyze intra-band coders that use quadtrees to decompose subbands into non-overlapping blocks of dyadically-decreasing sizes and then encode the minimum block size using context-based adaptive arithmetic coding. The initial subbands are hierarchically split in K quadtree levels, with blocks at quadtree level K having the smallest size. If a block at quadtree level k, $2 \le k \le K$, has n coefficients, its parent block at level $k-1$ has $4n$ coefficients. We define the significance test of a block of n coefficients with respect to a threshold T_b as $\mathrm{sig}(\upsilon_b, n) = \{0,1\}$. We also define the newly significance test as $\mathrm{newsig}(\upsilon_b, n) = \{0,1\}$, which returns one if the block was found to be significant at bitplane b and insignificant at bitplane $b+1$, i.e. $\mathrm{sig}(\upsilon_b, n) = 1$ and $\mathrm{sig}(\upsilon_{b+1}, n) = 0$. For notational abbreviation, the probability of a block being significant or newly-significant at bitplane b is indicated by $\chi^{\mathrm{band}}_{\upsilon_b,n}$ and $\delta^{\mathrm{band}}_{\upsilon_b,n}$, respectively, with $\mathrm{band} = \{\mathrm{low},\mathrm{high}\}$ indicating the frequency subband that the block belongs to.

III. Approximation of Block Significance Probabilities in Quadtree Decompositions

Under the stochastic modeling framework, we derive several important probabilities of significance for quadtree decompositions of quantized spatio-temporal subbands. These probabilities form the core of the complexity estimation as they provide the means of establishing the percentage of blocks that are expected to be coded or decoded at a given terminating SAQ threshold $T_{B_{\min}}$. In addition, the percentage of significant blocks within the spatio-temporal subbands along with the percentage of non-zero coefficients are the two features that express the complexity of the inverse DWT [4].

A. Probability of Block Significance at Bitplane b

Let us first consider a high-frequency spatio-temporal subband, which may be any subband of an error (H) frame, or any high-frequency subband of an L frame. Assuming the variance of the subband coefficients to be σ^2, we have:

$$\Pr\{\mathrm{sig}(\upsilon_b, n) = 1\} = 1 - \Pr\{\mid \mathbf{X} \mid_\infty \le T_b\} \quad (4)$$

where $\mathbf{X} = (X_1, X_2, ..., X_n)$ is a length-n random vector of all the coefficient random variables X_i ($1 \le i \le n$) of a block, and $\mid \bullet \mid_\infty$ is the $L^{\inf}$ norm. Considering that block sizes are generally small enough to capture local variances, we follow the doubly stochastic model in equation (3) to derive the conditional joint distribution of $\mathbf{X}$:

$$\begin{aligned} p(\mathbf{X}) &= \int_{0+}^{\infty} p(\mathbf{X} \mid \theta)p(\theta)d\theta \\ &= \int_{0+}^{\infty} 1/\sigma^2 \cdot e^{-\frac{1}{\sigma^2}\theta} \cdot (2\pi\theta)^{-n/2} e^{-\frac{1}{2\theta}(x_1^2 + x_2^2 + ... + x_n^2)} d\theta \end{aligned} \quad (5)$$

Proposition 1: The probability that a block of size n is significant compared to threshold T_b can be approximated by:

$$\chi^{\mathrm{high}}_{\upsilon_b,n} \triangleq \Pr\{\mathrm{sig}(\upsilon_b, n)=1\} \cong \exp\left\{\upsilon_b^2/k(n)\right\} \quad (6)$$

with $k(n) = \ln(n)^{1.296} + 0.166$.

Proof: See [8]. ■

For low-frequency subbands, the probability of block significance is simply the n-dimensional Gaussian tail probability along one of the orthogonal axes:

$$\chi^{\mathrm{low}}_{\upsilon_b,n} \triangleq \Pr\{\mathrm{sig}(\upsilon_b, n) = 1\} \cong \left[\mathrm{erfc}\left(\upsilon_b/\sqrt{2}\right)\right]^n \quad (7)$$

B. Probability of a Newly Significant Block at Bitplane b

In order to model the number of operations performed during the quadtree significance pass at each bitplane, it is necessary to derive the probability that a block is found significant at bitplane b, but not at any higher bitplanes.

