The Rate Variability-Distortion (VD) Curve of Encoded Video and its Impact on Statistical Multiplexing

Patrick Seeling and Martin Reisslein

Abstract

Encoded video is expected to contribute a significant portion of the load on future communication systems and networks, which often employ statistical multiplexing. In such systems, the number of video streams that can be supported depends both on the mean bit rate as well as bit rate variability of the video streams. At the same time, the utility (revenue) earned from video streaming depends both on the number of supported video streams as well as their quality level. In this paper we examine the interplay between video quality, traffic variability, and utility for open-loop encoded video. We introduce the *rate variability-distortion (VD) curve* which relates the bit rate variability to the quality level of an encoded video. We find that the VD curve generally exhibits a characteristic "hump" behavior of first increasing, peaking, and subsequently decreasing variability for increasing quality. We examine the impact of video content characteristics, encoding parameters, and traffic smoothing on the VD behavior. We describe a methodology for assessing (i) the set of the video streams that can be supported with a statistical quality of service requirement, and (ii) the utility earned from video streaming over a link. This methodology is based on the rate-distortion and rate variability-distortion characteristics of the videos. We find that the statistical multiplexing gain and the utility as a function of the video quality level typically exhibit a "hump" similar to the VD curve.

Index Terms

Network utility, statistical multiplexing, variable bit rate video, video content, video quality, video traffic, video streaming.

I. Introduction

Video streaming is expected to play a dominant role in future multimedia applications, including in multimedia applications that are offered over communication systems and networks. For the transport over communications systems and networks the video is typically compressed (encoded). Generally, video can be encoded (i) in an open loop with a fixed quantization scale, which results in fairly consistent video quality but variable bit rate (VBR) video traffic, or (ii) in a closed loop by adjusting the quantization scale, which can keep the bit rate close to a fixed target bit rate but typically results in quality variations in the video [1]. We note that video can also be encoded with scalability into multiple layers. Broadly speaking, a given layer is either encoded with an open loop to give fairly constant quality and variable bit rates, or with a closed loop to give a close to constant bit rate and variable quality. Also, the video can be encoded with fine granularity scalability, which permits the fine granular scaling of the video bit rate

Please direct correspondence to M. Reisslein.

This work was supported in part by the National Science Foundation through Grant No. Career ANI-0133252 and Grant No. ANI-0136774.

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: {patrick.seeling, reisslein}@asu.edu, web: http://www.fulton.asu.edu/~mre.

and quality. In this paper we focus primarily on non-scalable (single layer) video encoded with a fixed quantization scale. However, our methodology and results are also applicable to variable bit rate layers of scalable encodings.

In order to transport the variable bit rate video traffic at reasonably high levels of network utilization, the video streams are typically transported with some sort of statistical transport scheme, which may occasionally drop (lose) some of the video traffic. These statistical transport schemes employ *statistical multiplexing*, i.e., they exploit the fact that the peaks in the traffic of the simultaneously ongoing streams do typically not collude. The number of simultaneous streams that can be supported by a given network depends on the statistical characteristics of the video traffic and the tolerable loss rate. Importantly, the *utility* (i.e., revenue) earned from streaming videos over a given network typically not only depends on the *number* of simultaneously supported streams, but also their *quality*. Clearly, the utility increases by increasing the number as well as the quality of the streams that can be simultaneously supported with a given network capacity.

Assessing the utility of a network providing a video streaming service with statistical multiplexing thus requires the joint consideration of the number of supported streams as well as their quality levels. Importantly, the number of streams that can be supported with statistical multiplexing depends on the mean (average) bit rate as well as the variability of the bit rate. For given mean bit rates the number of supported streams decreases as the variability of the traffic increases. Intuitively, with higher variability it is more likely that colluding traffic peaks exceed the network capacity and result in losses.

Rate-distortion curves, which have been intensively studied (see Section II) relate the size (in bit) of an encoded video frame to its quality and can also be used to relate the average bit rate of a sequence of video frames to the average video quality[1]. Considering mean bit rates and quality levels, however, is insufficient to assess the utility of video streaming with statistical multiplexing since the number of supported streams also depends critically on the bit rate variability.

To facilitate the assessment of the utility of video streaming with statistical multiplexing we introduce and examine in this paper the *rate variability-distortion (VD) curve* which relates the bit rate variability of an encoded video sequence to its average quality level. We find that for a variety of transform video coders, the VD curve exhibits a characteristic “hump” behavior, i.e., the bit rate variability first increases, peaks, and subsequently decreases as a function of the video quality, as illustrated in Figs. 1 and 2 which show the coefficient of variation of the frame sizes (in bit) as a function of the PSNR video quality for scenes from the movie *Terminator* and a video of a football game. We also find that this hump is most pronounced for low motion video scenes (MC I in the figures, as detailed shortly) and for long video sequences consisting of many scenes. We study the impact of the VD behavior on the statistical multiplexing of video streams and the utility obtained from a given network capacity. We find that the statistical multiplexing gain and the utility typically reach a maximum at a quality level that is in the vicinity of the quality level where the VD curve peaks. Thus the existence of the hump phenomenon is of significance for the communication systems and networking domain, as well as for content providers

[1]Strictly speaking, rate-distortion curves relate bit rate to distortion (whereby distortion is inversely related to quality), but it is quite common to refer to the curves relating bit rate to quality as rate-distortion curves and we will follow this practice in this paper.

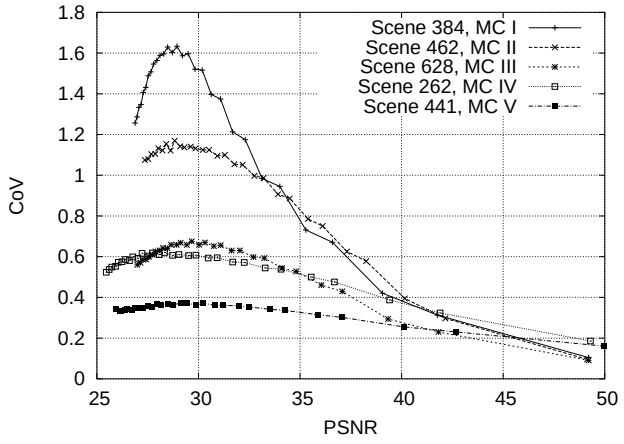


Fig. 1. VD curves for MPEG-4 coded scenes from *The Terminator*.

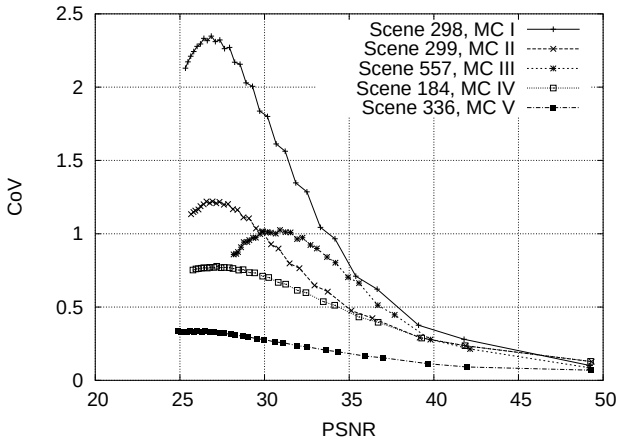


Fig. 2. VD curves for MPEG-4 coded scenes from *Football*.

who have to make a rate–quality trade-off decision.

This paper is organized as follows. In the following section we review related work. In Section III we describe the set-up of our study of the VD characteristics of encoded video and introduce our notations. We present a detailed study of the VD characteristics of open-loop encoded video in Section IV. We first examine the effects of intra coding and inter coding, as well as the effects of the different frame types and Group of Picture (GoP) patterns on the VD characteristics of scenes of different levels of motion. We then examine the effects of video traffic smoothing within individual scenes and over full length videos on the VD characteristics. In Section V we examine piecewise approximation models which give an estimate of the full VD curve from a few sample encodings. The impact of the VD characteristics on the statistical multiplexing of video streams over a bufferless link is studied in Section VI. In Section VII we consider the utility earned from the streaming with statistical multiplexing over the bufferless link considered in the previous section. Our main conclusions are summarized in Section VIII.

II. Related Work

Video streaming over networks has received a great deal of attention over past two decades, see for instance [2], [3]. Our study relates to the following three main lines of research: (i) research on the rate-distortion (RD) characteristics of encoded video, (ii) research on the analysis and modeling of video traffic, and (iii) research on video traffic management mechanisms. Research on the RD characteristics of encoded video examines the relationship between the (mean) bit rate and the video quality (and encoder quantization scale), see [4], [5] for tutorial overview of this area of research. The two main approaches that have been employed in RD research are analytical modeling and empirical modeling. Analytical modeling, such as pursued in [6] attempts to derive mathematical formulas for the RD behavior in terms of the statistics of the source video and the properties of the encoding mechanism. Empirical modeling, as studied in [7], [8], strives to approximate the RD curve by interpolating between a set of sample points. A unified RD analysis framework, which builds on an analysis of the percentage of zeros in the transformed video frames in conjunction with rate curve modeling is developed in [9]. The modelled RD characteristics are typically used to control the mean bit rate of video encoders [10], [11] and can also

be used for allocating mean bit rates to video streams for network transport [12], [13]. Our study differs from this literature on the RD characteristics in that we examine the relationship between the variability of the bit rate on the one hand, and the video quality (and quantization scale) on the other hand. In other words, the existing RD studies have focused on the first order statistic of the video traffic, whereas our focus is on the second order statistic (which we study in an empirical manner).

The statistical analysis of video traffic and the development of video traffic models has also received significant interest, see for instance [14]–[22]. This line of work is primarily focused on obtaining insights into the statistical properties of the traffic (including the bit rate variability) of a given video encoding (for a given, typically fixed quality level or quantization scale) and developing analytical models for the observed statistical properties. In particular, the focus is on finding parsimonious models that allow for the characterization of video traffic with a small number of model parameters. In contrast, in this paper we examine the bit rate variability of the video as a function of the quality level/quantization scale and demonstrate that the quality level of the video has a profound impact on the video traffic statistics, which to the best of our knowledge has not been reported in detail before.

The related area of research on video traffic management mechanisms may be viewed in terms of the following three different problem sub-areas: (a) traffic management for a single video stream, (b) traffic management for a group of multiplexed streams of fixed quality (quantization scale), and (c) traffic management that adapts the video quality while considering multiple multiplexed streams. The transmission of a single video stream is considered by the studies [23]–[27], which strive to maximize the video quality of the considered single stream subject to the available network bandwidth and suffered packet losses. The studies [28]–[36] represent a sample of the large body of literature on the sub-area of traffic management for a group of multiplexed streams of fixed quality. The main focus of this sub-area is to maximize the number of video streams with a given (known) traffic pattern that can be supported subject to a given amount of network bandwidth and a maximum permissible loss. Whereby the considered video streams are pre-encoded with a particular quality level or quantization scale that are not considered as independent variables in these studies. In contrast, in this study we examine the multiplexing of video streams as a function of the quality level/quantization scale. Our investigations indicate that the quality level of the video critically affects the multiplexing behavior; this effect has been largely ignored to date.

The sub-area of traffic management that adapts the quality while considering a group of multiplexed streams is generally more closely related to our work in that we also consider multiple multiplexed streams and examine the multiplexing behavior as a function of the video quality. The early studies [37]–[39] considered this problem sub-area at a conceptual level without considering in detail the dependency of the traffic statistics and multiplexing behavior on the video quality which is the focus of our study. Error-resilience mechanisms that mitigate the drop in quality due to multiplexing losses are developed in [40]. Joint source-channel coding in the context of multiplexing losses is examined [41]. The study [42] presents a system architecture for a streaming service which adaptively changes the quality level of the streamed video in response to user feedback while the streaming is ongoing and examines the achieved multiplexing performance and user perception rating. Our study of the statistical multiplexing and the utility of variable bit rate video streaming complements this literature by uncovering the fundamental relationships between the rate variability produced by the open-loop encoder for different quantization scales and the associated

achievable multiplexing gains and network utilities.

We note that the conceptual aspects of the pricing of video services and the utility of video streaming are discussed in [43]–[47]. Also, a recent study [48] examined the maximization of the utility a given user obtains from receiving a video stream. Our utility study in Section VII differs from [48] in that we consider the utility that a service provider earns from multiplexing multiple video streams over a given network bandwidth.

III. Methodology of Rate Variability-Distortion Study

For our evaluation of the rate-variability-distortion characteristics of encoded video sequences, we employ the setup illustrated in Figure 3. As example video scenes for this study we selected scenes

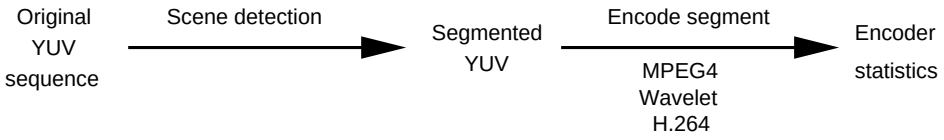


Fig. 3. Outline of the evaluation setup.

from the two movies *Star Wars IV* and *The Terminator*, and a *Football* game recording in QCIF format (176×144 pixels). In addition, we consider three of the well–known test sequences, namely *Carphone* and *Claire* in the QCIF format and *Paris* in the CIF (352×288 pixels) format.

We used publicly available scene detection software for the determination of scene boundaries. Scene boundaries were detected based on (i) director cuts (i.e., the abrupt change of scene content between two consecutive frames), and (ii) fades between scenes (i.e., the dissolving of one scene into the following scene using several frames). We visually verified the selected scenes for correctness of scene boundaries.

In general, video content can be classified according to several criteria, see for instance [49]– [50]. To illustrate the impact of the video content on the VD characteristics we consider the level of motion (content dynamics) in the video, which is widely considered a key characteristic of video content. We classify the content of a video scene according to the level of motion into five *motion classes* ranging from motion class I for a low level of motion to motion class V for a high level of motion. For each motion class we selected a representative scene from each video for our study. An overview of the selected scenes is presented in Table I.

We consider a variety of Group of Pictures (GoP) patterns, which are shown in Table II. Note that pattern 7 corresponds to the widely adopted 'standard' pattern. For the video coding, we use the reference implementations of the MPEG4 encoder, the H.264/AVC development version 6.1e [51], and the wavelet encoder presented in [52]. We used the single layer encoding and the simple profile of MPEG4, encoding each individual frame as single video object.

A. Definition of Traffic and Quality Metrics

In this section we provide the definitions of the traffic and quality metrics used throughout this paper. Let D_x and D_y denote the number of pixels in the horizontal and vertical directions in a given video frame (e.g., for QCIF $D_x = 176$ and $D_y = 144$). Let $n,\ n = 0, \ldots, N-1$, denote the position of the

TABLE I

OVERVIEW OF CONSIDERED SCENES.

Video	Scene #	Length (frames)	Motion Class	Description
Football	298	227	I	Intel logo with moving background (trailer).
Star Wars IV	274	443	I	Princess Leia's hologram pleas for help.
Terminator	384	520	I	Talk between two humans in a tunnel.
Football	299	367	II	NFL logo animation.
Star Wars IV	117	391	II	Slow zoom on R2D2 and C3PO marching in the desert.
Terminator	462	275	II	Terminator makes a call.
Football	557	266	III	Subway train GAP commercial.
Star Wars IV	115	112	III	Darth Vader gives orders and walks away.
Terminator	628	140	III	Gas truck makes U-turn.
Football	184	111	IV	Robot slides to camera (trailer).
Star Wars IV	165	89	IV	R2D2 is captured by sandpeople.
Terminator	262	69	IV	Car breaks through entrance doors.
Football	336	86	V	Camera follows running player.
Star Wars IV	632	46	V	Fight in a bar on Tatooine.
Terminator	441	253	V	A picture burns.
Carphone	1	382	III	Talking man inside car.
Claire	1	494	I	Woman talks in front of blue screen.
Paris (CIF)	1	1065	I	Man and woman are talking.

TABLE II

EVALUATED GOP PATTERNS FOR DCT–BASED ENCODINGS.

GoP pattern	Length	Pattern
0	1	II ...
1	12	IPP ...
2	30	IPP ...
3	scene	IPP ...
4	12	IBP ...
5	30	IBP ...
6	scene	IBP ...
7	12	IBBP ...
8	21	IBBP ...
9	scene	IBBP ...
10	12	IBBBP ...
11	12	IBBBBBP...

given video frame in a given video sequence that consists of N frames. Each video frame is divided into M macroblocks (MB) of 16×16 pixels (e.g., $M = 99$ for QCIF) for DCT based encoding. The DCT transformation is performed on a block (i.e., subdivision of a macroblock) of 8×8 pixels. We denote the quantization scale of the encoder by q. The possible values for q vary from $q = 1, 2, \ldots, 31$ for MPEG4 and $q = 1, 2, \ldots, 51$ for H.264/AVC.

1) Traffic Metrics: We denote the size (in bit) of the nth encoded video frame for quantization scale q by X_n^q. The mean frame size of an encoded video sequence is defined as

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

We let σ_q denote the standard deviation of the frame size defined as

$$\sigma_q = \sqrt{\frac{1}{(N-1)} \sum_{n=0}^{N-1} (X_n^q - \overline{X}_q)^2}. \tag{2}$$

We primarily employ the coefficient of variation [53] defined as

$$CoV_q = \frac{\sigma_q}{\overline{X}_q} \tag{3}$$

as the measure of variability of the frame sizes (the variability of the bit rate of the encoded video). We note that an alternative measure of the variability of the frame sizes is the peak-to-mean frame size ratio defined as

$$\frac{\max_{0\le n\le N-1} X_n^q}{\overline{X}_q}. \tag{4}$$

The peak-to-mean ratio, however, can be affected by a single very large video frame (i.e., an outlier). This problem is avoided by the coefficient of variation of the frame sizes CoV_q, which gives the normalized averaged deviation of the individual frame sizes from the mean frame size and is therefore widely employed in performance evaluation [53], [54]. We denote the maximum coefficient of variation for a given video sequence as

$$CoV_{\max} = \max_q CoV_q \tag{5}$$

and denote $q_{\max}$ for the quantization scale that attains this maximum, i.e., $q_{\max} = \arg\max_q CoV_q$. We similarly define $CoV_{\min}$ and $q_{\min}$.

We define the aggregated frame size trace with aggregation level $a,\ a \ge 1$, as

$$X_q^{i,a} = \frac{1}{a} \sum_{j=na}^{(n+1)a-1} X_q^j, \quad \text{for} \ \ i = 0, \ldots, N/a - 1, \tag{6}$$

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

When different frame types exist, these can be combined to form the group of pictures (GoP). Frame types are intra coded (I) frames, predicted (P) frames, and bi-directional predicted (B) frames. The latter two frame types are also referred to as inter coded frame types. Whenever we refer to an explicit frame type, we denote the type by its abbreviation in the subscript (e.g., X_P for the size of a P frame).

2) Quality Metrics: We denote an individual pixel value (an 8–bit value for the luminance) in the original (uncompressed) video frame by $F_n(x, y)$ and its encoded and decoded counterpart by $f_n^q(x, y)$, with $0 \le x \le D_x - 1$ and $0 \le y \le D_y - 1$. We use the peak signal to noise ratio (PSNR) as a measure for the objective video quality. The PSNR quality Q_n^q of frame n encoded with quantization scale q is defined as

$$Q_n^q = 10 \cdot \log_{10} \frac{255^2 \cdot D_x \cdot D_y}{\sum_{x=0}^{D_x-1} \sum_{y=0}^{D_y-1} \left[F_n(x, y) - f_q^q(x, y)\right]^2}. \tag{7}$$

We apply the PSNR only to the luminance (Y) component of an individual frame, as the human eye is most sensitive to this component [55]. We calculate the average objective quality for a given video sequence of N individual frames as

$$Q_q = \frac{1}{N} \sum_{n=0}^{N-1} Q_n^q \tag{8}$$

We note that several more sophisticated approaches for the calculation of the objective video quality (i.e., algorithms that take the human visual system into consideration) exist, but the PSNR is widely used as it is computationally simple a gives generally a reasonably good measure for the perceived quality [56].

3) Entropy: As measure for the complexity of an individual video frame, we employ the Shannon entropy. The complexity for each frame n is calculated as entropy of the individual byte values $F_n(x, y)$, given as

$$H_n = -\sum_{F_b=0}^{255} P[F_b] \cdot \log_2 P[F_b], \tag{9}$$

where $P[F_b]$ denotes the sample probability of byte value F_b, with $F_b \in \{0, 1, \ldots, 255\}$, for all $D_x \cdot D_y$ byte values. We denote the coefficient of variation of the entropy values of the frames $n,\ n = 1, \ldots, N$, of a video sequence by CoV_H.

We use the coefficient of correlation as measure of (linear) dependency, defined as

$$\rho_{x,y} = \frac{\frac{1}{M}\sum_{m=1}^{M} x_m y_m - \bar{x}\bar{y}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}(x_m - \bar{x})^2}\sqrt{\frac{1}{M}\sum_{m=1}^{M}(y_m - \bar{y})^2}} \tag{10}$$

for a set of data points denoted by x_m and $y_m,\ m = 1, \ldots, M$.

IV. Rate Variability-Distortion Characteristics of Open–Loop Encoded Video

In this section, we examine the coefficient of variation as a function of the video quality, i.e., the VD curve, of open-loop encoded video. Current video coders are mostly based on the Discrete Cosine Transform (DCT). The transform coefficients are quantized to further reduce the size of the video frames at the expense of the loss of video information (and thus reduced video quality). The combination of these two encoding mechanisms is commonly referred to as intra coding or texture coding. In addition to intra coding, current video coders typically exploit temporal dependencies between consecutive video frames by encoding only movements or differences between consecutive frames, which is referred to as motion estimation and compensation or inter coding. To examine the VD curve in detail, we first focus on intra coding and subsequently expand our study to include inter coding.

A. Intra Coding

In this section we examine the rate variability-distortion (VD) behavior of intra coded video and study the correlation between the VD behavior and the entropy of the video. In Figs. 6 and 7 we plot the *VD curve*s, i.e., the coefficient of variation of the frame sizes CoV_q as a function of the PSNR video quality, for the *Terminator* and *Football* scenes. We observe that the intra coded videos exhibit a characteristic hump in the VD curve, although the magnitude of the CoV as well as the change in the CoV are smaller compared to the encoding employing both intra coding and inter coding (see Figs. 1 and 2). Nevertheless, the CoV varies approximately between 0.1 and 0.22 for the motion class IV scene from *The Terminator* and between 0.14 and 0.3 for the motion class II scene from *Football*. As illustrated in Fig. 8 for scenes from *The Terminator*, we also observe a hump in the VD curve for wavelet based encoding, although the overall level of the variability is lower than with the DCT based encoding.

To examine the origin of the hump phenomenon in the intra coded videos, we study the correlation between the rate variability and the entropy of the raw (uncompressed) videos. More, specifically, we study the correlation between the coefficient of variation of the frame entropies CoV_H and elementary statistics of the coefficient of variation of the frame sizes CoV_q. We consider the following elementary statistics:

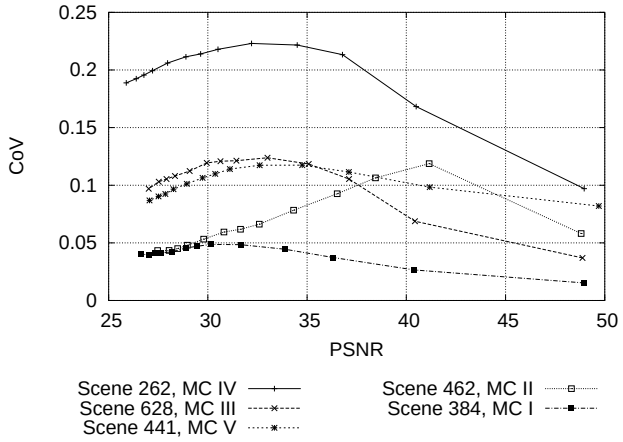


Fig. 4. VD curves for intra coded scenes using DCT based MPEG-4 from *The Terminator*.

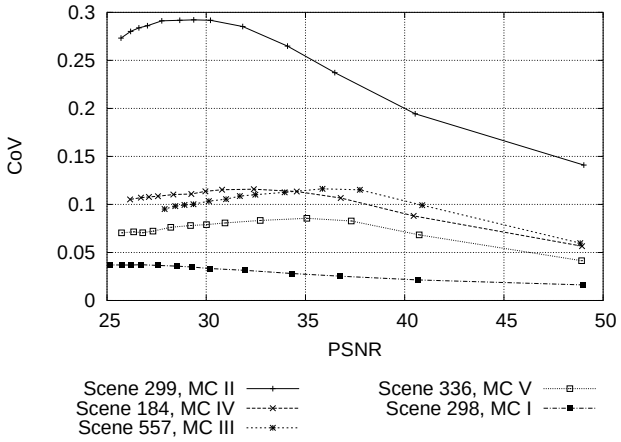


Fig. 5. VD curves for intra coded scenes using DCT based MPEG-4 from *Football*.

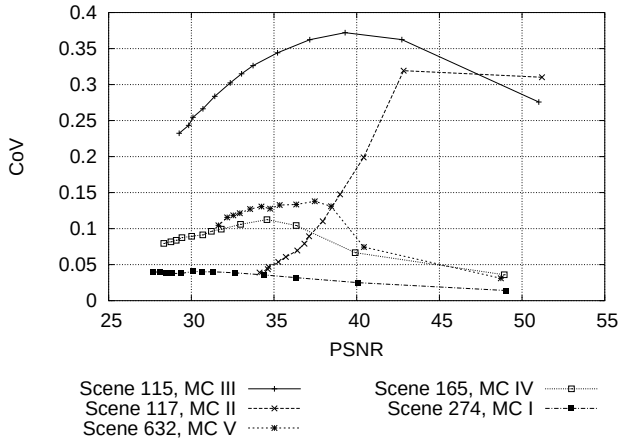


Fig. 6. VD curves for intra coded scenes using DCT based MPEG-4 from *The Terminator*.

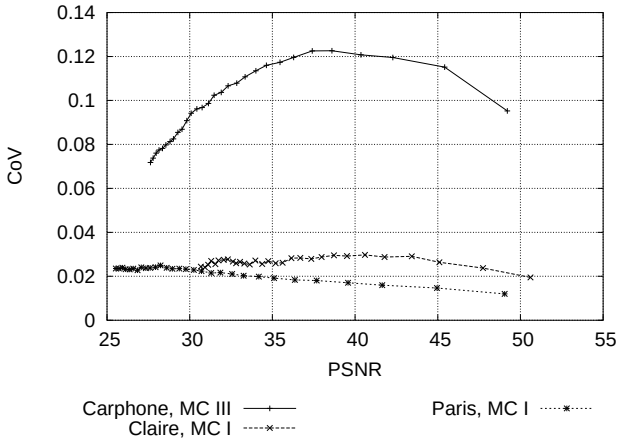


Fig. 7. VD curves for intra coded test sequences using DCT based MPEG-4 from *Football*.

(i) the largest CoV_q, i.e., $CoV_{\max}$ as defined in (5), (ii) the range of the CoV_q, i.e., $CoV_{\max} - CoV_{\min}$, and the standard deviation of the CoV_q. The coefficients of correlation between each of the columns with the elementary statistics of the CoV_q and the CoV_H column (calculated according to (10)) are all above 0.84. This indicates a relatively strong dependence of the variability of the encoded frame sizes on the variability of entropy of the uncompressed source video. We may thus conclude that the origin for the VD behavior of the intra coded video lies in the complexity of the individual video frames.

To visualize this dependence we plot in Fig. 9 the frame sizes of *Carphone* encoded with $q = 15$ and the corresponding frame entropy for *Carphone*. We observe that typically video frames with a high entropy result in large encoded frame sizes and that the variability of the entropy of the source video is reflected in the variability of the encoded video frames.

B. Inter Coding

In this section we examine the VD behavior of video encoded using intra coding together with inter coding. We initially study the VD behavior for the GoP pattern 7, which is widely considered in video studies. We also initially consider encodings where all three frame types are encoded with the same

TABLE III

ELEMENTARY STATISTICS OF THE VARIATION OF THE INTRA CODED FRAME SIZES CoV AND THE COEFFICIENT OF VARIATION OF THE ENTROPY OF THE UNCOMPRESSED VIDEO FRAMES CoV_H. THE CoV STATISTICS ARE HIGHLY CORRELATED TO THE ENTROPY VARIATION CoV_H.

Movie	Scene	MC	$CoV_{\max}$	$CoV_{\max} - CoV_{\min}$	Std. Dev. CoV_q	CoV_H
Football	298	I	0.037	0.021	0.007	0.007
	299	II	0.292	0.151	0.046	0.078
	557	III	0.116	0.057	0.014	0.057
	184	IV	0.116	0.059	0.016	0.019
	336	V	0.086	0.044	0.011	0.019
Star Wars IV	274	I	0.041	0.027	0.008	0.003
	117	II	0.319	0.280	0.097	0.232
	115	III	0.372	0.140	0.048	0.204
	165	IV	0.113	0.077	0.020	0.013
	632	V	0.138	0.107	0.030	0.031
The Terminator	384	I	0.049	0.033	0.009	0.004
	462	II	0.119	0.075	0.025	0.048
	628	III	0.124	0.087	0.025	0.02
	262	IV	0.223	0.126	0.033	0.069
	441	V	0.117	0.035	0.012	0.075
Test Sequence	*Carphone*	III	0.123	0.051	0.017	0.01
	Claire	I	0.030	0.010	0.002	0.007
	Paris	I	0.025	0.013	0.003	0.001

quantization scale q. We then examine the impact of different GoP patterns and different quantization scales for the different frame types.

