Comparison of Traffic and Quality Characteristics of Rate-Controlled Wavelet and DCT Video

Beshan Kulapala, Patrick Seeling, Martin Reisslein
Dept. of Electrical Engineering
Arizona State University
Tempe, Arizona 85287–5706
{beshan|patrick.seeling|reisslein}@asu.edu

***Abstract*— Wavelet-based encoding is now emerging as an efficient way to encode video for streaming over the Internet and for wireless applications. Wavelet-based video coding has been recently added to the JPEG-2000 video standards. As wavelet encoded video emerges as the next generation video encoding method, it is vital to compare the efficiency of wavelet encoded video against the widely used DCT-based MPEG encoded video. However, due to the lack of long wavelet encoded video streams, most research has so far been based on short video traces. This paper presents a comparison study on MPEG vs wavelet encoded video traces for one hour movie excerpts with rate control. These long video sequences allow for the evaluation of long range dependency and self similarity of the generated video traffic, which has not been studied before in the context of comparing DCT and wavlet-based encoding. We focus on the elementary as well as self-similar traffic characteristics of the encoded video. A *hump* behavior for the variability of frame sizes is observed for increasing video bit rates for both wavelet and MPEG encoded video. In addition, the quality characteristics of the encoded video is examined and related to the traffic. Our results indicate that the wavelet encoded video results in higher video quality than MPEG encoded video. For the frame size variability we find different characteristics depending on the aggregation level for a given data rate. The results also indicate that the variation of quality resulting from the wavelet encoding is lower than for the MPEG encoded video.**

I. Introduction

Video streaming over the Internet and for wireless applications is now exploding as the next great forefront of networking. Different video compression techniques have been standardized to compress the video data for more efficient transport over the network. Recently, wavelet transform based image and video coding techniques have emerged. The wavelet transform has many advantages over the DCT transform. The most obvious of them is the compact-support feature which makes it possible to translate a time-domain function into a representation that is not only localized in frequency, but in time as well. The net result of this is that the wavelet transform can be conducted over the entire image with reasonable computation and bit budgets. The DCT based transform, on the other hand, requires to window the source data typically into blocks of 8x8 pixels to meet similar budgets. Thus, the obvious visual advantage is that *block artifacts* common in DCT based transforms are eliminated in the wavelet transform.

As wavelet encoded video emerges as the next generation video encoding method, it is vital to compare the efficiency of wavelet encoded video against the currently widely used DCT-based MPEG encoded video. MPEG encoded video typically exploits temporal redundancy by employing motion estimation and compensation techniques across frames. Similar approaches have been developed in the wavelet domain, but currently lack standardization. This inter coding introduces dependencies of subsequent frames, as well as buffer requirements and error-spreading across multiple frames that reference a damaged video frame. Intra coding, on the other hand, only codes individual video frames without referencing other video frames. Intra coding typically results in larger frame sizes than inter coding, but size differences between frame types and frame size variability are reduced.

With the increased utilization of wireless networks, low complexity and high error robustness are desired features of a video (de-)compression and transmission system [1]. By employing only intra coding (and no inter coding), error propagation and hardware complexity are minimized. An additional application scenario for the streaming of intra coded video is the streaming of encoded video for later transcoding. In order to evaluate the impact of the different video encoding methods, the traffic and quality characteristics for long video sequences have to be evaluated. However, due to the lack of long wavelet encoded video streams, most research has so far been based on short video sequences for wavelet encoded video. Long video sequences of actual movie content, on the other hand, have not been studied to date. The traffic and quality characteristics of long video sequences in general are different from short scene-based sequences, as they represent the aggregation of different scenes and movie content. Most importantly it is impossible to study the effects of long range dependency and self similarity of video traffic with short video traces. In this paper we present a comparison study on DCT-based MPEG-4 vs. wavelet-based MC-3DEZBC encoded video characteristics based on one hour long video traces generated from two movies.

This paper is structured as follows. In the remainder of this section, we introduce related work. In Section II, we describe the generation of the encoded video and video traces which were used in the statistical analysis. In Section III we examine the traffic characteristics for both encoders. In Section IV, we compare the quality statistics of the two encoding methods. We conclude summarizing in Section V.