Proposition 2: The probability that a block of n coefficients in a high-frequency subband is found significant at bitplane b, but it is insignificant at bitplane $b+1, \ldots, B_{\max}$ is:

$$\delta^{\mathrm{high}}_{\upsilon_b,n} \cong \chi^{\mathrm{high}}_{\upsilon_b,n}(1 - \chi^{\mathrm{high}}_{\upsilon_b,n}) \quad (8)$$

Proof: See [8]. ■

Proposition 3: The probability that a block of n coefficients in a low-frequency subband is found significant at bitplane b, but it is insignificant at bitplane $b+1,\ldots,B_{\max}$ is:

$$\delta^{\text{low}}_{\upsilon_b,n} \cong (1-\chi^{\text{low}}_{\upsilon_{b+1},n})\chi^{\text{low}}_{\upsilon_b,n} \tag{9}$$

Proof: The approximation of (9) is a straightforward result of independent Gaussian coefficients. ■

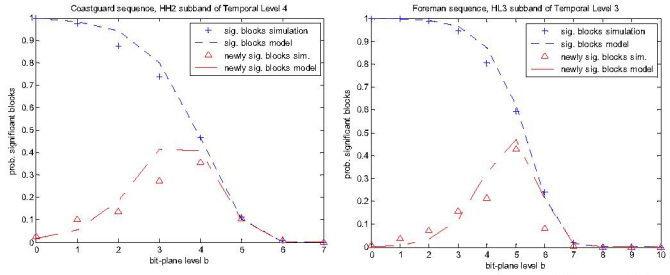

Figure 2. Simulation and model prediction of $\chi^{\text{high}}_{\upsilon_b,16}$, $\delta^{\text{high}}_{\upsilon_b,16}$ (4×4 blocks in high-frequency spatio-temporal subbands).

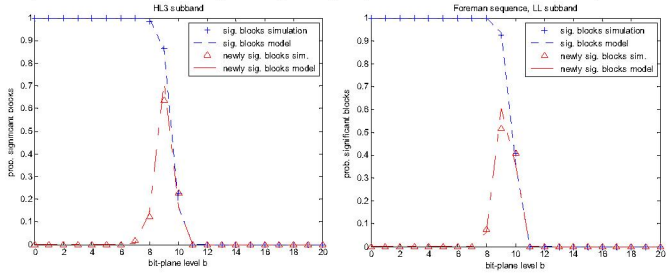

Figure 3. Simulation and model prediction of $\delta^{\text{low}}_{\upsilon_b,16}$, $\delta^{\text{low}}_{\upsilon_b,16}$ (4×4 blocks in LL subbands of L frames).

Figure 2 and Figure 3 demonstrate the accuracy of the model prediction of probability of block significance and newly-significance (with $n=16$) for several subbands belonging to two MCTF-decomposed frames of video sequences. Note that the high-frequency subbands exhibit a heavy-tail for significance (and newly-significance) between bitplanes $b=\{1,6\}$; hence the complexity of high-frequency subband coding is significant at finer quantization levels.

IV. Complexity of Entropy Decoding and IDWT

We model the complexity of decoding in terms of the number of symbols read from the entropy decoder (Subsection A and B) since predicting the symbol encoding/decoding operations captures the complexity of entropy coding implementations in real processor-based designs in an accurate manner [4]. Similarly, the complexity of inverse spatial DWT is modeled as a decomposition to a pair of functions relating to the sparsity of each subband's decoded wavelet coefficients (Subsection C).

A. Quadtree Decoding Complexity

The complexity of decoding the quadtree significance at bitplane b depends on the size of the quadtree before the significance pass. The significance of a block in the quadtree decomposition may be encoded in two cases: *i*) If the block is found newly significant at bitplane b, its significance will be encoded at that moment and it will never be encoded again; *ii*) if the block's parent is found to be significant at bitplane b even though the block itself is non-significant, it will be coded continuously until it is found newly significant. Notice that condition (*ii*) is added in most state-of-the-art coders to exploit the property of intra-band spatial correlation of coefficients.

