

Characteristics of Traffic and Quality Traces of Wavelet Encoded Video. ^{*†}

Beshan Kulapala Patrick Seeling Martin Reisslein [‡]

<http://trace.eas.asu.edu/wavelet/indexwavelet.htm>

December 8, 2003

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. However, due to the lack of long wavelet encoded video streams, most research has so far been based on short video traces. This thesis paper presents a public library of traces of long, as well as short, wavelet encoded videos. Currently, all traces in this library are for intra-coded video sequences. The library includes traces for over 10 one-hour movies, encoded with a wavelet-based codec, as well as several test sequences both in the Quarter Common Intermediate Format (QCIF) and the Common Intermediate Format (CIF) format. Both the frame sizes as well as the frame qualities (Peak Signal to Noise Ratio (PSNR) values) in the traces are included in this library. The statistical characteristics of the traces, including their long range dependence have also been studied. A hump like behavior of the coefficient of variation (CoV_X) of frame sizes is observed for increasing video bit rates. A similar hump behavior is observed for the peak to mean ratio of the frame sizes as well. From the video trace analysis, it is clear 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.

Keywords: Wavelet Video Traces; Quality Statistics; Subband; Traffic Statistics; Video Traces;

1 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

^{*}Supported in part by the National Science Foundation under Grant No. Career ANI-0133252 and Grant No. ANI-0136774. Supported in part by the State of Arizona through the IT301 initiative. Supported in part by a matching grant and a special pricing grant from Sun Microsystems.

[†]Please direct correspondence to M. Reisslein.

[‡]B. Kulapala, P. Seeling, and M. Reisslein are with the Dept. of Electrical Engineering, Arizona State University, Goldwater Center, MC 5706, Tempe AZ 85287-5706, Phone: (480)965-8593, Fax: (480)965-8325, (email: {beshan, patrick.seeling, reisslein, }@asu.edu, web: <http://trace.eas.asu.edu>).

to compress the video data for more efficient transport over the network. With a higher compression ratio in wavelet based compression schemes over discrete cosine transform (DCT) based schemes, the wavelet encoded video streams have a bright future in providing means of decreasing video traffic and providing scalable video streams to provide better quality at lower bit rates.

The wavelet transform has many advantages over the DCT transform. The most obvious of them, is the compact-support feature. This compact-support allows 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 occur over the entire image with reasonable computation and bit budgets. The DCT based, on the other hand, requires to window the data 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. The wavelet transform codec used in this study, is the motion compensated 3D embedded zerotree block coder (MC-3DEZBC) [12].

The traces have been generated from over 10 videos of one hour each. Each video is encoded using intra frame encoding and then the encoded bit stream is truncated to fit a given bit budget. Thus, we form a set of 10 encoded streams per movie at different bit rates. Each of these rate controlled streams consists of 5 sub-streams as explained in Section 2. We have provided two types of traces for each truncated stream. First the aggregated frame size and the PSNR quality of each frame, while the second trace is the size of each individual sub-stream component per frame.

1.1 Related Work

Mallat [18] introduced the wavelet transform to effectively analyze the information in images. He defined a decomposition of information using a set of orthonormal basis functions as a wavelet decomposition. Then Antonini *et al.* [3] successfully used the wavelet transform to code images in both the space and frequency domain. They decomposed the original image at different scales using a pyramidal algorithm, decomposing the image both in the vertical and horizontal directions. Further developments resulted in a variety of wavelet based still image coders, see for instance [21] [28] [17] [29]. A range of video coders [20] [32] [33] have also been studied that utilized the wavelet transform. Furthermore the video coders have been extended to a 3-D video codec [10] [12] [23] [16], decomposing the frames spatially and temporally. Nanda *et al.* [22] 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

estimation and the *MV (Motion Vector) Encoder* sections are bypassed. First the input frames are divided into group of pictures (GoPs). For our analysis we have used 16 frames as the GoP size. For intra frame encoding, the GoP is used for rate control as explained later. Each frame undergoes a four-stage spatial decomposition which is recursively performed on the low frequency subband. The first stage of a filter bank structure used for the spatial decomposition is illustrated in Figure 2.

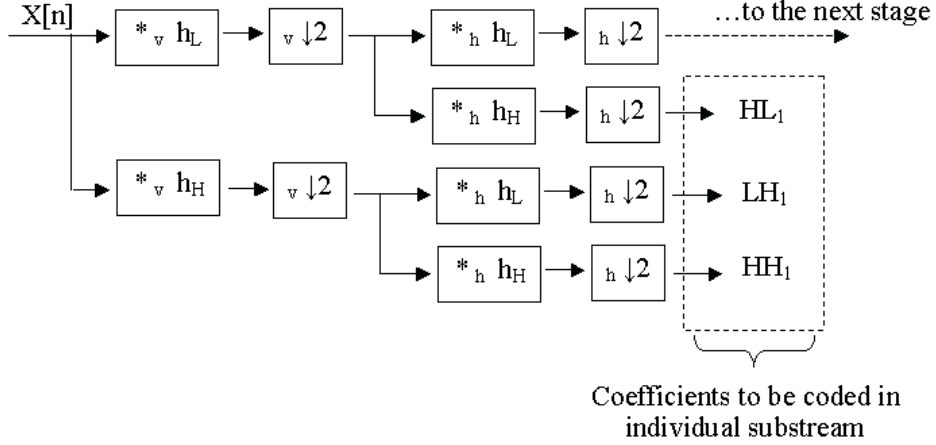


Figure 2: First wavelet decomposition stage

Here X_n is the input image. $*_v$ and $*_h$ represent convolution in the vertical direction and convolution in the horizontal direction, respectively. The impulse response of the low pass filter and high pass filter are represented by h_L and h_H , respectively. An arrow pointing downwards and followed by the number 2, represents subsampling by 2 in the horizontal or vertical direction (represented by the subscript preceding the arrow). HL_1 , LH_1 , and HH_1 represent the outputs of the filters of the first decomposition stage. Each stage creates 3 subbands, while the fourth (which is the lowest frequency subband in both the horizontal and the vertical dimensions) is fed into the next stage of the spatial decomposition. The four stage decomposition provides 13 subbands as illustrated in Figure 3. The 13 subbands obtained from the four decomposition stages are then coded individually using the 3-D version of the embedded zerotree block coding algorithm *3D-EZBC* [12]. This is an extension of the embedded zerotree block coding (EZBC) algorithm developed in [11]. The resulting bit streams are then bit plane encoded and combined to form one sub-stream as illustrated in Figure 4. For easier illustration, each sub-stream in Figure 4 is color coded such that it matches with the corresponding color in Figure 3. All sub-streams of each frame and all frames in the corresponding GoP are then combined to

create a hierarchical code stream [13]. Each GoP is coded as a separate message with context-dependent arithmetic coding. Each message is embedded, thus the bitstream can be truncated at any point to a given bit budget. Rate control is implemented on each GoP with the bit budget given by $R_g = N_g \cdot r / F$ (bits), where N_g denotes the number of frames in a GoP, r the given bit rate in bits/sec, and F denotes the frame rate of the image sequence in frames/sec.

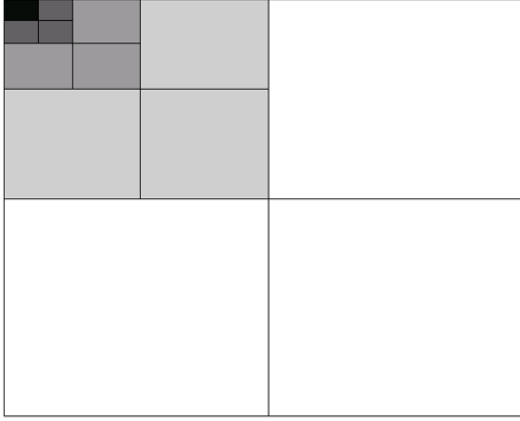


Figure 3: Passband structure for MC-3DEZBC [13]

Compressed substream 1	
Compressed substream 2	
Compressed substream 3	
Compressed substream 4	
Compressed substream 5	

Figure 4: Individually coded sub-bitstreams corresponding to Figure 3 [13]

We observed that even though [12] describes that the rate control is done at the GoP level, the GoP sizes for a long video varied heavily, indicating that the individual GoPs were not holding to the given bit budget N_g . But interestingly enough we also noticed that when taking into consideration the entire video stream, the specified bit rate is achieved almost perfectly. To verify that the encoder indeed takes the entire stream length into account in its the rate control, we encoded a video stream of 480 frames from *Lady and the Tramp* at 16 frames per GoP (30 GoPs in stream). Then we also encoded the same video frames, but as 30 individual streams, with each stream having 16 frames (1 GoP per stream). We observed that the 480 frame-long stream gives varying GoP sizes, in contrast to the almost constant GoP sizes for each of the 30 individual streams. This observation is illustrated in Figure 5. The straight line indicates the 30 different streams (1 GoP each) which when concatenated give 480 frame long video stream.

3 Structure And Generation of Video Traces

This section of the report describes how the video was played from its source, how it was captured, and a brief description of the video genres that we have in the our library. A similar

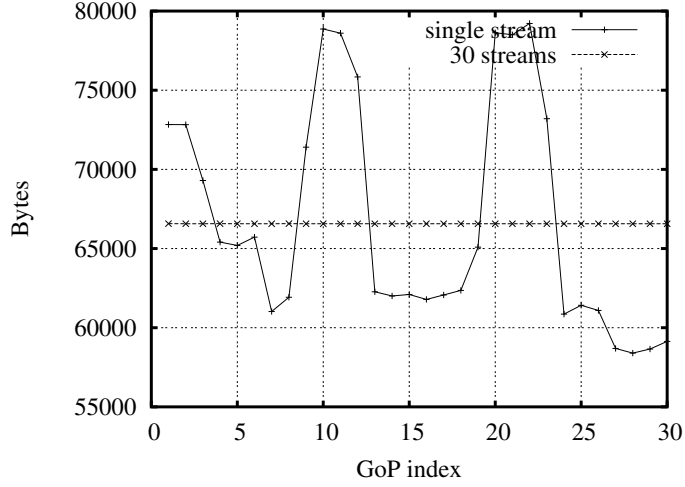


Figure 5: GoP size in Bytes as a function of GOP index for (i) one stream of 480 frames, and (ii) 30 streams of 16 frames each.

study and a trace library for MPEG-4 was created and published in [24].

First, the video sequences in Table 1 were played from a video cassette recorder (VCR). The analog video signal (uncompressed YUV) was then captured through a PC video capture card using the **bttvgrab** (version 0.15.10) software [31]. The computer used for this purpose is a high performance dual Intel Pentium III 933 MHz processors with 1 GB RAM and an 18 GByte high-speed SCSI hard disc. All the videos studied were grabbed at 30 frames per second with 4:2:0 chrominance subsampling and quantization into 8 bits. The YUV information was grabbed at the National Television Standards Committee (NTSC) frame rate of 30 frames per second, and captured in the QCIF (176x144 pels) format. We captured the 60 minute (108,000 frame) QCIF sequences, in two segments of 30 minutes (54,000 frame) each. This prevented buffer build-up, and we were able to capture all frames without any frame drops. In the QCIF format with 4:2:0 chroma subsampling with 8 bit quantization and 30 frames per second the bit rate of uncompressed QCIF video is 9,123,840 bit/sec. The file size of 1 hour of uncompressed QCIF video is 4,105,728,000 Byte.

Table 1: Overview of studied video sequences.

Movies (rental tapes)	Format	Length (min)	Frames Dropped
<i>Terminator</i>	QCIF	60	0
Cartoons (rental tapes)	Format	Length (min)	Frames Dropped
<i>Lady and the Tramp</i>	QCIF	60	0
Sports Events (recorded from Broadcast TV)	Format	Length (min)	Frames Dropped
<i>Foot Ball w/c</i>	QCIF	60	0
Other TV sequences (recorded from Broadcast TV)	Format	Length (min)	Frames Dropped
<i>Tonight Show w/c</i>	QCIF	60	0

We have covered a wide range of video genres including action movies, cartoons, sports, talk shows as well as some of test sequences as described in Table 1¹. Since the video characteristics (traffic and quality) strongly depend on the video content, it was important to cover a wide range of video genres for an effective study. We also encoded the talk shows sequences and the sports sequences to study how commercials effect the video characteristics. Note that encoded sequences with commercials is depicted as *w/c*. These sequences were recorded with a VCR from the regular TV broadcast. For all the movies in Table 1, we started the video capture process at the beginning of the feature presentation. Previews, trailers, or commercials preceding the feature presentation were not included.

3.1 Overview of Metrics and Notations

Let N denote the number of video frames in a given trace. Let X_n , $n = 0, \dots, N-1$, denote the frame size (number of byte) of the encoded (compressed) video frame frame n . Let Q_n^Y , $n = 0, \dots, N-1$, denote the quality of the luminance component in terms of the Peak Signal to Noise Ratio (PSNR) of the encoded (and subsequently decoded) video frame n (in dB).

3.2 Limitations

Due to a software limitation in the MC-3DEZBC codec [12] we had to split the 1 hour, i.e., 108,000 video frames into 150 separate encodings of 720 frames each, and then concatenate

¹To avoid any conflict with copyright laws, we emphasize that all image processing, encoding, and analysis was done for scientific purposes. The encoded video sequences have no audio stream and are not publicly available. We make only the frame size traces available to researchers.

the traces. We also found that 25 kbps is a good lower bound for video rates, since the codec fails at lower video rates for most sequences.

3.3 Trace Generation for Intra Video encodings

First, the raw YUV frames captured from the video capture mechanism discussed in Section 3 are used as the input of the encoder software. The encoder software produces an intra frame encoded video stream. The encoder software also contains the following variables that can be controlled for intra frame encodings.

1. **Denoise** : When coding in the intra frame mode, this variable must be set to ”-denoise NO”
2. **tPyrLev** : The GoP size is inferred from this argument as: $GoP_{size} = 2^{tPyrLev}$.

Then this encoded video stream is truncated at 10 different bit rate budgets, providing 10 individual streams at bit rates of 25, 75, 100, 300, 600, 800, 1000, 1200, 1400, and 1600 kbps. During the truncation the truncating software also provides the frame size of the individual sub-streams, described in Section 2 and illustrated in Figure 3. Finally, the individual encoded streams are passed through the decoder which produces the decoded video frames in YUV format. Also the decoder software produces the trace file which contains the frame number, aggregated frame size, and the PSNR of the decoded frame compared to the original frame. Note that the aggregated frame size is 10 bytes larger than the addition of the individual sub-streams. This is due to the fact that there is an overhead of 10 bytes in the aggregated frame size to incorporate the 5 individual sub-stream sizes. i.e., 2 bytes per sub-stream.

4 Organization of the Web Site

In this section we explain how the wavelet trace file web site is organized. The URL for the trace library web site for wavelets is as <http://trace.eas.asu.edu>. At the root of the web site is the main table. Different movies are ordered in rows, while the columns indicate different traces and figures for that particular video. Each video has the following columns:

1. **Stats-Figs** : The Stat-Figs include plots that have been drawn with the PSNR as the x-axis. The plots include the bit rate, mean, standard deviation, variance, coefficient of variation, and peak to mean ratio of the frame sizes. These above plots are included for aggregated sub-streams and for individual sub-streams.

2. **Stats-Data :** The Stats-Data field includes the data files for the plots in the Stats-Figs section. These ASCII files are tab delimited, with the column description in the first row.
3. **Bit Rate Traces :** These 10 fields i.e., 25, 75, 100, 300, 600, 800, 1000, 1200, 1400, and 1600 kbps include the trace files and plots for each of the truncated encoded video streams for a given bit budget. We have included two trace files for each bit rate. The first trace file is for the aggregated sub-streams, which includes the frame number, frame size in bytes, PSNR of Y, PSNR of U, and PSNR of V. The second trace file is for the individual sub-streams. This trace file includes the frame number, sub-stream 1 size in bytes to sub-stream 5 size in bytes. Also, this field includes plots of frame size and frame quality vs, frame number for aggregated sub-streams. And also includes plots of the frame size vs the frame number of each individual sub-streams.

5 Analysis of Video Traffic and Quality

In this section we conduct a statistical analysis on the intra frame, wavelet encoded video. The analysis includes the video traffic and the video quality characteristics based on the frame size and the quality of the luminance component of the video.

5.1 Video Traffic Metrics

First, we review the statistical definitions and methods used in the analysis. For further details and clarification we refer to [15, 6]. Recall from Section 3.1 that we denote N for the number of frames in a given trace, and X_n , $n = 0, \dots, N - 1$, for the size of frame n in bytes.

Mean, Coefficient of Variation, and Autocorrelation

The sample mean \bar{X} of a frame size trace is defined as

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

The sample variance S_X^2 of a frame size trace is defined as

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

The coefficient of variation CoV_X of the frame size trace is defined as

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

The maximum frame size X_{\max} is defined as

$$X_{\max} = \max_{0 \leq n \leq N-1} X_n. \quad (4)$$

The autocorrelation coefficient $\rho_X(k)$ for lag k , $k = 0, 1, \dots, N-1$, is estimated as

$$\rho_X(k) = \frac{1}{N-k} \sum_{n=0}^{N-k-1} \frac{(X_n - \bar{X})(X_{n+k} - \bar{X})}{S_X^2}. \quad (5)$$

We define the aggregated frame size trace with aggregation level a as

$$X_n^{(a)} = \frac{1}{a} \sum_{j=na}^{(n+1)a-1} X_j, \quad \text{for } n = 0, \dots, N/a - 1, \quad (6)$$

i.e., the aggregate frame size trace is obtained by averaging the original frame size trace X_n , $n = 0, \dots, N-1$, over non-overlapping blocks of length a .