A. Related Work

The wavelet transform has become popular in both, image and video coding domains. Several video coders that utilize the wavelet transform have been implemented and evaluated. Furthermore the video coders have been extended to a 3-D video codec, decomposing the frames spatially and the sequence of frames temporally. Nanda *et al*. [2] have evaluated the three short sequences *Clair, Suzie,* and *Misam,* and the results show the compression ratio and the PSNR only for the 15th frame of those three sequences. They compressed the three sequences using inter frame encoding in the wavelet domain. The motion vectors were coded using a zerotree, while the wavelet coefficients were coded using Set Partitioning in Hierarchical Trees (SPIHT). In [3] and [4], *Mobile Calendar* and *Flower Garden* have been used to evaluate the codecs, and the results only show the average PSNR for the entire test sequence. It has been common in studies and evaluations of wavelet transform codecs to only use short sequences to test the codecs and evaluate the codec in terms of the video quality, for example both [5] and [6] have used *Akiyo, Foreman, Coastguard, News*, and *Hall Monitor* sequences; [7] used *Mall* and *MIT* sequences (results show only up to 90 frames); [8] used *Clair, Foreman* and *Miss America* sequences; [9] used *Mobile Calendar, Table Tennis* and *Flower Garden*. In [5], the authors compare the quality for DCT- and wavelet-based video coding, but results are only provided for short sequences of 96 frames and for low target bit rates of 16 to 48kbps, while our comparison examines a wide range of target bit rates for long video traces. An image coding comparison of MPEG-4 and JPEG2000 has been performed by Moccagatta *el at.* [10] for *Face, Bike, and Natura* images. The authors compare the PSNR values of these images using the JPEG *baseline*, JPEG2000, and MPEG-4 *vtc*. For an extensive study on wavelet encoded long video sequences of one hour using the 3D-EZBC encoder, we refer to [11]. Chai *et al*. [12] have studied on the performance of the MPEG-4 encoder using the *Foreman* test sequence. In their study the authors only analyze the frame size and PSNR values for an exemplary 100 frames. A broad study on long MPEG encoded video streams has been done by Reisslein *et al*. [13], [14] where the video sequences studied consists of one hour long movie excerpts.

II. Setup and Trace Generation

In this section we give an overview of the evaluation setup and the generation of the video traces. Compared to our previous work, see, e.g., [15], we generate the source video from a DVD source. DVDs contain MPEG-2 compressed video streams, which may exhibit quality losses compared to the original video as it was filmed. We reason, however, that the typical quality of DVD video is superior to the grabbed VHS raw video, as no problems such as frame drops, buffer overflows, or noise degrade the raw video data. In addition, we note that the video on DVDs typically is in high resolution and has to be down-sampled. This down-sampling reduces the possible influence of visual impairments of the MPEG-2 encoded DVD video. For the generation of the uncompressed video source, we use the publicly available `ffmpeg` [16] program to generate the uncompressed video with the NTSC frame rate of 30 frames per second and in QCIF format (176x144 pixels) directly from the DVD MPEG-2 video. For this study we used two movie excerpts of 60 minutes (108,000 frame) length from the movies *The Terminator* and *The Lady and The Tramp*.

We encoded the thus generated raw video with the motion compensated 3D embedded zerotree block coder (MC-3DEZBC) [3]. For this study we use only the Intra frame encoder/decoder capability of the codec. We truncated the original wavelet encoded video stream at 10 different target bit rates of 25, 75, 100, 300, 600, 800, 1000, 1200, 1400 and 1600kbps. We refer to [11] for a more in-depth discussion on the generation of the wavelet encoded video streams using the MC-3DEZBC encoder. The MPEG encodings were conducted with the MPEG-4 reference encoder [17]. For the MPEG encodings, we use only intra coding with TM5 rate control [18] to match the above target bit rates used for the wavelet encoding for performance comparison. We note that the employed rate control schemes are different for each coder, but are the defaults for each encoding algorithm (i.e., wavelet or DCT). For each video frame, we use the encoder output to generate the trace file with the video frame size and quality.