Under the above-stated two conditions, the probability that the significance of a block of size n is coded at bitplane b and that its parent's significance is coded at bitplane $b+r$, $r\geq 0$, (which means that the block significance will be coded a total of $r+1$ times) can be formulated as:

$$C_{\text{block\_sig}}(\text{newsig}(\upsilon_b,n)) = \sum_{r=0}^{B_{\max}-b}\Pr\{\text{sig}(\upsilon_{b+r},4n)=1 \mid \text{newsig}(\upsilon_b,n)=1\} \tag{10}$$

Averaging over all bitplanes $b+r, b+r-1,\ldots,b$, we get the following rate estimate:

$$C_{\text{block\_sig}}(\upsilon_b,n) = \sum_{\beta=b}^{B_{\max}}\delta^{\text{band}}_{\upsilon_\beta,n}\sum_{r=0}^{B_{\max}-\beta}\Pr\{\text{sig}(\upsilon_{\beta+r},4n)=1 \mid \text{newsig}(\upsilon_\beta,n)=1\} \tag{11}$$

where band $=\{\text{low,high}\}$ depending on the frequency subband of interest. $\Pr\{\text{sig}(\upsilon_{b+r},4n)=1 \mid \text{newsig}(\upsilon_b,n)=1\}$ can be obtained using Bayes' rule:

$$\Pr\{\text{sig}(\upsilon_{b+r},4n)=1 \mid \text{newsig}(\upsilon_b,n)\} = \Pr\{\text{newsig}(\upsilon_b,n)=1 \mid \Theta\}\chi_{\upsilon_{b+r},n}\big/\delta_{\upsilon_b,n} \tag{12}$$

where:

$$\Theta \sim p(\theta \mid \text{sig}(\upsilon_{b+r},4n)) = 1\big/\chi^{\text{band}}_{\upsilon_{b+r},4n}\cdot\left(1-\left[\text{erf}\left(T_{b+r+1}\big/\sqrt{2\theta}\right)\right]^{4n}\right)1\big/\sigma^2\cdot e^{-\frac{\theta}{\sigma^2}} \tag{13}$$

Using similar approximations as in [8], we obtain the following (details can be found in [8]):

$$\Pr\{\text{newsig}(\upsilon_b,n)=1 \mid \Theta\} \triangleq \gamma^{\text{band}}_{\upsilon_b,n,r} \tag{14}$$

$$\gamma^{\text{band}}_{\upsilon_b,n,r} \cong \begin{cases}\dfrac{1}{\delta^{\text{band}}_{\upsilon_b,n}}\left(\chi^{\text{band}}_{\upsilon_{b+r},4n}-\chi^{\text{band}}_{\upsilon_{b+1},n}\right)^+, & \dfrac{T_b^2}{k(n)}\leq\dfrac{T_{b+r}^2}{k(4n)}\\ 1, & \text{otherwise}\end{cases} \tag{15}$$

Combining (11)–(15) together, we obtain the final expression:

$$C_{\text{block\_sig}}(\upsilon_b,n) = \sum_{\beta=b}^{B_{\max}}\sum_{r=0}^{B_{\max}-\beta}\chi^{\text{band}}_{\upsilon_{\beta+r},n}\gamma^{\text{band}}_{\upsilon_\beta,n,r} \tag{16}$$

The average rate per coefficient is $C_{\text{block\_sig}}(\upsilon_b,n)/n$. If we let n be the smallest block size, then summing up the rates for K levels of the quadtree decomposition gives the total rate for quadtree encoding within the subband:

$$C_{\text{quadtree}}(\upsilon_b) = \sum_{k=0}^{K}C_{\text{block\_sig}}(\upsilon_b,4^k n)\big/\left(4^k n\right) \tag{17}$$

B. Block and Refinement Decoding Complexity

We group together the number of symbols read from significance coding and refinement, since a coefficient will be significance-coded or refined at bitplane b as long as the it is in a significant block at bitplane b or higher. The sign is also encoded once when a coefficient is significant, which occurs with probability ρ_b at bitplane b. Summing up all symbols read in the passes down to $B_{\min}$:

$$C_{\text{coef}}(\upsilon_{B_{\min}}) = 4^K n\left(\rho^{\text{low}}_{B_{\min}}+\sum_{b=B_{\min}}^{B_{\max}}\chi^{\text{low}}_{\upsilon_b,n}\right) + 4^K n\left(\rho^{\text{high}}_{B_{\min}}+\sum_{b=B_{\min}}^{B_{\max}}\chi^{\text{high}}_{\upsilon_b,n}\right) \tag{18}$$