We define the GoP size trace as

$$Y_m = \sum_{n=mG}^{(m+1)G-1} X_n, \quad \text{for } m = 0, \dots, N/G - 1, \quad (7)$$

where G denotes the number of frames in a GoP (for this study GoP=16). Note that $Y_m = G \cdot X_n^{(G)}$

Variance–Time Test

The variance time plot [4, 5, 14] is obtained by plotting the normalized variance of the aggregated trace $S_X^{2(a)}/S_X^2$ as a function of the aggregation level (“time”) a in a log–log plot. We refer the interested reader to [24] for the algorithm used for the computation of the variance time plot.

Traces with long range dependence for large a , decrease linearly with a slope larger than -1 in the variance time plot. While the traces without long range dependence decrease at a slope of -1 . The aggregation levels are multiples of the GoP size (16 frames) to avoid intra–GoP correlations. For referencing, we plot a line from the origin with a slope of -1 . Using a least squares fit, we estimate the Hurst parameter by estimating the slope of the linear part of the variance time plot.

We consider the aggregation levels $a \geq 384$ in this estimation since our variance time plots are typically linear for these aggregation levels. The Hurst parameter is then estimated as $H = \text{slope}/2 + 1$.

R/S Statistic

The R/S statistic provides an heuristic graphical approach for estimating the Hurst parameter H . As before we use the R/S statistic [19, 4, 5] to study the long range dependence of the video traces. The R/S statistic is characterized by $E[R(n)/S(n)] \sim cn^H$ as $n \rightarrow \infty$ (where c is some positive finite constant). The Hurst parameter H is estimated as the slope of a log–log plot of the R/S statistic. For the algorithm used for the computation we refer the interested reader to [24]

Periodogram

We estimate the Hurst parameter H using the heuristic least squares regression in the spectral domain, see [4, Sec. 4.6] for details. This approach relies on the periodogram $I(\lambda)$ as approximation of the spectral density, which near the origin satisfies

$$\log I(\lambda) \approx \log c_f + (1 - 2H) \log \lambda_k + \log \xi_k. \quad (8)$$

To estimate the Hurst parameter H we plot the periodogram in a log–log plot. For the interested reader we once again refer to [24] for the algorithms used for estimating the hurst parameter in this method.

Multiscale Diagram

We study the multifractal scaling properties [1, 2, 7, 8, 9, 26] using the wavelet based framework [1]. The q th order scaling exponent α_q is estimated based on the q th order logscale diagram, i.e., a plot of

$$\log_2(\mu_j^{(q)}) = \log_2 \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j, k)|^q \quad (9)$$

as a function of $\log_2 j$. The multiscale diagram is then obtained by plotting $\zeta(q) = \alpha_q - q/2$ as a function of q . The linear multiscale diagram, a variation of the multiscale diagram, is obtained by plotting $h_q = \alpha_q/q - 1/2$ as a function of q . We use the approach from Abry and Veitch’s logscale diagram Matlab code [30] to determine the range of scales (octaves) for the estimation of the scaling parameters.

5.2 Video Quality Metrics

The Mean Squared Error (MSE) is defined as the mean of the squared differences between the luminance values of the video frames in two video sequences I and \tilde{I} . The MSE for an individual video frame n is defined as

$$M_n = \frac{1}{D_x \cdot D_y} \sum_{x=1}^{D_x} \sum_{y=1}^{D_y} [I(n, x, y) - \tilde{I}(n, x, y)]^2, \quad (10)$$

where D_x and D_y denote the horizontal and vertical dimensions. $I(n, x, y)$, $n = 0, \dots, N-1$; $x = 1, \dots, D_x$; $y = 1, \dots, D_y$, denotes the luminance value of the pixel at location (x, y) in video frame n .

The mean MSE for a sequence of N video frames is

$$\bar{M} = \frac{1}{N} \sum_{n=0}^{N-1} M_n. \quad (11)$$

The Peak Signal to Noise Ratio (PSNR) in decibels (dB) is generally defined as $\text{PSNR} = 10 \cdot \log_{10}(p^2/\text{MSE})$, where p denotes the maximum luminance value of a pixel (255 in 8-bit pictures). We define the *quality* (in dB) of a *video frame* n as

$$Q_n = 10 \cdot \log_{10} \frac{p^2}{M_n}. \quad (12)$$

We define the *average quality* (in dB) of a *video sequence* consisting of N frames as

$$\bar{Q} = 10 \cdot \log_{10} \frac{p^2}{\bar{M}}. \quad (13)$$

Note that the averaging is conducted with the MSE values and the video quality is given in terms of the PSNR (in dB).

We also define an *alternative average quality* (in dB) of a video sequence as

$$\bar{Q}' = \frac{1}{N} \sum_{n=0}^{N-1} Q_n, \quad (14)$$

where the averaging is conducted over the PSNR values directly.

We define the MSE sample variance S_M^2 of a sequence of N video frames as

$$S_M^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} (M_n - \bar{M})^2, \quad (15)$$

and the MSE standard deviation S_M as

$$S_M = \sqrt{S_M^2}. \quad (16)$$

We define the *quality standard deviation* S_Q of a video sequence as

$$S_Q = 10 \cdot \log_{10} \frac{p^2}{S_M}. \quad (17)$$

We define the *coefficient of quality variation* $CoQV$ of a video sequence as

$$CoQV = \frac{S_Q}{\bar{Q}}. \quad (18)$$

We define an *alternative quality standard deviation* as

$$S'_Q = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (Q_n - \bar{Q}')^2}, \quad (19)$$

and the alternative coefficient of quality variation as

$$CoQV' = \frac{S'_Q}{\bar{Q}'}. \quad (20)$$

We define the *quality range* (in dB) of a video sequence as

$$Q_{\min}^{\max} = \max_{0 \leq n \leq N-1} Q_n - \min_{0 \leq n \leq N-1} Q_n. \quad (21)$$

We define the MSE autocorrelation coefficient $\rho_M(k)$ for lag k , $k = 0, \dots, N-1$, as

$$\rho_M(k) = \frac{1}{N-k} \sum_{n=0}^{N-k-1} \frac{(M_n - \bar{M})(M_{n+k} - \bar{M})}{S_M^2}. \quad (22)$$

We now define the qualities for groups of a frames. Note that $a = G$ gives the aggregated qualities for a GoP.

Let $M_m^{(a)}$, $m = 0, \dots, N/a - 1$, denote the MSE of the m th group of frames, defined as

$$M_m^{(a)} = \frac{1}{a} \sum_{n=ma}^{(m+1)a-1} M_n. \quad (23)$$

Let $Q_m^{(a)}$, $m = 0, \dots, N/a - 1$, denote the corresponding PSNR quality (in dB), defined as

$$Q_m^{(a)} = 10 \cdot \log_{10} \frac{p^2}{M_m^{(a)}}. \quad (24)$$

We define the MSE sample variance $S_M^{2(a)}$ of a sequence of groups of a frames each as

$$S_M^{2(a)} = \frac{1}{N/a-1} \sum_{n=0}^{N/a-1} (M_n^{(a)} - \bar{M})^2, \quad (25)$$

and the corresponding MSE standard deviation $S_M^{(a)}$ as

$$S_M^{(a)} = \sqrt{S_M^{2(a)}}. \quad (26)$$

We define the quality standard deviation $S_Q^{(a)}$ of a sequence of groups of a frames each as

$$S_Q^{(a)} = 10 \cdot \log_{10} \frac{p^2}{S_M^{(a)}}. \quad (27)$$

We define the coefficient of quality variation $CoQV^{(a)}$ of a sequence of groups of a frames each as

$$CoQV^{(a)} = \frac{S_Q^{(a)}}{\bar{Q}}. \quad (28)$$

We define the *alternative quality standard deviation* for groups of a frames each as

$$S_Q'^{(a)} = \sqrt{\frac{1}{N/a-1} \sum_{n=0}^{N/a-1} (Q_n'^{(a)} - \bar{Q}')^2}, \quad (29)$$

where $Q_n'^{(a)} = \frac{1}{a} \sum_{n=ma}^{(m+1)a-1} Q_n$. We define the corresponding alternative coefficient of quality variation as

$$CoQV'^{(a)} = \frac{S_Q'^{(a)}}{\bar{Q}'}. \quad (30)$$

We define the quality range (in dB) of a sequence of groups of a frames each as

$$Q_{\min}^{\max(a)} = \max_{0 \leq n \leq N/a-1} Q_n^{(a)} - \min_{0 \leq n \leq N/a-1} Q_n^{(a)}. \quad (31)$$

We estimate the MSE autocorrelation coefficient for groups of a frames $\rho_M^{(a)}$ for lag k , $k = 0, a, 2a, \dots, N/a - 1$ frames as

$$\rho_M^{(a)}(k) = \frac{1}{N/a-k} \sum_{n=0}^{N/a-k-1} \frac{(M_n^{(a)} - \bar{M})(M_{n+k}^{(a)} - \bar{M})}{S_M^{(a)}}. \quad (32)$$

5.3 Correlation between Frame Sizes and Qualities

We define the covariance between the frame size and the MSE frame quality as

$$S_{XM} = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})(M_n - \bar{M}), \quad (33)$$

and the *size-MSE quality correlation coefficient* as

$$\rho_{XM} = \frac{S_{XM}}{S_X \cdot S_M}. \quad (34)$$

We define the covariance between the frame size and (PSNR) frame quality as

$$S_{XQ} = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})(Q_n - \bar{Q}'), \quad (35)$$

and the *size-quality correlation coefficient* as

$$\rho_{XQ} = \frac{S_{XQ}}{S_X \cdot S_Q'}. \quad (36)$$

Once again we define the covariance between the aggregated frame sizes $X_n^{(a)}$, $n = 0, \dots, N/a - 1$, and the aggregated MSE qualities $M_n^{(a)}$, $n = 0, \dots, N/a - 1$, as

$$S_{XM}^{(a)} = \frac{1}{N/a - 1} \sum_{n=0}^{N/a-1} (X_n^{(a)} - \bar{X})(M_n^{(a)} - \bar{M}), \quad (37)$$

and the corresponding correlation coefficient as

$$\rho_{XM}^{(a)} = \frac{S_{XM}^{(a)}}{S_X^{(a)} \cdot S_M^{(a)}}. \quad (38)$$

We define the covariance between aggregated frame size $X_n^{(a)}$, $n = 0, \dots, N/a - 1$, and the aggregated (PSNR) qualities $Q_n^{(a)}$, $n = 0, \dots, N/a - 1$, as

$$S_{XQ}^{(a)} = \frac{1}{N/a - 1} \sum_{n=0}^{N/a-1} (X_n^{(a)} - \bar{X})(Q_n^{(a)} - \bar{Q}), \quad (39)$$

and the corresponding correlation coefficient as

$$\rho_{XQ}^{(a)} = \frac{S_{XQ}^{(a)}}{S_X^{(a)} \cdot S_Q^{(a)}}. \quad (40)$$

5.4 Analysis of Video Traffic

Table 2 gives the the mean \bar{X} , the coefficient of variation CoV_X , and peak-to-mean ratio X_{\max}/\bar{X} of the frame sizes as well as the mean bit rates \bar{X}/T and the peak bit rates X_{\max}/T , as defined in Section 5.1.

From Table 2 we observe that the CoV_X increases as the encoded video rate increases from very low bit rates to medium bit rates, and then the CoV_X decreases as the encoded video rate increases further from the medium rate to very high rate. For example, for the video sequence *Foot Ball w/c* in Table 2, we observe that the CoV_X is 0.183 at 25 kbps and increases to 0.292 at 300 kbps. Then it starts to decrease back to 0.216 at 1600 kbps causing a *hump* like behavior. The causes for this phenomenon and its implications on channel utilization and buffer requirements will be explored in future work. This same phenomenon has been observed in [24] for MPEG-4 video traces. We observe from Table 2 that the peak to mean ratio of the frame sizes exhibits a similar hump behavior.

Table 2: Overview of frame statistics of intra-encoded video

Enc. M.	Video	Compression. ratio YUV:3D-EZBC	Frame Size			Bit Rate	
			Mean \bar{X} [kbyte]	$CoV_{\bar{X}}$ $S_{\bar{X}}/\bar{X}$	Peak/Mean X_{\max}/\bar{X}	Mean \bar{X}/T [Mbps]	Peak X_{\max}/T [Mbps]
25	<i>Terminator</i>	367.724	0.103	0.144	1.944	0.025	0.048
75		121.982	0.312	0.265	3.831	0.075	0.287
100		91.392	0.416	0.293	5.753	0.100	0.574
300		30.434	1.249	0.312	5.483	0.300	1.644
600		15.212	2.499	0.296	4.850	0.600	2.909
800		11.408	3.332	0.281	3.985	0.800	3.187
1000		9.126	4.166	0.263	3.948	1.000	3.947
1200		7.604	4.999	0.247	3.377	1.200	4.051
1400		6.518	5.833	0.225	2.940	1.400	4.116
1600		5.703	6.666	0.197	3.022	1.600	4.834
25	<i>Lady and the Tramp</i>	367.757	0.103	0.123	2.119	0.025	0.053
75		121.982	0.312	0.222	2.445	0.075	0.183
100		91.365	0.416	0.239	2.483	0.100	0.248
300		30.434	1.249	0.239	2.441	0.300	0.732
600		15.212	2.499	0.214	2.141	0.600	1.284
800		11.408	3.332	0.195	2.154	0.800	1.722
1000		9.126	4.166	0.175	1.899	1.000	1.898
1200		7.605	4.999	0.161	1.867	1.200	2.239
1400		6.518	5.832	0.145	1.764	1.400	2.470
1600		5.703	6.666	0.125	1.627	1.600	2.604
25	<i>Foot Ball w/c</i>	367.653	0.103	0.183	2.679	0.025	0.066
75		121.979	0.312	0.280	2.519	0.075	0.188
100		91.425	0.416	0.291	2.434	0.100	0.243
300		30.434	1.249	0.292	2.382	0.300	0.714
600		15.212	2.499	0.286	2.497	0.600	1.498
800		11.408	3.332	0.276	2.316	0.800	1.852
1000		9.126	4.166	0.262	2.315	1.000	2.315
1200		7.605	4.999	0.249	2.180	1.200	2.616
1400		6.518	5.832	0.232	2.030	1.400	2.842
1600		5.703	6.666	0.216	1.904	1.600	3.046
25	<i>Tonight Show w/c</i>	367.754	0.103	0.135	2.012	0.025	0.050
75		121.987	0.312	0.254	3.225	0.075	0.241
100		91.426	0.416	0.267	3.093	0.100	0.309
300		30.433	1.249	0.280	3.521	0.300	1.056
600		15.212	2.499	0.259	3.012	0.600	1.807
800		11.408	3.332	0.241	2.516	0.800	2.012
1000		9.126	4.166	0.219	2.487	1.000	2.486
1200		7.605	4.999	0.203	2.239	1.200	2.686
1400		6.518	5.833	0.186	1.990	1.400	2.786
1600		5.703	6.666	0.168	1.954	1.600	3.125

Table 3 gives the the mean \bar{Y} , the coefficient of variation CoV_Y , and peak-to-mean ratio Y_{\max}/\bar{Y} of the GoP sizes as well as the mean bit rates \bar{Y}/GT and the peak bit rates Y_{\max}/GT , as defined in Section 5.1. We observe that the CoV_Y is smaller for the GoP level compared to the frame level depicted on Table 2. Here too we observe the *hump* phenomenon of increasing CoV_Y from low bit rates to mid bit rates and then decreasing from mid bit rates to high bit rates.