In Figs. 1 and 2 we have plotted the VD curves of the *Terminator* and *Football* scenes encoded in MPEG-4. Similarly in Figs. 10 and 11 we plot the VD curves of the test sequences encoded in MPEG-4 and H.264. The maximum coefficient of variation $CoV_{\max}$, the quantization scale attaining this maximum $q_{\max}$, and the average PSNR quality at this quantization scale $Q_{q_{\max}}$ for the MPEG-4 encodings are summarized in Table IV. From a close inspection of Figs. 1, 2, 10, and 11, as well as Table IV, we observe that the scenes with motion class I exhibit by far the most pronounced peak of the coefficient of variation CoV, whereas the scenes with motion class V exhibit the smallest peak. The other motion classes

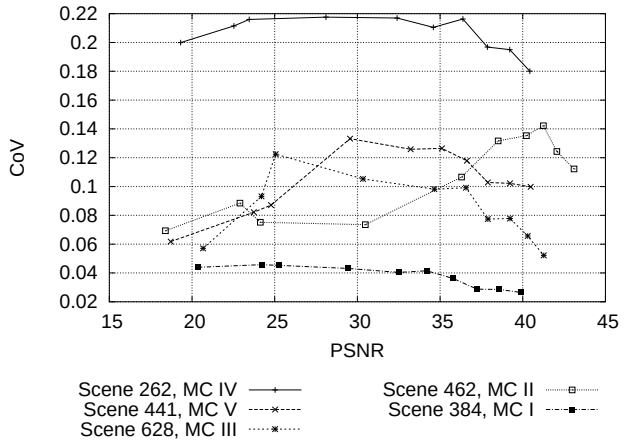
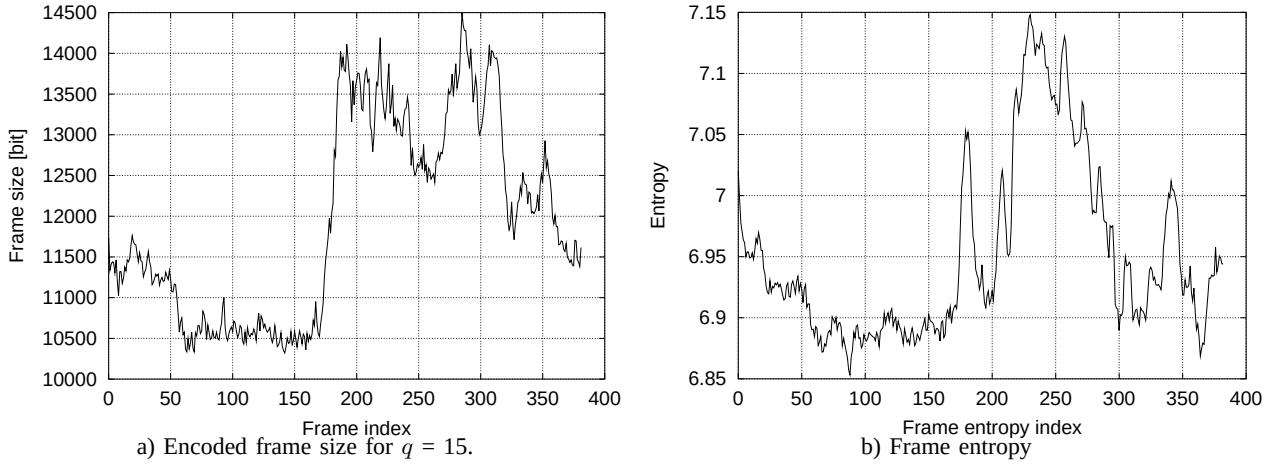


Fig. 8. VD curves for intra coded scenes using wavelet based coding from *The Terminator*.



a) Encoded frame size for q = 15. b) Frame entropy

Fig. 9. Encoded frame sizes and entropy of the raw video for *Carphone*.

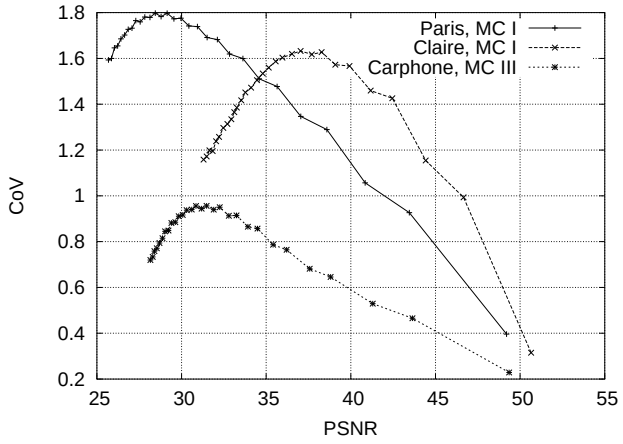


Fig. 10. VD curves for MPEG4 encoded test sequences.

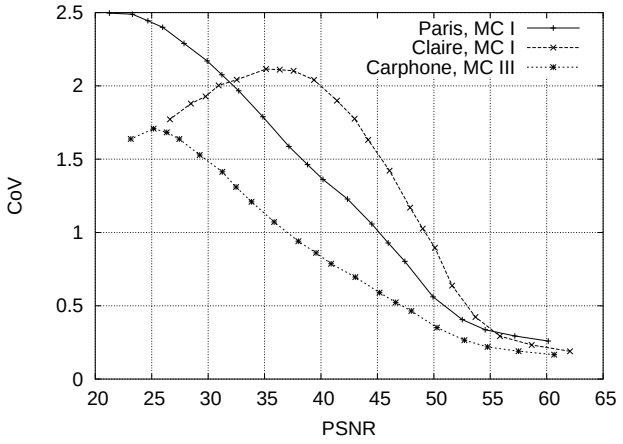


Fig. 11. VD curves for H.264/AVC encoded test sequences.

lie in between these two extremes. For these other motion classes there is again an ordering where the lower motion classes attain higher CoV peaks, although the differences between these peaks are relatively small compared to the wide gap between the peaks for the motion classes I and V. All VD curves exhibit the characteristic hump behavior whereby the variability increases with decreasing quality, reaches a peak, and then drops off as the quality decreases further. These typical behaviors of the VD curve are further confirmed by the VD curves of the test sequences in Figs. 10 and 11, which demonstrate that both the MPEG-4 and H.264 encoders produce the characteristic hump behavior of the VD curve. We observe that the VD curve for H.264 is smoother than the curve for MPEG-4, which may be due to the various enhancements in H.264 over MPEG-4.

To explain these VD characteristics of inter coded video, we have first examined whether there is strong connection between the entropy of the uncompressed video and the coefficient of variation of the frame sizes. Specifically, we have encoded the video scenes using GoP pattern 3, i.e., the first frame of the scene is intra coded and all subsequent frames are inter coded with respect to the preceding frame. We compared this variability of the frame sizes with the entropy of the difference in the pixel values between successive frames. That is, we calculated the entropy of the differences $F_n(x,y) - F_{n-1}(x,y)$ for each of

TABLE IV

MAXIMUM COEFFICIENT OF VARIATION OF FRAME SIZE $CoV_{\max}$, QUANTIZATION SCALE ATTAINING MAXIMUM $q_{\max}$, AND QUALITY AT MAXIMUM $Q_{q_{\max}}$ FOR MPEG-4 ENCODINGS WITH GOP PATTERN 7 AND SAME QUANTIZATION SCALE FOR ALL FRAME TYPES.

Movie	Scene	MC	GoP Pattern 7		
			$CoV_{\max}$	$q_{\max}$	$Q_{q_{\max}}$
Football	298	I	2.35	22	26.85
	299	II	1.22	22	26.93
	557	III	1.03	16	30.92
	184	IV	0.78	22	27.18
	336	V	0.33	22	26.49
Star Wars IV	274	I	1.26	16	30.08
	117	II	0.78	24	34.77
	115	III	0.96	18	31.37
	165	IV	0.59	14	30.83
	632	V	0.42	9	35.75
The Terminator	384	I	1.63	16	28.94
	462	II	1.17	22	28.83
	628	III	0.68	16	29.67
	262	IV	0.62	16	28.33
	441	V	0.37	14	29.42
Test sequences	*Carphone*	III	0.96	14	31.48
	Claire	I	1.63	10	37.04
	Paris	I	1.8	18	28.45

the frames $n,\ n = 1, \ldots, N-1$, and then calculate the coefficient of variation of these individual frame difference entropies. This calculation of the entropy is motivated by the inter coding, which primarily encodes the differences between successive frames. We have observed that there is no strong correlation between the coefficient of variation of the frame sizes and entropy differences. From this observation we may conclude that the inter coding introduces additional variation, which is not explained by the frame complexity measured as entropy.

To obtain insight into the origins of the VD behavior we proceed to examine the coefficient of variability more closely and then examine the encoding of the different frame types. The slope of the VD curve is given by

$$\frac{dCoV_q}{dq} = \frac{\sigma'_q \cdot \bar{X}_q - \sigma_q \cdot \bar{X}'_q}{\bar{X}_q^2}, \tag{11}$$

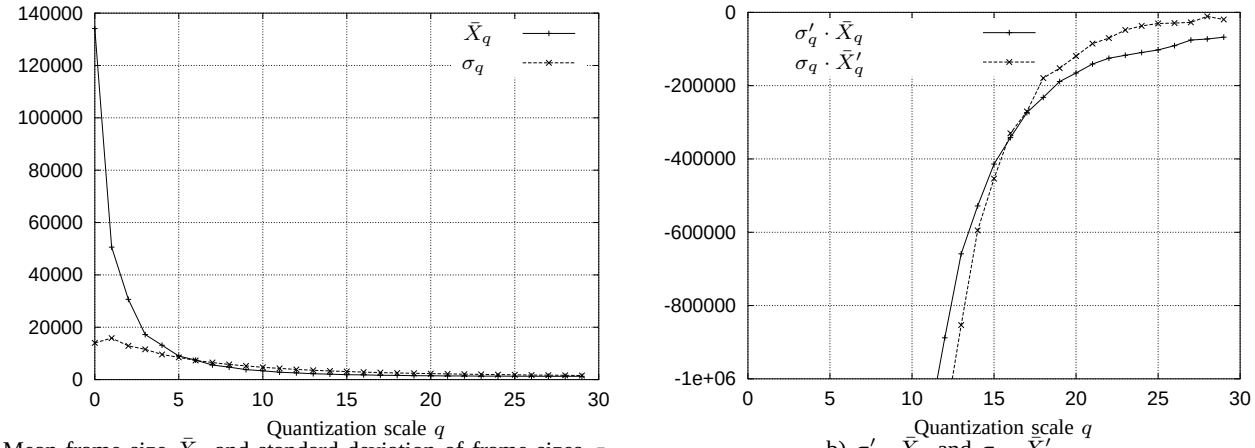
where y' denotes the derivative of y with respect to q, The VD curve is increasing in q if

$$\sigma'_q \cdot \bar{X}_q > \sigma_q \cdot \bar{X}'_q, \tag{12}$$

and decreasing in q if this inequality is reversed. We illustrate this behavior for Scene 384 (MC I) of *The Terminator* in Fig. 12.

We observe that for $q < 17$, with increasing q the standard deviation of the frames sizes is dropping relatively slower than the mean frame size such that inequality (12) holds and the VD curve is increasing. For $q > 17$, the trend is reversed, i.e., the standard deviation of the frame sizes is dropping relatively faster than the mean frame size with increasing q, resulting in the decrease of the VD curve.

In the following, we further examine the different factors influencing the VD behavior for GoP pattern 7.



a) Mean frame size $\bar{X}_q$ and standard deviation of frame sizes σ_q.

b) $\sigma'_q \cdot \bar{X}_q$ and $\sigma_q \cdot \bar{X}'_q$

Fig. 12. Mean and standard deviation of frame size and derivative terms as a function of quantization scale q for scene 384 from *The Terminator* encoded with GoP pattern 7.

a) Average Sizes of Frame Types: Due to the fixed GoP structure for the encoding, the different frame types Intra (I), Predicted (P), and Bi–directional predicted (B) occur at fixed positions in the encoded video. The differences of the mean frame sizes of the three frame types thus contribute to the variability of the encoded video. In Figs. 13,14, and 15, we illustrate the mean frame sizes by frame type for the scenes from *Football*,*Star Wars IV*, and *The Terminator*.

We observe that the difference between the different frame types is more pronounced in the low motion scene 384 than in the high motion scene 441. In addition, we observe that the pronounced $CoV_{\max}$ observed in scene 384 is located in the region where the mean sizes of the two predicted frame types begin to remain roughly constant compared to the intra coded frame types. For scene 441, where we do not observe a pronounced hump in the VD curve, we do not observe such behavior, as the mean frame sizes of all three frame types constantly decrease as the quantization scale increases.

The previous observations were made based on the mean frame type sizes without regard of their contribution to the total video traffic. In Figs. 16, 17, and18 we plot the fraction of the total video traffic that is from a specific frame type.

We observe that in scene *384* the fraction of encoded video data that is stored in the different frame types decreases for the two predicted frame types, reaches a minimum where the $CoV_{\max}$ is located on the quantization scale, and slightly increases afterwards. In the high motion scene from *The Terminator*, we do not observe such behavior. Instead, the fractions of data stored in each frame type remain approximately constant.

b) Variabilities of Sizes of Frame Types: In addition to the mean sizes, the variability of the sizes of the different frame types (measured as CoV) influences the VD curve characteristics. In Figures 19, 20, and 21 we plot the coefficients of variation of the different frame types for scenes from *Football*, *Star Wars IV*, and *The Terminator*.

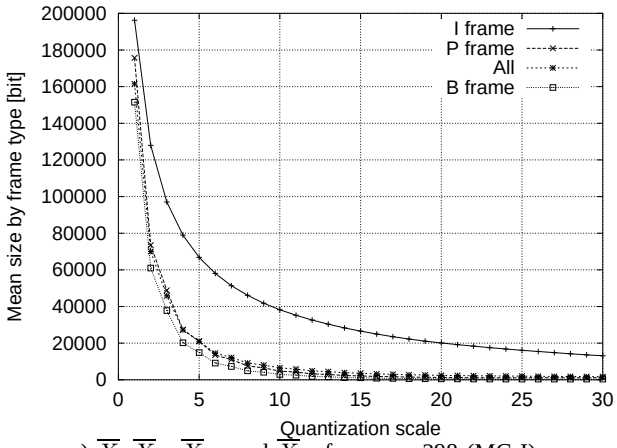
From the plots shown here we observe that the I frame variability is lower than the variability of the predicted frame types. We also observe that for P frames the maximum variability typically decreases from motion class I to V. We furthermore observe that the overall $CoV_{\max}$ of the scene is mainly governed by

the P and B frames.

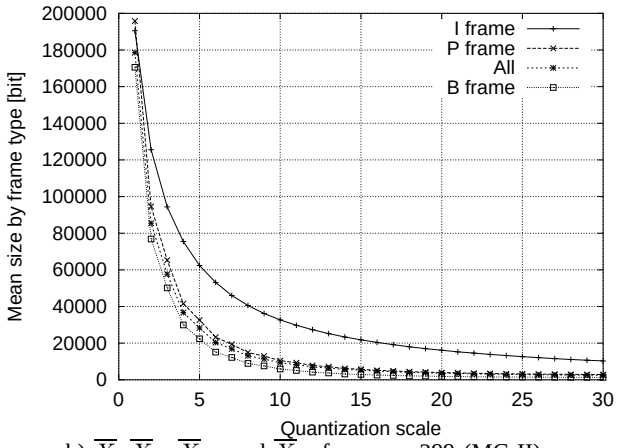
C. Impact of Different Quantization Scales for Different Frame Types on VD Characteristics

In this section we examine the impact of different quantization scales q for the different frame types. We denote q_I, q_P, and q_B, for the quantization scales used for I, P, and B frames, respectively. Initially, we consider a constant offset between the quantization scales, specifically, $q_P = q_I + 5$ and $q_B = q_P + 5$ for $q_I = 5, 6, \ldots, 20$. We plot the resulting VD curves for the scenes from *Football*, *Star Wars IV*, *The Terminator*, and the test sequences in Figure 22. We observe that with these different quantization scales, the variability is overall steadily increasing with increasing video quality. This behavior is in contrast to the hump behavior observed in Figs. 1 and 2 for identical quantization scales for the different frame types. We also observe that with both, identical and different quantization scales, motion class I has by far the highest variability and motion class V has the smallest variability, with the variabilities of the other motion classes lying in between. In addition, we observe that the frame size variabilities with different quantization scales are higher than with identical quantization scales. This is because with the larger quantization scales for the P and B frames, these frames have now even smaller sizes compared to the I frame sizes, which results in an overall increased variability of the frame sizes.

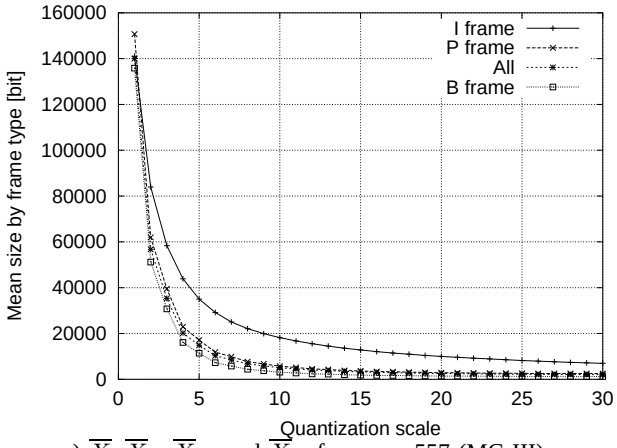
Next, we examine the impact of the magnitude of the differences between the quantization scales for the different frame sizes. Specifically, we consider the differences 1, 3, 5, 7, and 10 (i.e., $q_P = q_I + 1$, $q_P = q_I + 3$, $q_P = q_I + 7$, and $q_P = q_I + 10$, and analogously for q_B). In Figs. 23, 24, and 25 we plot the resulting VD curves for the different motion classes I to V and scenes from all three movies. We observe that with larger differences of the quantization scales, the variability increases significantly for higher video quality. The general level of the variability is higher for the low motion scenes than for the high motion scenes, which corroborates our earlier observations. In addition we note the difference in the VD behavior for low and high motion scenes, i.e., the increase in the frame quantization differences results in an almost exponentially increasing VD curve for high motion scenes, whereas the VD curve for low motion scenes becomes linearly increasing.



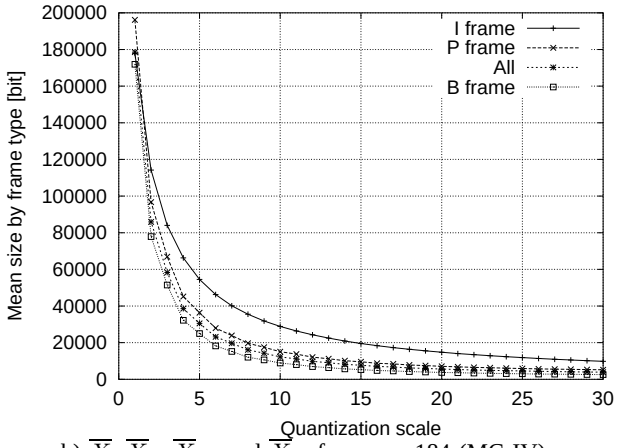
a) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 298 (MC I)



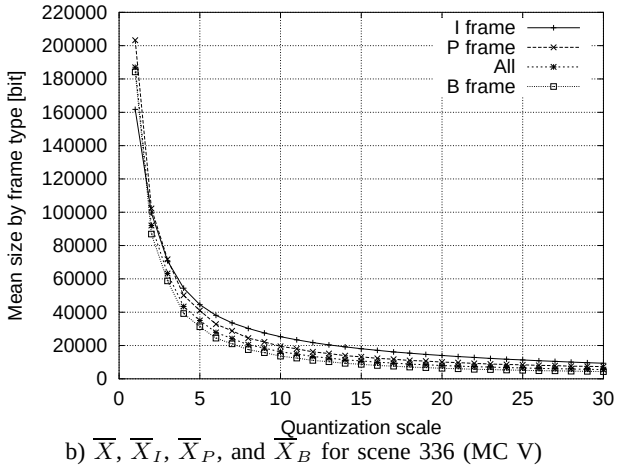
b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 299 (MC II)



a) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 557 (MC III)

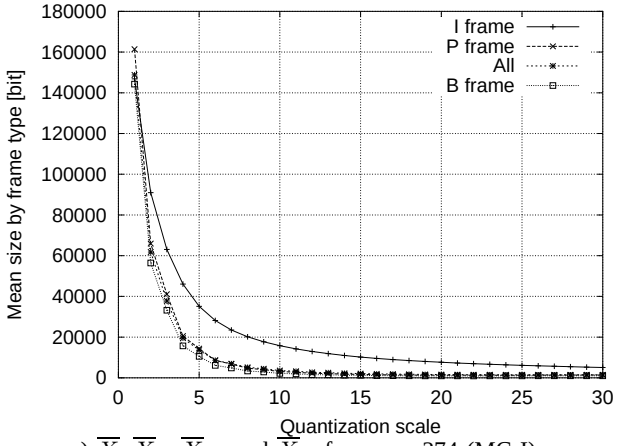


b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 184 (MC IV)

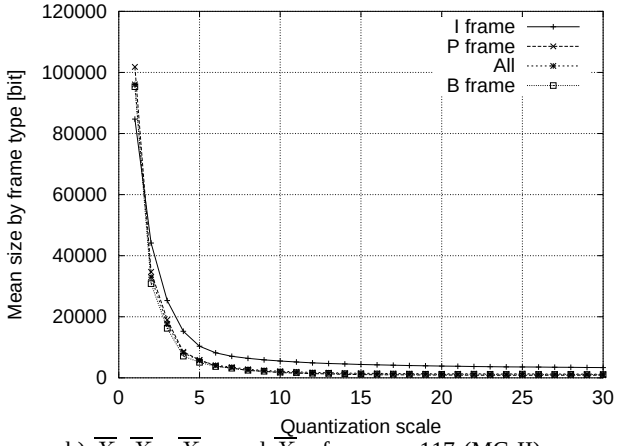


b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 336 (MC V)

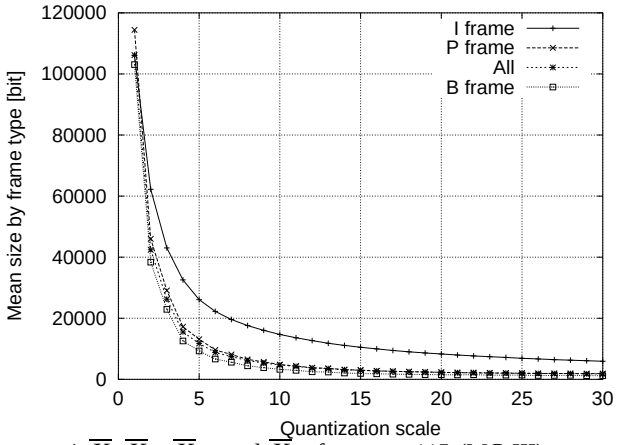
Fig. 13. Mean frame sizes by frame type for scenes from *Football* encoded with GoP pattern 7.



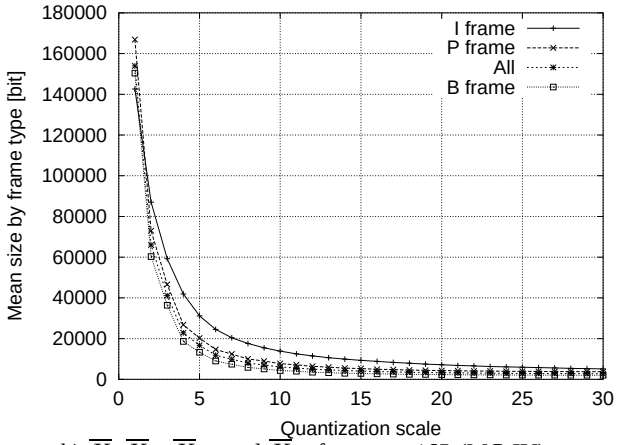
a) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 274 (MC I)



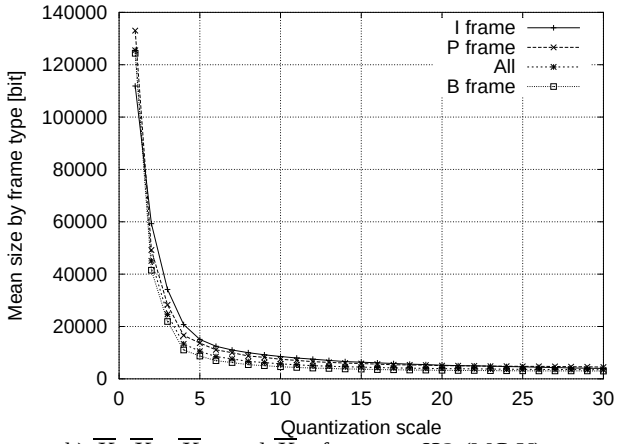
b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 117 (MC II)



a) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 115 (MC III)

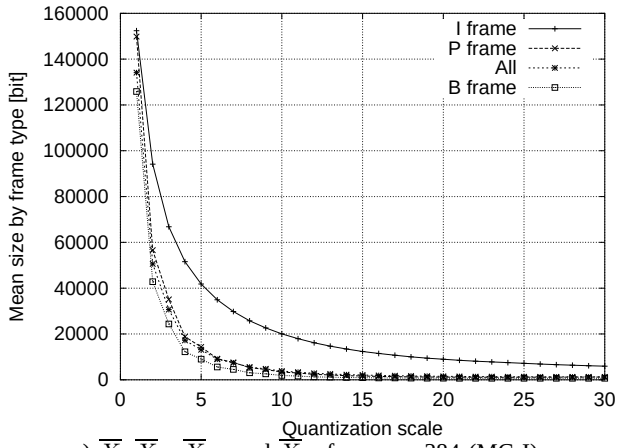


b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 165 (MC IV)

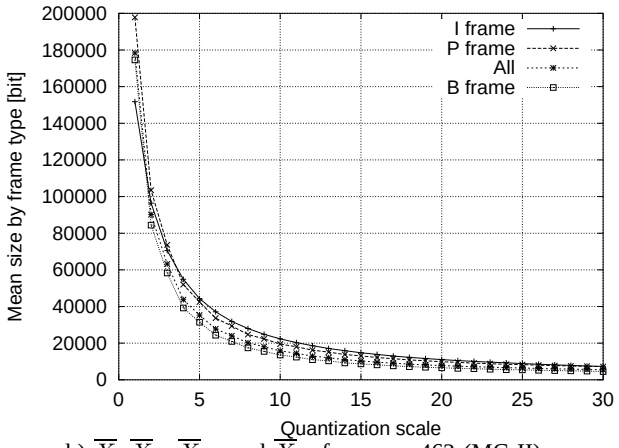


b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 632 (MC V)

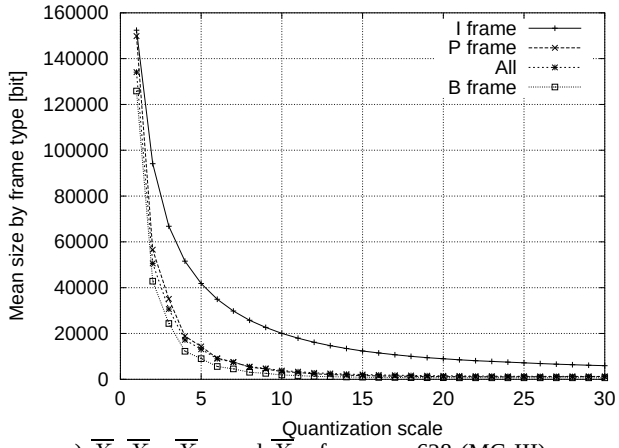
Fig. 14. Mean frame sizes by frame type for scenes from *Star Wars IV* encoded with GoP pattern 7.