Since each subband is encoded independently, the complexity metrics must first be estimated for each subband and then summed in the same weighted fashion as the rate calculation. In

other words, for a given frame i, $1 \le i \le N$, we have:

$$C^i_{\mathrm{op}}(v_{B_{\min}}) = 4^{-J} C_{\mathrm{op},J,0}(v_{B_{\min}}) + \sum_{j=1}^{J}\sum_{k=1}^{3} 4^{-j} C_{\mathrm{op},j,k}(v_{B_{\min}}) \quad (19)$$

where $\mathrm{op} = \{\mathrm{quadtree}, \mathrm{coef}\}$ and $C_{\mathrm{op},j,k}(v_{B_{\min}})$ is the quadtree and block coding complexity for each subband at spatial resolution j. A linear combination of block and quadtree complexity metrics $a_1 C^i_{\mathrm{quadtree}}(v_{B_{\min}}) + a_2 C^i_{\mathrm{block}}(v_{B_{\min}})$ estimates the total number of RS operations for frame i.

C. Complexity of the Inverse Spatial DWT

We model the transform-related complexity of a coding system that processes N video frames by expressing it as a decomposition into two functions relating to: *i*) the percentage of non-zero coefficients for a given SAQ threshold T_b (function $\mathcal{T}_{\mathrm{nonzero}}$); *ii*) the sum of run-lengths of zero wavelet coefficients (function $\mathcal{T}_{\mathrm{runlen}}$). The motivation behind (*i*) is that the number of non-zero multiply-accumulate (MAC) operations in the synthesis filter-bank is directly proportional to the percentage of non-zero coefficients. Moreover, the distribution of zero run-lengths within the transform subbands affects the number of consecutive filtering operations that can be avoided altogether. The complexity of the inverse spatial DWT (non-zero MAC operations) can be formulated as:

$$\mathrm{FC}^N = \mathbf{C}^N_{\mathrm{nonzero}} \boldsymbol{\mathcal{T}}^N_{\mathrm{nonzero}} + \mathbf{C}^N_{\mathrm{runlen}} \boldsymbol{\mathcal{T}}^N_{\mathrm{runlen}} + \mathbf{C}^N_{\mathrm{dec\_const}} \mathbf{1} \quad (20)$$

with $\boldsymbol{\mathcal{T}}^N_{\mathrm{nonzero}}$ and $\boldsymbol{\mathcal{T}}^N_{\mathrm{runlen}}$ the N-element vectors of the corresponding functions. The parameter vectors $\mathbf{C}^N_{\mathrm{nonzero}}$ and $\mathbf{C}^N_{\mathrm{runlen}}$ can be estimated based on linear MMSE fitting over the actual number of MAC operations and model-based $\boldsymbol{\mathcal{T}}^N_{\mathrm{nonzero}}$ and $\boldsymbol{\mathcal{T}}^N_{\mathrm{runlen}}$. $\mathcal{T}_{\mathrm{nonzero}}$ for the high-frequency spatio-temporal subbands is derived by (3), while for the low-frequency spatio-temporal subbands it is derived by:

$$\mathcal{T}_{\mathrm{nonzero}} = \mathrm{erfc}\left(\frac{v_{B_{\min}}}{\sqrt{2}}\right) \quad (21)$$

In addition, $\mathcal{T}_{\mathrm{runlen}}$ is derived by the percentage of non-significant blocks for a certain SAQ threshold $T_{B_{\min}}$, given by:

$$\mathcal{T}_{\mathrm{runlen}} = \Pr\{\mathrm{sig}(v_{B_{\min}}, n) = 0\} = 1 - \chi^{\mathrm{band}}_{v_{B_{\min}}, n} \quad (22)$$

with $\chi^{\mathrm{band}}_{v_{B_{\min}}, n}$ estimated by (6) for the high-frequency temporal subbands and by (7) for the LL subband of the L frames. Following the lifting dependencies of popular wavelet filter-pairs, we set an average of $n = 64$ since a window of 7×7 coefficients and 9×9 coefficients is used in the lifting steps of the inverse DWT [5].