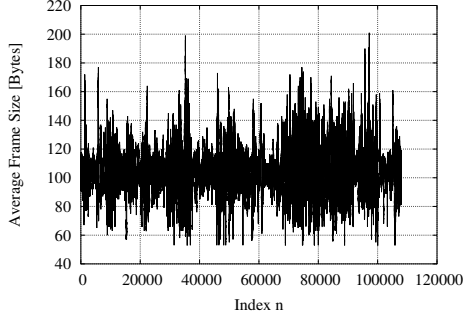
Table 3: Overview of GoP statistics of intra frame QCIF video

Enc. M.	Video	Encoded bit rate	GoP Size			Bit Rate	
			Mean \bar{Y} [kbyte]	CoV S_Y/\bar{Y}	Peak/Mean Y_{\max}/\bar{Y}	Mean $\bar{Y}/(Gt)$ [Mbps]	Peak $Y_{\max}/(Gt)$ [Mbps]
Intra Frame Video	<i>Terminator</i>	25	1.654	0.133	1.763	0.025	0.044
		75	4.986	0.248	2.351	0.075	0.176
		100	13.407	0.604	3.876	0.201	0.780
		300	19.983	0.294	2.837	0.300	0.850
		600	39.980	0.278	2.506	0.600	1.503
		800	53.311	0.264	2.328	0.800	1.862
		1000	66.643	0.247	2.206	1.000	2.206
		1200	79.974	0.232	2.091	1.200	2.508
		1400	93.305	0.211	1.926	1.400	2.696
		1600	106.637	0.185	1.744	1.600	2.790
	<i>Lady and the Tramp</i>	25	1.571	0.212	2.123	0.024	0.050
		75	4.637	0.337	2.522	0.070	0.175
		100	12.924	0.626	3.639	0.194	0.705
		300	18.434	0.373	2.551	0.277	0.705
		600	36.832	0.361	2.237	0.552	1.236
		800	49.097	0.351	2.195	0.736	1.616
		1000	61.363	0.340	2.048	0.920	1.885
		1200	73.627	0.334	2.012	1.104	2.222
		1400	85.892	0.327	1.887	1.288	2.432
		1600	98.157	0.318	1.718	1.472	2.530
	<i>Foot Ball w/c</i>	25	1.654	0.172	1.820	0.025	0.045
		75	4.986	0.266	2.151	0.075	0.161
		100	6.652	0.277	2.185	0.100	0.218
		300	19.983	0.278	2.314	0.300	0.694
		600	39.980	0.272	2.458	0.600	1.474
		800	53.312	0.262	2.296	0.800	1.836
		1000	66.643	0.249	2.293	1.000	2.292
		1200	79.974	0.237	2.163	1.200	2.595
		1400	93.305	0.220	1.950	1.400	2.729
		1600	106.637	0.205	1.889	1.600	3.021
	<i>Tonight Show w/c</i>	25	1.654	0.126	1.950	0.025	0.048
		75	4.986	0.240	2.919	0.075	0.218
		100	6.652	0.253	2.988	0.100	0.298
		300	19.983	0.265	3.392	0.300	1.017
		600	39.980	0.244	2.935	0.600	1.760
		800	53.311	0.227	2.465	0.800	1.971
		1000	66.643	0.206	2.440	1.000	2.439
		1200	79.974	0.191	2.200	1.200	2.639
		1400	93.306	0.176	1.934	1.400	2.706
		1600	106.637	0.159	1.920	1.600	3.072

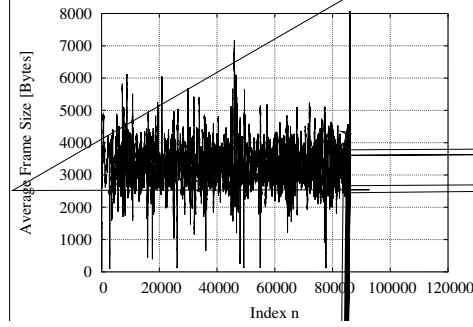
Next, we provide plots to illustrate the video traffic characteristics and statistical characteristics of the following video sequences: (a) *Terminator* encoded at 25 kbps, (b) *Terminator* encoded at 100 kbps, (c) *Lady and the Tramp* encoded at 300 kbps, (d) *Lady and the Tramp* encoded at 800 kbps, (e) *Foot Ball w/c* encoded at 1000 kbps, and (f) *Foot Ball w/c* encoded at 1600 kbps. The video sequences were chosen from the three different genres action, cartoon, and a TV talk show with commercials; to give a representation of different video content.

Figure 6 illustrates the behavior of the frame sizes (in bytes) as a function of the frame index n . We observed that *Terminator* encoded at 100 kbps is smoother than the *Terminator* encoded at 25 kbps. But *Lady and the Tramp* encoded at 300 kbps shows more variations than

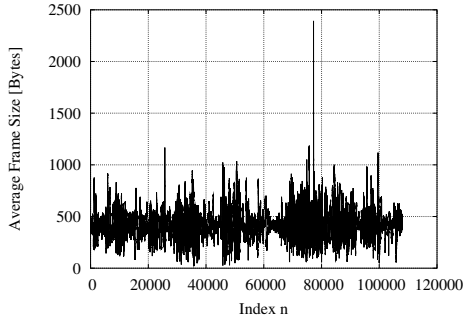
the *Terminator* at 100 kbps.



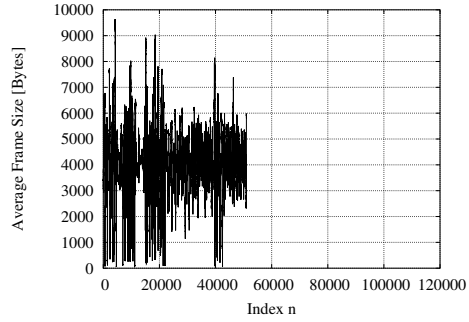
a) *Terminator* at 25 kbps



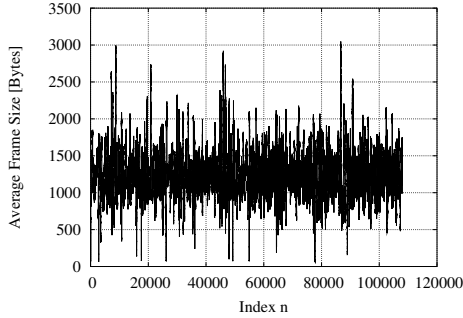
d) *Lady and the Tramp* at 800 kbps



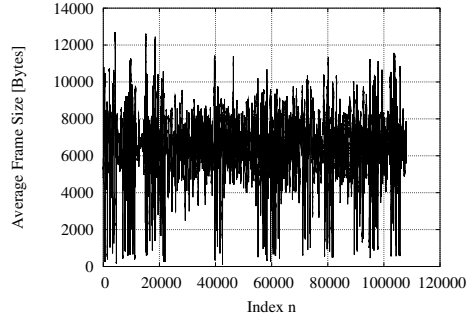
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps



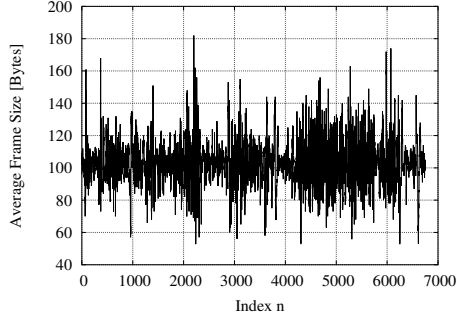
c) *Lady and the Tramp* at 300 kbps



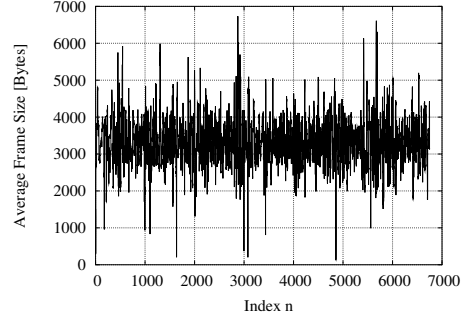
f) *Foot Ball w/c* at 1600 kbps

Figure 6: Frame size X_n as a function of the frame index n for intra frame QCIF video.

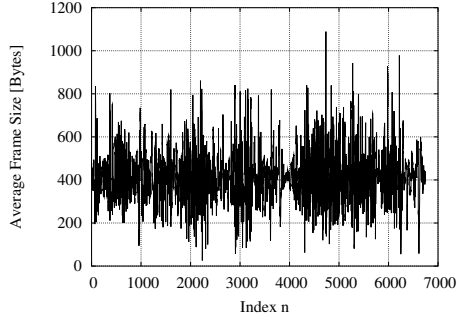
By visual inspection of Figure 6 *Foot Ball w/c* encoded at 1000 kbps and 1600 kbps both have almost the same variations, but obviously due to different bit rates they are centered at the corresponding frame sizes. For all bit rate encodings, we observed that some parts of the trace that had higher variations than the others, which correspond to different scenes of the video sequence. Next, we observe the behavior of the GoP sizes as a function of the GoP index m , illustrated in Figure 7. In Section 2 we described the behavior of the MC-3DEZBC encoder's rate control which gives an insight to the large variations observed in Figure 7.



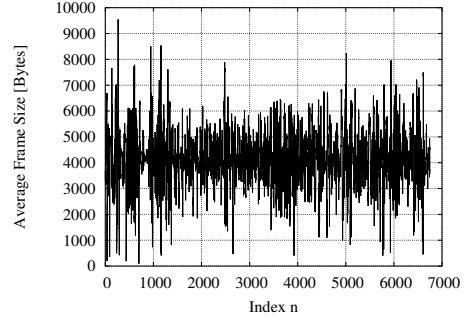
a) *Terminator* at 25 kbps



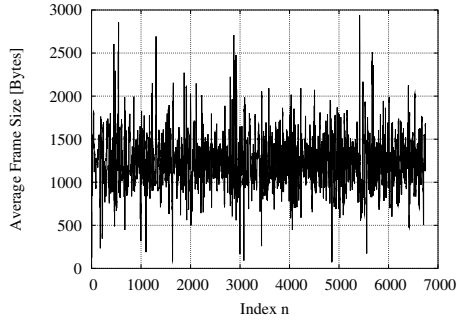
d) *Lady and the Tramp* at 800 kbps



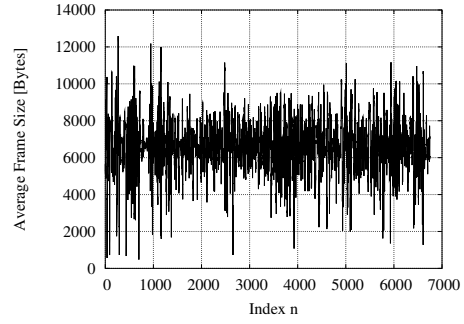
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps



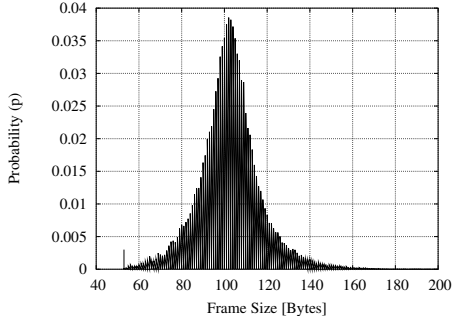
c) *Lady and the Tramp* at 300 kbps



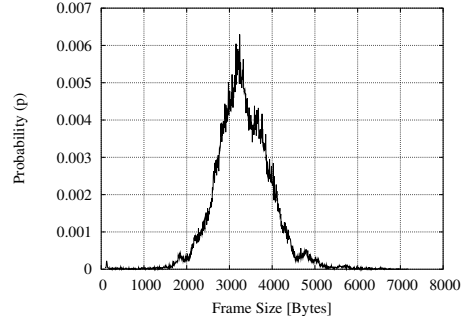
f) *Foot Ball w/c* at 1600 kbps

Figure 7: GoP size Y_m as a function of the index m for intra frame QCIF video.

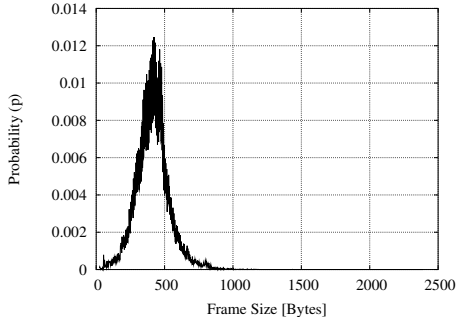
In contrast to the frame level, here we observe a much smoother plot due to the fact that when taking the aggregate of frame sizes over a GoP the variations get somewhat smoothed out. We have observed this behavior for different aggregation levels, not shown here due to space constraints. But at the GoP level we still observe different variations along the trace due to different scenes of the video. Figure 8 illustrates the histograms of the frame sizes. We observe a single peak with a relative smooth slope in contrast to the MPEG 4 traces where a double peak was observed [25].



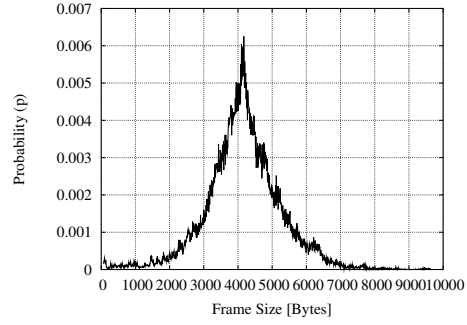
a) *Terminator* at 25 kbps



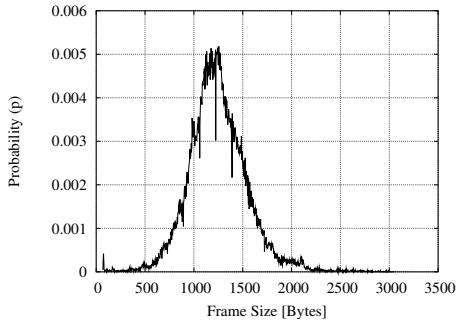
d) *Lady and the Tramp* at 800 kbps



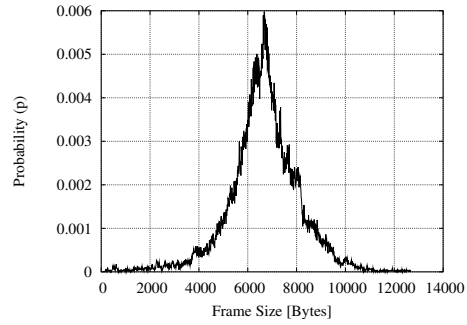
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps



c) *Lady and the Tramp* at 300 kbps



f) *Foot Ball w/c* at 1600 kbps

Figure 8: Frame size histograms for intra frame QCIF.

At the GoP level the histograms are much smoother relative to the frame level histograms,

as illustrated in Figure 9. A thorough investigation of the histograms is out of the scope of this technical report and will be investigated in the future work.

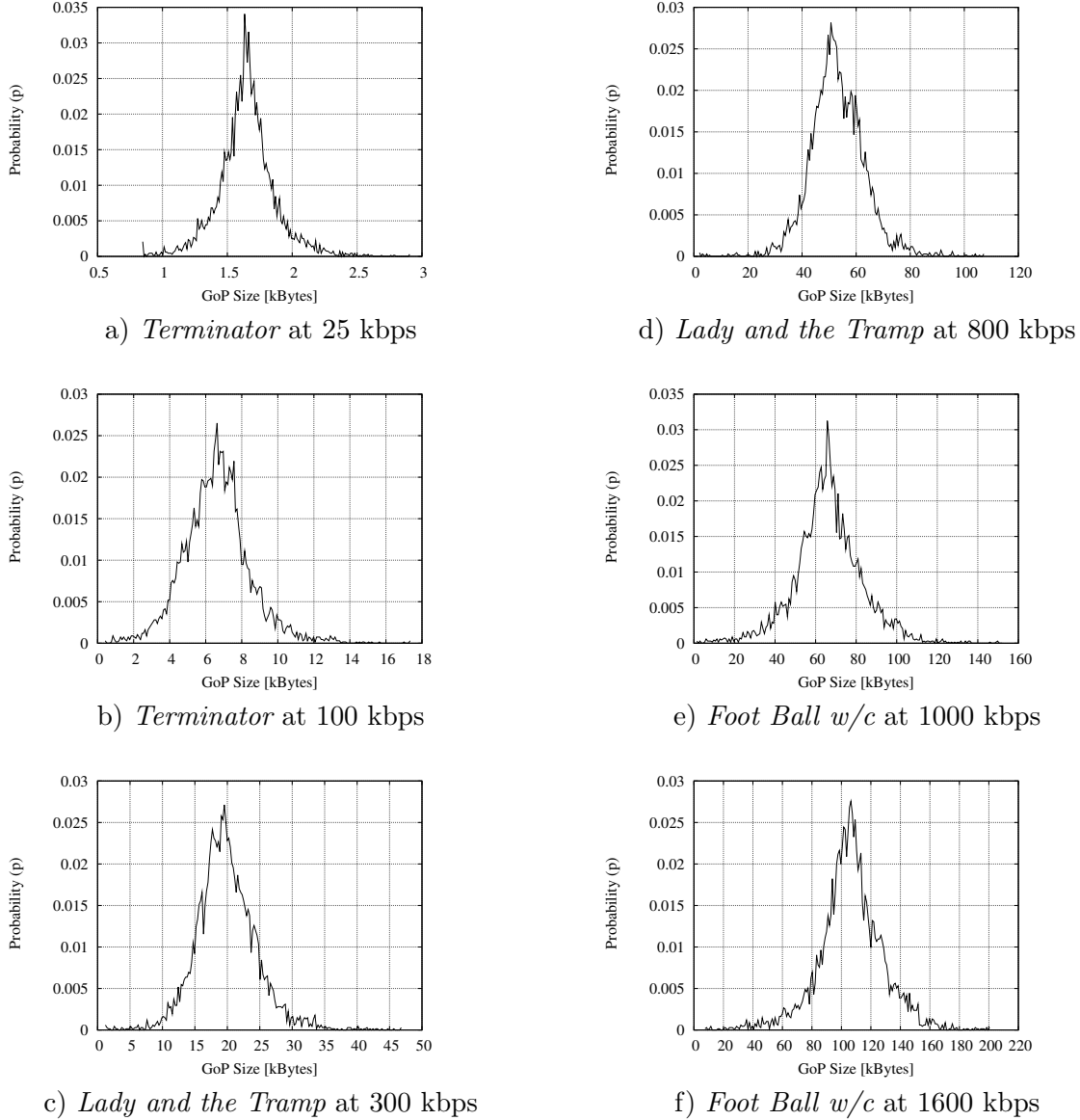


Figure 9: GoP size histograms for intra frame QCIF video.