III. Video Traffic Characteristics

In this section, we compare the statistical video traffic characteristics of the MC-3DEZBC wavelet encodings with those from the MPEG-4 intra encodings, which we obtained from the generated trace files. We focus in the presentation of our results on the movie *The Terminator*, as they are representative for both evaluated movies.[1]

Let N denote the number of video frames in a given trace. Let X_n, $n = 0, \ldots, N-1$, denote the frame size (number of byte) of the encoded (compressed) video frame frame n. The mean frame size $\bar{X}$ of a given video trace is calculated as

$$\bar{X} = \frac{1}{N} \sum_{n=0}^{N-1} X_n. \tag{1}$$

With the maximum frame size $X_{\max}$, defined as $X_{\max} = \max_{0 \le n \le N-1} X_n$, we determine the peak to mean ratio of the frame sizes as $P2M_X = X_{\max}/\bar{X}$. Although the peak to mean ratio is commonly used to evaluate the variability of the encoded frame sizes, it is highly prone to outliers. For the evaluation of long video sequences, it is therefore more reasonable to evaluate the variability of the video frame sizes in terms of the coefficient of variation (CoV). The CoV gives the normalized averaged deviation of the individual frame sizes from the mean frame size and is therefore widely employed in performance evaluation [19], [20]. With the standard deviation S_X of the frame sizes defined as

$$S_X = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})^2}, \tag{2}$$

[1]We refer to `http://trace.eas.asu.edu` for the trace files for both movies.

the coefficient of variation CoV_X of the frame sizes is calculated as

$$CoV_X = \frac{S_X}{\bar{X}}. \quad (3)$$

Table I gives the aforementioned basic statistics and the compression ratio (i.e., amount of data for the uncompressed frame size compared to the mean compressed frame size) for *The Terminator* encodings with different target bit rates. We observe

TABLE I
OVERVIEW OF FRAME STATISTICS FOR *The Terminator* ENCODED WITH WAVELET-BASED MC-3DEZBC AND DCT-BASED MPEG-4 ENCODER.

Target Rate [kbps]	Compress. ratio YUV:Enc	Mean $\bar{X}$ [kbyte]	CoV_X $S_X/\bar{X}$	Peak to Mean $X_{\max}/\bar{X}$
Wavelet				
25	367.696	0.103	0.198	2.979
75	121.979	0.312	0.322	3.911
100	91.421	0.416	0.334	3.826
300	30.434	1.249	0.340	4.173
600	15.212	2.499	0.321	3.336
800	11.408	3.332	0.307	3.096
1000	9.126	4.166	0.297	2.867
1200	7.605	4.999	0.284	2.766
1400	6.518	5.832	0.272	2.642
1600	5.704	6.665	0.259	2.435
MPEG-4				
25	74.186	0.512	0.319	3.101
75	74.183	0.512	0.319	3.101
100	74.149	0.513	0.319	4.061
300	30.399	1.251	0.338	12.455
600	15.203	2.501	0.474	6.229
800	11.403	3.334	0.623	7.134
1000	8.257	4.604	0.884	5.300
1200	7.602	5.001	0.826	4.879
1400	6.516	5.834	0.763	4.182
1600	6.362	5.975	0.809	4.501

that the wavelet encoder achieves a better match of the lower target bit rates than the MPEG encoder, which fails to match the lower target bit rates. For target bit rates from 25kbps to 100kbps the MPEG-4 encodings result in similar mean frame sizes of approximately 0.5kbyte. The target bandwidth of 100kbps is thus exceeded by approx. 26%. This behavior is due to the maximum quantization scale of 31 available in the reference encoder implementation. With this bound on the quantization scale, the TM5 algorithm is unable match the lower target bit rates. With data rates higher than 100kbps the compression ratios for both coding modes become very close. For 25 and 75kbps, the CoV_X and peak-to-mean ratio are identical for MPEG-4. For the encoding with 100kbps target rate, we observe that the peak-to-mean ratio for the MPEG encoding is no longer identical to that of the two lower target bit rates while the CoV_X is, which corroborates our previous reasoning in favor of the CoV_X as a robust measure of the traffic variability.

From Table I we observe additionally that the coefficient of variation increases as the encoded video rate increases, reaches a peak, and decreases as the encoded video rate increases further, building a *hump* of variability. The result is present for both, MPEG and wavelet encodings. The trend is much clearer, however, for wavelet encodings. Figure 1 illustrates this characteristic of the coefficient of variation for both movies. We observe that for wavelet encodings the peak is located at