V. Experimentation and Results

We validate the derived analytical complexity expressions above by presenting experiments with two common interchange format (CIF) resolution sequences ("Coastguard", "Foreman") that encapsulate a variety of motion and texture characteristics using the coder in [7]. Figure 4 and Figure 5 present our results for a variety of spatial (S) and temporal (T) decomposition levels. The results indicate that the proposed complexity modeling predicts the experimental behavior of the advanced MCTF-based wavelet video coder accurately for the different cases under investigation. Note that different coding parameters, such as the number of spatio-temporal levels, can lead to significant tradeoffs between entropy decoding and inverse transform complexity.

VI. Conclusions

This paper presents an analytical modeling framework that derives complexity predictions for wavelet-based video coders. By analytically deriving probabilities for block and coefficient significance according to the quantization threshold, we derived analytical models that approximate well the complexity behavior of a wide variety of modern wavelet-based video coder. In this way, this work bridges the gap between the operational measurements used in prior complexity modeling work and information-theoretic estimates common in rate-distortion modeling work.

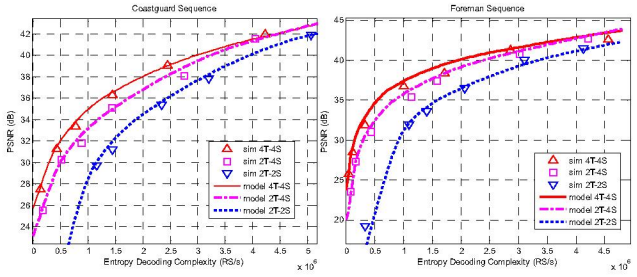

Figure 4. Entropy decoding complexity vs. distortion plots for different spatio-temporal decomposition parameters; "S" and "T" indicate the number of spatial and temporal levels (respectively).

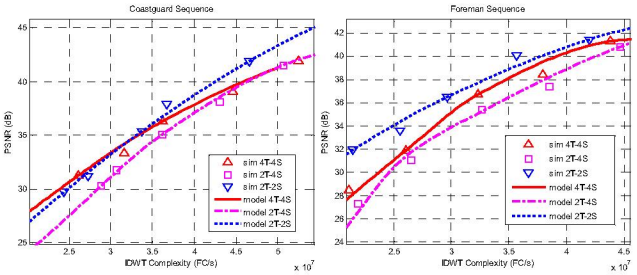

Figure 5. IDWT complexity vs. distortion plots for different spatio-temporal decomposition parameters; "S" and "T" indicate the number of spatial and temporal levels (respectively).

References

[1] J.-R. Ohm, "Advances in scalable video coding," *Proc. of the IEEE*, vol. 93, pp. 42-56, Jan. 2005.

[2] J. Valentim, P. Nunes and F. Pereira, "Evaluating MPEG-4 video decoding complexity for an alternative video verifier complexity model", *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 12, no. 11, pp. 1034-1044, Nov. 2002.

[3] M. Wang, M. van der Schaar, "Operational rate-distortion modeling for wavelet video coders," *IEEE Trans. on Signal Proc,*, to appear.

[4] M. van der Schaar and Y. Andreopoulos, "Rate-distortion-complexity modeling for network and receiver aware adaptation," *IEEE Trans. on Multimedia*, vol. 7, no. 3, pp. 471-479, June 2005.

[5] D. Taubman, M. W. Marcellin, JPEG 2000-Image Compression Fundamentals, Standards and Practice, *Kluwer Academic Publishers*, 2002.

[6] M. K. Mihçak, I. Kozintsev, K. Ramchandran, P. Moulin, "Low-complexity Image Denoising based on Statistical Modeling of Wavelet Coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300-303, 1999.

[7] Y. Andreopoulos, A. Munteanu, J. Barbarien, M. van der Schaar, J. Cornelis and P. Schelkens, "In-band motion compensated temporal filtering," *Signal Processing: Image Communication*, vol. 19, no. 7, pp. 653-673, Aug. 2004.

[8] B. Foo, Y. Andreopoulos, M. van der Schaar. "Analytical Rate-Distortion-Complexity Modeling of Wavelet-based Video Coders." Submitted to IEEE Trans. on Signal Processing.