Figure 10 illustrates the autocorrelation coefficient as a function of the frame lag k (in frames). Here we observe a smooth decaying curve. This is in contrast to the spiky autocorrelation coefficient behavior observed in MPEG-4 due to the three different frame types I, P and B.

In Figure 11 we see a different type of behavior of the autocorrelation coefficient as a function of the lag k (in GoPs).

Table 4 gives the Hurst parameter determined with the R/S method depicted in the first line of the given video and encoded bit rate, while the second line with the same video and bit rate depicts the Hurst parameter estimated with the periodogram as a function of the aggregation level a . We calculate the H values for $a = 1, 2, 4, 16, 24, 32, 64, 128, 256, 400, 480, 560, 640$, and 800.

Table 4: Hurst parameters estimated from pox diagram of R/S and periodogram as a function of the aggregation level a .

Enc. M.	Video	Aggregation level a [frames]											
		1	16	24	32	64	128	256	400	480	560	640	800
25	<i>Terminator</i>	0.697	0.521	0.511	0.480	0.424	0.364	0.307	0.326	0.287	0.268	0.351	0.262
25		1.065	0.973	0.893	0.706	0.454	0.108	-0.059	-0.210	-0.176	-0.248	-0.306	-0.306
75		0.673	0.498	0.489	0.463	0.412	0.364	0.273	0.300	0.226	0.264	0.310	0.233
75		1.080	0.998	0.930	0.746	0.507	0.166	-0.002	-0.158	-0.239	-0.172	-0.354	-0.354
100		0.674	0.496	0.479	0.457	0.406	0.346	0.266	0.294	0.227	0.250	0.296	0.323
100		1.075	1.015	0.931	0.762	0.526	0.170	-0.002	-0.222	-0.196	-0.239	-0.307	-0.307
300		0.665	0.484	0.462	0.443	0.384	0.327	0.256	0.270	0.245	0.221	0.301	0.247
300		1.073	1.002	0.930	0.747	0.511	0.178	0.003	-0.148	-0.157	-0.150	-0.314	-0.314
600		0.664	0.480	0.460	0.431	0.367	0.318	0.254	0.251	0.244	0.217	0.285	0.217
600		1.066	1.000	0.924	0.742	0.490	0.195	-0.022	-0.136	-0.132	-0.101	-0.303	-0.303
800		0.667	0.479	0.464	0.430	0.366	0.323	0.268	0.259	0.246	0.232	0.291	0.215
800		1.058	0.990	0.924	0.737	0.488	0.220	-0.032	-0.098	-0.172	-0.092	-0.292	-0.292
1000		0.671	0.477	0.459	0.429	0.366	0.320	0.260	0.261	0.230	0.227	0.284	0.206
1000		1.057	0.986	0.916	0.725	0.484	0.227	-0.062	-0.107	-0.173	-0.111	-0.306	-0.306
1200		0.678	0.481	0.464	0.432	0.370	0.325	0.268	0.271	0.244	0.243	0.289	0.195
1200		1.052	0.983	0.913	0.733	0.480	0.234	-0.015	-0.092	-0.187	-0.115	-0.309	-0.309
1400		0.683	0.478	0.465	0.433	0.371	0.325	0.268	0.274	0.230	0.254	0.295	0.191
1400		1.041	0.968	0.907	0.738	0.473	0.215	-0.050	-0.108	-0.215	-0.120	-0.332	-0.332
1600		0.682	0.477	0.461	0.431	0.369	0.330	0.260	0.266	0.207	0.249	0.301	0.184
1600		1.027	0.950	0.893	0.719	0.457	0.174	-0.098	-0.155	-0.230	-0.180	-0.373	-0.373
25	<i>Lady Tramp</i>	0.723	0.493	0.450	0.410	0.346	0.306	0.308	0.270	0.278	0.259	0.264	0.348
25		1.113	1.024	0.965	0.752	0.448	0.171	-0.033	-0.132	-0.202	-0.166	-0.174	-0.174
75		0.704	0.466	0.434	0.396	0.345	0.296	0.269	0.237	0.256	0.224	0.257	0.255
75		1.158	1.041	0.997	0.786	0.494	0.205	0.063	-0.120	-0.108	-0.237	-0.122	-0.122
100		0.697	0.465	0.429	0.393	0.342	0.303	0.255	0.245	0.252	0.262	0.289	0.245
100		1.149	1.021	1.002	0.767	0.498	0.176	0.008	-0.056	-0.063	-0.175	-0.087	-0.087
300		0.693	0.465	0.429	0.396	0.346	0.307	0.248	0.254	0.259	0.254	0.297	0.256
300		1.150	1.016	0.995	0.755	0.482	0.148	-0.038	-0.078	-0.060	-0.152	-0.059	-0.059
600		0.690	0.456	0.424	0.386	0.339	0.301	0.241	0.264	0.248	0.252	0.286	0.254
600		1.136	0.994	0.974	0.723	0.469	0.118	-0.095	-0.115	-0.091	-0.119	-0.068	-0.068
800		0.690	0.447	0.415	0.377	0.331	0.288	0.234	0.264	0.246	0.237	0.268	0.275
800		1.124	0.984	0.965	0.697	0.457	0.087	-0.113	-0.154	-0.116	-0.118	-0.063	-0.063
1000		0.679	0.441	0.411	0.370	0.324	0.282	0.238	0.276	0.246	0.230	0.263	0.270
1000		1.131	0.986	0.980	0.710	0.462	0.082	-0.126	-0.164	-0.127	-0.097	-0.066	-0.066
1200		0.685	0.443	0.413	0.373	0.330	0.292	0.240	0.281	0.240	0.237	0.256	0.265
1200		1.125	0.979	0.977	0.695	0.438	0.074	-0.128	-0.175	-0.130	-0.097	-0.054	-0.054
1400		0.683	0.447	0.415	0.377	0.335	0.296	0.250	0.302	0.259	0.241	0.247	0.288
1400		1.117	0.972	0.968	0.681	0.417	0.059	-0.154	-0.193	-0.142	-0.103	-0.028	-0.028
1600		0.687	0.452	0.417	0.381	0.335	0.301	0.274	0.325	0.260	0.235	0.226	0.312
1600		1.118	0.986	0.975	0.678	0.425	0.072	-0.159	-0.211	-0.159	-0.111	-0.070	-0.070
25	<i>Foot Ball w/c</i>	0.693	0.473	0.447	0.408	0.348	0.293	0.253	0.248	0.203	0.251	0.237	0.202
25		1.055	0.977	0.939	0.740	0.469	0.068	-0.181	-0.214	-0.134	-0.088	-0.263	-0.263
75		0.669	0.456	0.424	0.383	0.310	0.266	0.236	0.202	0.217	0.248	0.196	0.195
75		1.038	0.962	0.920	0.704	0.434	0.057	-0.146	-0.158	-0.197	-0.157	-0.240	-0.240
100		0.674	0.467	0.434	0.395	0.316	0.273	0.246	0.208	0.230	0.267	0.193	0.190
100		1.026	0.961	0.916	0.713	0.449	0.037	-0.119	-0.192	-0.123	-0.208	-0.227	-0.227
300		0.693	0.495	0.465	0.430	0.357	0.304	0.278	0.238	0.294	0.318	0.245	0.181
300		1.027	0.950	0.893	0.719	0.457	0.174	-0.098	-0.155	-0.230	-0.180	-0.373	-0.373

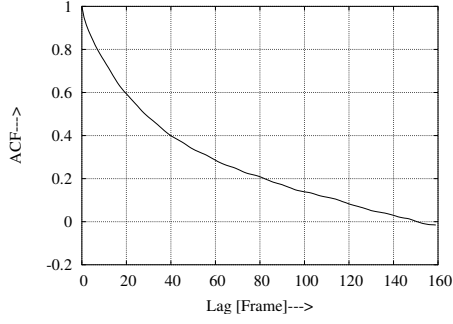
Table 4: *continued*

Bit rate	Video	Aggregation level a [frames]											
		1	16	24	32	64	128	256	400	480	560	640	800
300			1.012	0.973	0.918	0.737	0.498	0.174	-0.067	-0.133	-0.201	-0.192	-0.169
600		0.698	0.503	0.470	0.442	0.378	0.317	0.283	0.248	0.298	0.317	0.266	0.202
600			1.008	0.971	0.911	0.755	0.511	0.205	-0.018	-0.152	-0.157	-0.162	-0.071
800		0.693	0.500	0.468	0.442	0.377	0.312	0.274	0.237	0.275	0.311	0.259	0.202
800			1.010	0.976	0.903	0.750	0.498	0.191	-0.005	-0.157	-0.137	-0.193	-0.111
1000		0.694	0.503	0.472	0.444	0.384	0.319	0.285	0.241	0.275	0.301	0.271	0.203
1000			1.013	0.975	0.901	0.750	0.489	0.185	-0.006	-0.163	-0.132	-0.209	-0.147
1200		0.691	0.499	0.467	0.441	0.379	0.314	0.279	0.234	0.259	0.284	0.262	0.195
1200			1.018	0.974	0.917	0.742	0.486	0.166	-0.016	-0.172	-0.132	-0.230	-0.158
1400		0.696	0.503	0.469	0.444	0.387	0.318	0.283	0.240	0.254	0.295	0.271	0.205
1400			1.011	0.969	0.906	0.737	0.472	0.146	-0.033	-0.178	-0.142	-0.241	-0.179
1600		0.694	0.505	0.472	0.446	0.394	0.328	0.294	0.238	0.251	0.288	0.272	0.215
1600			1.013	0.967	0.899	0.730	0.474	0.128	-0.044	-0.195	-0.189	-0.261	-0.178
25	<i>Tonight</i> <i>Show w/c</i>	0.703	0.499	0.484	0.453	0.416	0.360	0.292	0.338	0.336	0.370	0.357	0.299
25			1.020	0.934	0.875	0.695	0.420	0.137	-0.191	-0.288	-0.099	-0.229	-0.335
75		0.703	0.500	0.476	0.454	0.417	0.387	0.321	0.300	0.299	0.395	0.330	0.318
75			1.022	0.949	0.899	0.734	0.474	0.211	-0.039	-0.101	-0.043	-0.249	-0.216
100		0.712	0.499	0.473	0.453	0.413	0.375	0.305	0.279	0.292	0.370	0.322	0.327
100			1.014	0.941	0.898	0.735	0.488	0.226	-0.092	-0.107	-0.086	-0.301	-0.242
300		0.700	0.494	0.464	0.437	0.405	0.371	0.318	0.251	0.342	0.359	0.323	0.265
300			0.981	0.917	0.885	0.736	0.481	0.247	-0.038	-0.071	-0.124	-0.158	-0.090
600		0.696	0.499	0.479	0.449	0.407	0.373	0.312	0.237	0.373	0.329	0.332	0.212
600			0.951	0.883	0.845	0.709	0.461	0.215	-0.056	-0.044	-0.158	-0.130	-0.171
800		0.693	0.504	0.484	0.459	0.412	0.379	0.337	0.271	0.382	0.317	0.345	0.213
800			0.955	0.877	0.838	0.707	0.477	0.208	-0.058	-0.051	-0.149	-0.079	-0.210
1000		0.691	0.505	0.483	0.462	0.410	0.370	0.330	0.281	0.386	0.301	0.354	0.212
1000			0.972	0.875	0.820	0.694	0.470	0.189	-0.088	-0.063	-0.130	-0.079	-0.217
1200		0.693	0.502	0.480	0.459	0.402	0.358	0.335	0.269	0.365	0.274	0.361	0.206
1200			0.976	0.899	0.850	0.680	0.465	0.195	-0.079	-0.048	-0.090	-0.077	-0.233
1400		0.691	0.506	0.484	0.466	0.409	0.365	0.339	0.287	0.391	0.270	0.368	0.200
1400			0.983	0.899	0.844	0.677	0.465	0.196	-0.070	-0.049	-0.102	-0.068	-0.251
1600		0.691	0.509	0.486	0.470	0.415	0.368	0.334	0.301	0.403	0.265	0.384	0.225
1600			1.001	0.908	0.843	0.668	0.483	0.187	-0.106	-0.040	-0.090	-0.059	-0.235

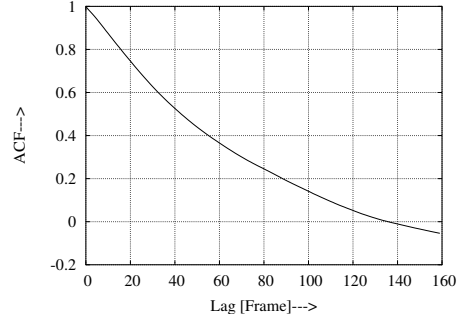
Table 5 gives the Hurst parameter estimated using the variance time plot. Also Table 5 provides the values of the scaling parameters c_f and α (the latter shown as $H = (1 + \alpha)/2$) estimated from the logscale diagram. Figure 13, 14, and 15 give the variance-time plots, the pox plots of R/S (for $a = 16$), and the periodogram (for $a = 16$). H estimates typically decrease as the aggregation level increase from $a = 1$ to around $a = 200$ and then are more or less stable [5]. We make similar observations here. The pox plot of R/S for $a = 1$ and the periodogram for $a \leq 64$ give H estimates larger than 0.5, which usually indicate the presence of long range dependence in the video traffic.

However the H estimates obtained for larger aggregation levels a are all below 0.5 indicates that there is in fact, no long range dependence in the traffic. All in all our investigations indicate that there is no significant long range dependence in the video traffic.

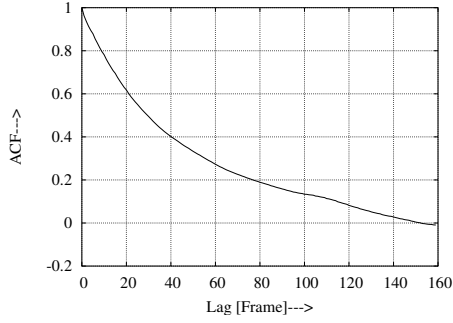
In Figure 12 we illustrate the logscale diagram of the H value estimates with a general trend of a increasing curve for lower octaves j , and then a decreasing trend for higher octaves j .



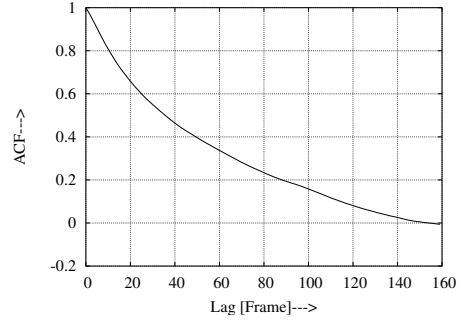
a) *Terminator* at 25 kbps



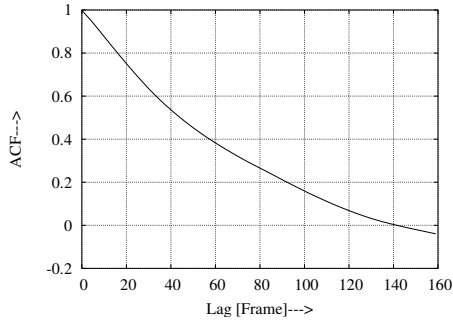
d) *Lady and the Tramp* at 800 kbps



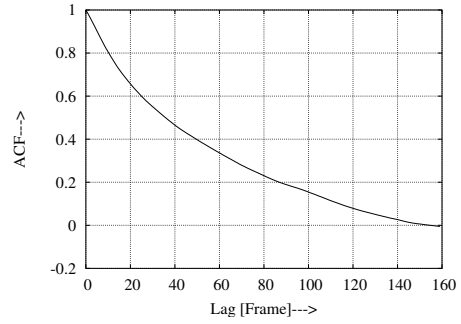
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

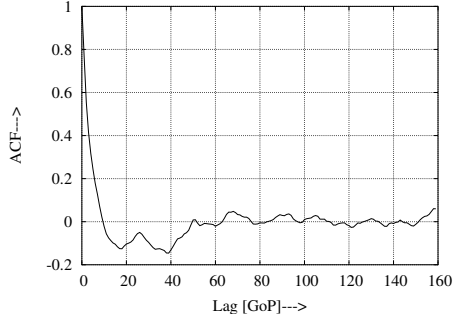


c) *Lady and the Tramp* at 300 kbps

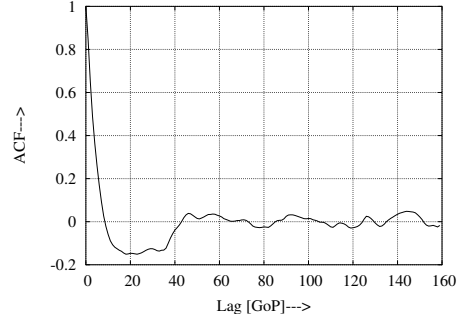


f) *Foot Ball w/c* at 1600 kbps

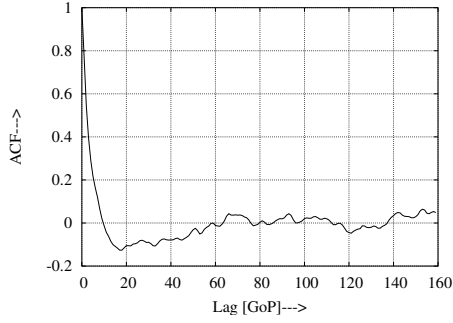
Figure 10: Frame size autocorrelations for intra frame QCIF video.