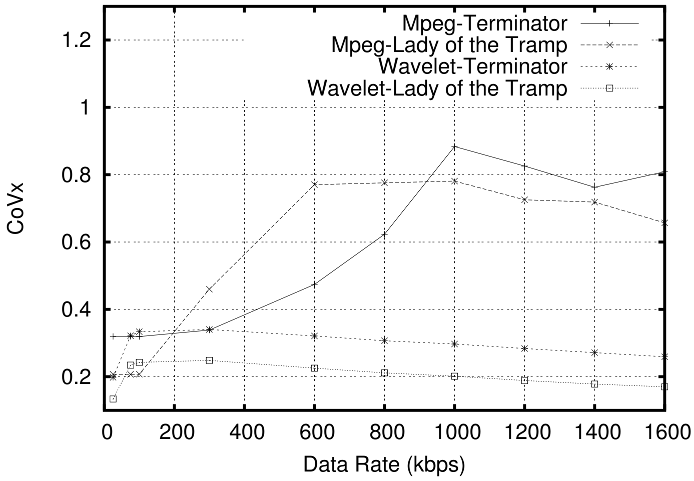

Fig. 1. Coefficient of variation as function of target bit rate for wavelet 3D-EZBC and MPEG-4 encodings.

300kbps. For MPEG-4 encodings, the peak is located at 1Mbps. We furthermore observe that the level of variability depends on the content (i.e., encoded movie) as well as on the encoding type. The MPEG-4 encodings tend to have a higher variability compared to the wavelet encodings and *The Terminator* encodings exhibit higher variability than the *The Lady and The Tramp* encodings. To study general characteristics without short-term effects, we average over non-overlapping blocks of a frames for an aggregation level a. We calculate the sequence of averaged aggregated frame sizes using the following notation. Let j denote the non-overlapping block out of N frames for which we want to average over a frames. The average aggregated frame size $X_a(j)$ is then calculated as

$$\overline{X}_a(j) = \frac{1}{a} \cdot \sum_{i=ja}^{(j+1)a-1} X_i, \text{ for } j = 0, \ldots, \frac{N}{a-1}. \quad (4)$$

In Figure 2 we exemplarily illustrate the aggregated frame size trace for *The Terminator* with a target bit rate of 300kbps and aggregation level of $a = 792$. We observe that the TM5 rate control algorithm used for the MPEG-4 encoding produces a generally close fit to the target bit rate with a limited number of exceptions. The TM5 algorithm matches target bit rates at the Group of Pictures (GoP) level. We note that the GoP length in our study equals a single frame. The TM5 therefore tries to match the target bit rate for individual frames. For higher aggregation levels the resulting average aggregated frame sizes therefore typically exhibit lower variability than the individual frame sizes, as can be obtained by comparing Figures 1 and 2. The MC-3DEZBC, on the other hand, produces more variable video frame sizes, but matches the target bit rate over longer time scales. As result, the traffic produced by the MC-3DEZBC encoder accurately fits the target bit rate overall, but produces more variable traffic over shorter time scales.

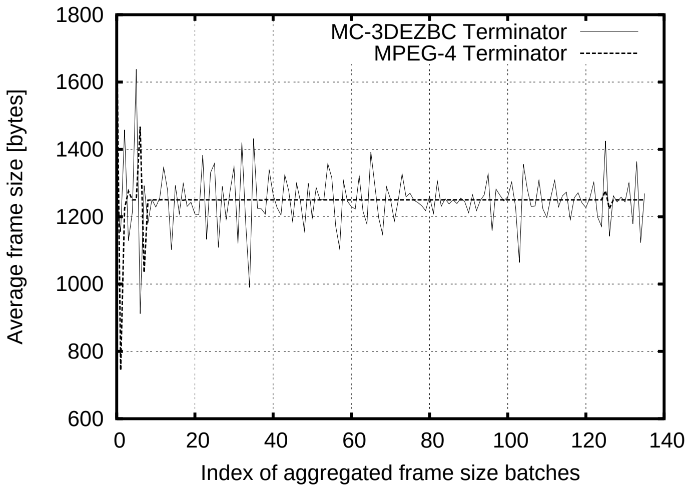

Fig. 2. Aggregated frame size trace ($a = 792$) with target bit rate 300kbps for *The Terminator* encodings.

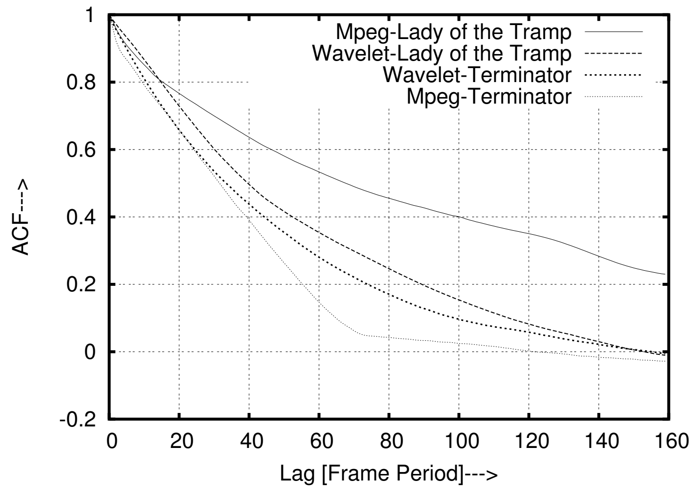

Fig. 3. Autocorrelation as function of lag k for wavelet 3D-EZBC and MPEG-4 encodings.