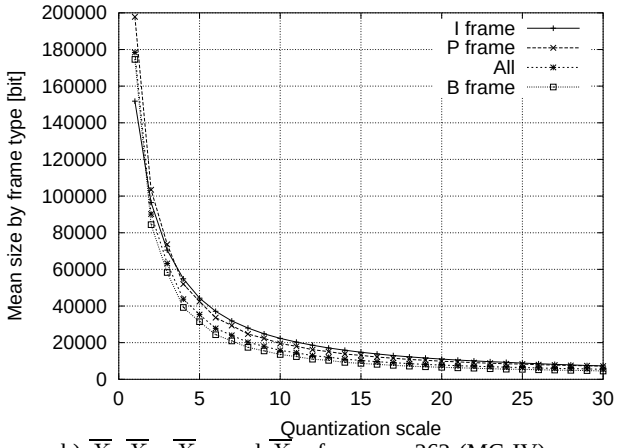
a) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 384 (MC I)



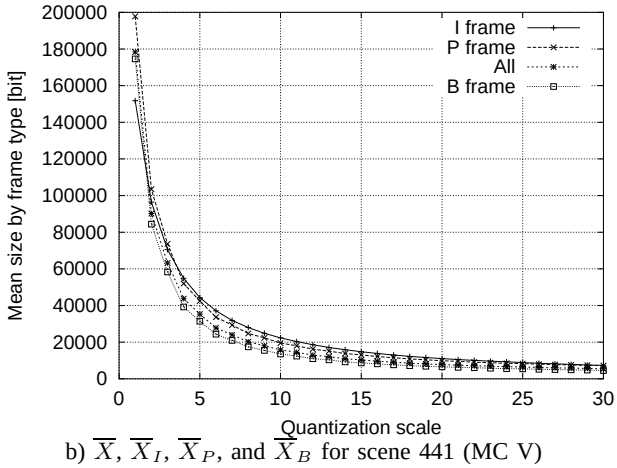
b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 462 (MC II)



a) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 628 (MC III)

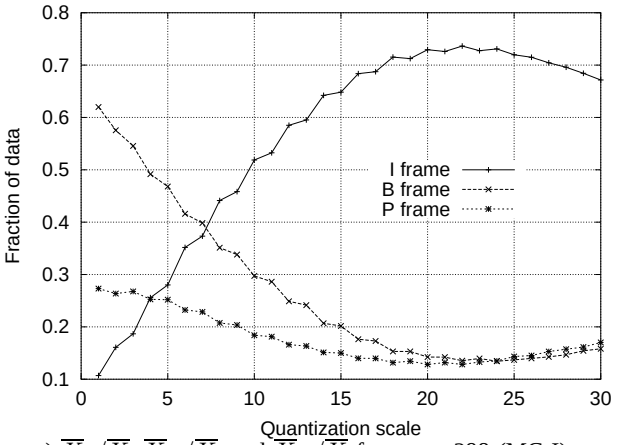


b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 262 (MC IV)

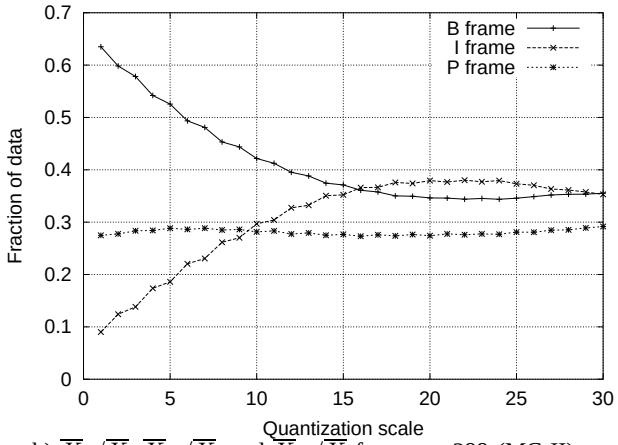


b) $\overline{X}$, $\overline{X}_I$, $\overline{X}_P$, and $\overline{X}_B$ for scene 441 (MC V)

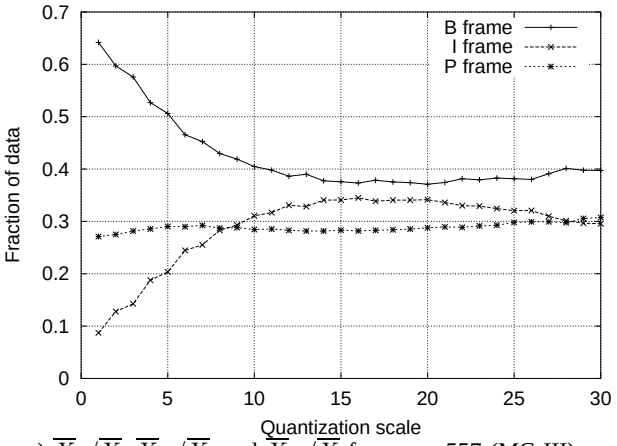
Fig. 15. Mean frame sizes by frame type for scenes from *The Terminator* encoded with GoP pattern 7.



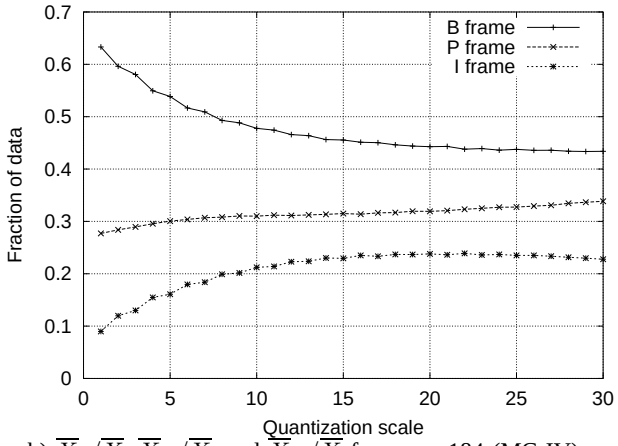
a) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 298 (MC I)



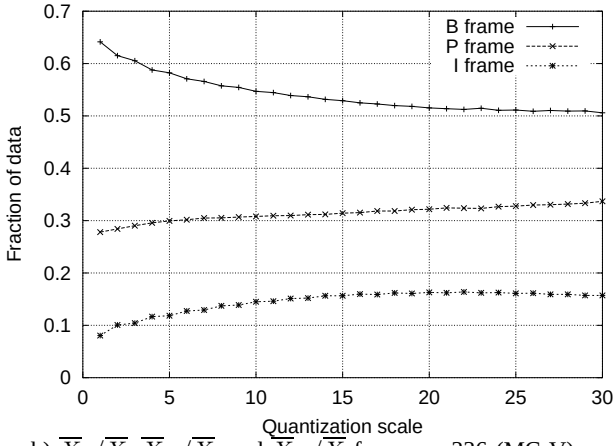
b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 299 (MC II)



a) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 557 (MC III)

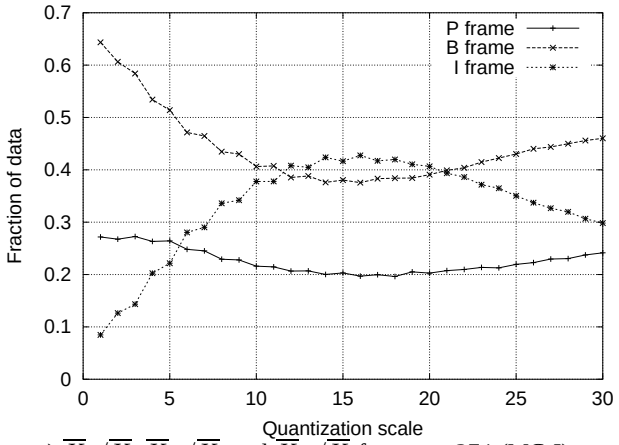


b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 184 (MC IV)

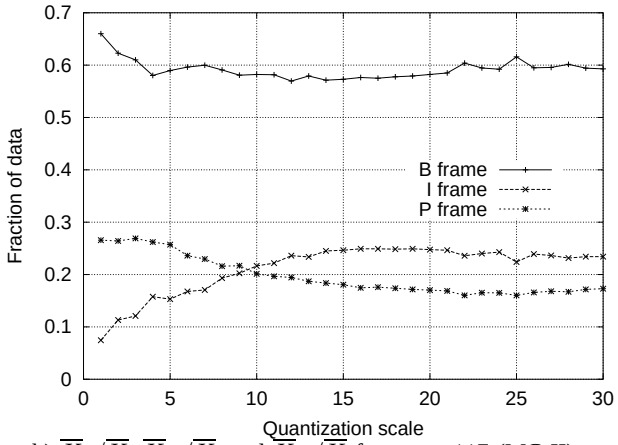


b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 336 (MC V)

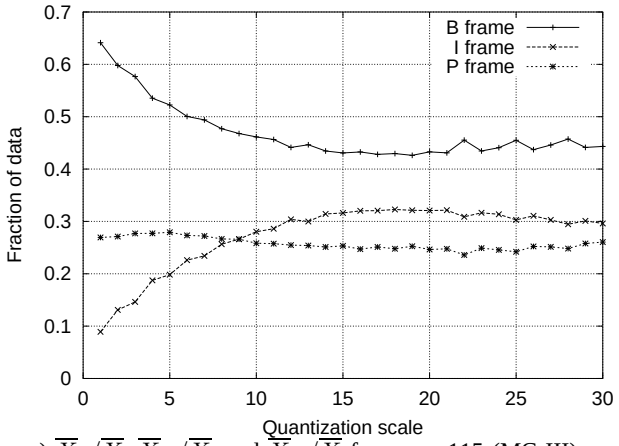
Fig. 16. Fractions of frame types for scenes from *The Terminator* encoded with GoP pattern 7.



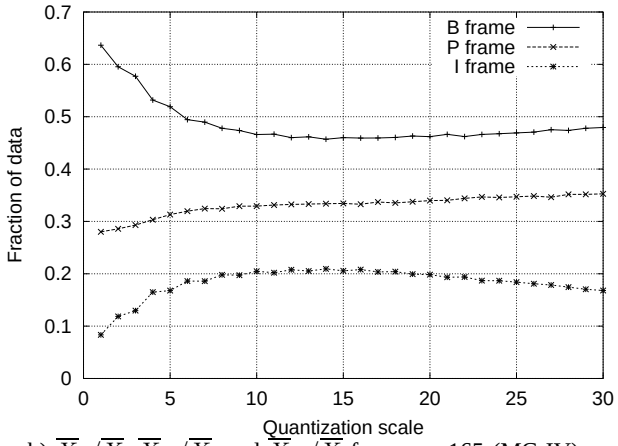
a) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 274 (MC I)



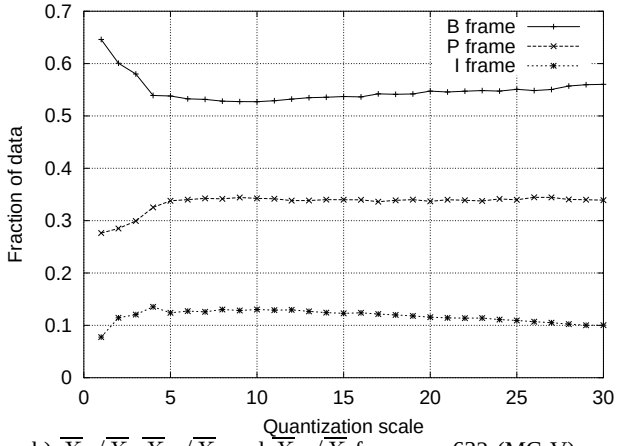
b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 117 (MC II)



a) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 115 (MC III)

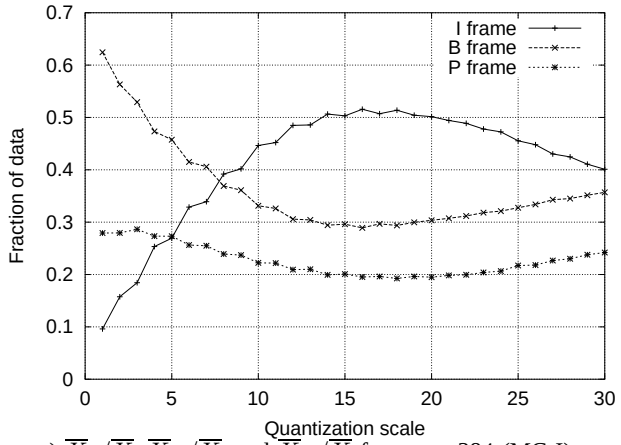


b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 165 (MC IV)

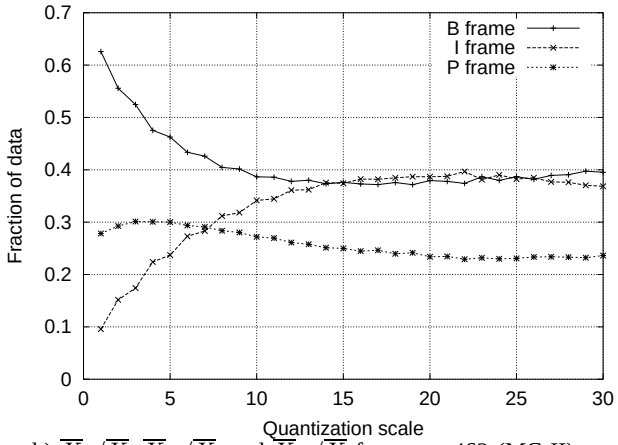


b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 632 (MC V)

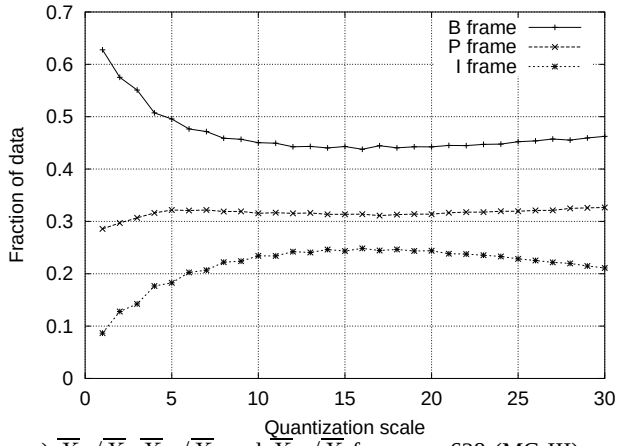
Fig. 17. Fractions of frame types for scenes from *Star Wars IV* encoded with GoP pattern 7.



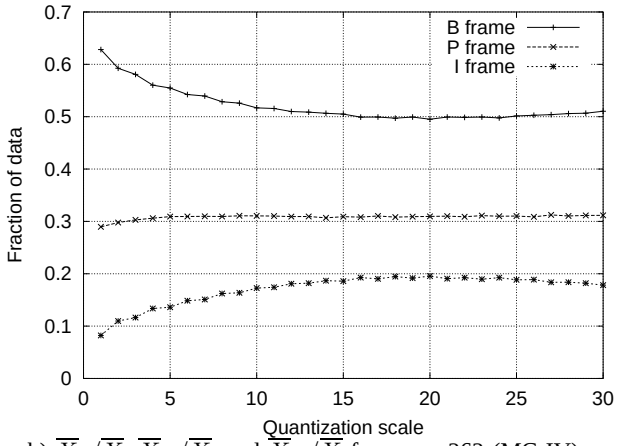
a) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 384 (MC I)



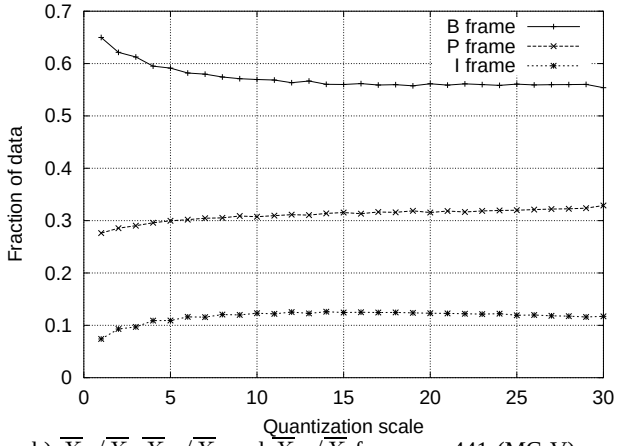
b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 462 (MC II)



a) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 628 (MC III)



b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 262 (MC IV)



b) $\overline{X}_I/\overline{X}$, $\overline{X}_P/\overline{X}$, and $\overline{X}_B/\overline{X}$ for scene 441 (MC V)

Fig. 18. Fractions of frame types for scenes from *The Terminator* encoded with GoP pattern 7.

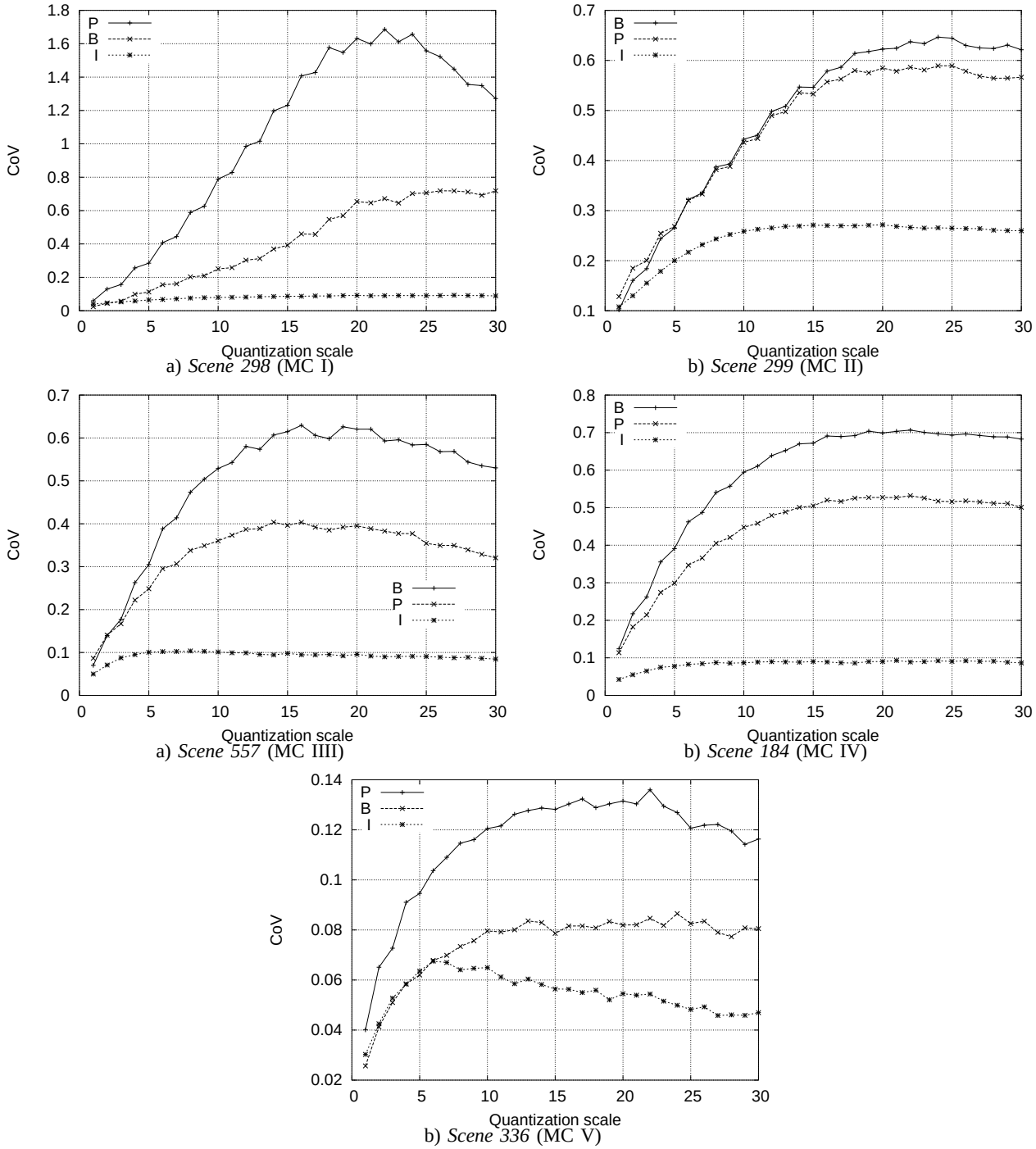
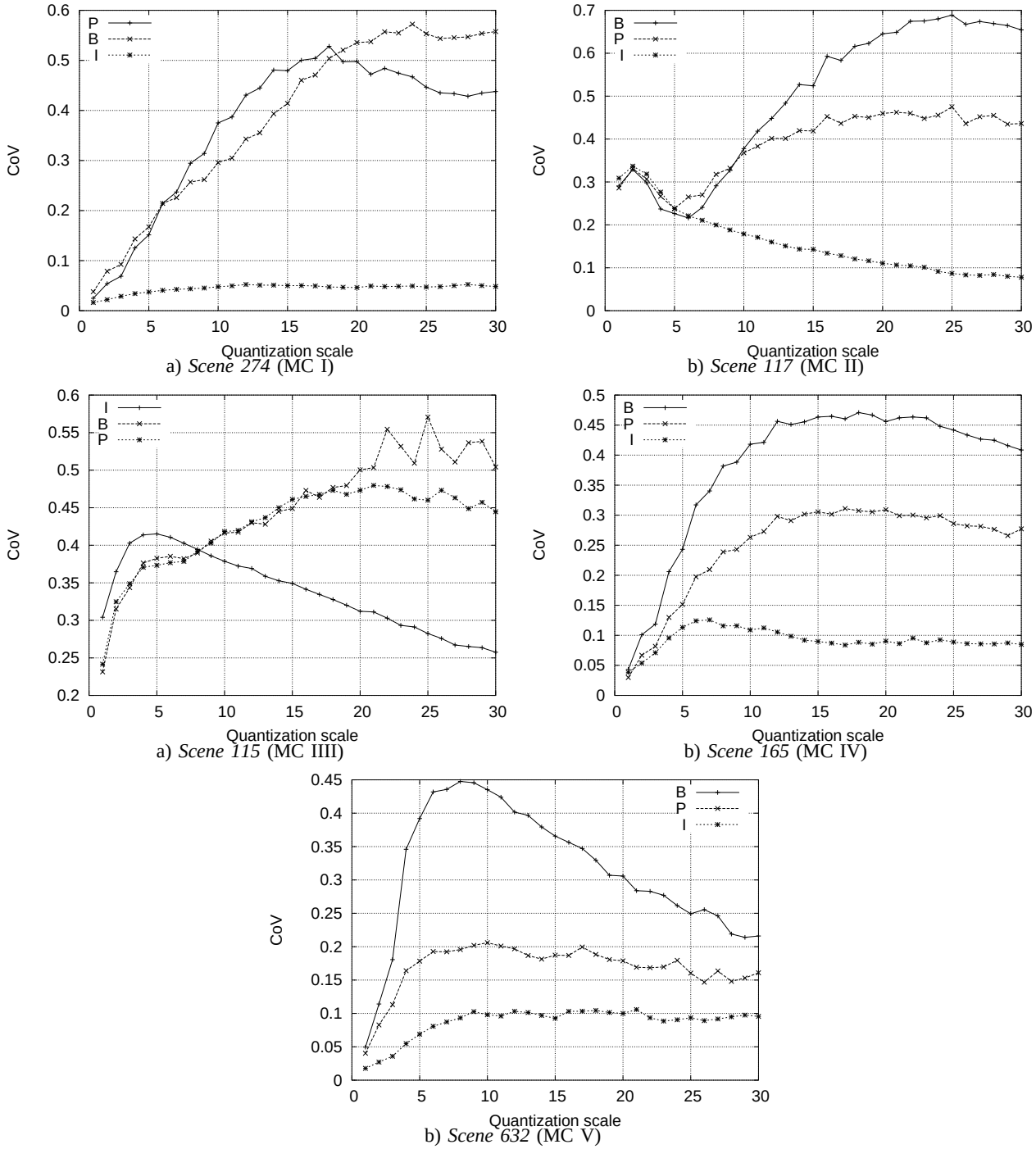


Fig. 19. Coefficient of variation CoV_q for different frame types for *Football* scenes.



a) *Scene 274* (MC I)

b) *Scene 117* (MC II)

a) *Scene 115* (MC IIII)

b) *Scene 165* (MC IV)

b) *Scene 632* (MC V)

Fig. 20. Coefficient of variation CoV_q for different frame types for *Star Wars IV* scenes.

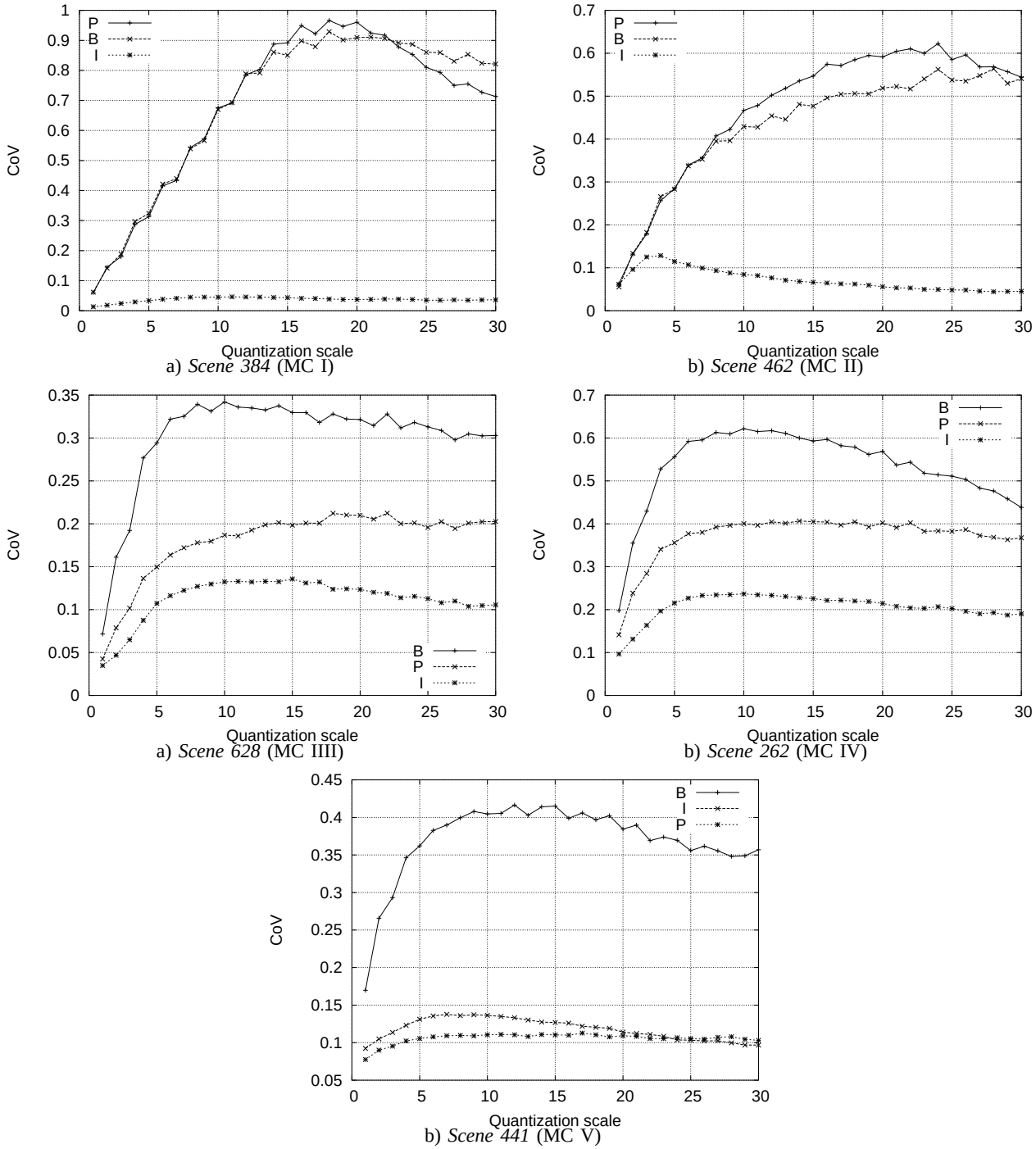
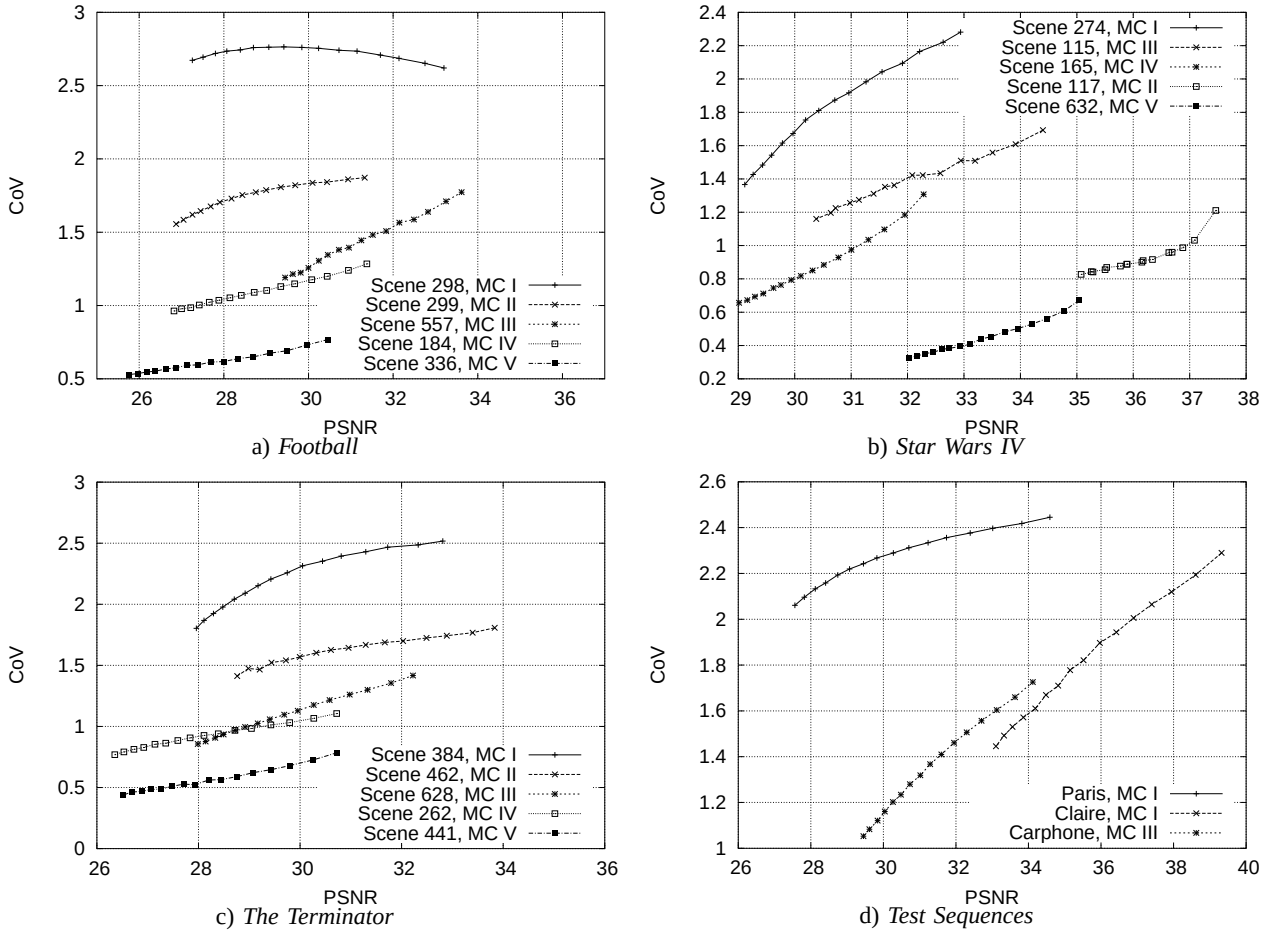