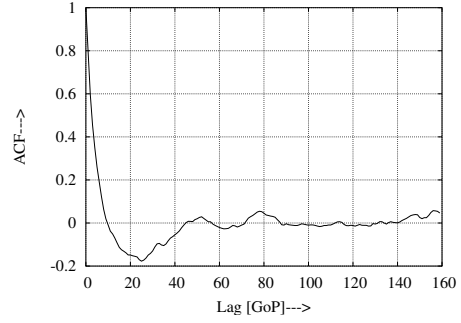
a) *Terminator* at 25 kbps



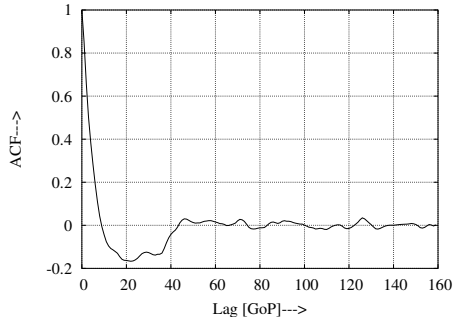
d) *Lady and the Tramp* at 800 kbps



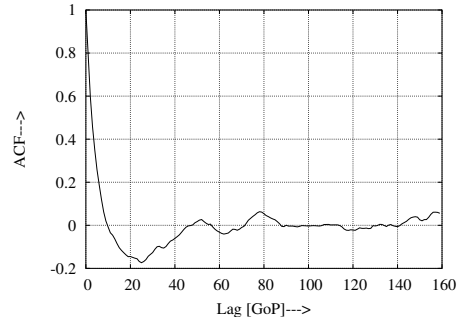
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

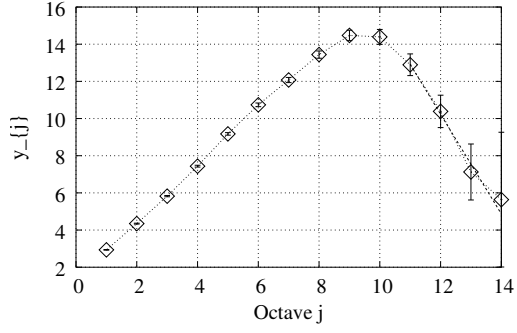


c) *Lady and the Tramp* at 300 kbps

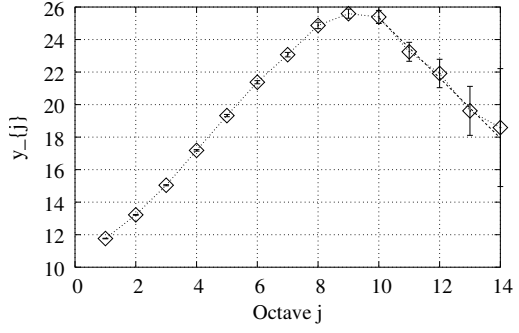


f) *Foot Ball w/c* at 1600 kbps

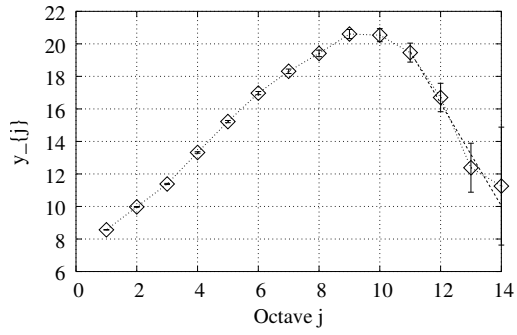
Figure 11: GoP size autocorrelations for intra frame QCIF video.



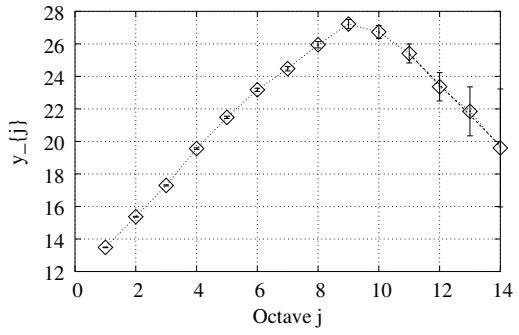
a) *Terminator* at 25 kbps



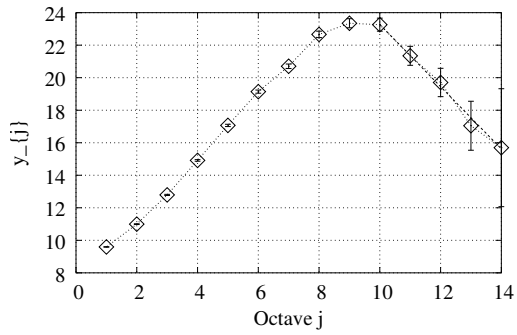
d) *Lady and the Tramp* at 800 kbps



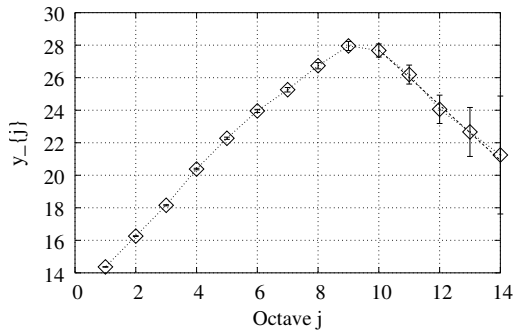
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

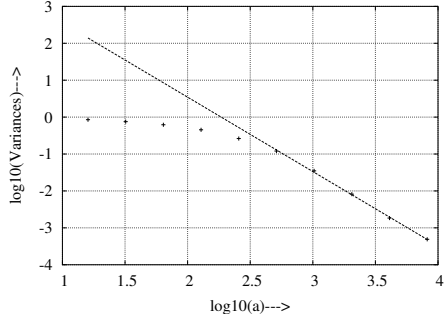


c) *Lady and the Tramp* at 300 kbps

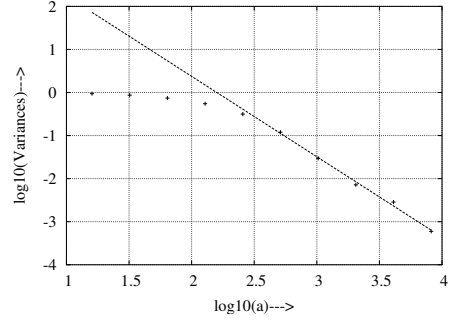


f) *Foot Ball w/c* at 1600 kbps

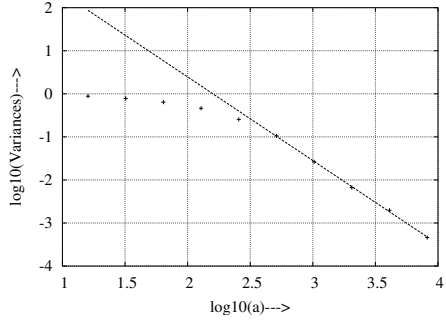
Figure 12: Logscale diagrams for intra frame QCIF video.



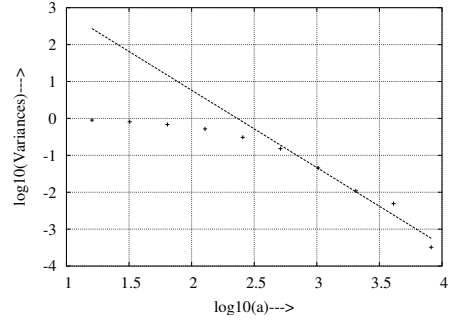
a) *Terminator* at 25 kbps



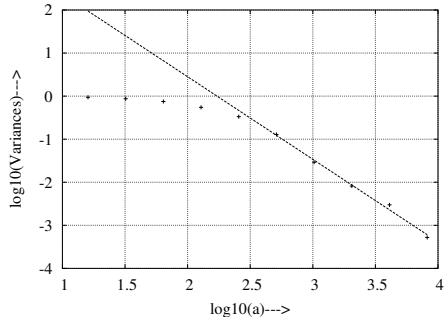
d) *Lady and the Tramp* at 800 kbps



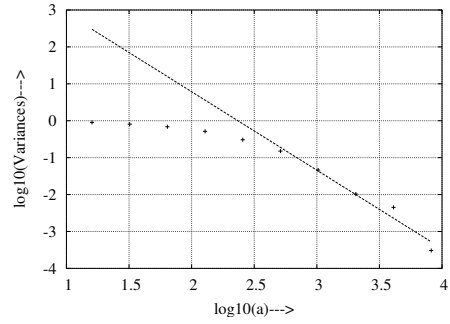
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

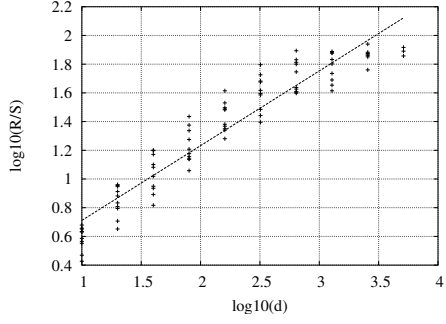


c) *Lady and the Tramp* at 300 kbps

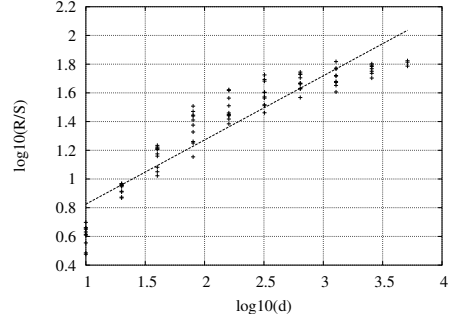


f) *Foot Ball w/c* at 1600 kbps

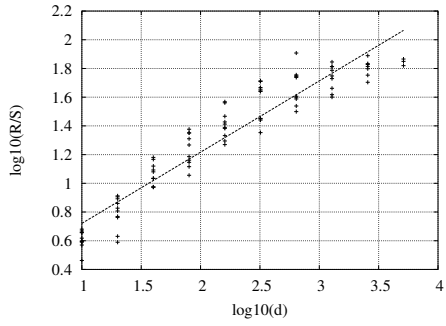
Figure 13: Variance time plots for intra frame QCIF video



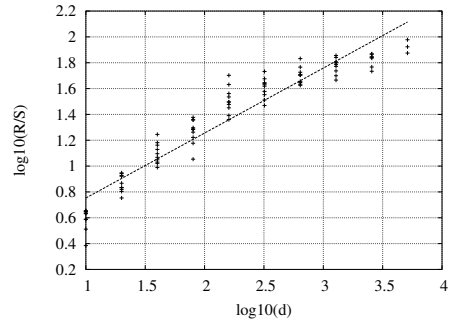
a) *Terminator* at 25 kbps



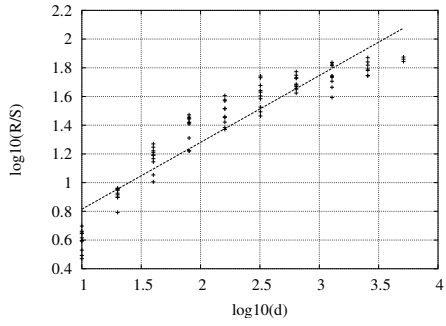
d) *Lady and the Tramp* at 800 kbps



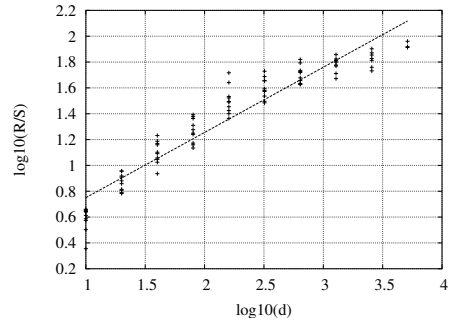
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

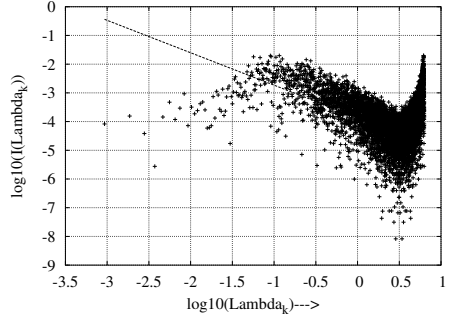


c) *Lady and the Tramp* at 300 kbps

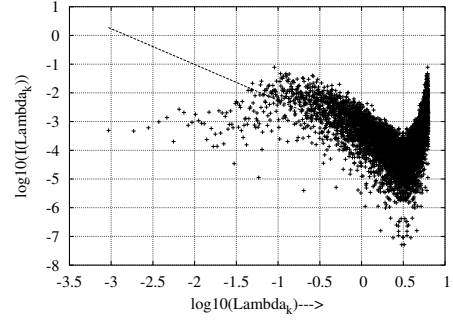


f) *Foot Ball w/c* at 1600 kbps

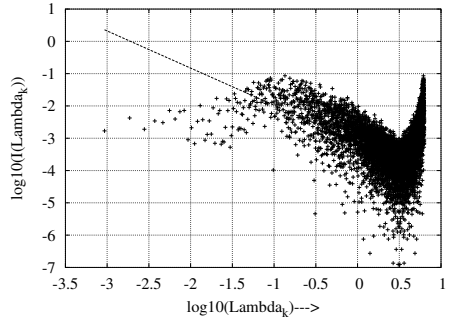
Figure 14: POX plots of R/S for aggregation level $a = 16$ for intra frame QCIF video.



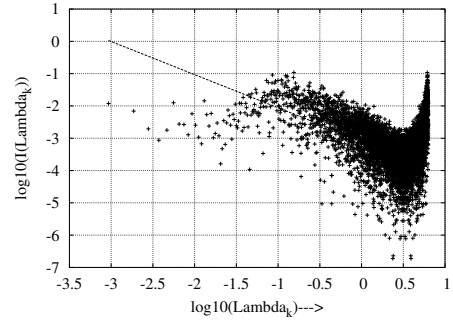
a) *Terminator* at 25 kbps



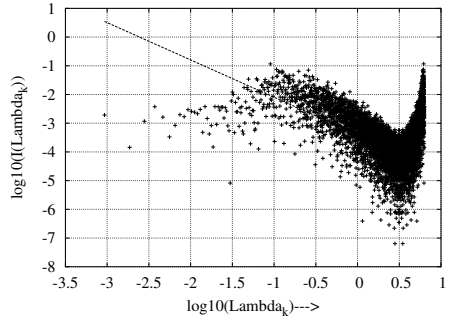
d) *Lady and the Tramp* at 800 kbps



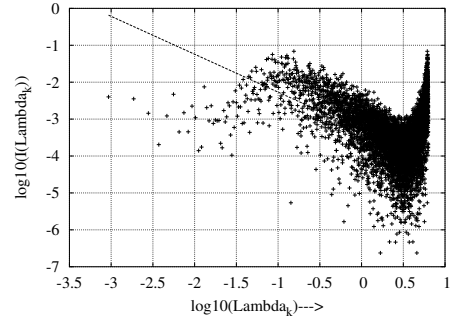
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps



c) *Lady and the Tramp* at 300 kbps



f) *Foot Ball w/c* at 1600 kbps

Figure 15: Periodogram for aggregation level $a = 16$ for intra frame QCIF video.

Table 5: Hurst parameters estimated from variance time plot, scaling parameters estimated from logscale diagram.