The autocorrelation function [21] can be used for the detection of non-randomness in data or identification of an appropriate time series model if the data are not random. One basic assumption is that the observations are equi-spaced. The autocorrelation is expressed as a correlation coefficient, referred to as autocorrelation coefficient (acc). Instead of the correlation between two different variables, the correlation is between two values of the same process (stream) at times X_t and X_{t+k}. When the autocorrelation is used to detect non-randomness, it is usually only the first (lag $k = 1$) autocorrelation that is of interest. When the autocorrelation is used to identify an appropriate time series model, the autocorrelations are usually plotted for a range of lags k. With our notation the acc can be estimated by

$$\rho_X(k) = \frac{1}{N-k} \cdot \sum_{i=1}^{N-k} \frac{\left(X_i - \overline{X}\right) \cdot \left(X_{i+k} - \overline{X}\right)}{S_X^2}, \qquad (5)$$

where $k = 0, 1, \ldots, N-1$. In Figure 3 we plot the frame size autocorrelation coefficients as a function of the lag k for a target bit rate of 300kbps. The autocorrelations of the MC-3DEZBC encodings drop sharply and are reduced to minuscule values for higher lags. The autocorrelation for *The Lady and The Tramp* encoded in MPEG-4, however, only drops off sharply at the beginning and levels out around 0.2. This outcome indicates that there is some correlation between relative distant frame sizes for *The Lady and The Tramp* MPEG-4 encoding. The autocorrelation for *The Terminator* encoded in MPEG-4, however, drops off faster than the two wavelet encodings,with no correlations for higher lags. The frame sizes for both MC-3DEZBC encodings exhibit no correlations at longer distances. The result of the autocorrelation comparison is thus that the DCT-based MPEG-4 encoding produces frame sizes with content-dependent autocorrelation, whereas the wavelet-based MC-3DEZBC seems to be more neutral in producing only minimally autocorrelated frame sizes.

IV. Video Quality Characteristics

In this section we compare the quality of 3D-EZBC and MPEG-4 encodings based on the peak signal to noise ratio (PSNR). Video quality evaluation can be done either subjectively (i.e., by viewing and evaluating the encoded video with a sufficient number of subjects) or objectively (i.e., by using fixed calculations and evaluation schemes). The PSNR is commonly used in video quality evaluation, as subjective tests are impractical for larger sets of videos. Recent studies have found that the PSNR is as good a measure of video quality as other more sophisticated objective quality metrics [22]. As the video quality depends mainly on the luminance (Y) component of the uncompressed video, we focus on the luminance PSNR values. Let $Q_n^Y,\ n = 0, \ldots, N-1$, denote the quality of the luminance component in terms of the PSNR of video frame n (in dB). For an individual frame, the difference between the original image I and its encoded (and subsequently decoded) counterpart $\tilde{I}$ is calculated as mean squared error (MSE)

$$MSE = \frac{1}{D_x \cdot D_y} \sum_{x=1}^{D_x} \sum_{y=1}^{D_y} \left[I(\ x,\ y) - \tilde{I}(\ x,\ y) \right]^2, \qquad (6)$$

where D_x and D_y denote the horizontal and vertical dimensions and (x, y) denotes the position of the pixel. The quality as PSNR for a given video frame n is then calculated as

$$Q_n = 10 \cdot \log_{10} \frac{255^2}{MSE_n}. \qquad (7)$$

The basic video quality statistics are similar to the video traffic characteristics in Section III defined as mean quality $\bar{Q}_n$ and coefficient of quality variation $CoQV$. In addition we define the quality range (in dB) of a video sequence as

$$Q_{\min}^{\max} = \max_{0 \le n \le N-1} Q_n - \min_{0 \le n \le N-1} Q_n. \qquad (8)$$

These basic statistics for *The Terminator* are given in Table II. We begin our observation for the target bit rates of 100kbps and

TABLE II

OVERVIEW OF QUALITY STATISTICS FOR *The Terminator* ENCODED WITH WAVELET-BASED MC-3DEZBC AND DCT-BASED MPEG-4 ENCODER.