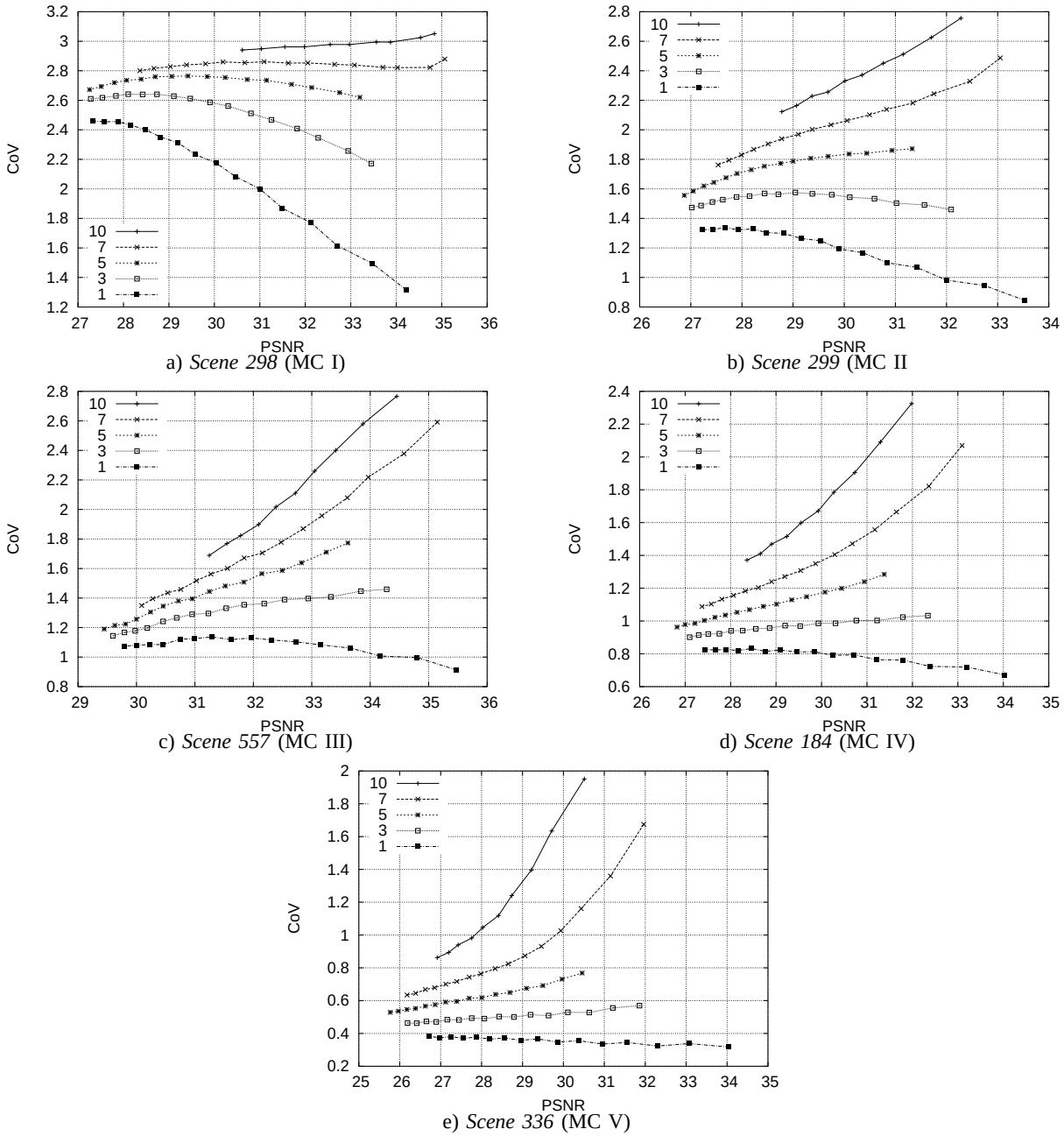
Fig. 21. Coefficient of variation CoV_q for different frame types for *The Terminator* scenes.



a) *Football* b) *Star Wars IV*

c) *The Terminator* d) *Test Sequences*

Fig. 22. VD curves for different quantization scales for different frame types with $q_P = q_I + 5$ and $q_B = q_P + 5$.



a) *Scene 298* (MC I)

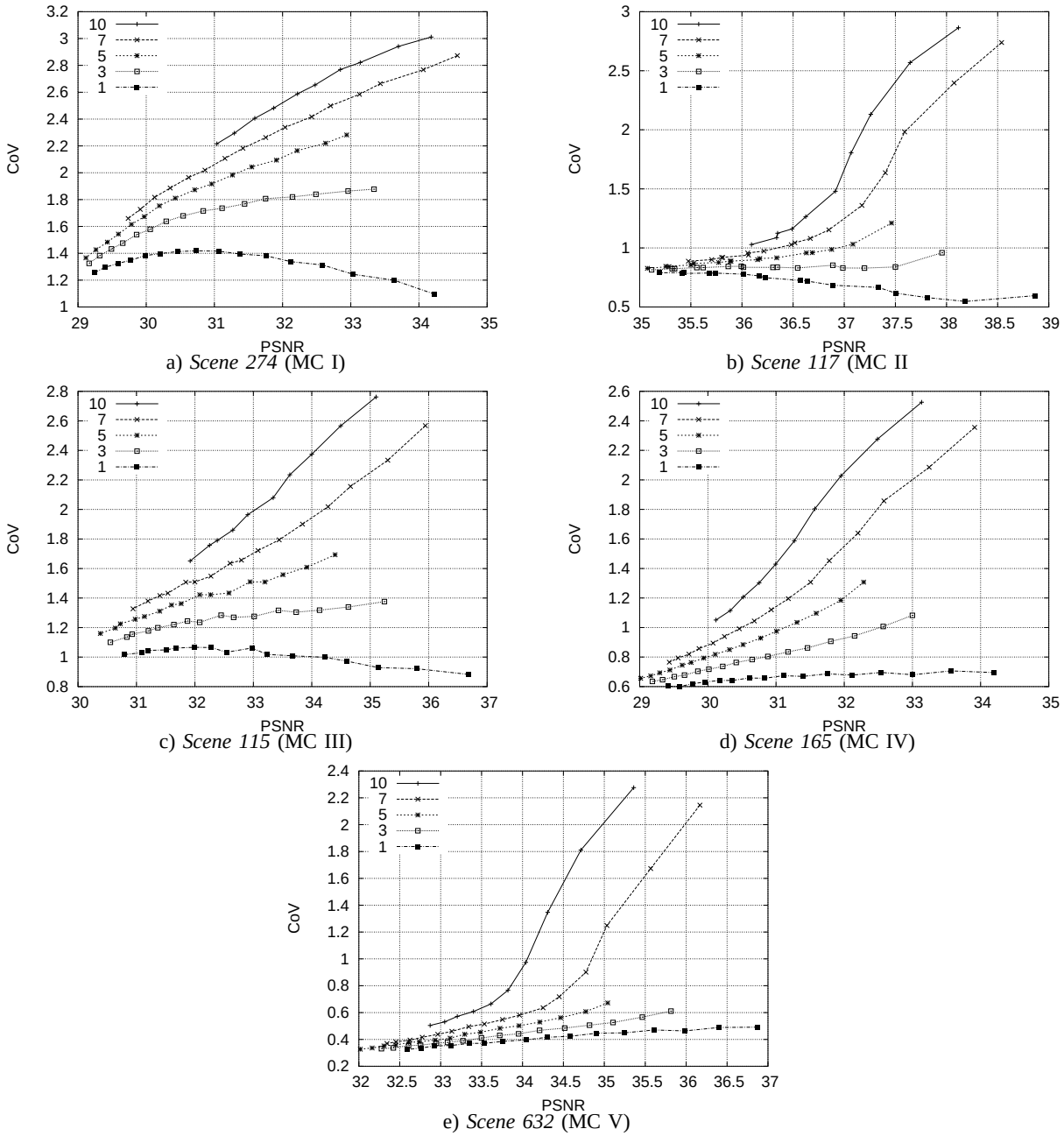
b) *Scene 299* (MC II

c) *Scene 557* (MC III)

d) *Scene 184* (MC IV)

e) *Scene 336* (MC V)

Fig. 23. Different magnitudes of differences between quantization scales for different frame types, $q_P = q_I + 1$, $q_P = q_I + 3$, $q_P = q_I + 5$, $q_P = q_I + 7$, and $q_P = q_I + 10$ for *Football*.



a) *Scene 274* (MC I)

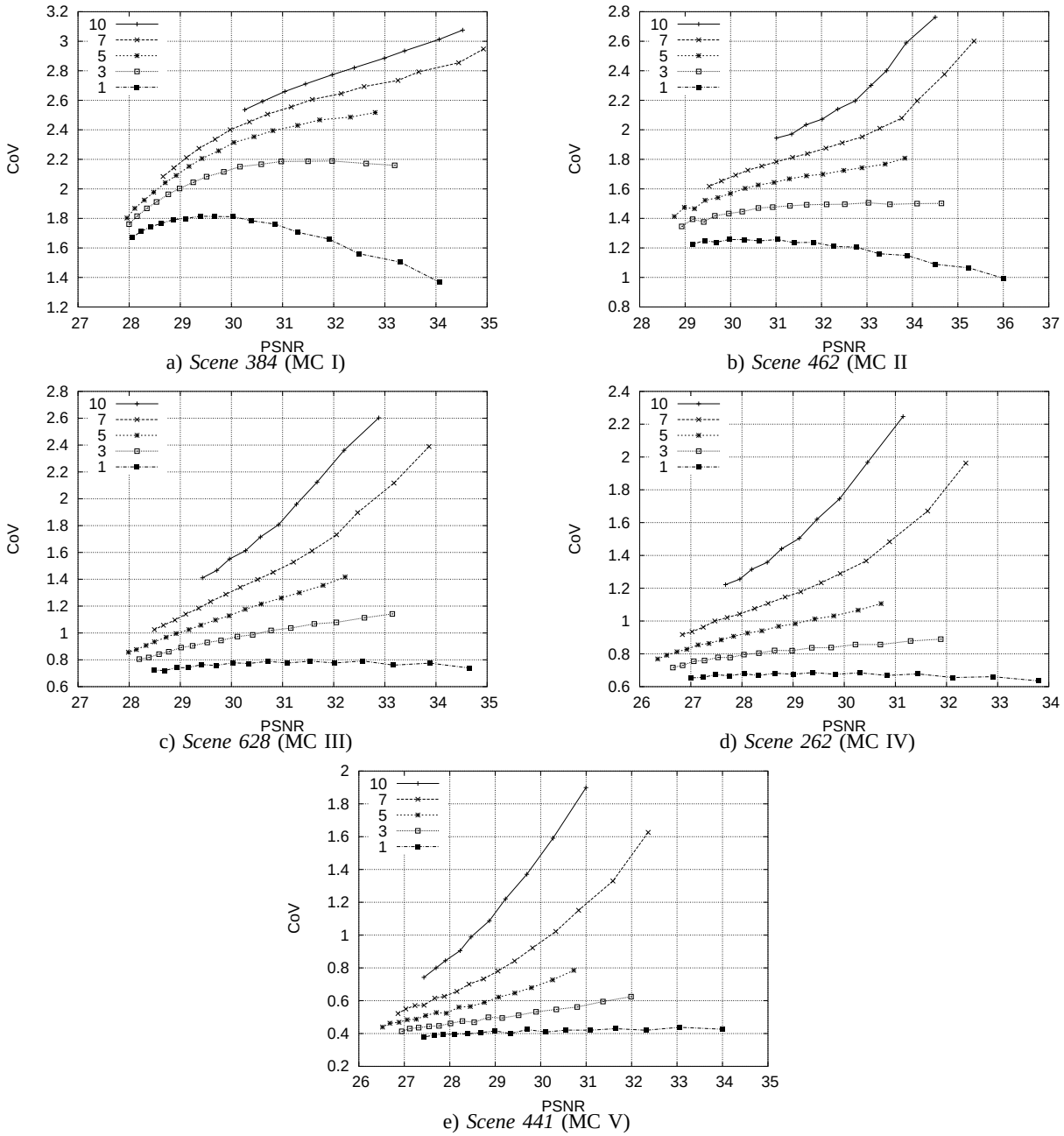
b) *Scene 117* (MC II

c) *Scene 115* (MC III)

d) *Scene 165* (MC IV)

e) *Scene 632* (MC V)

Fig. 24. Different magnitudes of differences between quantization scales for different frame types, $q_P = q_I + 1$, $q_P = q_I + 3$, $q_P = q_I + 5$, $q_P = q_I + 7$, and $q_P = q_I + 10$ for *Star Wars*.



a) *Scene 384* (MC I)

b) *Scene 462* (MC II

c) *Scene 628* (MC III)

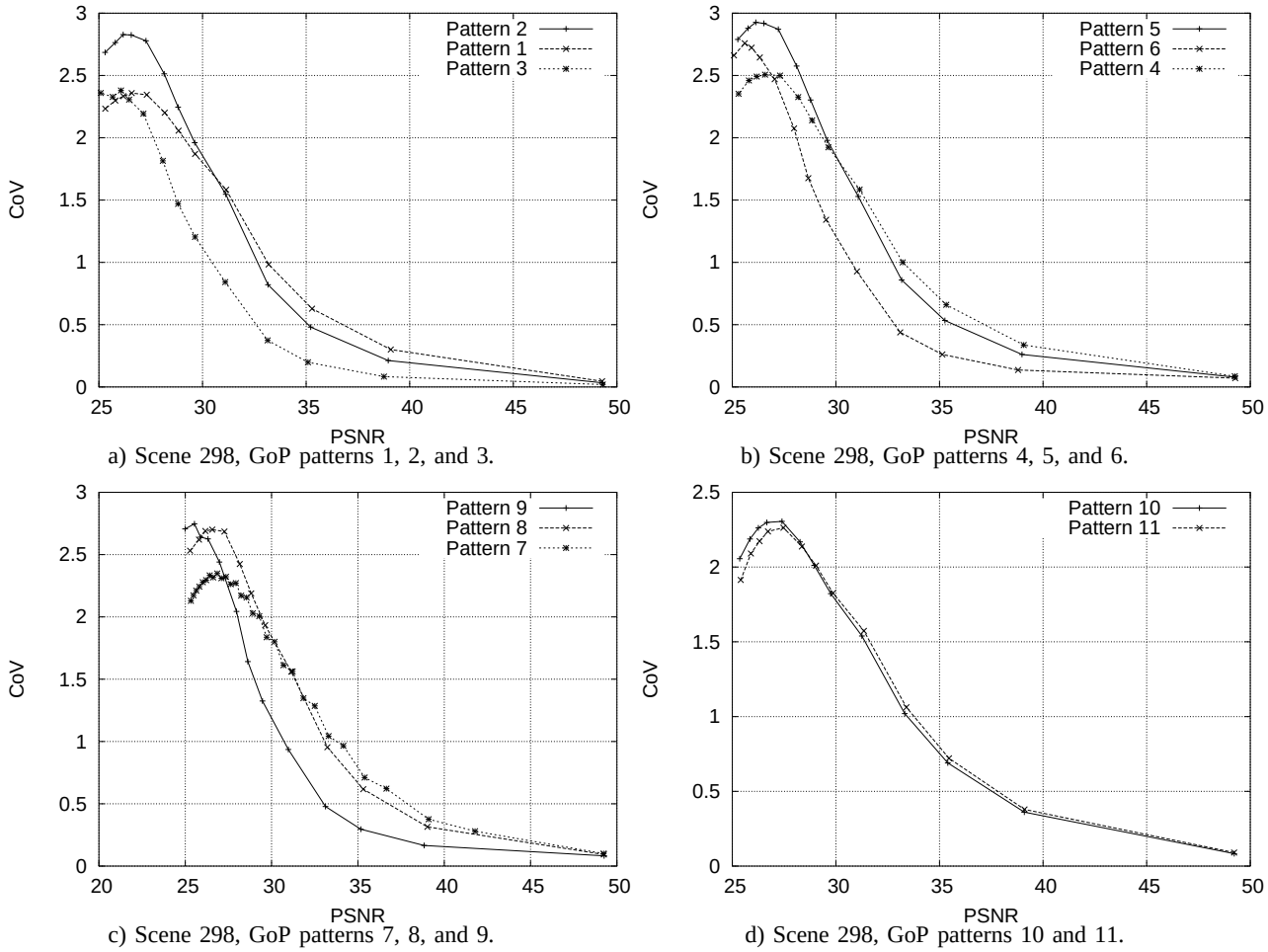
d) *Scene 262* (MC IV)

e) *Scene 441* (MC V)

Fig. 25. Different magnitudes of differences between quantization scales for different frame types, $q_P = q_I + 1$, $q_P = q_I + 3$, $q_P = q_I + 5$, $q_P = q_I + 7$, and $q_P = q_I + 10$ for *The Terminator*.

D. Impact of Different GoP Patterns on VD Characteristics

In this section we examine the impact of different GoP patterns on the VD curve. In Figs. 26, 27, and 28 we plot the VD curves for the motion classes I, III, and V for scenes from *Football* for the GoP patterns defined in Table II. In Figs. 29, 30, and 31 we plot the scenes with motion classes I, II, and III from *Star Wars IV*, and in Figs. 32, 33, and 34 for scenes from *The Terminator*. We employ identical quantization scales for all frame types. We observe that the VD curves for the medium motion activity scene from



a) Scene 298, GoP patterns 1, 2, and 3.

b) Scene 298, GoP patterns 4, 5, and 6.

c) Scene 298, GoP patterns 7, 8, and 9.

d) Scene 298, GoP patterns 10 and 11.

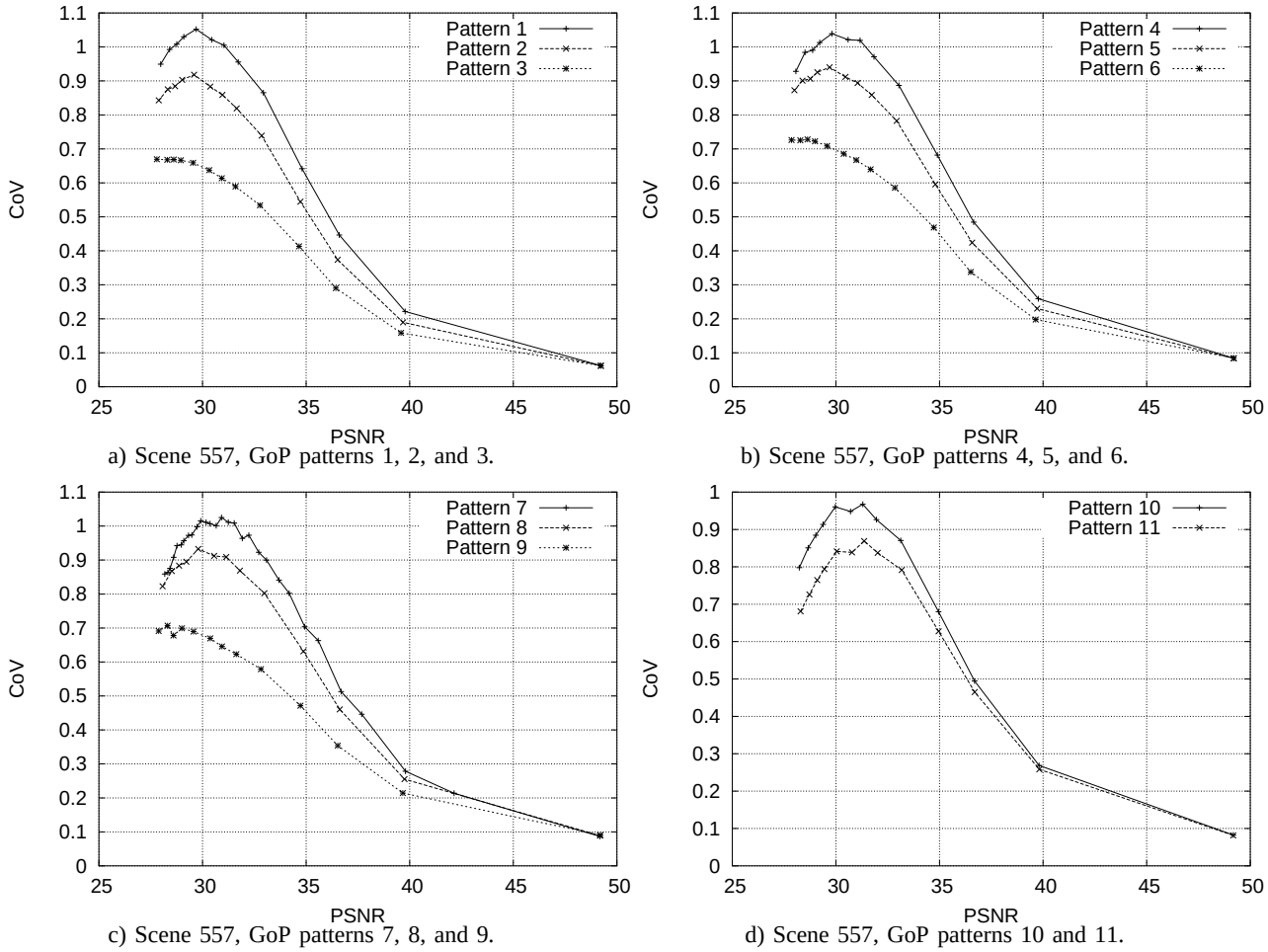
Fig. 26. Effect of different GoP patterns for motion class I scene from *Football*.

The Terminator are representative of the curves for the low and medium motion activity scenes. For the high motion scenes, the different GoP patterns have an overall negligible impact on the VD curves.

We observe from Fig. 33 that lengthening the GoP results in lower variability. This is primarily due to the decreased influence of the always large sized I frames, as the GoP includes more frames. By comparing the curves in Figs. 33 (a), (b), (c), and (d) we observe that introducing more B frames into the GoP pattern results in general in a minor reduction of the variability.

E. Effect of Frame Aggregation (Traffic Smoothing) on VD Characteristics

In this section we study the effect of frame aggregation, i.e., the smoothing of the video traffic over multiple frames as defined in (6). Note that all the preceding results were obtained for an aggregation



a) Scene 557, GoP patterns 1, 2, and 3.

b) Scene 557, GoP patterns 4, 5, and 6.

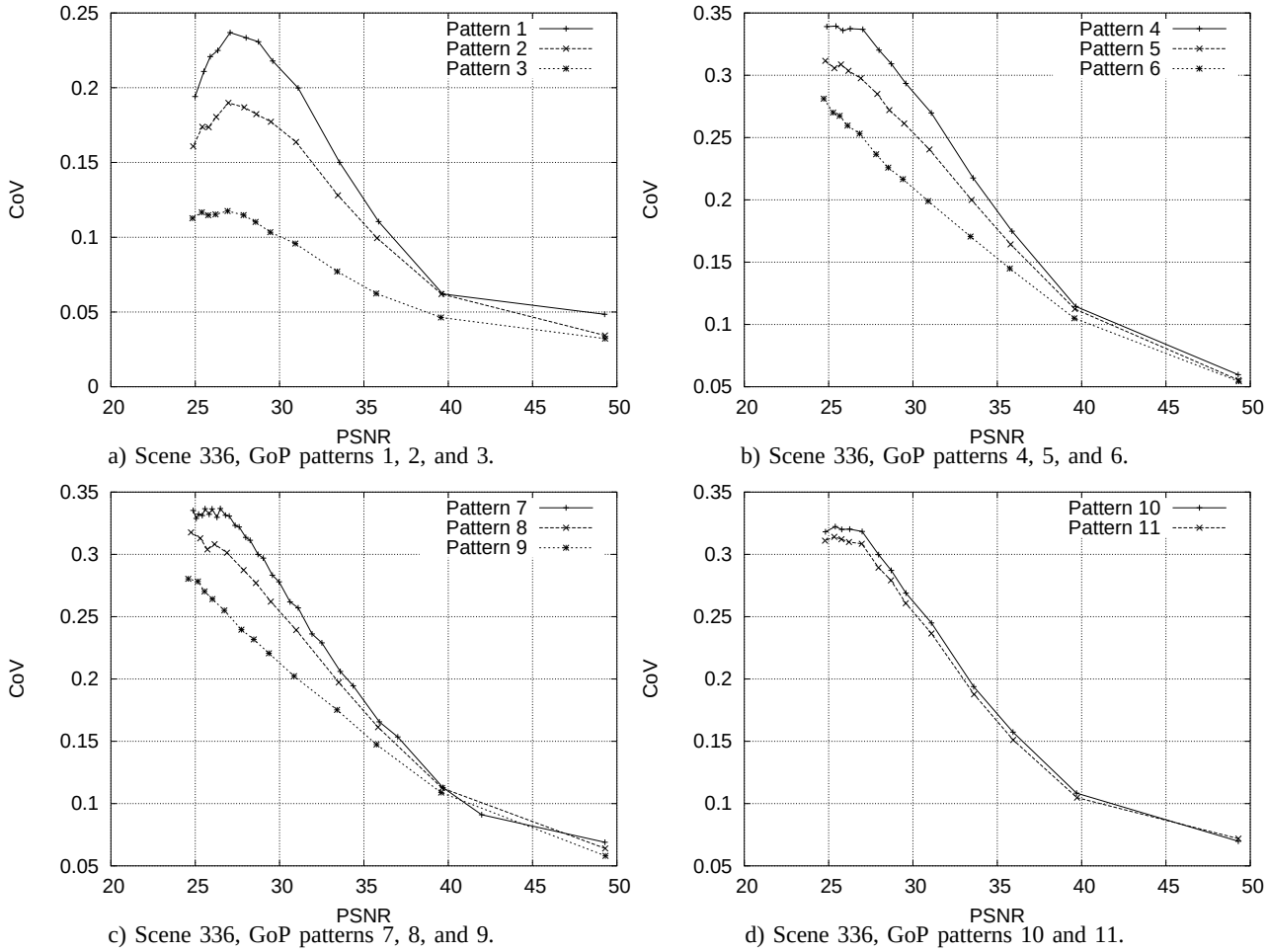
c) Scene 557, GoP patterns 7, 8, and 9.

d) Scene 557, GoP patterns 10 and 11.

Fig. 27. Effect of different GoP patterns for motion class III scene from *Football*.

level of $a = 1$, i.e., by considering individual frames. We now consider a range of aggregation levels a from 1 through 36. Throughout this section we consider GoP pattern 7 and apply the same quantization scale q to all frame types. In Fig. 35 we plot the VD curves for the MC I to MC V scenes from *Football* for different aggregation levels a. Similarly, we plot the VD curves for *Star Wars IV* and *The Terminator* in Figs. 36 and 37, respectively We observe that the variability is greatly reduced by smoothing the frames over an increasing number a of frames. With an increase from a = 1 to 3 frames, the differences between the I and P frames on the one hand, and the B frames on the other hand, are smoothed out. As a result the maximum variability is reduced from 1.633 to 0.853 for scene 384 and reduced from 0.373 to 0.186 for scene 441. Overall, our results indicate that smoothing over three frames approximately cuts the variability in half.