Enc. M.	Video	VT H	Logscale Diagram		
			c_f	α	H
25	<i>Terminator</i>	-0.007	6696783360.000	-2.684	-0.842
75		0.014	134226712199168.000	-4.100	-1.550
100		0.029	361141272576.000	-3.159	-1.080
300		0.047	322789810634752.000	-4.190	-1.595
600		0.047	478617935020032.000	-4.255	-1.627

Table 5: *continued*

Enc. M.	Video	VT H	Logscale Diagram		
			c_f	α	H
800		0.036	2104900845568.000	-3.392	-1.196
1000		0.041	3280063430656.000	-3.450	-1.225
1200		0.022	689201872896.000	-3.251	-1.125
1400		0.021	377870319616.000	-3.173	-1.086
1600		0.026	875160141824.000	-3.283	-1.141
25	<i>Lady and the Tramp</i>	0.022	213082080.000	-2.201	-0.601
75		0.028	489060224.000	-2.325	-0.662
100		0.033	22928542.000	-1.936	-0.468
300		0.041	19194778.000	-1.912	-0.456
600		0.063	9321051.000	-1.824	-0.412
800		0.067	10888958.000	-1.848	-0.424
1000		0.092	820040.312	-1.556	-0.278
1200		0.096	718594.750	-1.544	-0.272
1400		0.097	495879.500	-1.502	-0.251
1600		0.086	442595.625	-1.484	-0.242
25	<i>Foot Ball w/c</i>	-0.111	6687762.500	-1.759	-0.380
75		-0.077	17504038.000	-1.907	-0.453
100		-0.071	23999492.000	-1.955	-0.478
300		-0.042	36904152.000	-2.000	-0.500
600		-0.040	24528310.000	-1.944	-0.472
800		-0.047	13327088.000	-1.867	-0.434
1000		-0.048	15617054.000	-1.884	-0.442
1200		-0.044	12771494.000	-1.863	-0.431
1400		-0.050	3192834.500	-1.669	-0.334
1600		-0.062	4051244.250	-1.697	-0.349
25	<i>Tonight Show w/c</i>	-0.190	230368864.000	-2.258	-0.629
75		-0.174	675199.625	-1.486	-0.243
100		-0.154	748491.125	-1.493	-0.246
300		-0.374	165650.844	-1.295	-0.148
600		-0.432	213499472.000	-2.186	-0.593
800		-0.421	120589.367	-1.560	-0.280
1000		-0.403	156895.969	-1.587	-0.294
1200		-0.382	174308.781	-1.600	-0.300
1400		-0.373	73974.336	-1.501	-0.250
1600		-0.348	55982.273	-1.460	-0.230

Table 6 shows the behavior of the multiscaling parameter α_q for $q = 0.5, 1, 1.5, 2, 2.5, 3, 3.5$, and 4. Hurst parameter estimate is given by $H = \alpha_2/2$ for the employed estimation with c norm of one. We observe that α_q decreases with increasing q . Figure 16 illustrates the behavior of α_q as a function of q .

Table 6: Scaling parameters estimated from multiscale diagram.

Enc. M.	Video	Multiscale Diagram, α_q for							
		$q = 0.5$	$q = 1$	$q = 1.5$	$q = 2$	$q = 2.5$	$q = 3$	$q = 3.5$	$q = 4$
25	<i>Terminator</i>	-0.501	-0.995	-1.453	-1.882	-2.297	-2.712	-3.132	-3.561
75		-0.823	-1.651	-2.495	-3.364	-4.261	-5.181	-6.118	-7.067
100		-0.556	-1.133	-1.742	-2.379	-3.038	-3.717	-4.413	-5.122
300		-1.025	-1.926	-2.850	-3.814	-4.814	-5.842	-6.891	-7.955
600		-0.848	-1.772	-2.736	-3.732	-4.758	-5.810	-6.882	-7.969
800		-0.639	-1.272	-1.912	-2.565	-3.236	-3.927	-4.635	-5.360
1000		-0.636	-1.265	-1.895	-2.534	-3.187	-3.856	-4.540	-5.238
1200		-0.588	-1.195	-1.800	-2.398	-2.994	-3.591	-4.195	-4.807
1400		-0.578	-1.161	-1.743	-2.322	-2.901	-3.481	-4.066	-4.656
1600		-0.575	-1.170	-1.772	-2.382	-3.002	-3.637	-4.286	-4.950
25	<i>Lady and the Tramp</i>	-0.333	-0.708	-1.098	-1.491	-1.882	-2.272	-2.660	-3.048
75		-0.322	-0.703	-1.087	-1.463	-1.835	-2.207	-2.584	-2.970
100		-0.278	-0.549	-0.812	-1.078	-1.350	-1.633	-1.925	-2.226
300		-0.260	-0.503	-0.742	-0.990	-1.247	-1.513	-1.787	-2.070
600		-0.235	-0.451	-0.664	-0.885	-1.115	-1.354	-1.601	-1.854
800		-0.257	-0.473	-0.688	-0.915	-1.152	-1.398	-1.650	-1.907
1000		-0.333	-0.459	-0.545	-0.643	-0.768	-0.925	-1.118	-1.343
1200		-0.295	-0.455	-0.554	-0.646	-0.752	-0.880	-1.035	-1.218
1400		-0.310	-0.445	-0.523	-0.600	-0.692	-0.807	-0.949	-1.121
1600		-0.311	-0.441	-0.511	-0.577	-0.657	-0.762	-0.896	-1.061
25	<i>Foot Ball w/c</i>	-0.180	-0.341	-0.515	-0.712	-0.935	-1.181	-1.450	-1.739
75		-0.190	-0.380	-0.576	-0.785	-1.009	-1.248	-1.499	-1.759
100		-0.185	-0.374	-0.569	-0.778	-1.001	-1.239	-1.488	-1.745
300		-0.250	-0.496	-0.732	-0.953	-1.163	-1.367	-1.569	-1.774
600		-0.220	-0.461	-0.702	-0.934	-1.154	-1.363	-1.563	-1.757
800		-0.205	-0.430	-0.659	-0.882	-1.093	-1.294	-1.486	-1.673
1000		-0.226	-0.453	-0.681	-0.902	-1.109	-1.304	-1.489	-1.668
1200		-0.216	-0.438	-0.661	-0.878	-1.086	-1.283	-1.471	-1.654
1400		-0.147	-0.317	-0.507	-0.710	-0.921	-1.137	-1.358	-1.583
1600		-0.154	-0.330	-0.529	-0.744	-0.970	-1.204	-1.445	-1.693
25	<i>Tonight Show w/c</i>	-0.288	-0.587	-0.918	-1.270	-1.634	-2.011	-2.398	-2.796
75		-0.065	-0.188	-0.357	-0.562	-0.794	-1.047	-1.318	-1.603
100		-0.089	-0.228	-0.402	-0.604	-0.830	-1.075	-1.335	-1.607
300		-0.049	-0.141	-0.265	-0.420	-0.606	-0.822	-1.063	-1.325
600		-0.226	-0.565	-0.965	-1.403	-1.869	-2.360	-2.872	-3.402
800		-0.205	-0.511	-0.818	-1.122	-1.429	-1.744	-2.068	-2.398
1000		-0.215	-0.522	-0.838	-1.147	-1.453	-1.759	-2.068	-2.377
1200		-0.185	-0.462	-0.775	-1.113	-1.461	-1.811	-2.159	-2.502
1400		-0.178	-0.411	-0.675	-0.959	-1.247	-1.532	-1.811	-2.084
1600		-0.175	-0.398	-0.649	-0.922	-1.209	-1.502	-1.797	-2.092

Finally, we examine the behavior of the linear multiscale, illustrated in Figure 17. We observed that the linear multiscale diagram does not show a significant multi-fractal behavior since h_q does not change significantly as a function of q .

5.5 Analysis of Video Quality

In this section of the paper we analyze the video quality aspects of the video traces. Our main focus is on the PSNR and MSE values, defined in Section 5.2. For the PSNR values we only take into account the luminance component of the video traces. Mainly since the human visual system is more sensitive to the luminance component in contrast to the chrominance (color) components. We denote Q_n for Q_n^Y , and M_n for $p^2/10^{(Q_n/10)}$ for convenience.

Table 7 gives the average quality \bar{Q} , the coefficient of quality variation $CoQV$, the alternative coefficient of quality variation $CoQV'$, and the quality range Q_{\min}^{\max} for the video frames, while at the GoP aggregation level it gives the coefficients of variation $CoQV^{(G)}$, $CoQV'^{(G)}$ and the quality range $Q_{\min}^{\max(G)}$.

Here we observe that the low average video quality \bar{Q} is around 18–20 dB for 25 kbps video while for the 1600 kbps video the \bar{Q} is around 39–40 dB. As we observed in Table 2, the $CoQV$ shows a *hump* like behavior, it is increasing for the low bit rates and then coming to a peak around the mid bit rates, and gradually decreasing back for the higher bit rates. The $CoQV'$ on the other hand shows a gradual decreasing trend when the bit rate is increased. We observe that the Q_{\min}^{\max} decreases with the increasing bit rate as well. *Foot Ball w/c* shows a much larger Q_{\min}^{\max} than the other videos. Next, at a GoP level we observe similar results from Table 7. The $CoQV^{(G)}$ shows the *hump* like behavior while the $CoQV'^{(G)}$ and $Q_{\min}^{\max(G)}$ decreases with increasing video bit rates. This phenomenon of the *hump* behavior of the $CoQV$ and $CoQV^{(G)}$ will be explored in future work. The $CoQV^{(G)}$ we observed is relatively smaller than $CoQV$. Finally we note that the average PSNR can not be used as an absolute measure to measure the quality of a video, since the quality depends on the content, the visual composition of the scenes, background, lighting, movement, etc.

Table 7: Overview of quality statistics of single-layer traces

Enc. M.	Video	Encoded bit rate	Frame Level				GoP level		
			\bar{Q}	$CoQV$	$CoQV'$	Q_{\min}^{\max}	$CoQV^{(G)}$	$CoQV'^{(G)}$	$Q_{\min}^{\max(G)}$
Intra Frame Video	<i>Terminator</i>	25	19.256	0.529	0.128	23.910	0.518	0.128	22.706
		75	22.965	0.638	0.120	22.960	0.629	0.122	21.070
		100	24.576	0.793	0.154	29.360	0.785	0.162	23.602
		300	28.847	0.729	0.100	24.590	0.725	0.102	19.102
		600	33.126	0.720	0.081	24.280	0.717	0.082	19.749
		800	35.189	0.688	0.070	24.390	0.686	0.071	18.635
		1000	36.889	0.641	0.062	22.130	0.640	0.062	18.689
		1200	38.339	0.587	0.055	20.880	0.585	0.054	17.526
		1400	39.643	0.531	0.049	20.650	0.530	0.048	17.354
		1600	40.781	0.482	0.044	20.880	0.481	0.043	16.174
	<i>Lady and the Tramp</i>	25	18.449	0.395	0.106	18.430	0.389	0.106	17.131
		75	21.196	0.432	0.093	17.050	0.428	0.093	15.822
		100	23.509	0.673	0.111	22.190	0.671	0.114	15.852
		300	25.681	0.476	0.074	19.350	0.474	0.074	13.011
		600	28.737	0.468	0.060	21.290	0.467	0.060	11.655
		800	30.224	0.449	0.053	20.530	0.448	0.054	10.864
		1000	31.531	0.439	0.049	20.210	0.439	0.049	10.835
		1200	32.728	0.423	0.045	20.570	0.422	0.045	10.707
		1400	33.923	0.410	0.043	19.560	0.409	0.044	10.780
		1600	35.130	0.405	0.041	18.320	0.404	0.041	10.304
	<i>Foot Ball w/c</i>	25	18.490	0.443	0.139	68.090	0.431	0.130	33.760
		75	21.796	0.477	0.124	64.750	0.469	0.121	30.405
		100	22.730	0.484	0.121	63.870	0.477	0.118	29.641
		300	27.592	0.530	0.105	58.750	0.525	0.106	24.114
		600	31.862	0.527	0.088	54.300	0.524	0.089	20.235
		800	33.886	0.502	0.078	52.120	0.499	0.078	18.141
		1000	35.552	0.469	0.069	50.340	0.466	0.069	18.858
		1200	36.957	0.433	0.064	48.650	0.430	0.063	17.834
		1400	38.094	0.415	0.060	46.920	0.413	0.059	16.552
		1600	39.224	0.408	0.056	45.300	0.406	0.056	16.456

Table 7: *continued*

Enc. M.	Video	Encoded bit rate	Frame Level				GoP level		
			\bar{Q}	$CoQV$	$CoQV'$	Q_{\min}^{\max}	$CoQC^{(G)}$	$CoQV'^{(G)}$	$Q_{\min}^{\max(G)}$
<i>Tonight Show w/c</i>		25	18.427	0.401	0.104	29.000	0.390	0.099	21.654
		75	20.413	0.383	0.094	27.080	0.374	0.091	19.417
		100	21.014	0.381	0.093	25.910	0.374	0.091	18.276
		300	24.053	0.396	0.087	21.470	0.392	0.086	15.778
		600	27.044	0.402	0.076	18.190	0.400	0.076	15.589
		800	28.631	0.379	0.068	18.250	0.378	0.068	15.750
		1000	30.015	0.372	0.064	17.910	0.370	0.064	15.485
		1200	31.315	0.351	0.059	17.430	0.350	0.059	14.699
		1400	32.475	0.353	0.058	17.780	0.352	0.058	14.336
		1600	33.646	0.354	0.056	17.940	0.353	0.057	14.365

Figure 18 illustrates the behavior of the video quality in $PSNR$ as a function of the frame index n . Here we observe a relatively high variance of the video quality for the low bit rate videos, while the quality tends to smooth out as the bit rate is increased. Different sections of the trace tend to have different variations and an average video quality, which corresponds to the different scenes in the video sequence. We observed that the variations of the quality for the same bit rate of different videos also vary (not shown here because of space constraints) due to the different content of the video genres.

Figure 19 shows the histograms of the frame qualities. We observed that the histograms are wider for the low bit rate video, and much narrow for the high bit rate video. This is due to the fact that with large bit budgets, the encoder can encode frames with less loss, consistently, while at lower bit rates more detailed, complicated frames have a lower PSNR. *Terminator* encoded at 100 kbps behaves much differently illustrating a edgy histogram which is contrast to the other bit rates which show a smoother single peak histograms.

Figure 20 and Figure 21 show the autocorrelation coefficient as a function of lag k (in frames) and lag k (in GoPs) respectively. In Figure 20 we observe that the autocorrelation function is smooth and decaying slowly, this is again in contrast to the MPEG-4 encodings [24]. At the GoP level, in Figure 21 we observe a relatively sharper, less smoother decay.

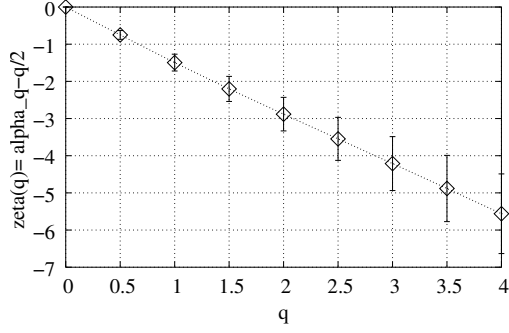
Figures 22 and 23 show the scatter plots of frame quality as a function of the video frame size and respectively. Here the interesting point is that higher bit size frames are not necessary have a high video quality. We observe that the frame quality levels tend to disperse horizontally for higher bit rates, while at lower bit rates the frame qualities tend to stay closer to the mean.

6 Correlation Between Frame Sizes and Qualities

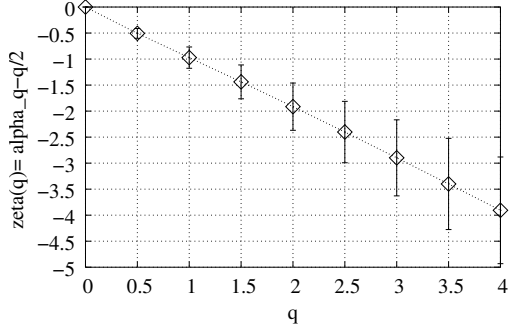
Table 8 gives the size–MSE quality correlation coefficient ρ_{XM} and the size–PSNR quality correlation coefficient ρ_{XQ} , as well as the corresponding correlation coefficients $\rho_{XM}^{(G)}$ and $\rho_{XQ}^{(G)}$ for the GoP aggregation. First at the frame level we observe from Table 8 that the ρ_{XM} decreases as the bit rate is increased. The ρ_{XQ} , on the other hand, decreases for increasing bit rates. This is natural since the PSNR and the MSE has an inverse relationship. For the bit rates in observation, the ρ_{XQ} stays negative. We see a similar trend in the GoP level where the $\rho_{XM}^{(G)}$ increases and the $\rho_{XQ}^{(G)}$ decreases for increasing bit rates.