Target Rate [kbps]	$\bar{Q}$	$CoQV$	$Q_{\min}^{\max}$
Wavelet			
25	25.13	0.145	72.570
75	29.76	0.123	69.070
100	31.01	0.118	68.030
300	36.86	0.102	62.480
600	41.65	0.090	57.940
800	43.93	0.083	55.480
1000	45.83	0.077	53.670
1200	47.39	0.071	51.700
1400	48.72	0.065	49.930
1600	49.84	0.058	49.160
MPEG-4			
25	30.18	0.081	29.688
75	30.18	0.081	37.507
100	30.19	0.081	37.454
300	35.41	0.134	38.628
600	39.18	0.139	65.120
800	40.41	0.160	65.120
1000	41.25	0.188	65.594
1200	41.90	0.190	64.356
1400	43.12	0.191	66.059
1600	42.55	0.221	66.039

up, as for the lower target bit rates, the bounded MPEG-4 quantization scale setting does not allow for a fair comparison. We observe that the average video quality for MPEG encoded video sequences is always lower than for the 3D-EZBC encodings. Earlier comparison studies in [5], where only the lowest target bit rates were evaluated, showed a difference of approximately 0.5dB in favor of DCT-based video encodings based on the PSNR of the luminance. In contrast we find that the quality difference increases with the target bit rate and even reaches a significant difference of more than 7dB, but for the wavelet based encodings. We also find that the video quality from wavelet based encodings is always higher than for DCT-based MPEG-4 encodings for target bit rates higher than 100kbps. Our results indicate that the quality difference between wavelet and MPEG encoded video increases faster than linear depending on the target bit rates. For network simulation studies the quality to bit rate relationship cannot be simply scaled. Our results furthermore show that for higher target bit rates, the wavelet-based 3D-EZBC clearly outperforms the DCT-based MPEG-4 encoding.

Figure 4 illustrates the average video quality for the two encoding methods and the two evaluated movies. The average qualities for both encoding methods increase over the whole bit rate scale, although the marginal return in terms of quality decreases with increasing target bit rate (i.e., for higher target bit rates, an increase in the bit rate results in a less than linear increase in quality). From Table II we observe the variation of the video quality $CoQV$ increases over the whole quantization scale for the MPEG-4 encoded *The Terminator*, whereas the $CoQV$ decreases over the whole quantization scale for the 3D-EZBC encoding. We illustrate the characteristic of the $CoQV$ in Figure 5. The quality range $Q_{\min}^{\max}$ follows the same

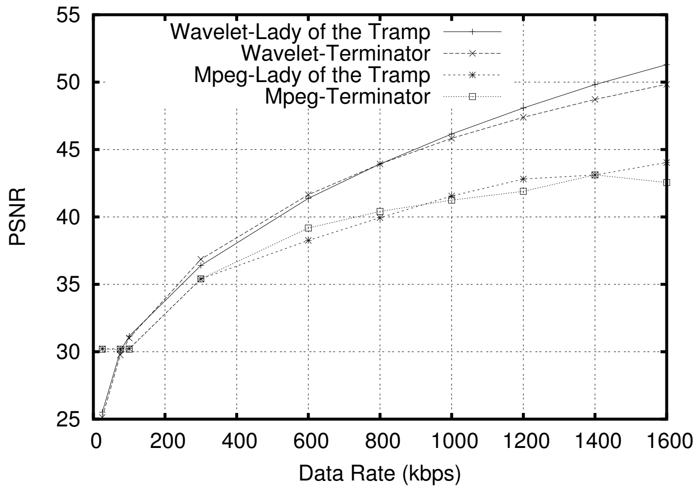

Fig. 4. Average video quality as function of encoding target bit rate for wavelet 3D-EZBC and MPEG-4 encoded movies.