When the aggregation level is increased further to one GoP (i.e., a = 12 with the considered GoP pattern 7), we observe again a decrease in the variability to about one half compared to an aggregation of three frames. This is because the significant frame size differences between the I frames and the P frames are smoothed out by the aggregation. If the aggregation level a is further increased from a single GoP to multiple GoPs there is typically a smaller reduction in variability, which diminishes with



a) Scene 336, GoP patterns 1, 2, and 3.

b) Scene 336, GoP patterns 4, 5, and 6.

c) Scene 336, GoP patterns 7, 8, and 9.

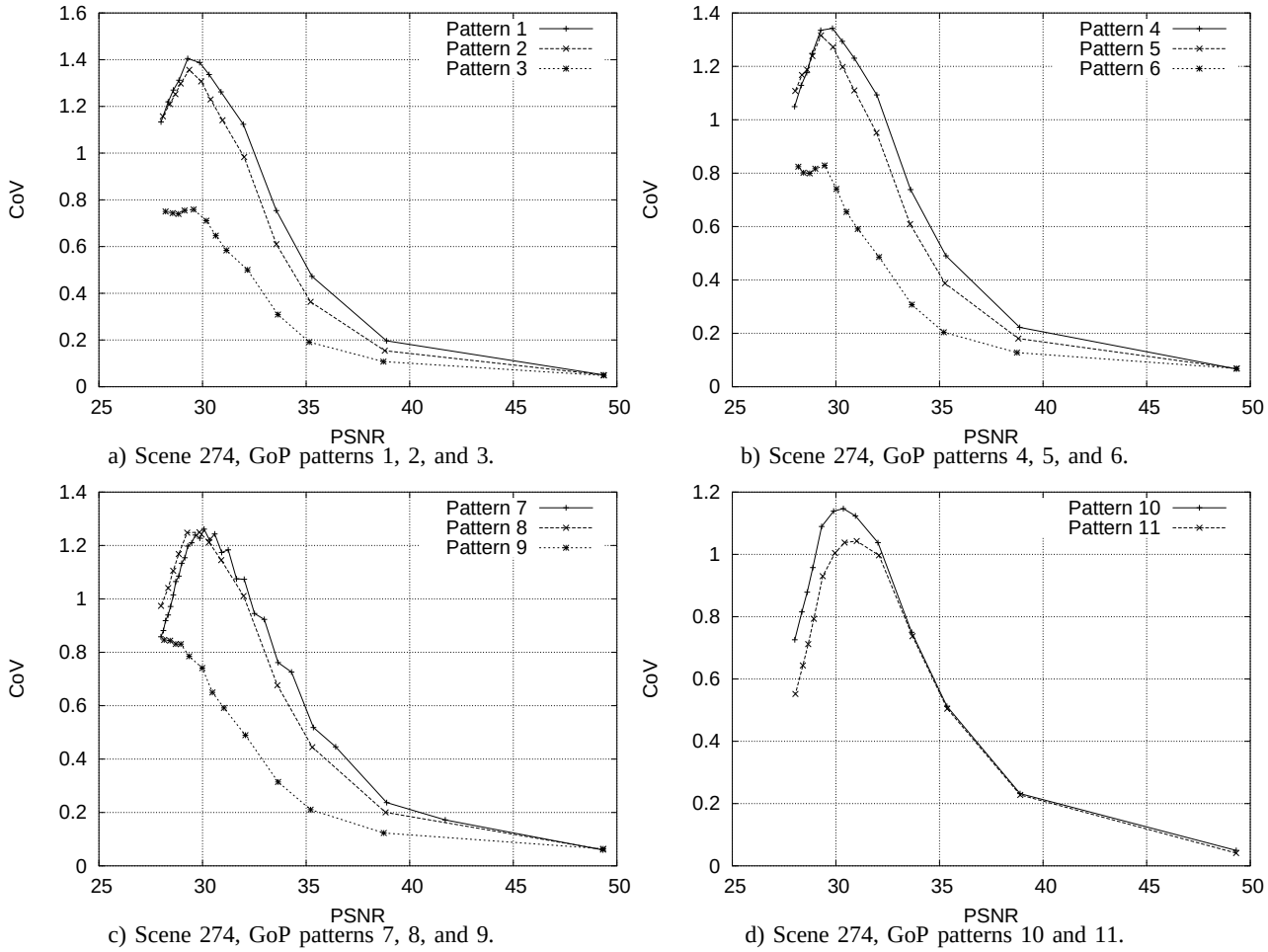
d) Scene 336, GoP patterns 10 and 11.

Fig. 28. Effect of different GoP patterns for motion class V scene from *Football*.

increasing aggregation level a. In summary, we find that the video traffic variability *within a given scene* is approximately cut in half by smoothing over 3 frames (i.e., the I or P frame and the two adjacent B frames in the considered GoP pattern 7). Smoothing over one GoP further cuts the variability roughly in half.

F. Effect of Scene Concatenation

In this section we examine the traffic variability across *multiple scenes*. Towards this end, we concatenate all five scenes from a given movie into one video segment and plot the VD curve for the thus obtained video segment for different frame aggregation levels a for the five scenes from *Football*, *Star Wars IV*, and *The Terminator* in Figs. 35, 36, and 37, respectively. We observe that the variability is reduced by smoothing over 3 frames. Smoothing over one GoP reduces the variability a little further. However, in contrast to the above results for individual scenes, the reduction in variability is relatively small for the video segment consisting of concatenated scenes. For an aggregation level of a = 24 frames, for instance, we observe from Fig. 37 that the individual scenes give a maximum variability smaller than 0.2, yet we observe from Fig. 44 that the maximum variability of the concatenated scenes is approximately 0.75. This high variability is mainly due to the variability of the frame sizes from scene to scene.



a) Scene 274, GoP patterns 1, 2, and 3.

b) Scene 274, GoP patterns 4, 5, and 6.

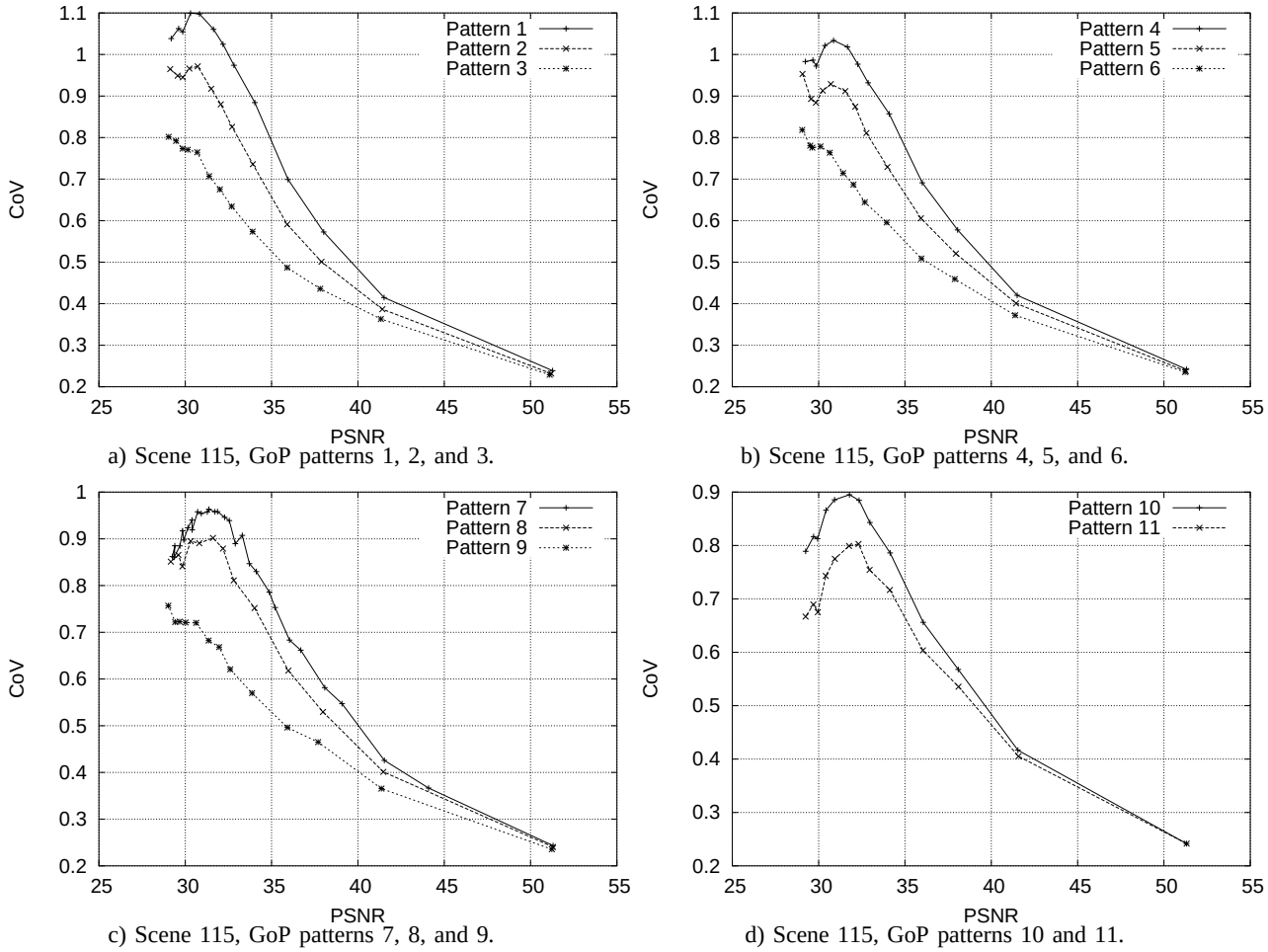
c) Scene 274, GoP patterns 7, 8, and 9.

d) Scene 274, GoP patterns 10 and 11.

Fig. 29. Effect of different GoP patterns for motion class I scene from *Star Wars IV*.

Interestingly, we also observe from the comparison of Fig. 44 with Fig. 37 that the VD curves for individual scenes flatten out as the aggregation level a increases, whereas the VD curve for the concatenation of scenes retains a pronounced hump and relatively steep slope even for large aggregation levels. This is due to the behaviors of the mean frame sizes and the standard deviation of the frames sizes which when evaluated over the segment of concatenated scenes give pronounced slopes of the VD curve (11).

Next, we consider the VD curves of the entire one-hour long excerpts from *Star Wars*, *The Terminator*, and *Football*, which we plot in Figs. 41, 42, and 43 for the frame aggregation level a = 1, the GoP aggregation level $a = 12$, and aggregation levels of multiple GoPs. We observe again the characteristic hump behavior of the VD curve, indicating that the behaviorial trends observed in Fig. 44 for the concatenation of 5 scenes extend to the concatenation of several hundred scenes in a 1-hour video. We observe also that the GoP smoothing tends to slightly shift the peak in the VD curve to higher quality levels. Overall, the GoP smoothing of the long videos roughly cuts the variability in half, compared to the individual scenes where GoP smoothing roughly cut the variability down to a quarter (cf. Fig. 37). This indicates that video traffic smoothing is quite effective in reducing the traffic variability within a given



a) Scene 115, GoP patterns 1, 2, and 3.

b) Scene 115, GoP patterns 4, 5, and 6.

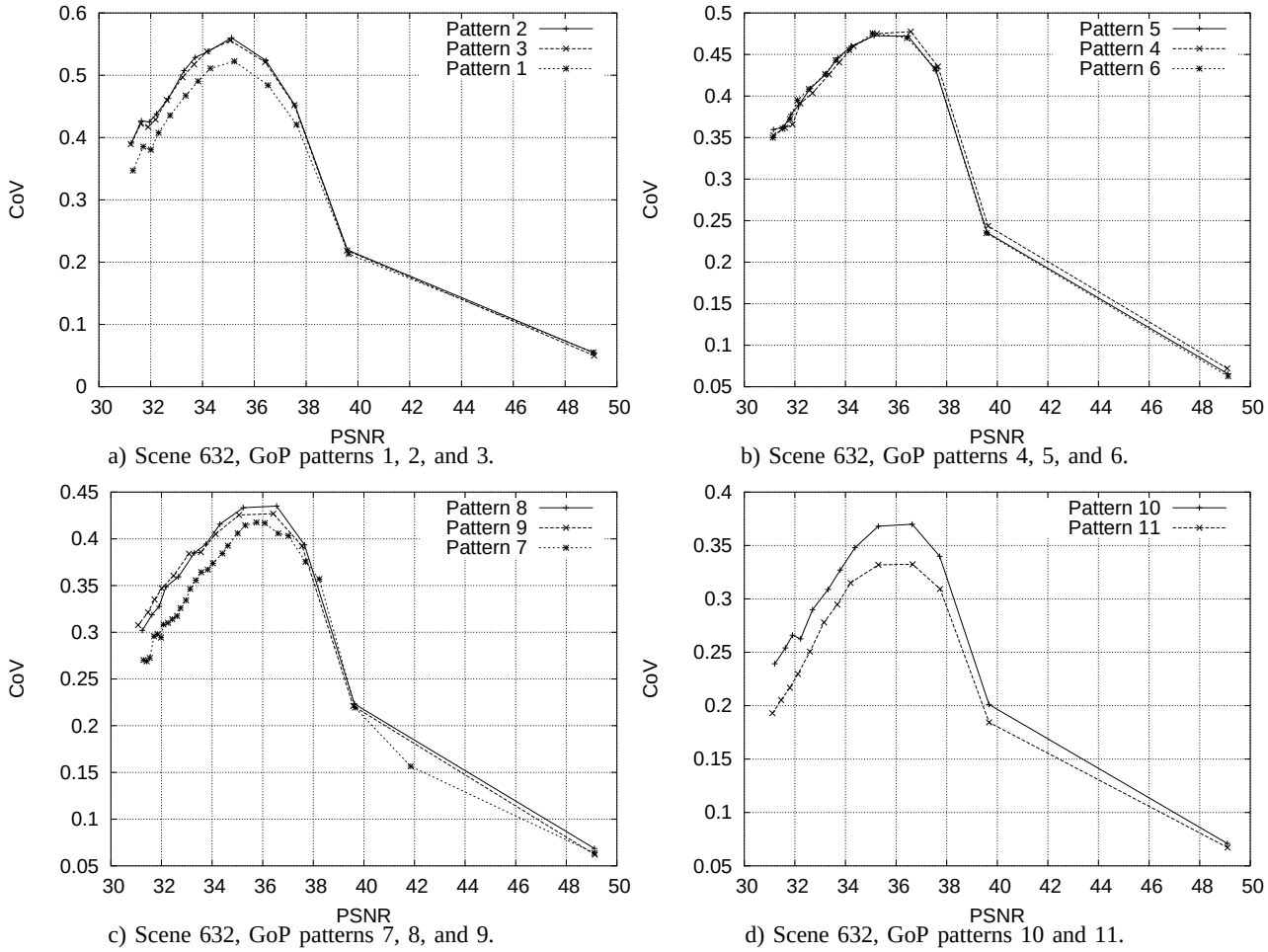
c) Scene 115, GoP patterns 7, 8, and 9.

d) Scene 115, GoP patterns 10 and 11.

Fig. 30. Effect of different GoP patterns for motion class III scene from *Star Wars IV*.

scene. On the other hand, traffic smoothing is less effective in reducing the traffic variability of a long video consisting of several scenes with distinct content characteristics. It appears hence beneficial to pay attention to the scene structure of the video when streaming smoothed video.

We finally examine the effect of considering PSNR values of the luminance as well as the chrominance components in assessing the video quality. We demonstrate that jointly considering luminance and chrominance PSNR values does not affect the overall tendencies and characteristic behaviors of the VD curve and the statistical multiplexing/utility. The only change is a scaling of the (horizontal) quality axis in the various plots. To illustrate this effect we plot in Figs. 44 and 45 the VD curve of the concatenated scenes from *The Terminator* and the entire video. We plot these VD curves using two different quality metrics on a common dB scale: (i) the PSNR values of the luminance component (denoted by `PSNR-Y` in the figures) as used in the paper, and (ii) a 4:1:1 weighing of the PSNR values of the luminance component and the two chrominance components (denoted by `PSNR-YUV` in the figures); the 4:1:1 weighing is inspired by the 4:1:1 chroma subsampling in the videos. We observe from the plots that independent of the considered quality metric, the VD curves exhibit the characteristic hump behavior. The different quality metrics only scale the curves in the horizontal direction. The specific scaling in the figures is due to the chrominance



a) Scene 632, GoP patterns 1, 2, and 3.

b) Scene 632, GoP patterns 4, 5, and 6.

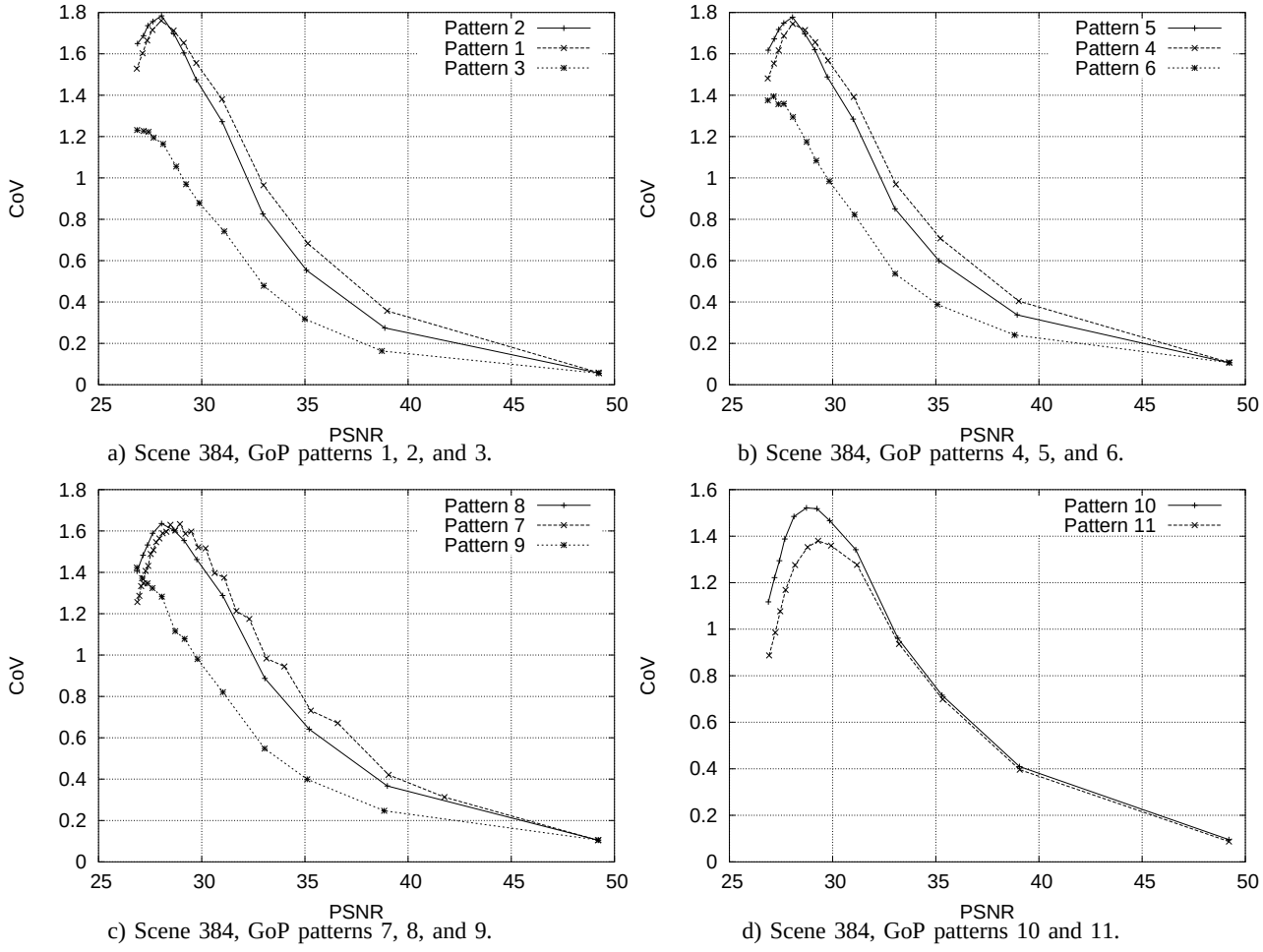
c) Scene 632, GoP patterns 7, 8, and 9.

d) Scene 632, GoP patterns 10 and 11.

Fig. 31. Effect of different GoP patterns for motion class V scene from *Star Wars IV*.

component having somewhat larger PSNR values than the luminance component for low quality levels, as the quality level increases this difference in the PSNR values diminishes.

Overall, we note that considering the PSNR values of the luminance component, as is common in video studies, captures the main effects and characteristics of the VD curve and the resulting statistical multiplexing/utility behaviors and is therefore a reasonable choice for our study.



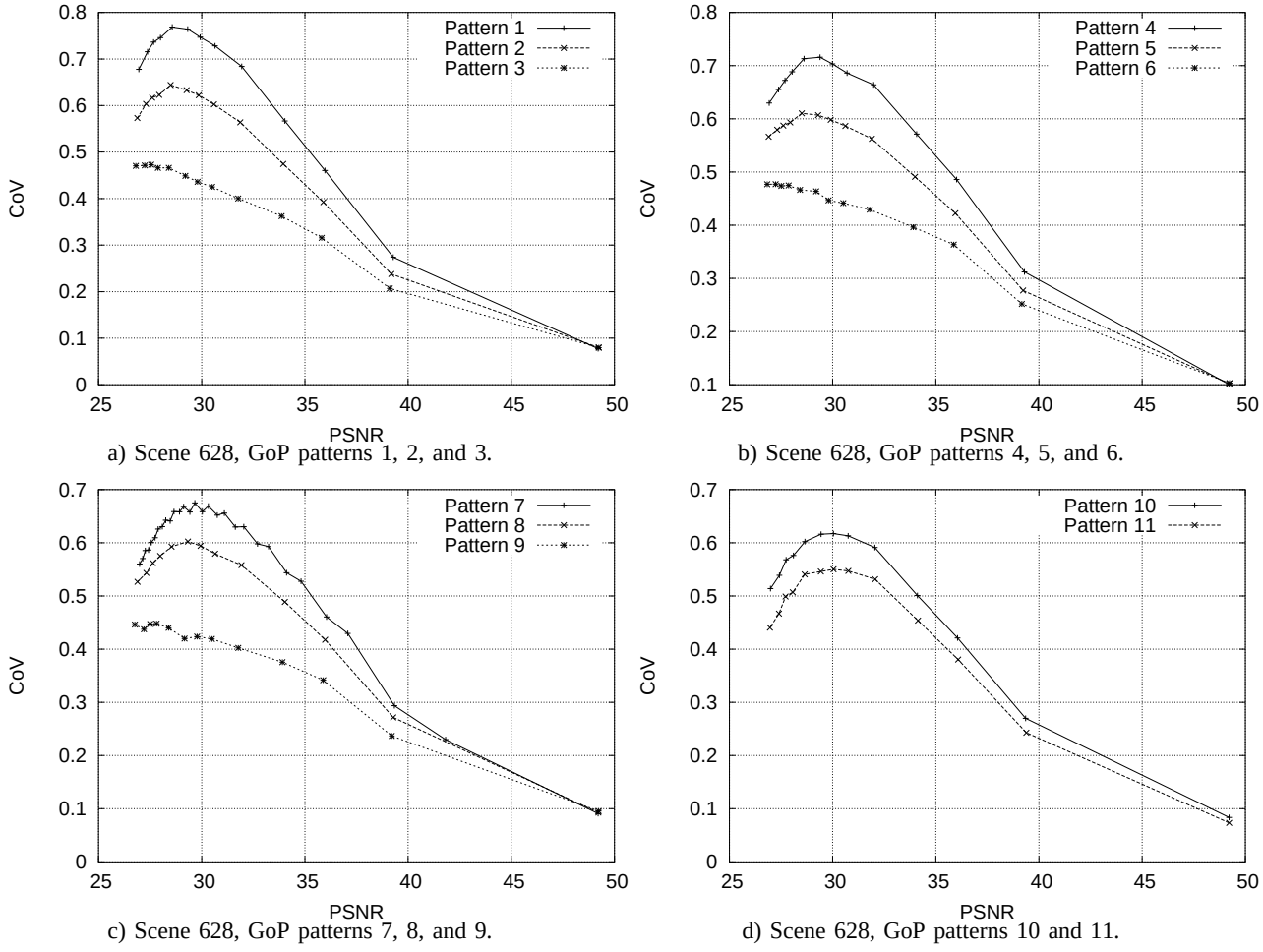
a) Scene 384, GoP patterns 1, 2, and 3.

b) Scene 384, GoP patterns 4, 5, and 6.

c) Scene 384, GoP patterns 7, 8, and 9.

d) Scene 384, GoP patterns 10 and 11.

Fig. 32. Effect of different GoP patterns for motion class I scene from *The Terminator*.



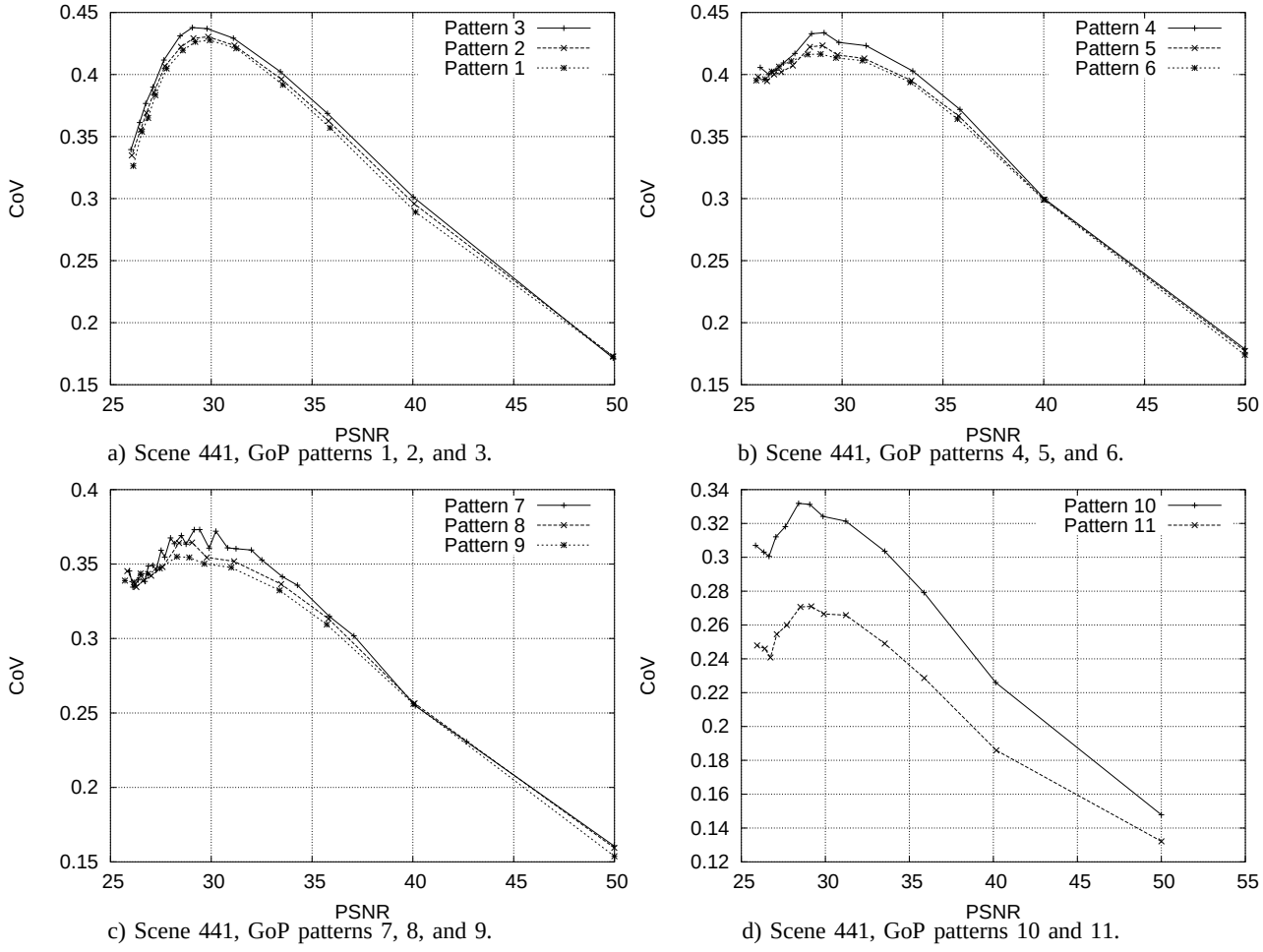
a) Scene 628, GoP patterns 1, 2, and 3.

b) Scene 628, GoP patterns 4, 5, and 6.

c) Scene 628, GoP patterns 7, 8, and 9.

d) Scene 628, GoP patterns 10 and 11.

Fig. 33. Effect of different GoP patterns for motion class III scene from *The Terminator*.