Table 8: Correlation between quality and traffic for single-layer traces

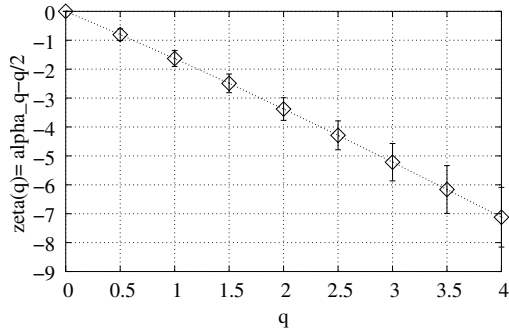
Enc. M.	Video	Frame Level		GoP level	
		ρ_{XM}	ρ_{XQ}	$\rho_{XM}^{(G)}$	$\rho_{XQ}^{(G)}$
25	<i>Terminator</i>	0.389	-0.481	0.399	-0.483
75		0.390	-0.484	0.382	-0.464
100		0.302	-0.322	0.292	-0.301
300		0.279	-0.382	0.270	-0.356
600		0.195	-0.286	0.187	-0.260
800		0.148	-0.224	0.141	-0.198
1000		0.115	-0.172	0.109	-0.144
1200		0.072	-0.107	0.066	-0.078
1400		0.034	-0.069	0.028	-0.035
1600		0.027	-0.075	0.019	-0.034
25	<i>Lady and the Tramp</i>	0.371	-0.414	0.390	-0.426
75		0.395	-0.425	0.397	-0.421
100		0.241	-0.271	0.237	-0.263
300		0.289	-0.315	0.284	-0.306
600		0.184	-0.210	0.179	-0.201
800		0.128	-0.146	0.124	-0.138
1000		0.080	-0.093	0.077	-0.086
1200		0.030	-0.028	0.028	-0.021
1400		-0.017	0.023	-0.017	0.028
1600		-0.025	0.017	-0.025	0.022
25	<i>Foot Ball w/c</i>	0.493	-0.505	0.501	-0.472
75		0.471	-0.508	0.465	-0.460
100		0.439	-0.484	0.429	-0.436
300		0.356	-0.419	0.347	-0.381
600		0.293	-0.359	0.285	-0.326
800		0.262	-0.329	0.254	-0.298
1000		0.233	-0.301	0.224	-0.270
1200		0.194	-0.261	0.187	-0.232
1400		0.162	-0.229	0.155	-0.201
1600		0.125	-0.206	0.118	-0.179
25	<i>Tonight Show w/c</i>	0.540	-0.554	0.546	-0.518
75		0.548	-0.537	0.545	-0.502
100		0.509	-0.512	0.499	-0.474
300		0.322	-0.382	0.309	-0.348
600		0.195	-0.258	0.186	-0.235
800		0.147	-0.194	0.140	-0.176
1000		0.101	-0.144	0.095	-0.131
1200		0.059	-0.095	0.056	-0.084
1400		0.013	-0.050	0.013	-0.043
1600		-0.012	-0.038	-0.012	-0.031



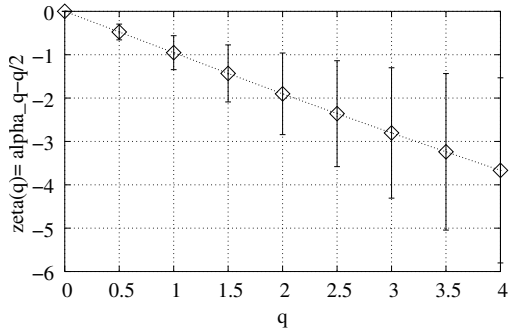
a) *Terminator* at 25 kbps



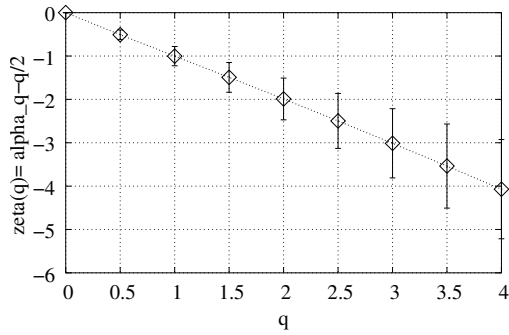
d) *Lady and the Tramp* at 800 kbps



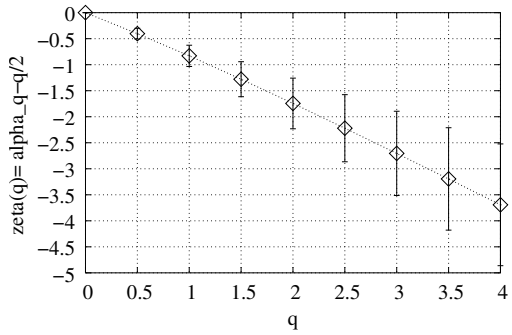
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

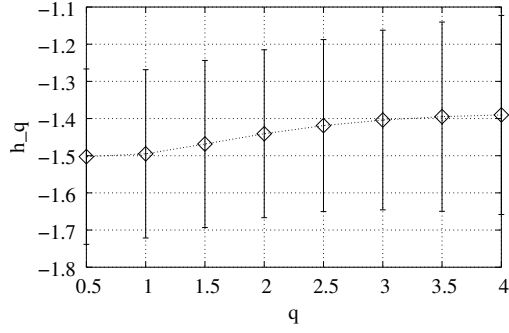


c) *Lady and the Tramp* at 300 kbps

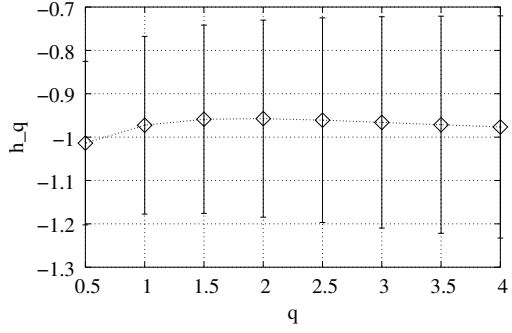


f) *Foot Ball w/c* at 1600 kbps

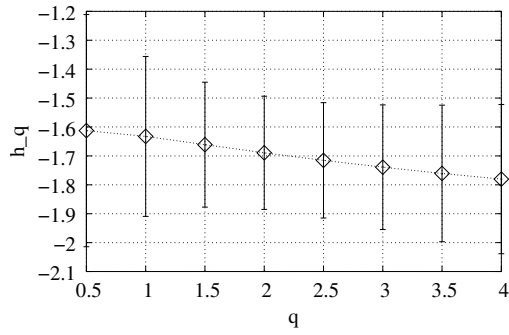
Figure 16: Multiscale diagrams for intra frame QCIF video.



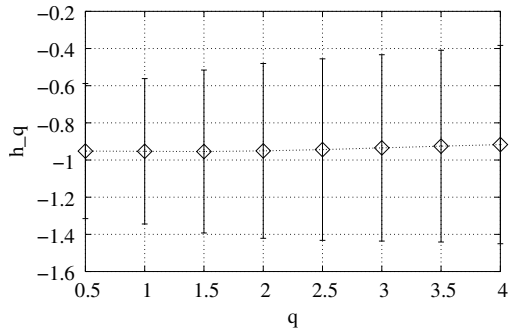
a) *Terminator* at 25 kbps



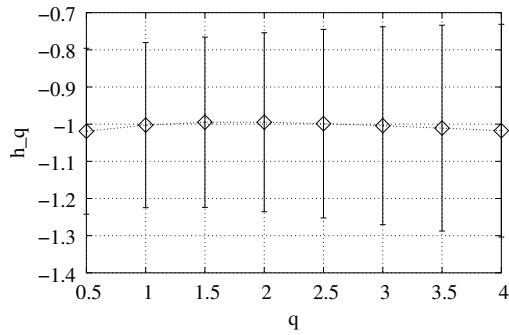
d) *Lady and the Tramp* at 800 kbps



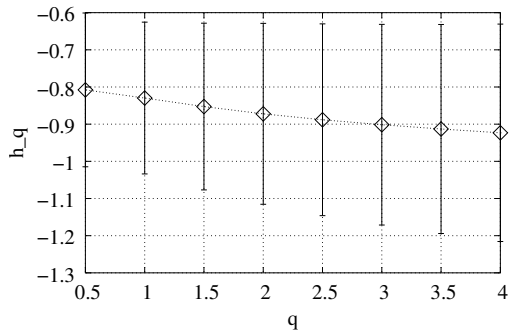
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

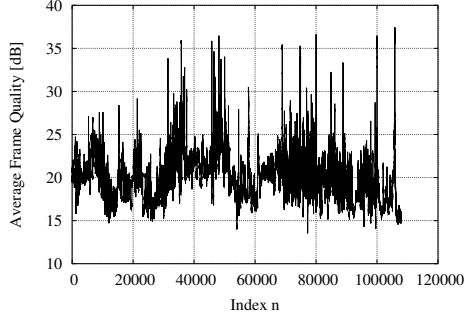


c) *Lady and the Tramp* at 300 kbps

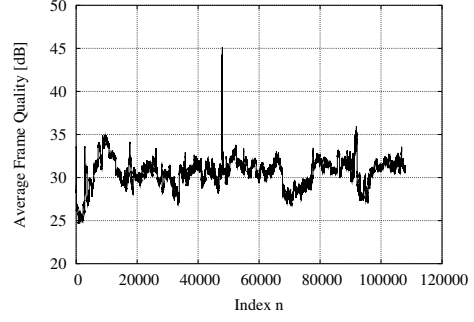


f) *Foot Ball w/c* at 1600 kbps

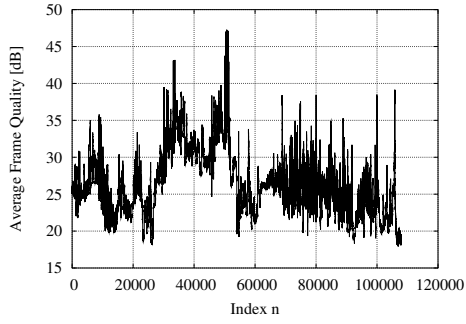
Figure 17: Linear multiscale diagrams for intra frame QCIF video.



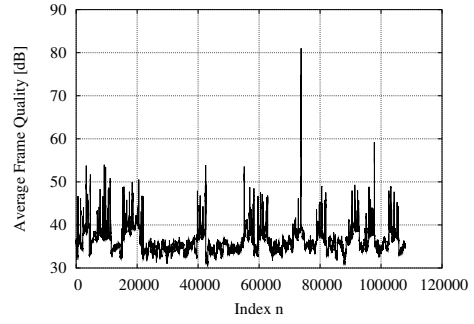
a) *Terminator* at 25 kbps



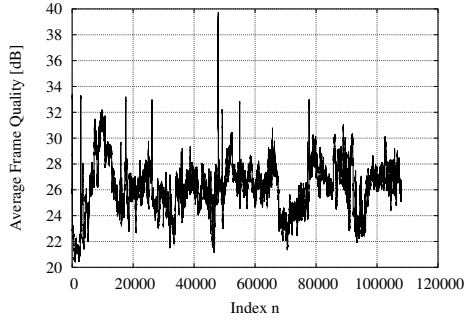
d) *Lady and the Tramp* at 800 kbps



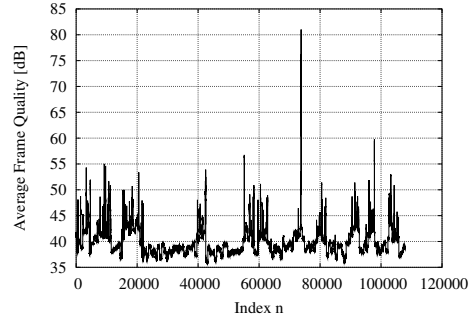
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

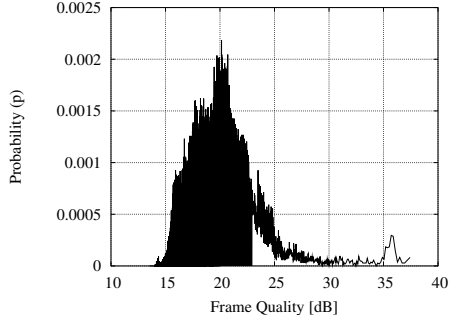


c) *Lady and the Tramp* at 300 kbps

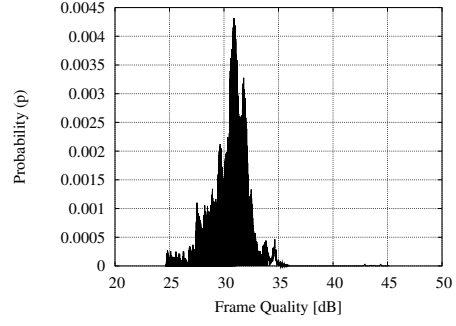


f) *Foot Ball w/c* at 1600 kbps

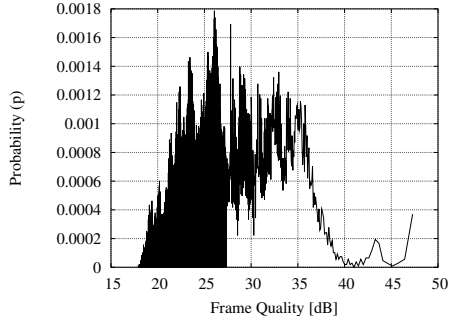
Figure 18: Video frame quality Q_n (in dB) as a function of the frame index n for intra frame QCIF video.



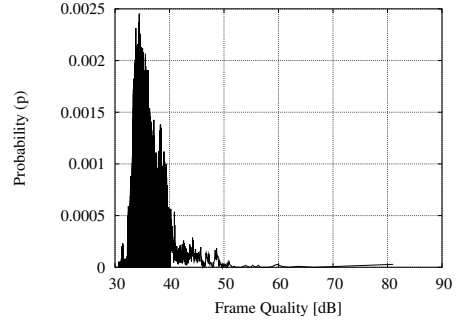
a) *Terminator* at 25 kbps



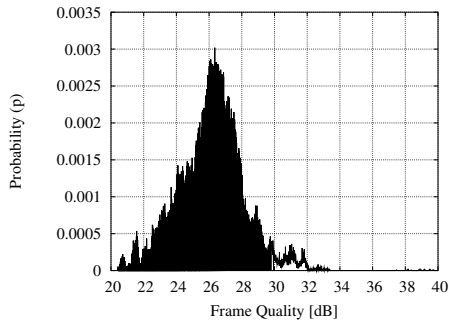
d) *Lady and the Tramp* at 800 kbps



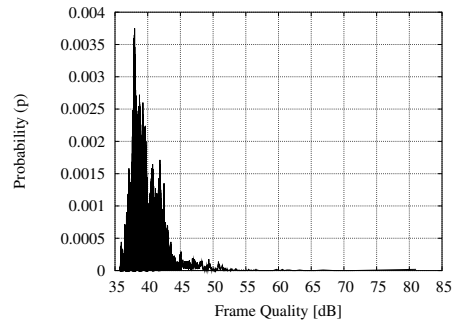
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

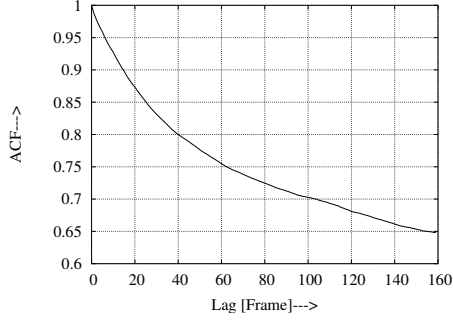


c) *Lady and the Tramp* at 300 kbps

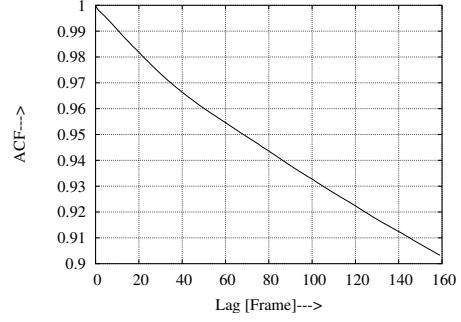


f) *Foot Ball w/c* at 1600 kbps

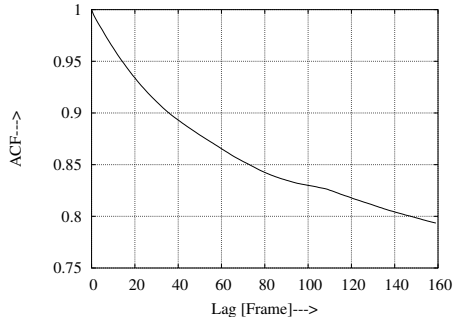
Figure 19: Histograms of video frame quality Q_n (in dB) of intra frame QCIF video.



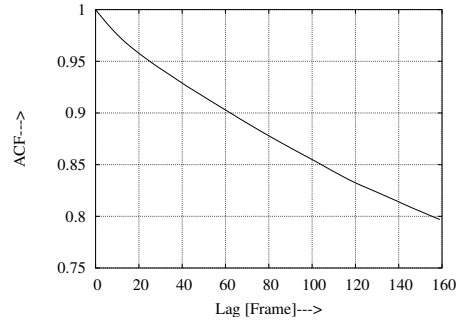
a) *Terminator* at 25 kbps



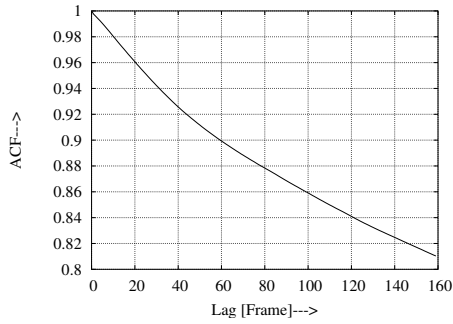
d) *Lady and the Tramp* at 800 kbps



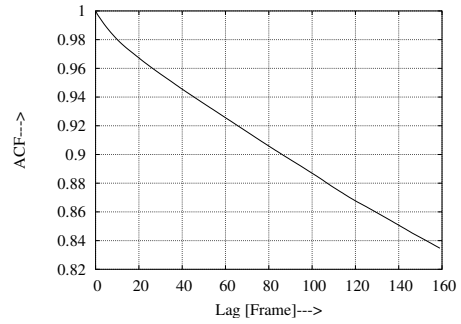
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

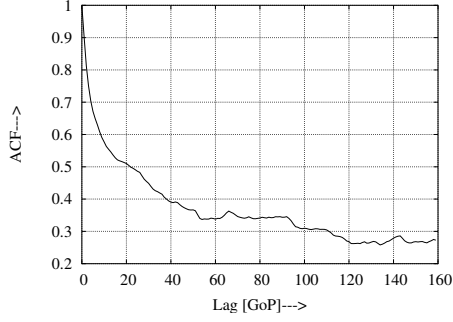


c) *Lady and the Tramp* at 300 kbps

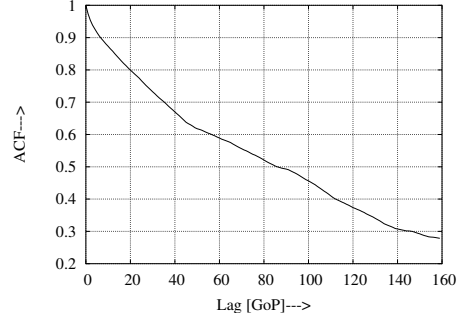


f) *Foot Ball w/c* at 1600 kbps

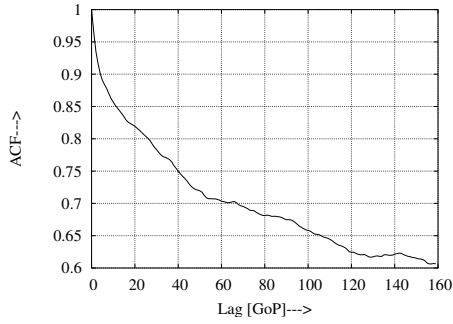
Figure 20: MSE autocorrelation coefficient $p_M(k)$ as a function of the lag k (in frames) for intra frame QCIF video.