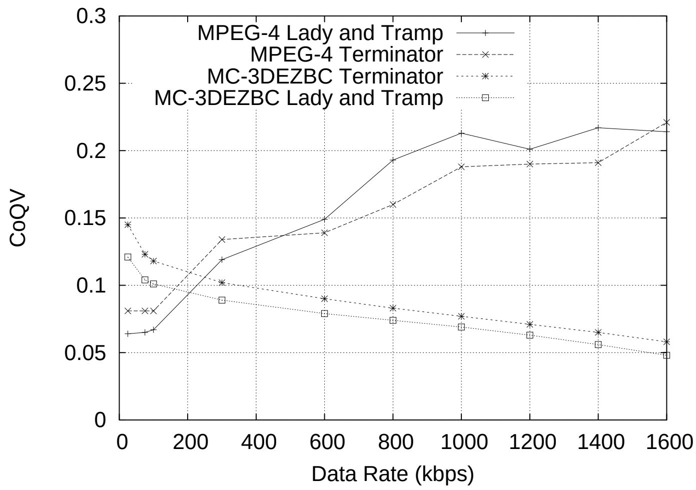

Fig. 5. Coefficient of quality variation for wavelet 3D-EZBC and MPEG-4 encoded movies.

trend of decreasing in value for the 3D-EZBC encoding, while increasing in value for MPEG-4 encoding.

For the transmission of video, the encoded video quality and video traffic have to be taken into account. We use the coefficient of correlation as measure of (linear) dependency, defined as

$$\rho_{x,y} = \frac{\left(\sum (x_i - \overline{x})^2 \cdot (y_i - \overline{y})^2\right)^2}{\sum (x_i - \overline{x})^2 \cdot \sum (y_i - \overline{y})^2} \tag{9}$$

for a set of data points by x_i and y_i. We start by comparing the correlation of mean frame sizes and mean frame qualities for target bit rates greater than or equal to 100kbps. For *The Lady and The Tramp* as well as *The Terminator*, we obtain a correlation of 0.9 between quality and size for 3D-EZBC and MPEG-4 encodings. This indicates a strong correlation between the quality and the size of the encoded video frames for different target bit rates. The correlation between the coefficient of variation for frame sizes $CoVx$ and the coefficient of variation for video qualities $CoQV$, also calculated starting from 100kbps target bit rate, are similarly pronounced and above 0.85 for both considered video encoding methods and both evaluated video sequences.

These findings indicate that frame quality and frame size are strongly dependent. In addition we observe that video quality variability and video traffic variability are highly correlated for 3D-EZBC and MPEG-4 encodings.

V. Conclusion and Outlook

In this comparison of rate-controlled wavelet and MPEG video for one hour movie excerpts, we have shown the major differences between wavelet-based encodings with the MC-3DEZBC encoder and the DCT-based MPEG-4 reference software encoder. In contrast to other studies that focused on results for single pictures or short video sequences encoded with only a limited set of very low target bit rates, we focus on the comparison of video quality and video traffic resulting from one hour encodings of actual movie content for a wide range of target bit rates.

We find very different traffic characteristics for the two evaluated encoders, which primarily result from the rate adaption techniques involved. We find a characteristic *hump* behavior of the coefficient of variation of the video frame sizes as function of the target bit rate. This *hump* was found to be more pronounced for the MPEG encoded video than for the wavelet encoded video. We furthermore observed the self-similarity of the wavelet produced video traffic to be less significant and mostly content independent compared to the DCT-based MPEG-4 encodings. We clearly see from the generated video statistics that the video frame sizes can not be scaled for simulations, as the scaling of frame sizes would not change the variations observed at different bit rates.

For the video quality of 3D-EZBC based encodings, we find an increasing advantage over MPEG-4 based encodings with increasing target bit rates above 100kbps. This increase was furthermore found to be more than linear. This behavior indicates that scaling of video frame qualities for simulations has to take this characteristic behavior into account. The coefficients of variability of the video qualities were found to exhibit different characteristics for MPEG and 3D-EZBC. More, specifically, we found that the quality variation increases for MPEG-4 and decreases for MC-3DEZBC encodings with increasing target bit rates. We found a strong correlation between the variability of the frame sizes and the variability of the frame qualities for the DCT-based MPEG-4 and the wavelet-based 3D-EZBC encodings.

Future research could be directed at incorporating motion estimation and compensation into the DCT and wavelet domains and to evaluate the performance of both encoding methods concerning traffic, quality, and error-resilience.

Acknowledgment

Supported in part by the National Science Foundation under Grant No. Career ANI-0133252 and Grant No. ANI-0136774.

References

[1] B. Belzer, J. Liao, and J. Villasenor, “Adaptive video coding for mobile wireless networks,” in *Proc. of IEEE Int’l Conf. on Image Processing*, vol. 2, Austin, TX, Nov. 1994, pp. 972–976.