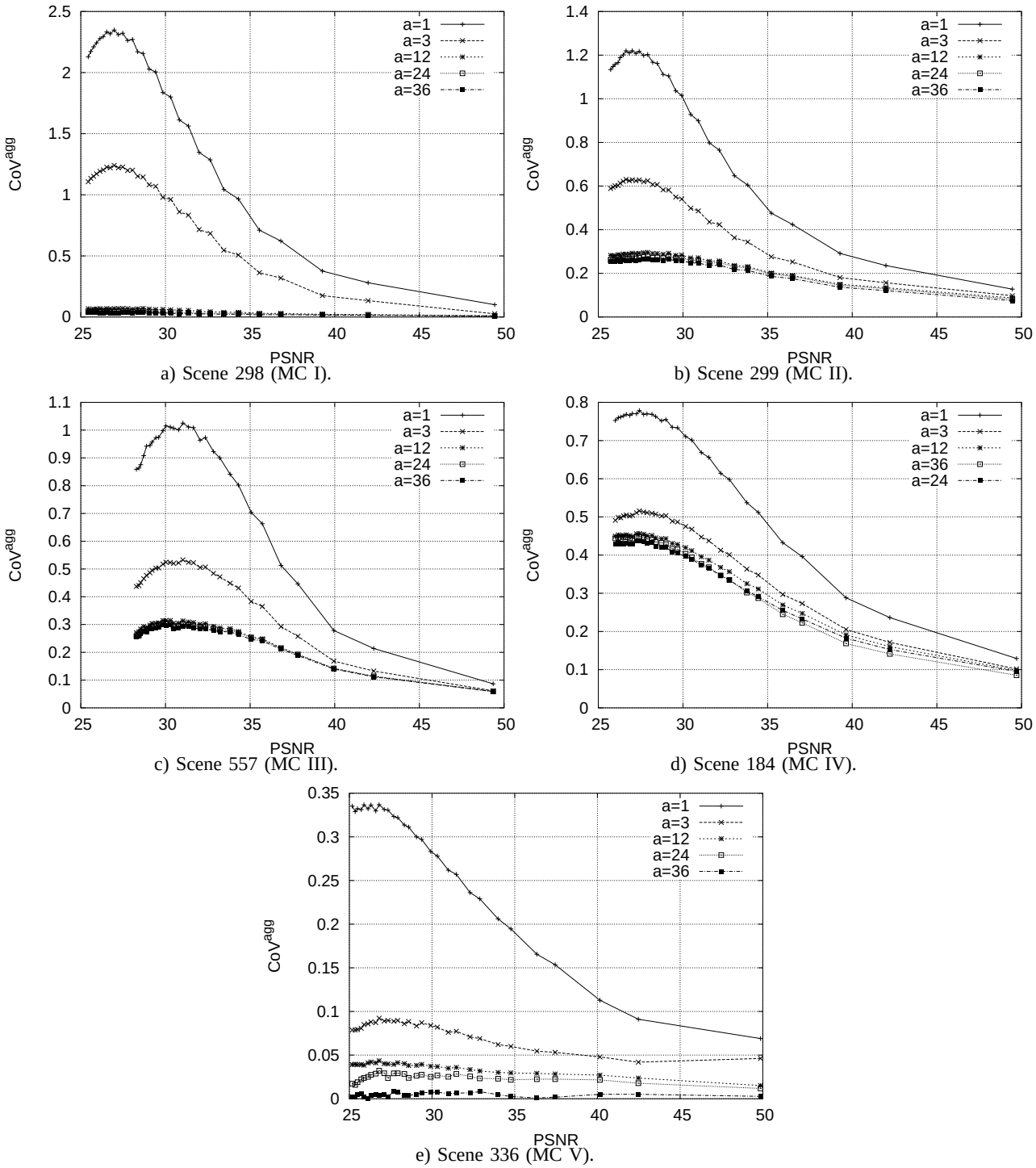
a) Scene 441, GoP patterns 1, 2, and 3.

b) Scene 441, GoP patterns 4, 5, and 6.

c) Scene 441, GoP patterns 7, 8, and 9.

d) Scene 441, GoP patterns 10 and 11.

Fig. 34. Effect of different GoP patterns for motion class V scene from *The Terminator*.



a) Scene 298 (MC I).

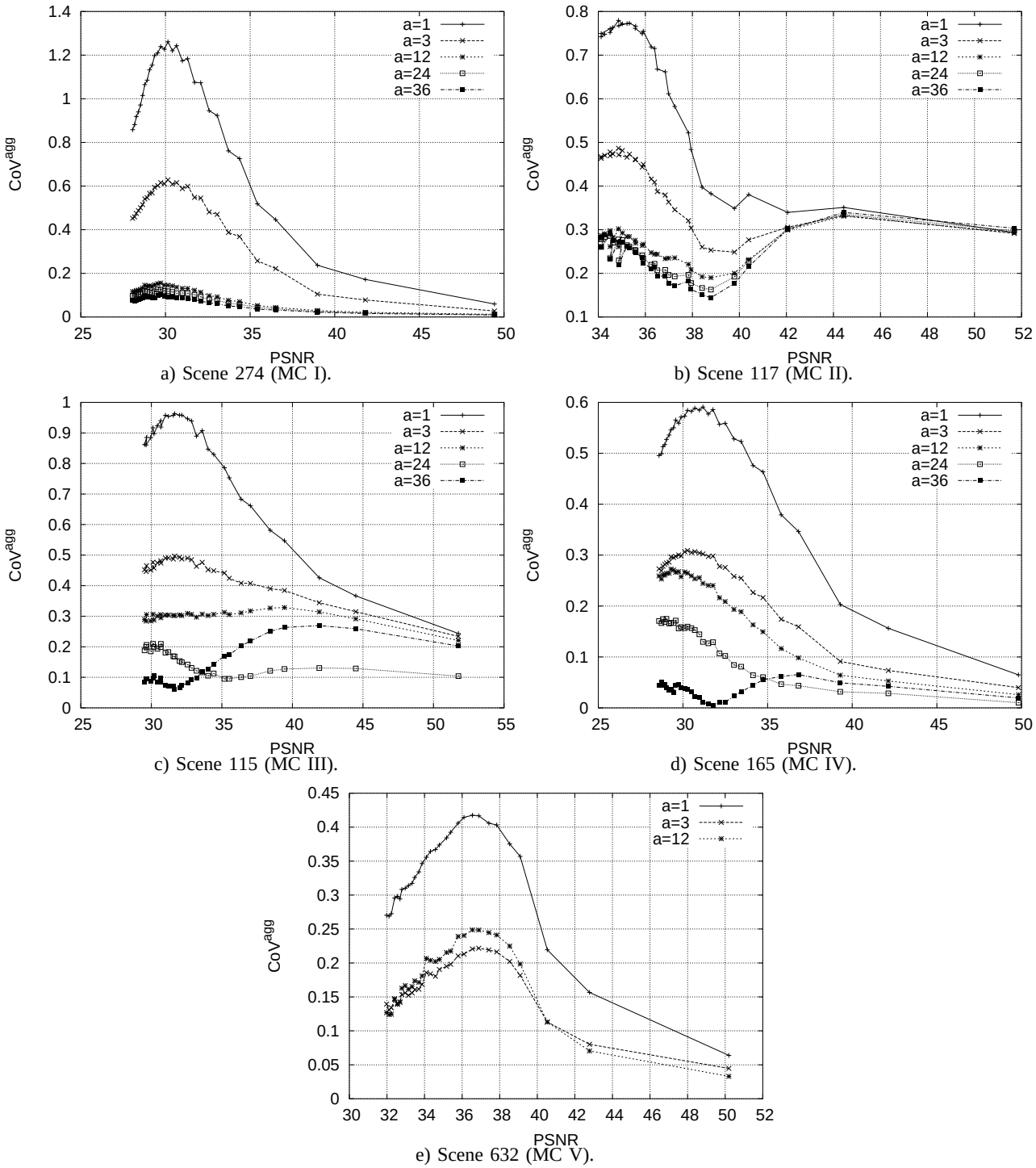
b) Scene 299 (MC II).

c) Scene 557 (MC III).

d) Scene 184 (MC IV).

e) Scene 336 (MC V).

Fig. 35. Effect of frame aggregation (smoothing) on CoV_q for *Football* scenes.



a) Scene 274 (MC I).

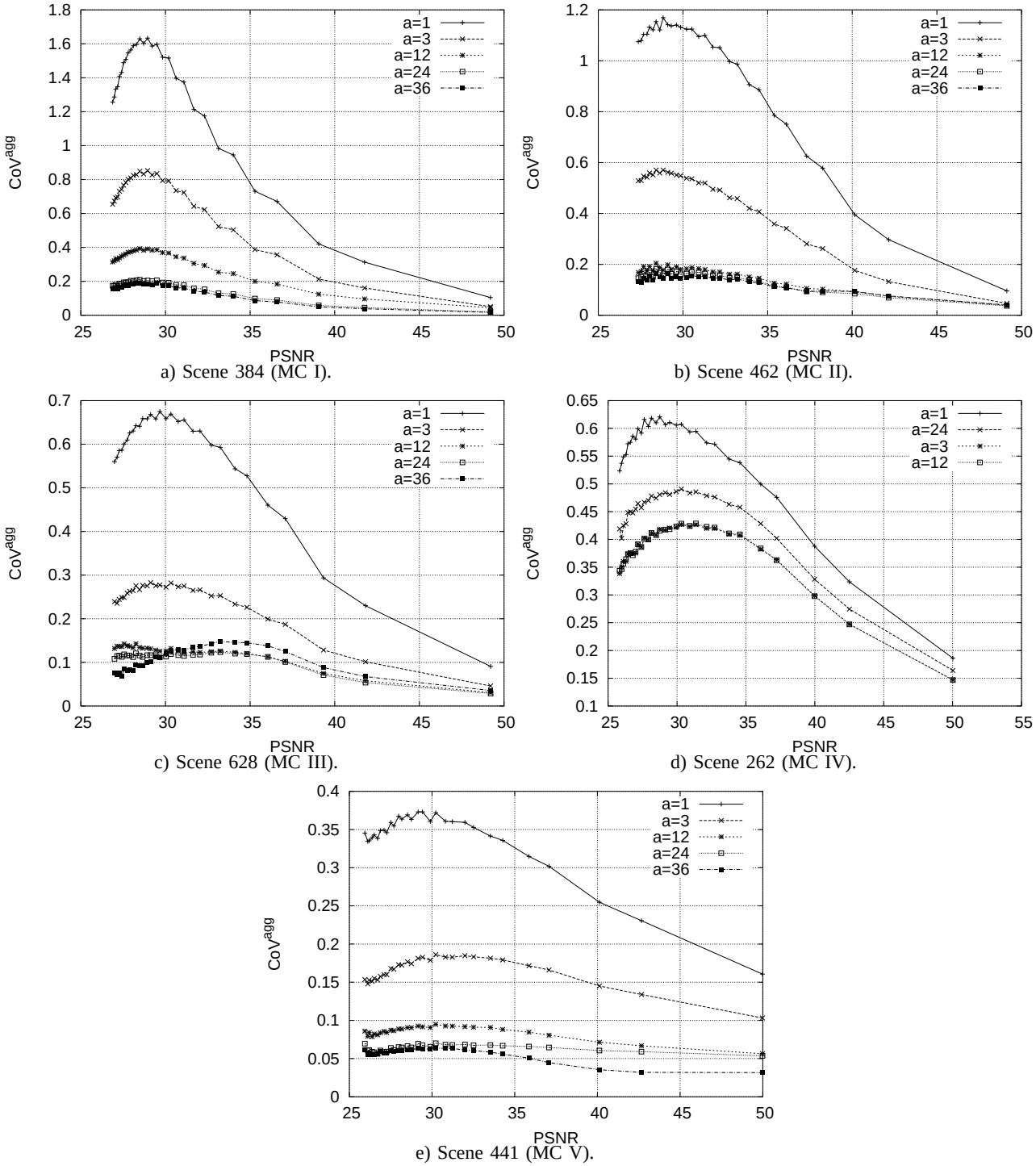
b) Scene 117 (MC II).

c) Scene 115 (MC III).

d) Scene 165 (MC IV).

e) Scene 632 (MC V).

Fig. 36. Effect of frame aggregation (smoothing) on CoV_q for *Star Wars IV* scenes.



a) Scene 384 (MC I).

b) Scene 462 (MC II).

c) Scene 628 (MC III).

d) Scene 262 (MC IV).

e) Scene 441 (MC V).

Fig. 37. Effect of frame aggregation (smoothing) on CoV_q for *The Terminator* scenes.

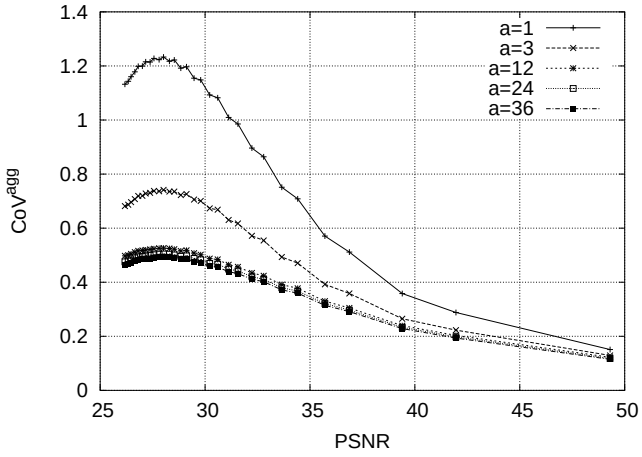


Fig. 38. Effect of frame aggregation (smoothing) on CoV_q for concatenated scenes (298, 299, 557, 184, and 336) from *Football*.

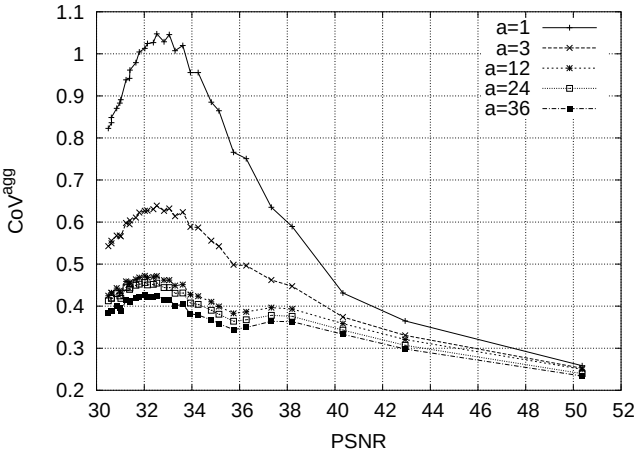


Fig. 39. Effect of frame aggregation (smoothing) on CoV_q for concatenated scenes (274, 117, 115, 165, and 632) from *Star Wars IV*.

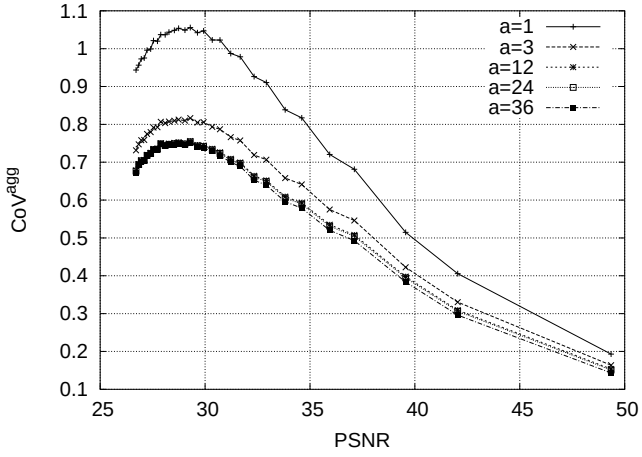


Fig. 40. Effect of frame aggregation (smoothing) on CoV_q for concatenated scenes (384, 462, 628, 262, and 441) from *The Terminator*.

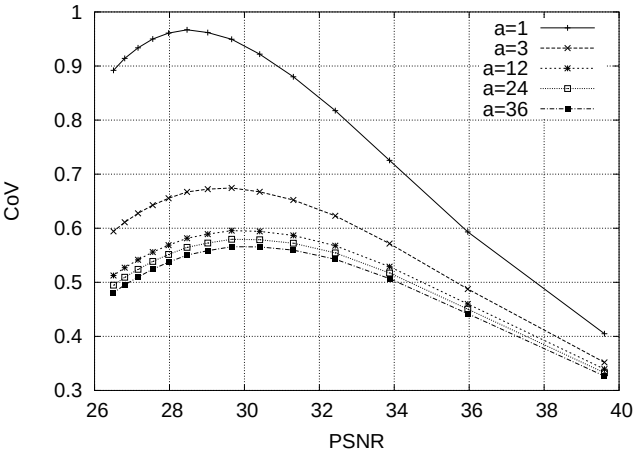


Fig. 41. CoV_q for entire video *Football*.

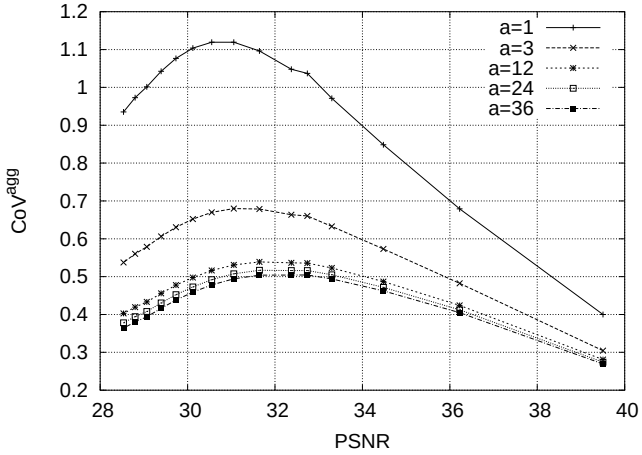


Fig. 42. CoV_q for entire video *Star Wars IV*.

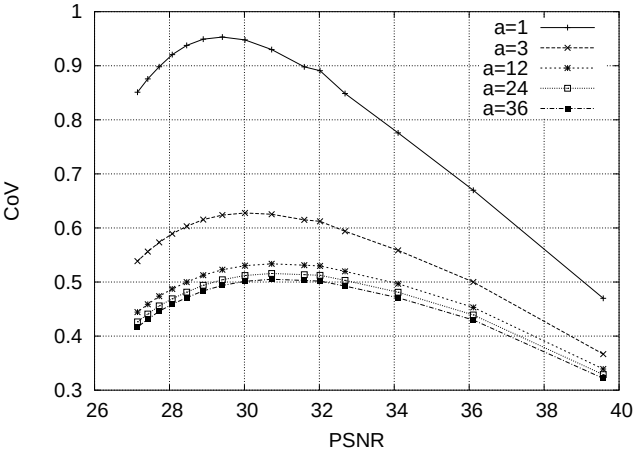


Fig. 43. CoV_q for entire video *The Terminator*.

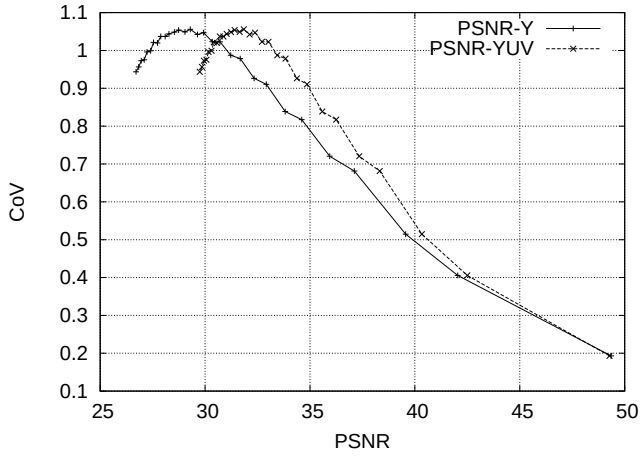


Fig. 44. Effect of considering different quality metrics on CoV_q for concatenated scenes (384, 462, 628, 262, and 441) from *The Terminator*.

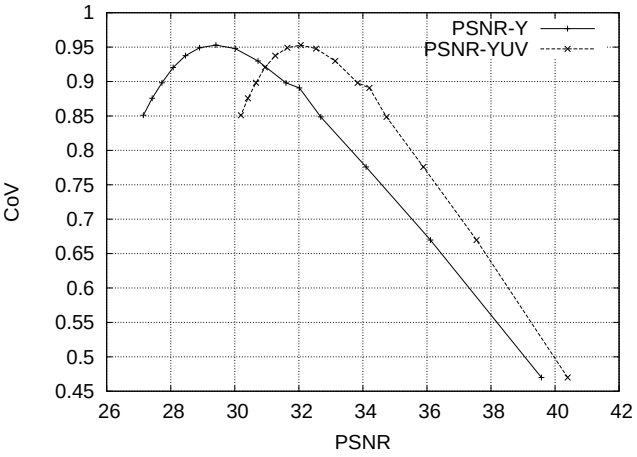


Fig. 45. CoV_q for entire *The Terminator* video.

V. VD Approximation Framework

In this section, we introduce a piecewise approximation framework for the VD curve. The purpose of the piecewise approximation is to estimate the VD curve for the entire range of quantization scales from sample encodings for a small number of quantization scales.

A. Piecewise Approximation Model

Our method employs piecewise power curve fitting for the estimation of the relationships between (i) size of encoded video frame and quantization scale, (ii) standard deviation of size of encoded video frame and quantization scale, and (iii) PSNR of encoded and subsequently decoded frame and quantization scale. Given the estimation of these three relationships, the approximated VD curve can be calculated. Our power curve fitting method is inspired by [7], [8], where similar techniques are employed to approximate rate-distortion curves.

Let q denote a quantization scale setting for which the encoded video data is available (i.e., interpolation points or samples) and $\tilde{q}$ denote the quantization scale settings for which we use the approximation framework to estimate the values needed for the calculation of the VD values (i.e., intermediate points). The power curve representation of an arbitrary function $f(q)$ (e.g., mean, standard deviation, or PSNR) for a quantization scale setting q is given as

$$f(q) = \alpha \cdot q^{\beta}. \tag{13}$$

Between a set of two consecutive interpolation points q_i and q_j we calculate the parameters α and β as

$$\alpha_{i,j} = \frac{f(q_i)}{q_i^{\beta}} \quad \text{and } \beta_{i,j} = \log_{\frac{q_i}{q_j}} \frac{f(q_i)}{f(q_j)}. \tag{14}$$

The estimated values between the interpolation points q_i and q_j are then given as

$$f(\tilde{q}) = \alpha_{i,j} \cdot (\tilde{q})^{\beta_{i,j}}, \quad \text{where } q_i < \tilde{q} < q_j. \tag{15}$$

We consider three different levels of estimation. On the *aggregated* level, we estimate the mean frame size $\overline{X}_{\tilde{q}}$, standard deviation $\sigma_{\tilde{q}}$, and mean PSNR value $Q_{\tilde{q}}$ from the given interpolation points. On the *frame* level, we first estimate the individual frame sizes $X_n^{\tilde{q}}$ and the corresponding frame qualities $Q_n^{\tilde{q}}$ for the intermediate quantization scale settings $\tilde{q}$. From the thus estimated frame size and quality values we calculate the (estimated) mean, standard deviation, and PSNR values as given in Eqs. (1), (2), and (8) for each q and intermediate $\tilde{q}$. Thirdly, we consider the *macroblock* level, where we estimate the size of each individual macroblock $i,\ i = 0, \ldots, M-1$, of frame n as $y_{n,i}^{\tilde{q}}$. We estimate the frame overhead of frame n as $\delta_n^{\tilde{q}}$, with

$$\delta_n^q = X_n^q - \sum_{i=0}^{M-1} y_{n,i}^q \tag{16}$$

for the given interpolation points. The frame sizes of the *macroblock* level are then calculated as

$$X_n^{\tilde{q}} = \sum_{i=0}^{M-1} y_{n,i}^{\tilde{q}} + \delta_n^{\tilde{q}}. \tag{17}$$

From these frame sizes we determine the coefficient of variation as in the two previous levels, but use the frame level quality estimation for the VD curve estimation.

Intuitively, the accuracy of the piecewise power curve fitting approach varies with the number of available interpolation points, i.e., more available encodings yield better approximations. We follow the reasoning in [8] and assume in the following that extremely high or extremely low quantization scale settings, i.e., $q < 5$ or $q > 25$, will most likely not be needed for general applications. We subsequently focus in our evaluation only on the region of $5 \leq q \leq 25$. We consider three different numbers of available interpolation points as outlined in Table V and examine their suitability for an approximation of the VD

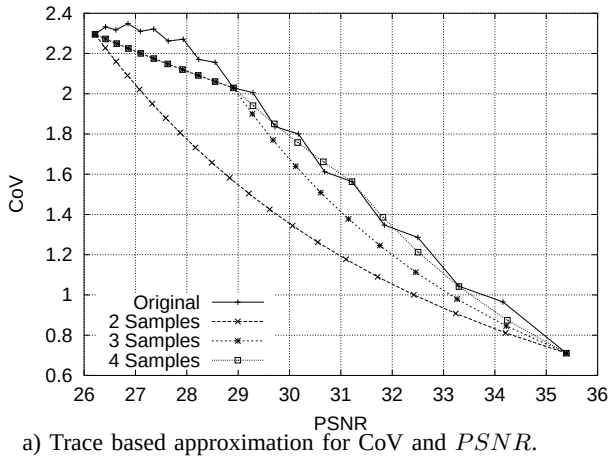
TABLE V
INTERPOLATION POINTS FOR THE PIECEWISE POWER CURVE ESTIMATION.

Samples	Interpolation points (q_i)
2	$q_1 = 5$, $q_2 = 25$
3	$q_1 = 5$, $q_2 = 15$, $q_3 = 25$
4	$q_1 = 5$, $q_2 = 10$, $q_3 = 17$, $q_4 = 25$

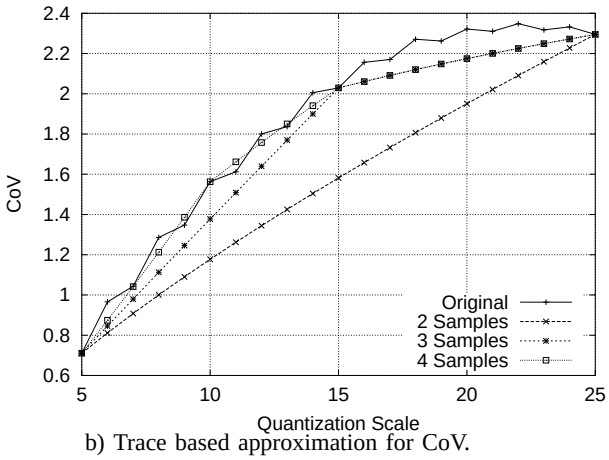
curve. We base our evaluation on the common GoP structure (i.e., pattern 7 in Table II) and uniform quantization scale settings for all frame types.

B. VD Approximation Results

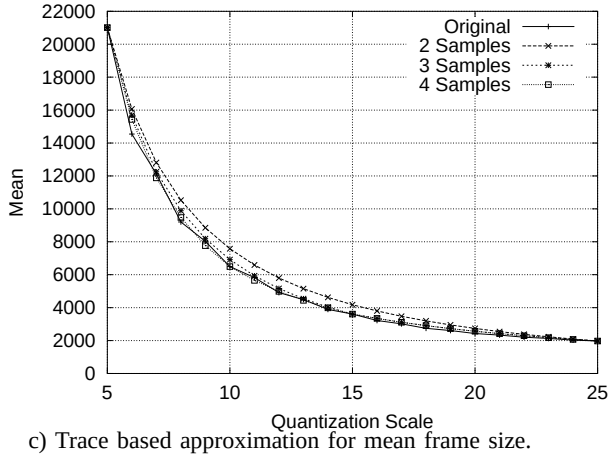
In this section we present the VD approximation results for the individual scenes from *Football*, *Star Wars IV*, and *The Terminator* for all three scenarios specified in Table V. In the following, we illustrate the approximation for (i) the coefficient of variation as function of the approximated quality, (i) the coefficient of variation as function of the quantization scale, (iii) the mean frame size, (iv) the frame size variance, and (v) the mean frame quality as $PSNR$ for the *aggregated* level. In Figs. 46, 47, 48, 49, and 50 the approximation results for the scenes from *Football* are illustrated. Similarly, the results for *Star Wars IV* and *The Terminator* are illustrated in Figs. 51, 52, 53, 54, 55 and Figs. 56, 57, 58, 59, 60.



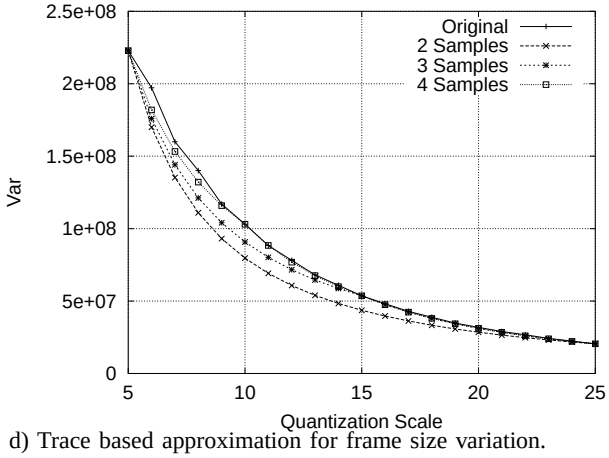
a) Trace based approximation for CoV and $PSNR$.