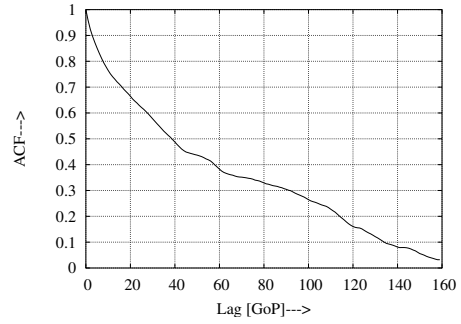
a) *Terminator* at 25 kbps



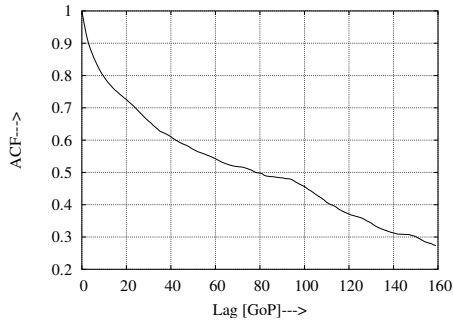
d) *Lady and the Tramp* at 800 kbps



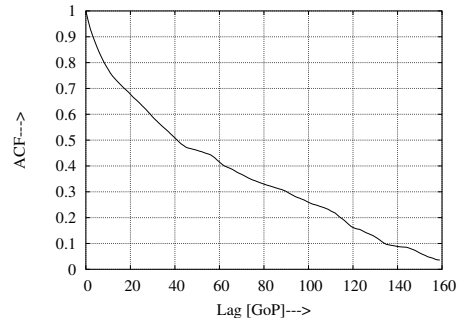
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

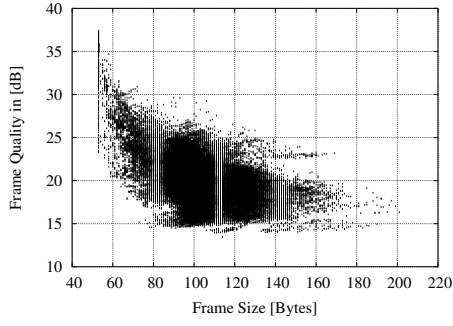


c) *Lady and the Tramp* at 300 kbps

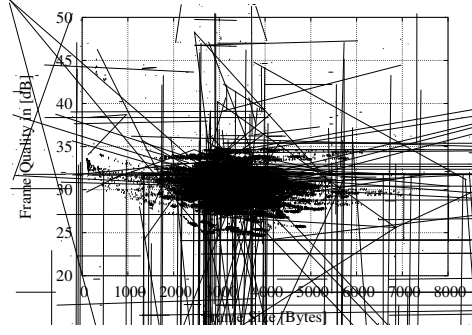


f) *Foot Ball w/c* at 1600 kbps

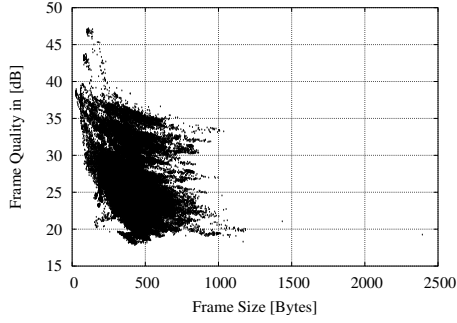
Figure 21: MSE autocorrelation coefficient $p_M^G(k)$ as a function of the lag k (in GoPs) for intra frame QCIF video.



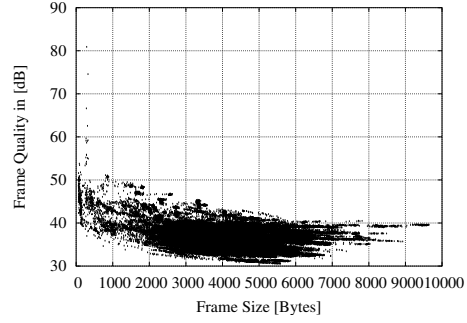
a) *Terminator* at 25 kbps



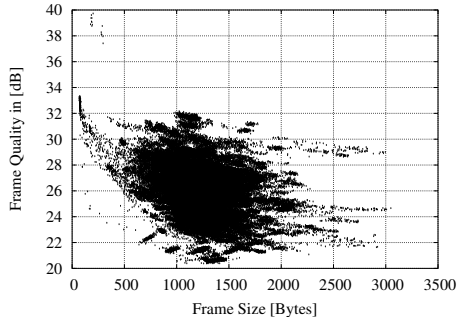
d) *Lady and the Tramp* at 800 kbps



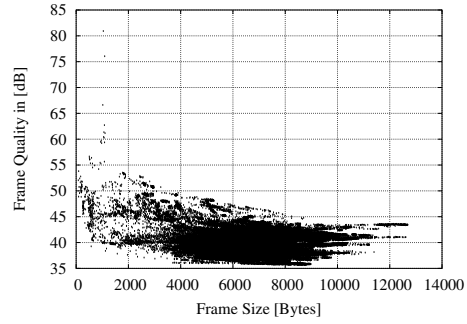
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps

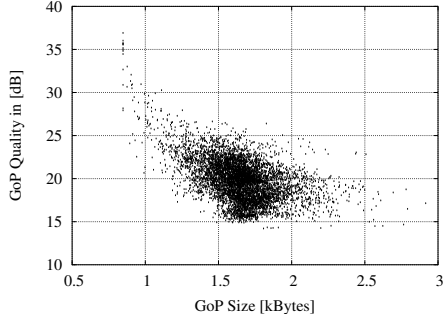


c) *Lady and the Tramp* at 300 kbps

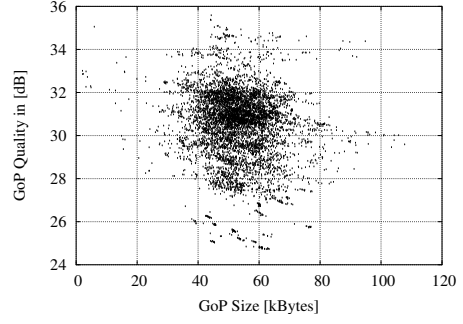


f) *Foot Ball w/c* at 1600 kbps

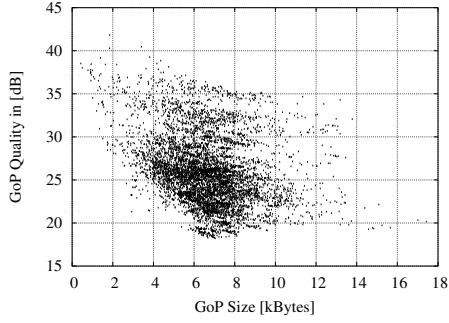
Figure 22: Scatter plots of frame size and frame quality for intra frame QCIF video.



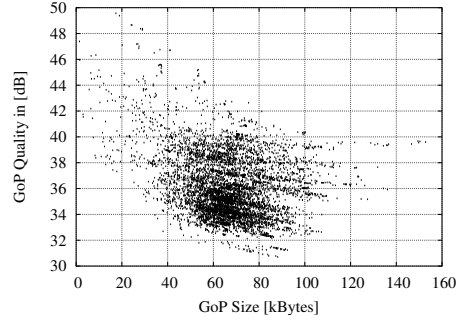
a) *Terminator* at 25 kbps



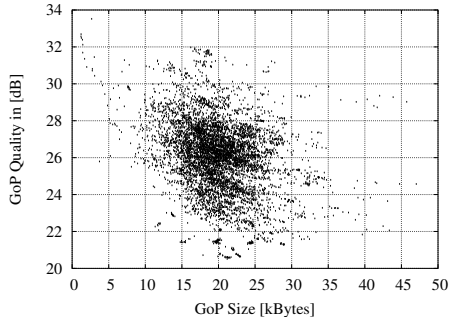
d) *Lady and the Tramp* at 800 kbps



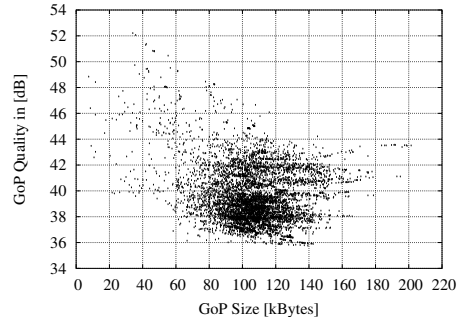
b) *Terminator* at 100 kbps



e) *Foot Ball w/c* at 1000 kbps



c) *Lady and the Tramp* at 300 kbps



f) *Foot Ball w/c* at 1600 kbps

Figure 23: Scatter plots of GoP size and average GoP quality for intra frame QCIF video.

7 Conclusion

In this study of the wavelet encoded video traces of frame sizes and frame qualities we have observed several interesting properties and phenomena. For example we observed the *hump* like behavior of the CoV_X , for the wavelet transformed video encoder. The observations made parallel to the MPEG-4 analysis clearly show the differences of most of the statistical characteristics. For instance the autocorrelation coefficients behave very differently and so do the magnitudes of the coefficient of variation where we observed that the wavelet encoded videos have much lower CoV_X . We clearly see from the video traces 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.

8 Acknowledgment

We are indebted to Shih-Ta Hsiang and Prof. John W. Woods at the Rensselaer Polytechnic Institute for providing us with the 3D-EZBC codec and explaining its operation.

We are grateful to Zhen Liu and Prof. Lina Karam of Arizona State University for providing us with background on wavelet based video coding.

9 Appendix A

Table 9: Table of acronyms

Acronym	Definition
JPEG	Joint Photographic Experts Group
QCIF	Quarter Common Intermediate Format
CIF	Common Intermediate Format
PSNR	Peak Signal to Noise Ratio
DCT	Discrete Cosine Transform
MC-3DEZBC	Motion Compensated 3D Embedded Zerotree Block Coder
SPHIT	Set Partitioning in Hierarchical Trees
GoP	Group of Pictures
MC	Motion Compensated
MV	Motion Vector
3D-EZBC	3D Embedded Zerotree Block Coder
EZBC	Embedded Zerotree Block Coder
VCR	Video Cassette Recorder
MPEG	Moving Picture Experts Group
NTSC	National Television Standards Committee
MSE	Mean Squared Error

References

- [1] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. Wavelets for the analysis, estimation and synthesis of scaling data. In K. Park and W. Willinger, editors, *Self Similar Network Traffic Analysis and Performance Evaluation (Wiley)*, 2000.
- [2] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. Self-similarity and long-range dependence through the wavelet lens. In Doukhan, Oppenheim, and Taqqu, editors, *Long Range Dependence: Theory and Applications*, 2002.
- [3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2):205–220, April 1992.
- [4] J. Beran. *Statistics for long-memory processes*. Chapman and Hall, 1994.

- [5] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger. Long-range dependence in variable-bit-rate video traffic. *IEEE Transactions on Communications*, 43(2/3/4):1566–1579, February/March/April 1995.
- [6] C. Chatfield. *The Analysis of Time Series: An Introduction*. Chapman and Hall, fourth edition, 1989.
- [7] A. Feldmann, A. C. Gilbert, W. Willinger, and T.G. Kurtz. The changing nature of network traffic: Scaling phenomena. *Computer Communication Review*, 28(2), April 1998.
- [8] J. Gao and I. Rubin. Multiplicative multifractal modeling of long-range-dependent network traffic. *International Journal of Communication Systems*, 14:783–801, 2001.
- [9] A. C. Gilbert, W. Willinger, and A. Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):971–991, April 1999.
- [10] S.-T. Hsiang and J.W. Woods. Invertible three-dimensional analysis/synthesis system for video coding with half-pixel-accurate motion compensation. In *Proceedings of The International Society for Optical Engineering (SPIE) on Visual Communications and Image Processing*, volume 3653, pages 537–546, San Jose, CA, January 1999.
- [11] S.-T. Hsiang and J.W. Woods. Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling. In *Proceedings of the 2000 IEEE International Symposium on Circuits and Systems, ISCAS 2000*, volume 3, pages 662–665, Geneva, Switzerland, May 2000.
- [12] S.-T. Hsiang and J.W. Woods. Embedded video coding using invertible motion compensated 3-d subband/wavelet filter bank. *Signal Processing: Image Communications*, 16:705–724, May 2001.
- [13] Shih-Ta Hsiang. *Thesis - Highly scalable subband/wavelet image and video coding*. PhD thesis, Rensselaer Polytechnic Institute, NY, May 2002.
- [14] M. Krunz. On the limitations of the variate-time test for inference of long-range dependence. In *Proceedings of IEEE Infocom 2001*, pages 1254–1260, Anchorage, Alaska, April 2001.
- [15] A. M. Law and W. D. Kelton. *Simulation, Modeling and Analysis*. McGraw Hill, third edition, 2000.
- [16] A. S. Lewis and G. Knowles. Video compression using 3D wavelet transforms. *Electronics Letters*, 26(6):396–398, September 1990.
- [17] S. M. LoPresto, K. Ramchandran, and M. T. Orchard. Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework. volume DCC '97, pages 221–230, March 1997.
- [18] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

- [19] B. B. Mandelbrot and M. S. Taqqu. Robust R/S analysis of long-run serial correlations. In *Proceedings of 42nd Session ISI, Vol. XLVIII, Book 2*, pages 69–99, 1979.
- [20] S. A. Martucci, I. Sodagar, T. Chiang, and Ya-Qin Zhang. A zerotree wavelet video coder. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):109–118, February 1997.
- [21] F. G. Meyer, A. Z. Averbuch, and J. O. Stromberg. Fast adaptive wavelet packet image compression. *IEEE Transactions on Image Processing*, 9(5):792–800, May 2000.
- [22] B.S. Nanda and N. Kaulgud. Effect of quatization on video compression. In *In Proceedings of 2002 IEEE International Conference on Industrial Technology, 2002. IEEE ICIT '02*, volume 2, pages 764–768, Bangkok, Thailand, December 2002.
- [23] J. R. Ohm. Three-dimensional subband coding with motion compensation. *IEEE Transactions on Image Processing*, 3(5):559–571, 1994.
- [24] M. Resslein, 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, part 1: Overview and definitions. *Arizona State University, Dep. of Electrical Engineering, Tech. Rep.*, August 2003.
- [25] M. Resslein, 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, part 2: Statistical analysis of single-layer encoded video. *Arizona State University, Dep. of Electrical Engineering, Tech. Rep.*, August 2003.
- [26] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk. A multifractal wavelet model with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):992–1018, April 1999.
- [27] A. Said and W.A. Pearlman. A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. on Circuits and Systems for Video Technology*, 6:243–250, June 1996.
- [28] J. M. Shapiro. An embedded hierarchical image coder using zerotrees of wavelet coefficients. In *Proceedings of the Data Compression Conference*, volume DCC '93, pages 214–223, Monterey, CA , USA, March 1993.
- [29] D. Taubman. High performance scalable image compression with ebcot. *IEEE Transactions on Image Processing*, 9(7):1158–1170, July 2000.
- [30] D. Veitch and P. Abry. A wavelet based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory*, 45(3):878–897, April 1999. Matlab code available at <http://www.emulab.ee.mu.oz.au/~darryl>.
- [31] J. Walter. bttvgrab. <http://www.garni.ch/bttvgrab/>.
- [32] J. W. Woods and G. Lilienfield. A resolution and frame-rate scalable subband/wavelet video coder. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(9):1035–1044, September 2001.

- [33] S. H. Yoon and S. S. Rao. A scalable wavelet video coder for hybrid communication channels. *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, 1:382–386, 1997.