[2] B. S. Nanda and N. Kaulgud, “Effect of quantization on video compression,” in *Proc. of 2002 IEEE Int’l Conf. on Industrial Technology*, vol. 2, Bangkok, Thailand, Dec. 2002, pp. 764–768.

[3] S.-T. Hsiang and J. Woods, “Embedded video coding using invertible motion compensated 3-d subband/wavelet filter bank,” *Signal Processing: Image Comm.*, vol. 16, pp. 705–724, May 2001.

[4] ——, “Invertible three-dimensional analysis/synthesis system for video coding with half-pixel-accurate motion compensation,” in *Proc. of The Int’l Society for Optical Eng. (SPIE) on Visual Comm. and Image Processing*, vol. 3653, San Jose, CA, Jan. 1999, pp. 537–546.

[5] Z. Xiong, K. Ramchandran, M. T. Orchard, and Y. Zhang, “A comparative study of DCT and wavelet based coding,” in *Proc. of the 1998 IEEE Int’l Symp. on Circuits and Systems*, vol. 4, Monterey, CA , USA, May 1998, pp. 273–276.

[6] S. A. Martucci, I. Sodagar, T. Chiang, and Y.-Q. Zhang, “A zerotree wavelet video coder,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 109–118, Feb. 1997.

[7] J. W. Woods and G. Lilienfield, “A resolution and frame-rate scalable subband/wavelet video coder,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 9, pp. 1035–1044, Sept. 2001.

[8] S. H. Yoon and S. S. Rao, “A scalable wavelet video coder for hybrid communication channels,” *Conf. Record of the Thirty-First Asilomar Conf. on Signals, Systems and Computers*, vol. 1, pp. 382–386, 1997.

[9] J.-R. Ohm, “Three-dimensional subband coding with motion compensation,” *IEEE Trans. on Image Processing*, vol. 3, no. 5, pp. 559–571, 1994.

[10] I. Moccagatta, M. Coban, and H. Chen, “Wavelet-based image coding: comparison of MPEG-4 and JPEG-2000,” in *Proc. of the Thirty-Third Asilomar Conf. on Signals, Systems, and Computers*, vol. 2, Pacific Grove, CA USA, Oct. 1999, pp. 1178–1182.

[11] B. Kulapala, P. Seeling, and M. Reisslein, “Characteristics of traffic and quality traces of wavelet encoded video,” Arizona State Univ., Dept. of Electrical Eng., Tech. Rep., Nov. 2003. [Online]. Available: http://trace.eas.asu.edu

[12] D. Chai and K. Ngan, “A performance study on MPEG-4 coder,” in *Proc. of the Fifth Int’l Symp. on Signal Processing and Its Applications*, vol. 2, Brisbane, Qld. Australia, Aug. 1999, pp. 821–824.

[13] M. Reisslein, J. Lassetter, S. Ratnam, O. Lotfallah, F. H. Fitzek, and S. Panchanathan, “Traffic and quality characterization of scalable encoded video: A large-scale trace based study,” Arizona State Univ., Dept. of Electrical Eng., Tech. Rep., Aug. 2003. [Online]. Available: http://trace.eas.asu.edu

[14] P. Seeling, M. Reisslein, and B. Kulapala, “Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial,” *IEEE Comm. Surveys and Tutorials*, vol. 6, no. 3, 2004.

[15] F. Fitzek and M. Reisslein, “MPEG–4 and H.263 video traces for network performance evaluation,” *IEEE Network*, vol. 15, no. 6, pp. 40–54, November/December 2001. [Online]. Available: http://trace.eas.asu.edu

[16] F. Bellard, “FFmpeg,” Apr. 2004. [Online]. Available: http://ffmpeg.sf.net

[17] ISO/IEC 14496, “Video Reference Software, Microsoft–FDAM1–2.3–001213.”

[18] Test Model Editing Committee, “MPEG–2 Video Test Model 5, ISO/IEC JTC1/SC29WG11 MPEG93/457,” Apr. 1993.

[19] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling.* Wiley, 1991.

[20] A. M. Law and W. D. Kelton, *Simulation, Modeling and Analysis*, 3rd ed. McGraw Hill, 2000.

[21] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.

[22] A. M. Rohaly and et al., “Video quality experts group: Current results and future directions,” in *SPIE Visual Comm. and Image Processing*, Perth, Australia, June 2000, pp. 742–453.