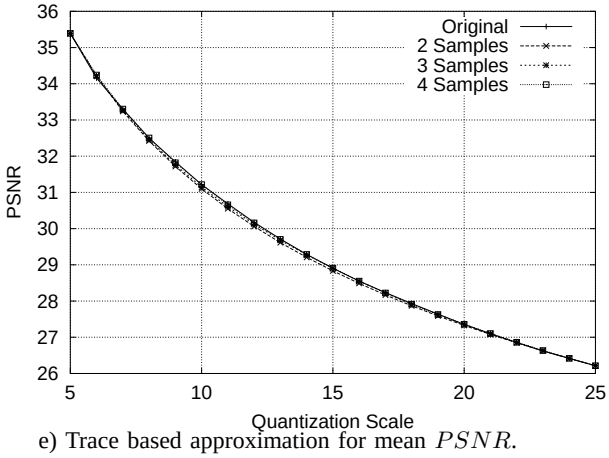
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

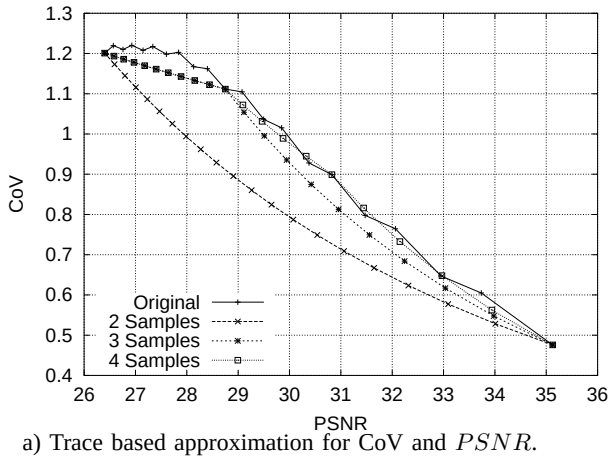


d) Trace based approximation for frame size variation.

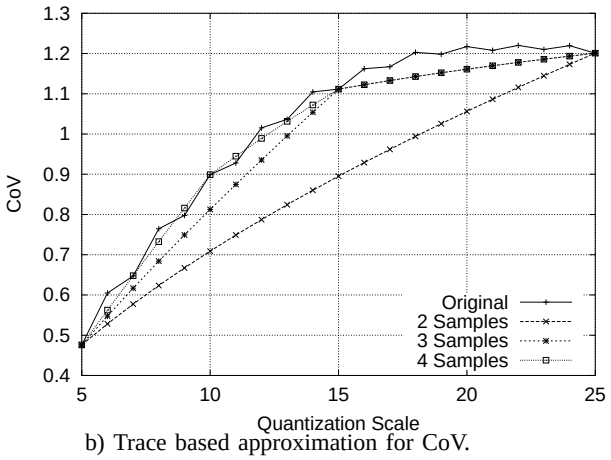


e) Trace based approximation for mean $PSNR$.

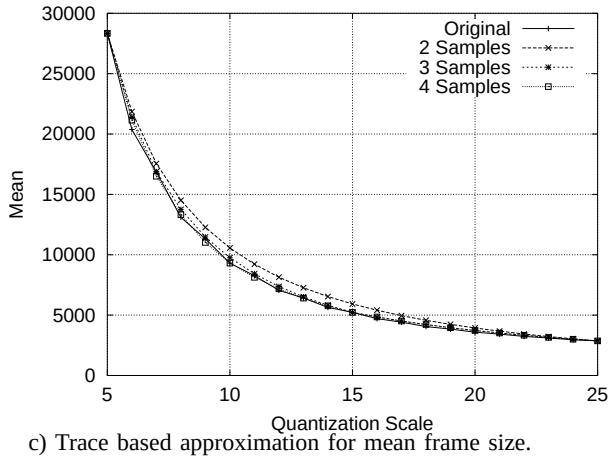
Fig. 46. *Aggregated* level approximation results for scene *298* (MC I) from *Football* .



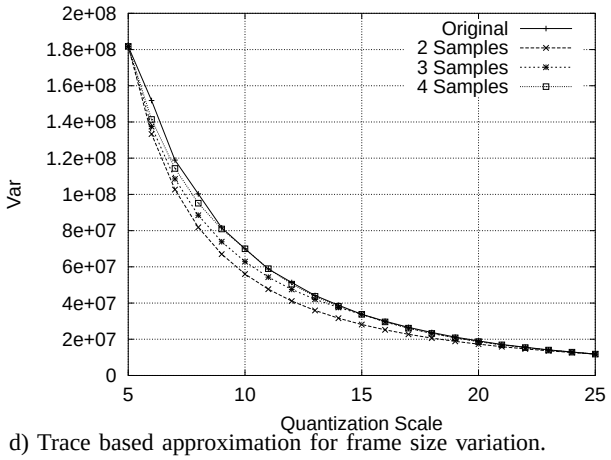
a) Trace based approximation for CoV and $PSNR$.



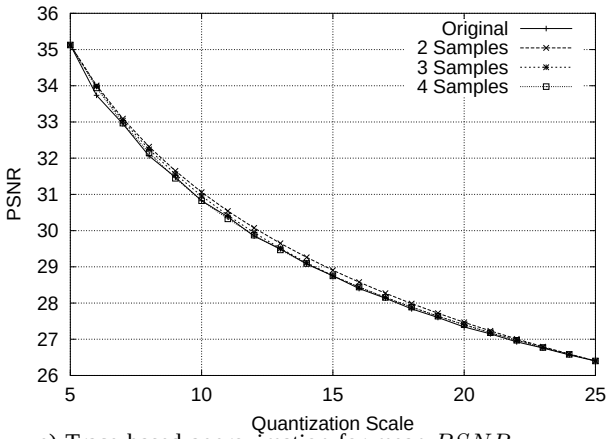
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

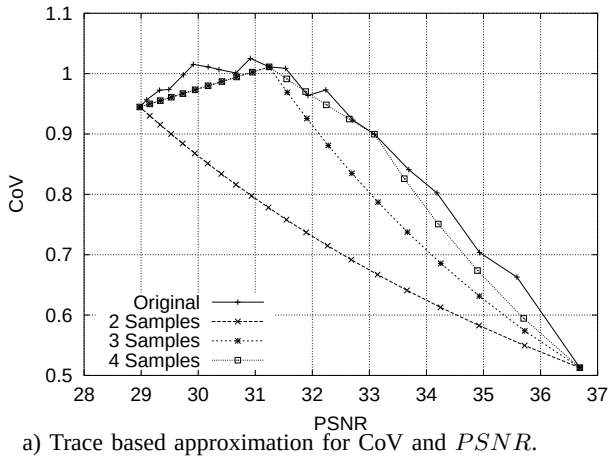


d) Trace based approximation for frame size variation.

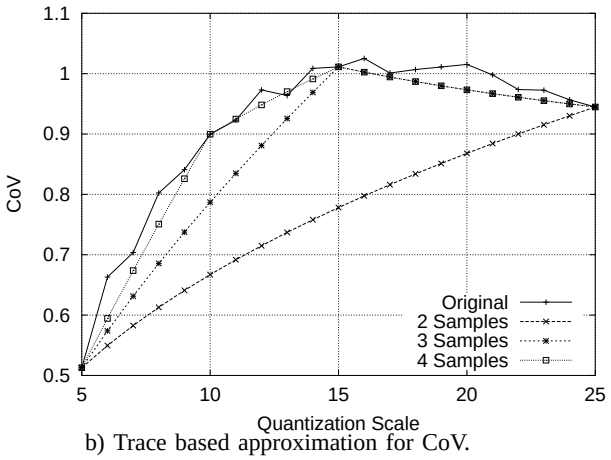


e) Trace based approximation for mean $PSNR$.

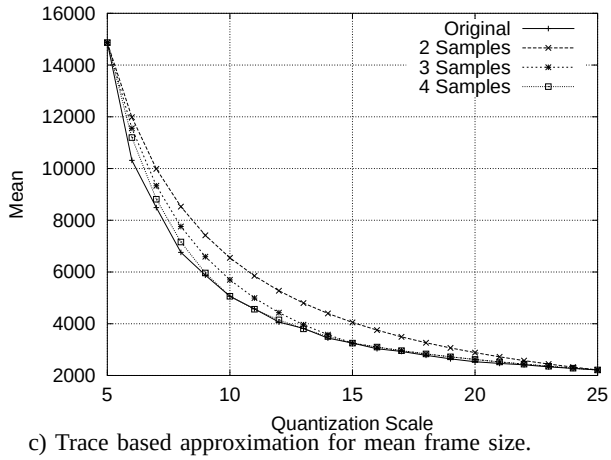
Fig. 47. *Aggregated* level approximation results for scene *299* (MC II) from *Football* .



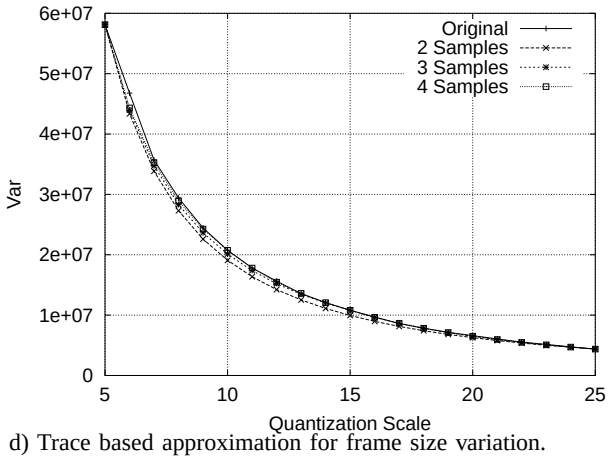
a) Trace based approximation for CoV and $PSNR$.



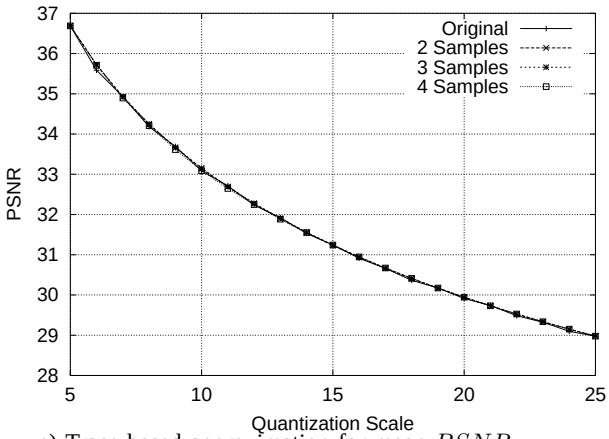
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

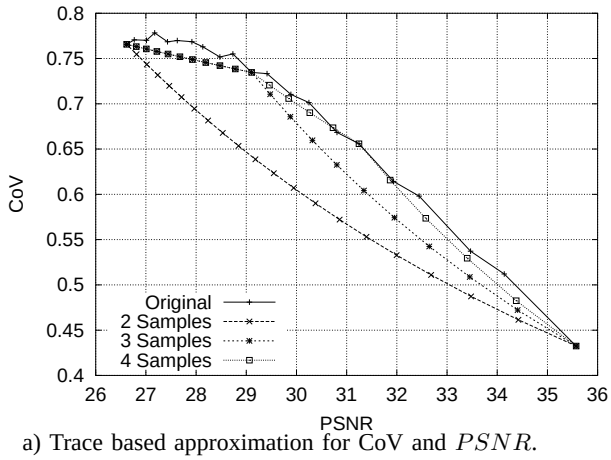


d) Trace based approximation for frame size variation.

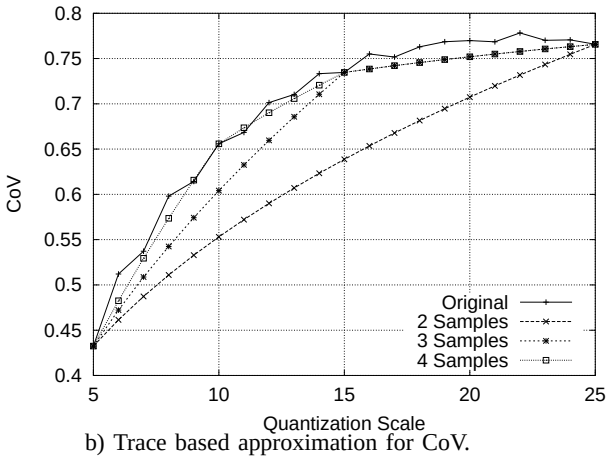


e) Trace based approximation for mean $PSNR$.

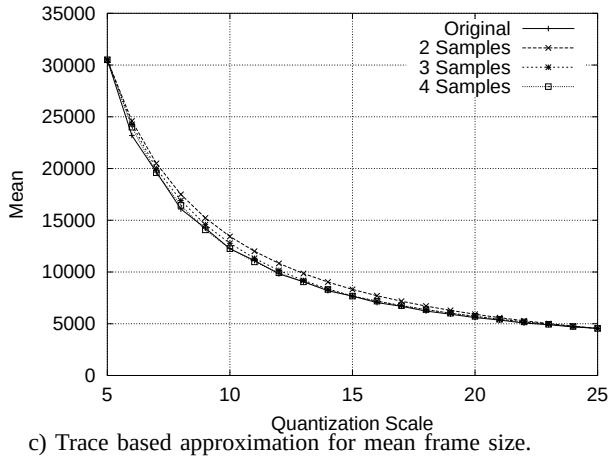
Fig. 48. *Aggregated* level approximation results for scene *557* (MC III) from *Football* .



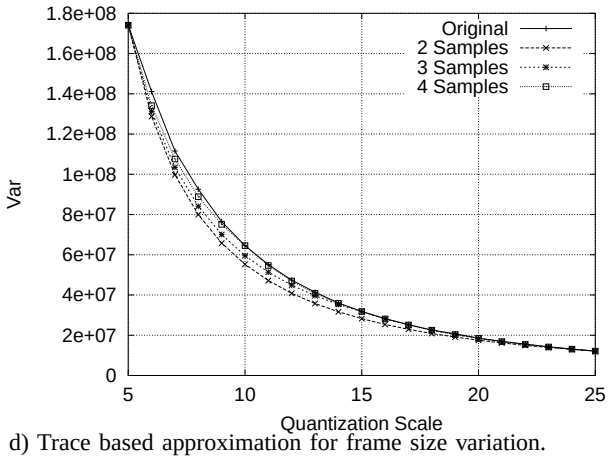
a) Trace based approximation for CoV and $PSNR$.



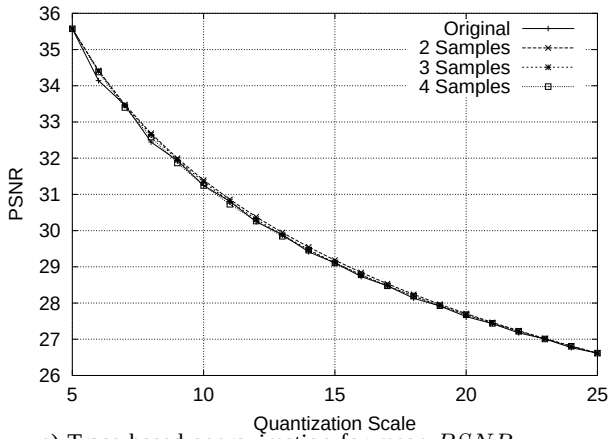
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

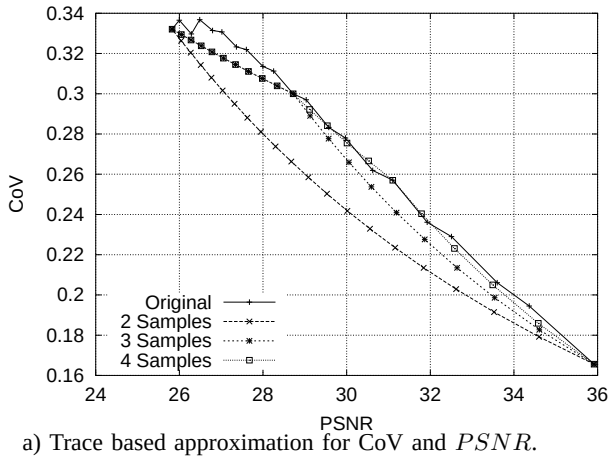


d) Trace based approximation for frame size variation.

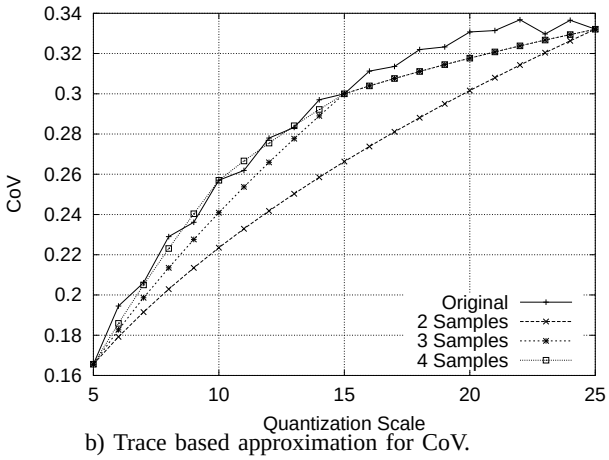


e) Trace based approximation for mean $PSNR$.

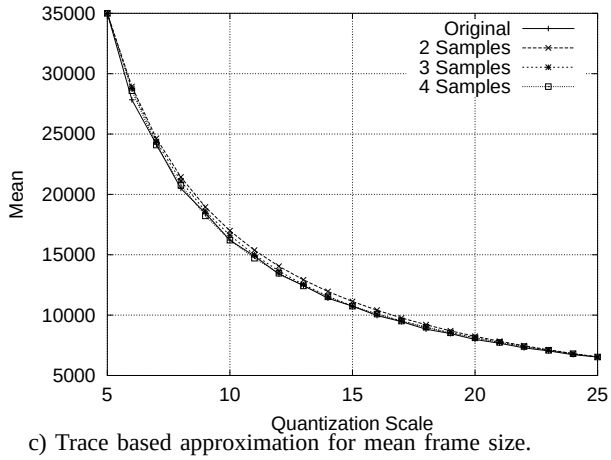
Fig. 49. *Aggregated* level approximation results for scene *184* (MC IV) from *Football* .



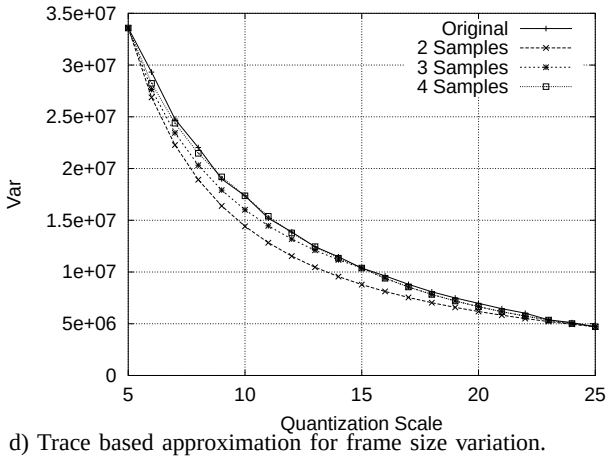
a) Trace based approximation for CoV and $PSNR$.



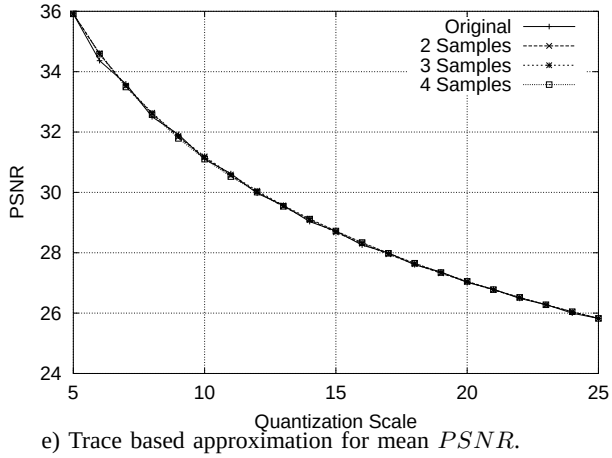
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

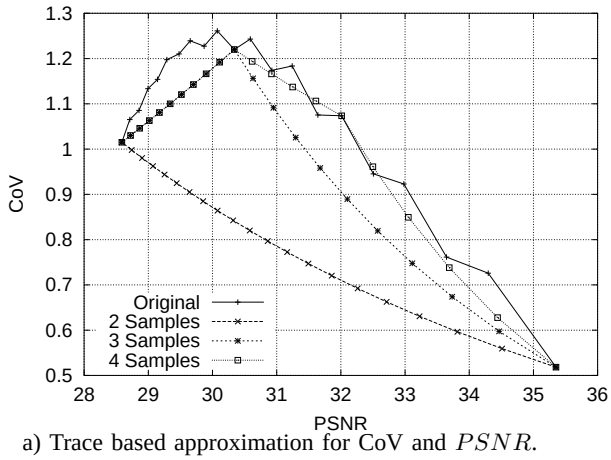


d) Trace based approximation for frame size variation.

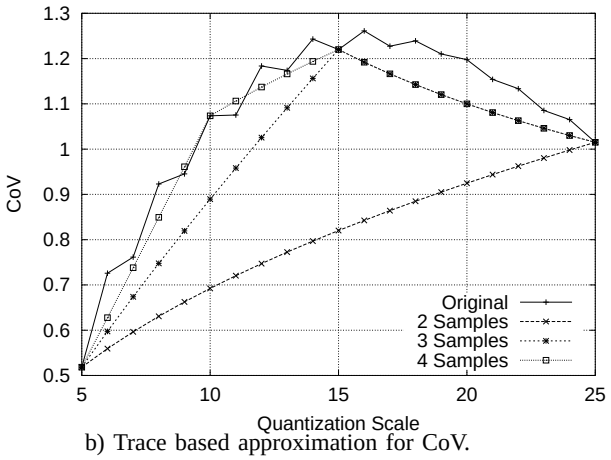


e) Trace based approximation for mean $PSNR$.

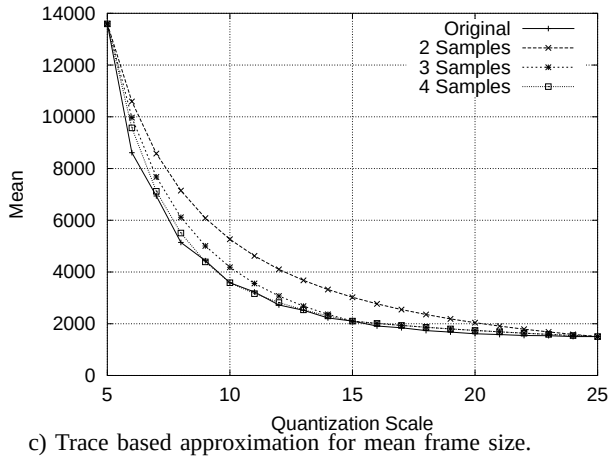
Fig. 50. *Aggregated* level approximation results for scene *336* (MC V) from *Football* .



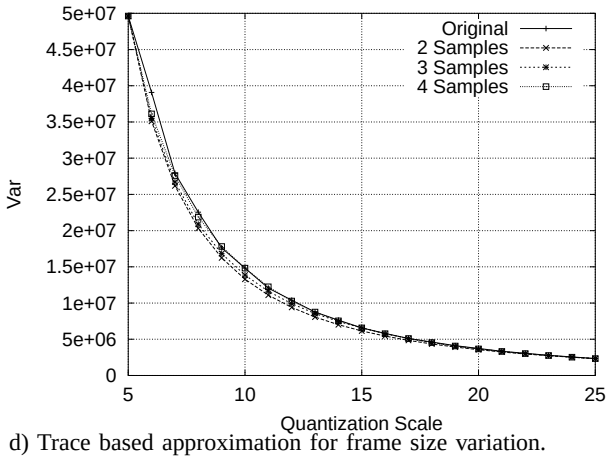
a) Trace based approximation for CoV and $PSNR$.



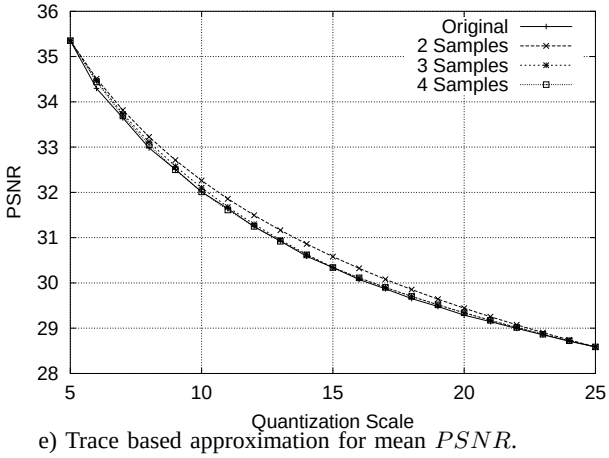
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

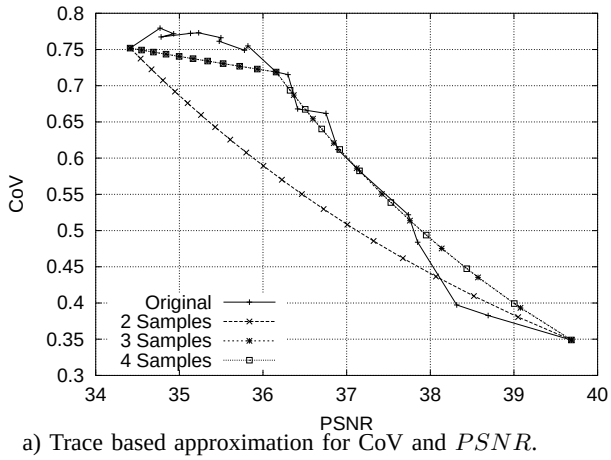


d) Trace based approximation for frame size variation.

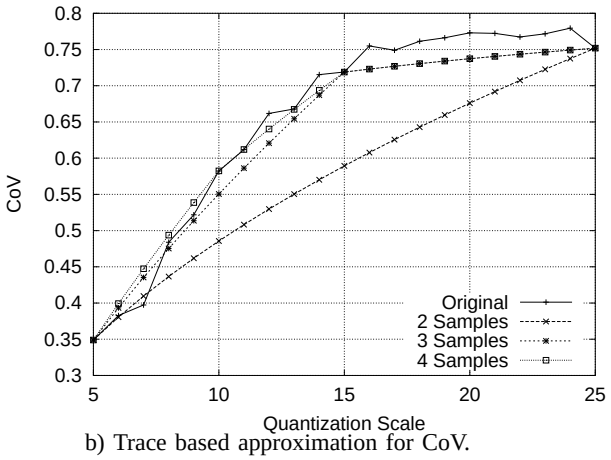


e) Trace based approximation for mean $PSNR$.

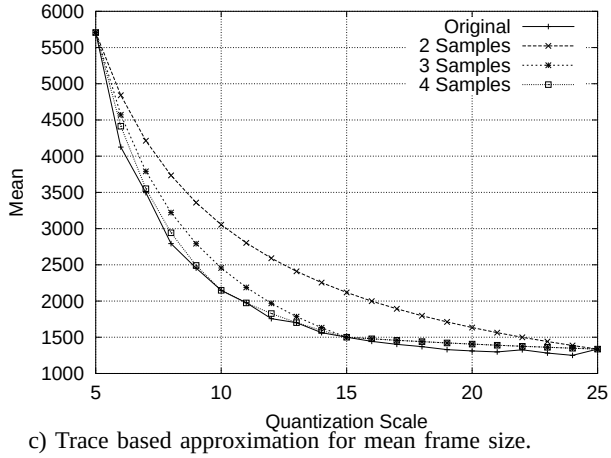
Fig. 51. *Aggregated* level approximation results for scene *274* (MC I) from *Star Wars IV* .



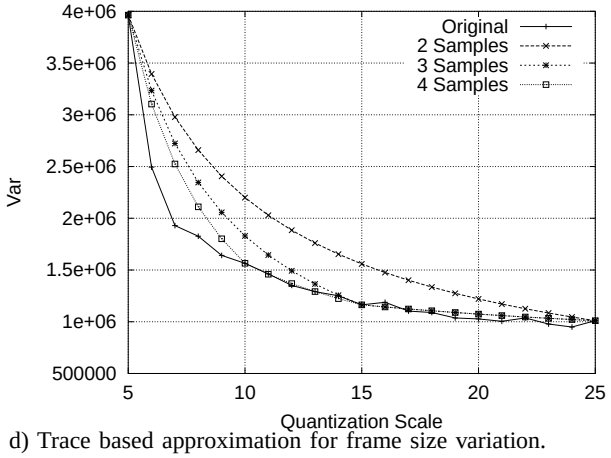
a) Trace based approximation for CoV and $PSNR$.



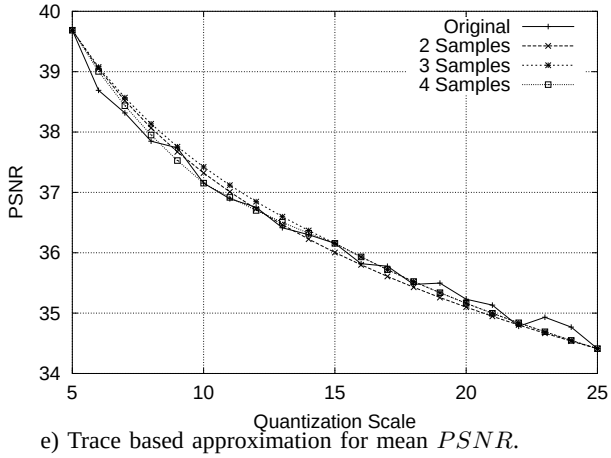
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

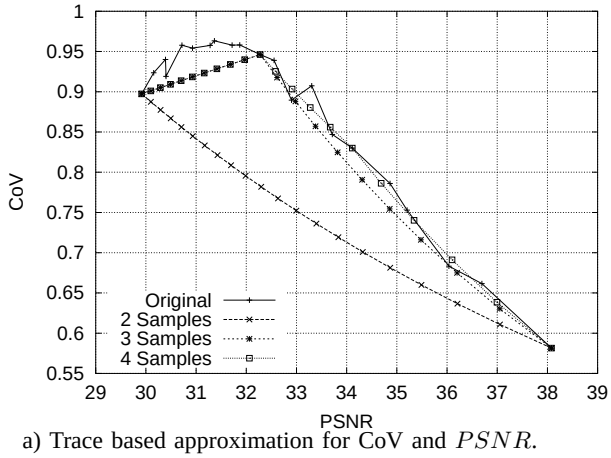


d) Trace based approximation for frame size variation.

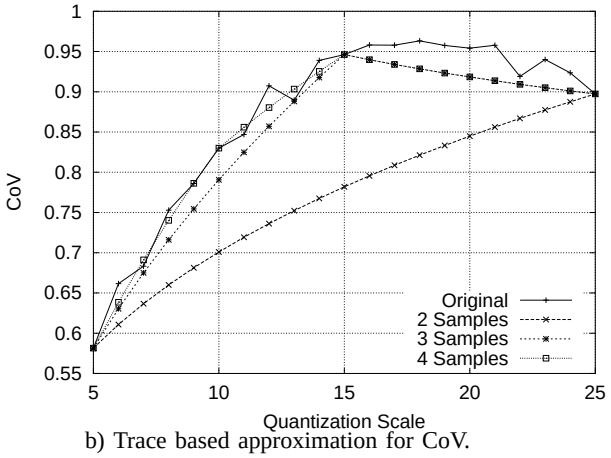


e) Trace based approximation for mean $PSNR$.

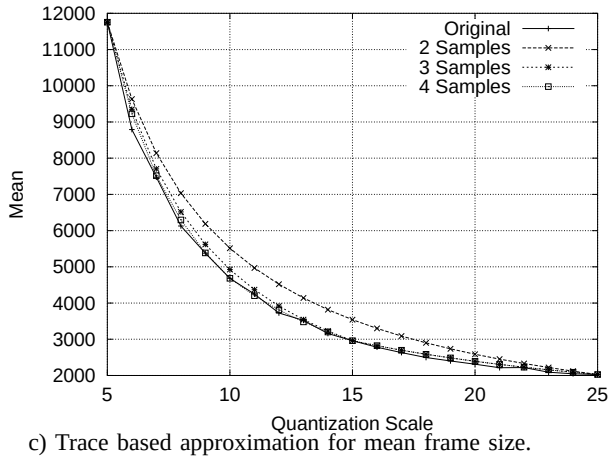
Fig. 52. *Aggregated* level approximation results for scene *117* (MC II) from *Star Wars IV* .



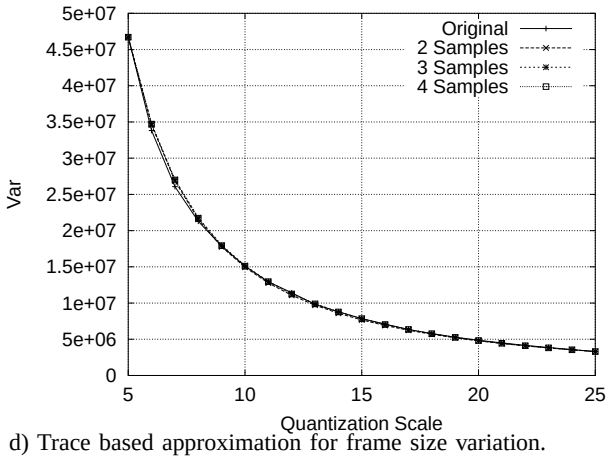
a) Trace based approximation for CoV and $PSNR$.



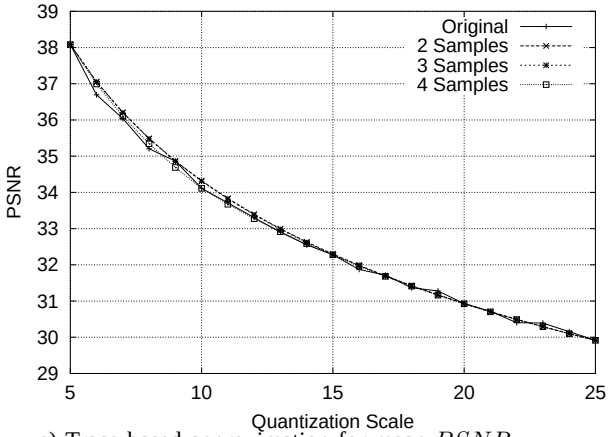
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

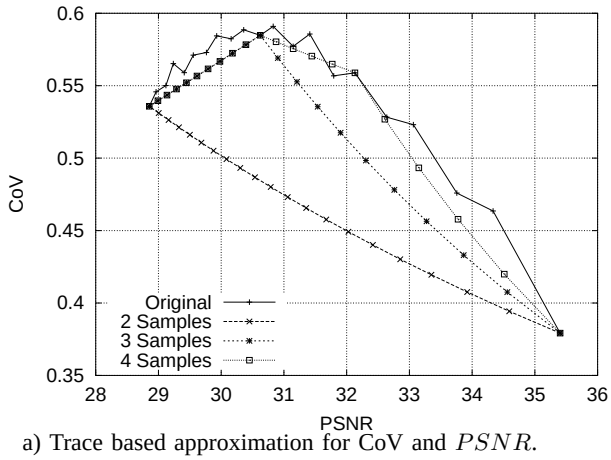


d) Trace based approximation for frame size variation.

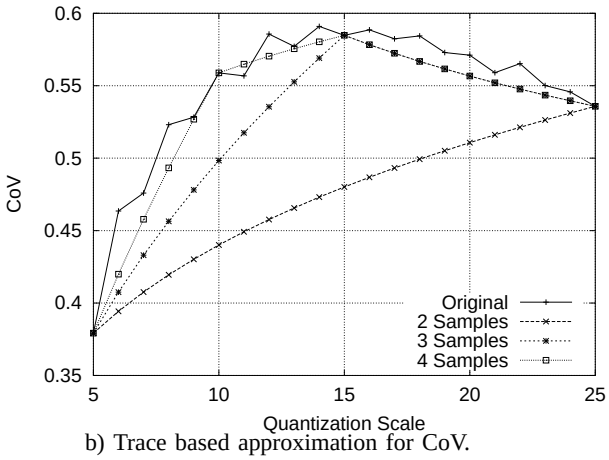


e) Trace based approximation for mean $PSNR$.

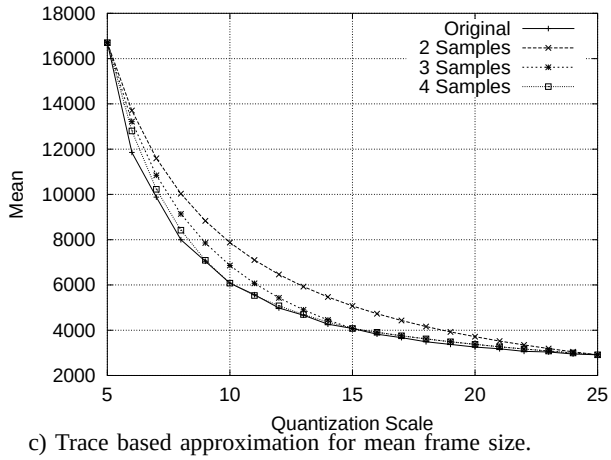
Fig. 53. *Aggregated* level approximation results for scene *115* (MC III) from *Star Wars IV* .



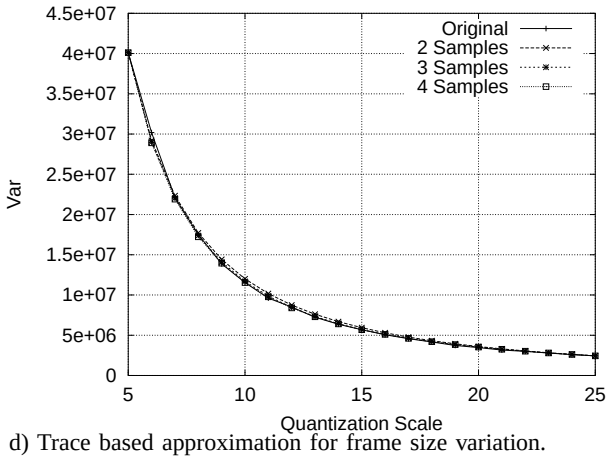
a) Trace based approximation for CoV and $PSNR$.



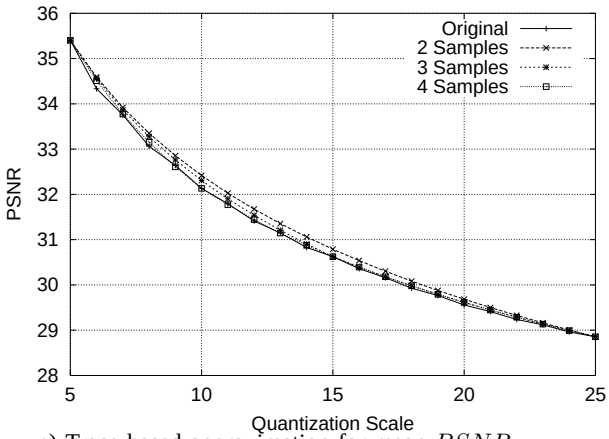
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

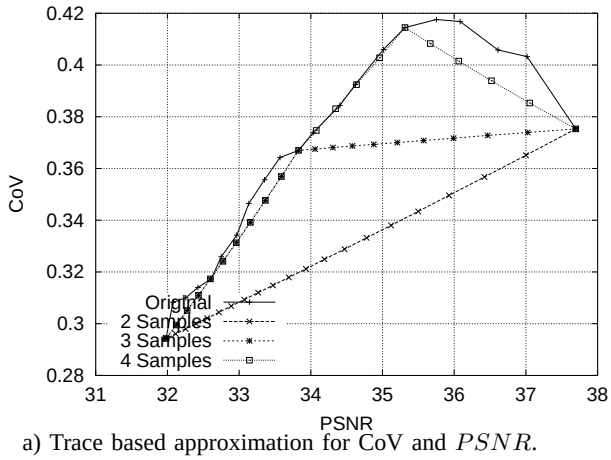


d) Trace based approximation for frame size variation.

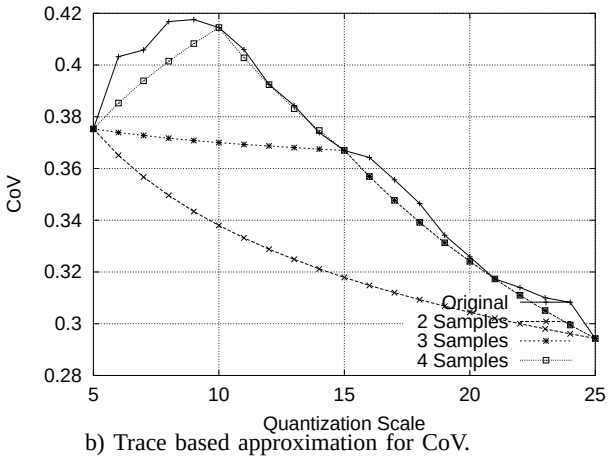


e) Trace based approximation for mean $PSNR$.

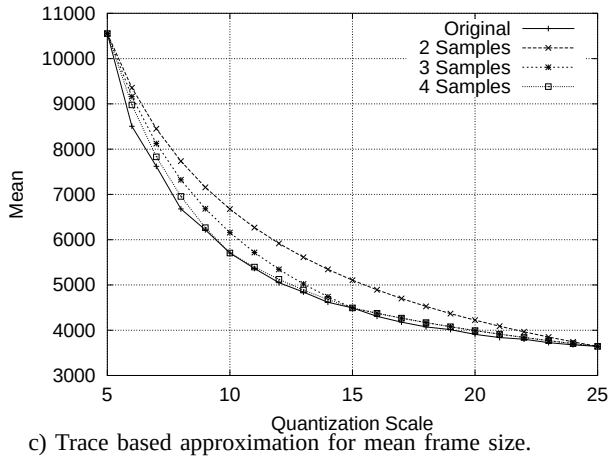
Fig. 54. *Aggregated* level approximation results for scene *165* (MC IV) from *Star Wars IV* .



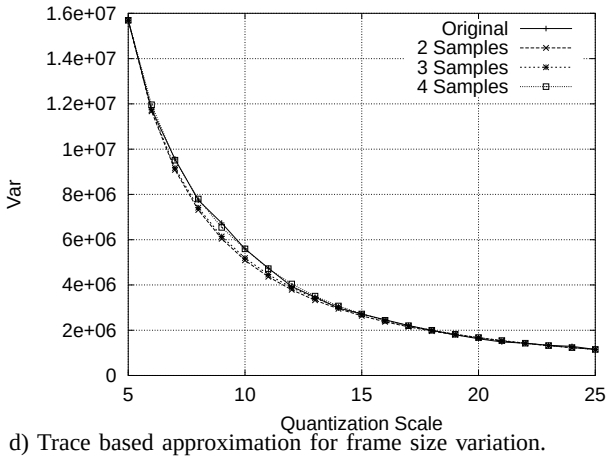
a) Trace based approximation for CoV and $PSNR$.



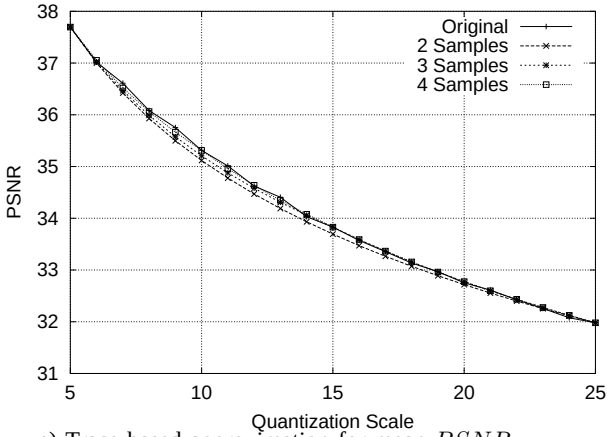
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

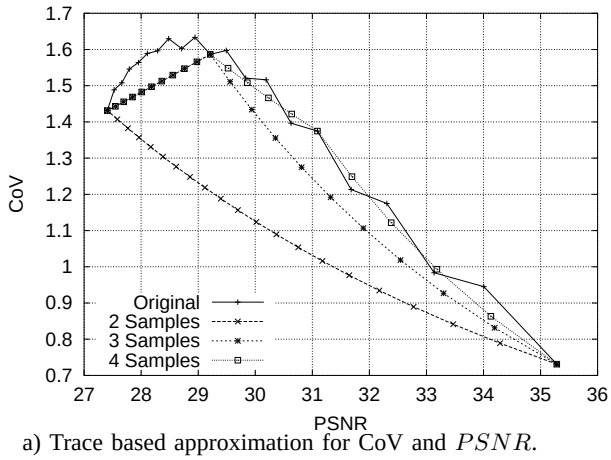


d) Trace based approximation for frame size variation.

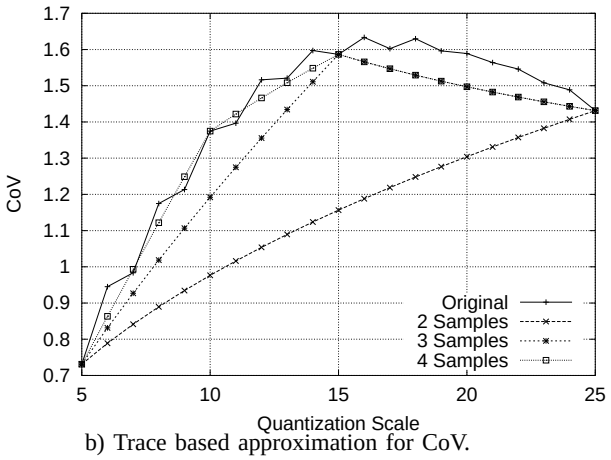


e) Trace based approximation for mean $PSNR$.

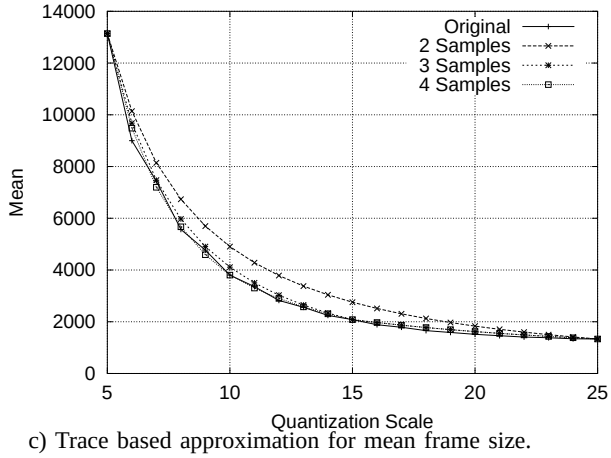
Fig. 55. *Aggregated* level approximation results for scene *632* (MC V) from *Star Wars IV* .



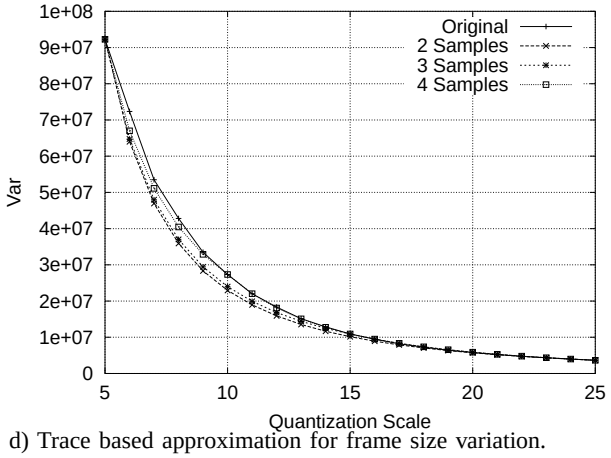
a) Trace based approximation for CoV and $PSNR$.



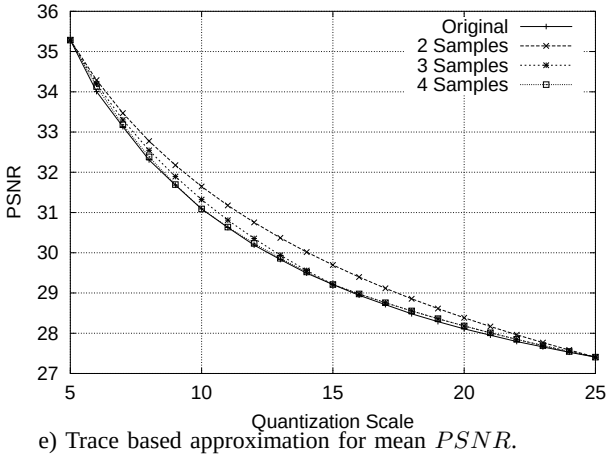
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

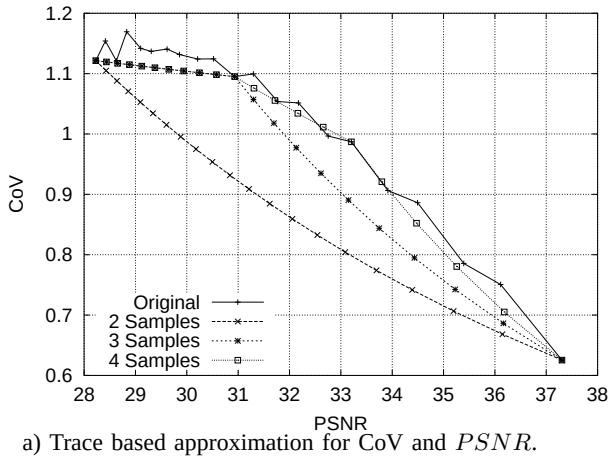


d) Trace based approximation for frame size variation.

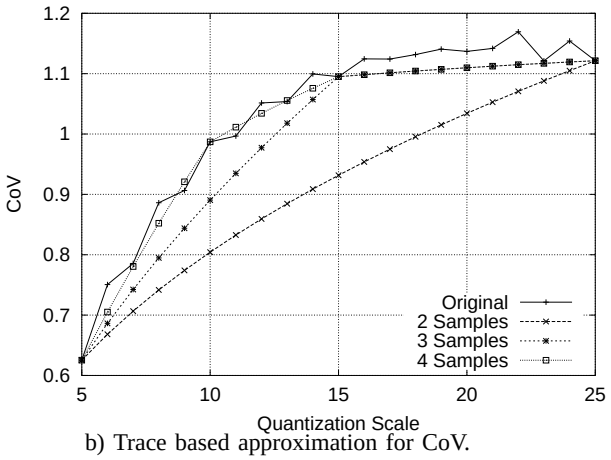


e) Trace based approximation for mean $PSNR$.

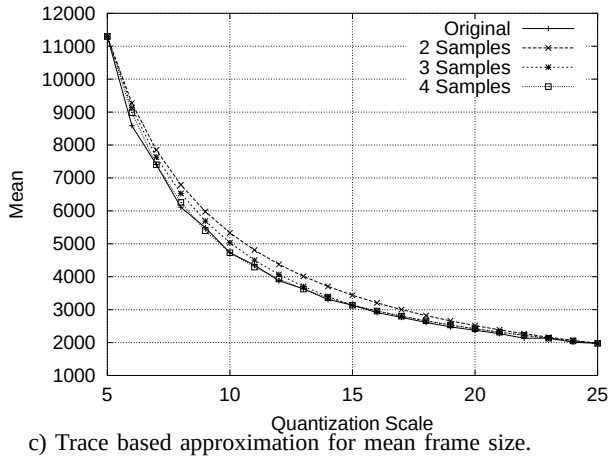
Fig. 56. *Aggregated* level approximation results for scene *384* (MC I) from *The Terminator* .



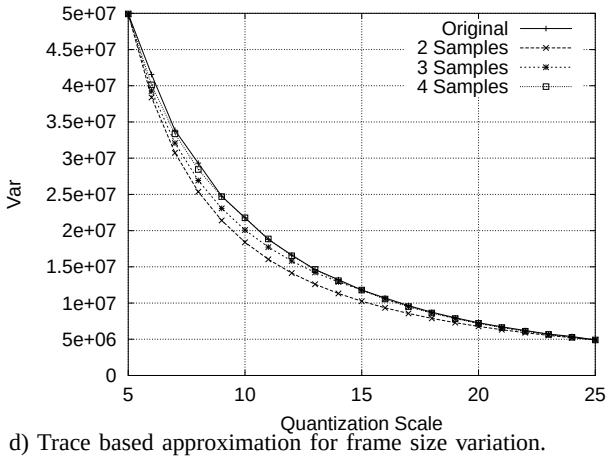
a) Trace based approximation for CoV and $PSNR$.



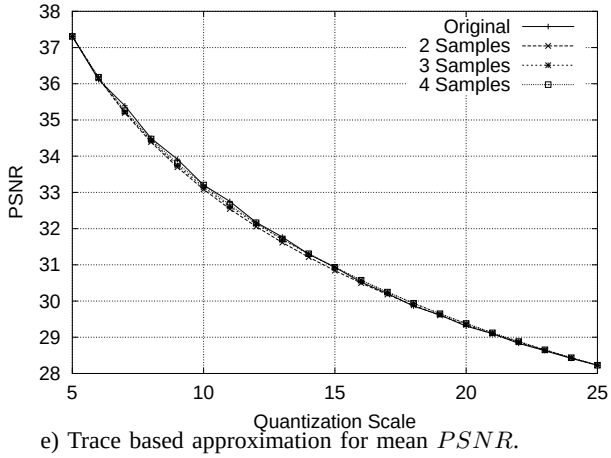
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

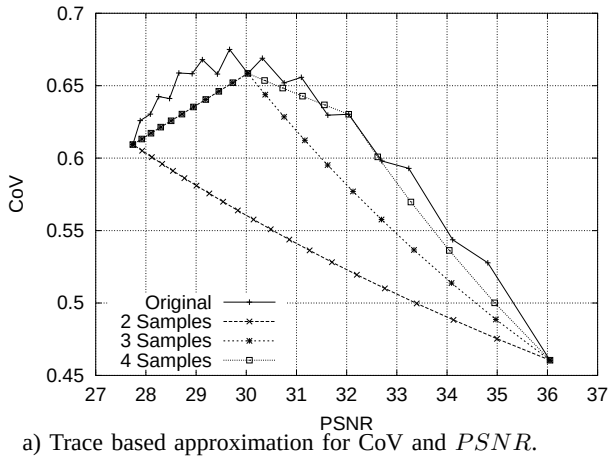


d) Trace based approximation for frame size variation.

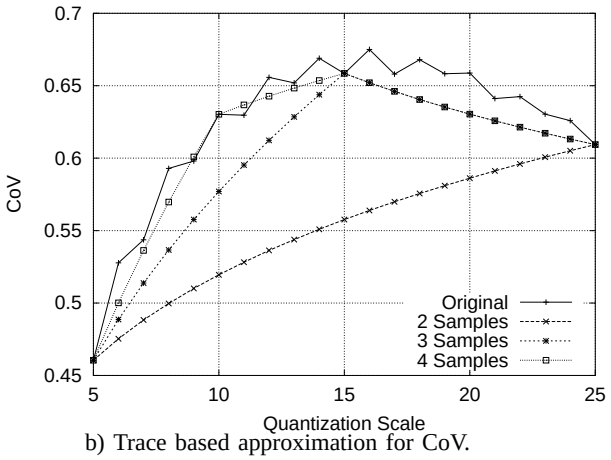


e) Trace based approximation for mean $PSNR$.

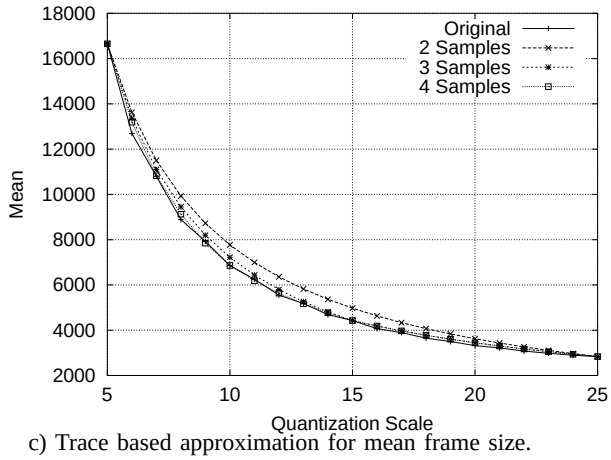
Fig. 57. *Aggregated* level approximation results for scene *462* (MC II) from *The Terminator* .



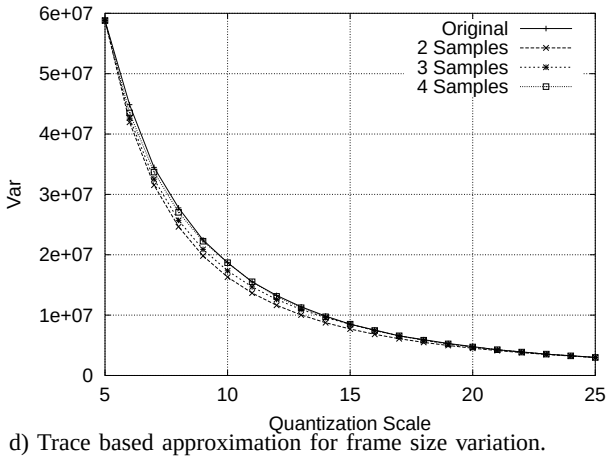
a) Trace based approximation for CoV and $PSNR$.



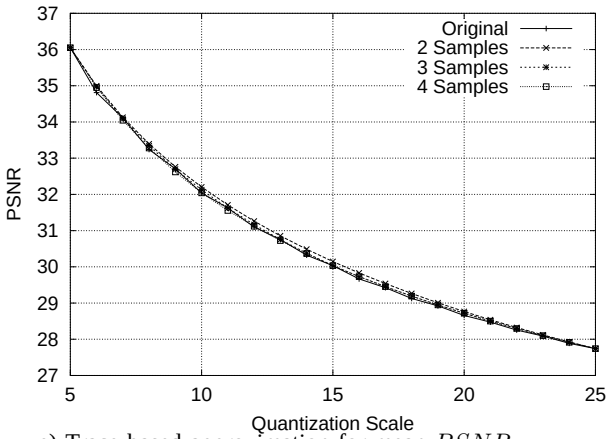
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

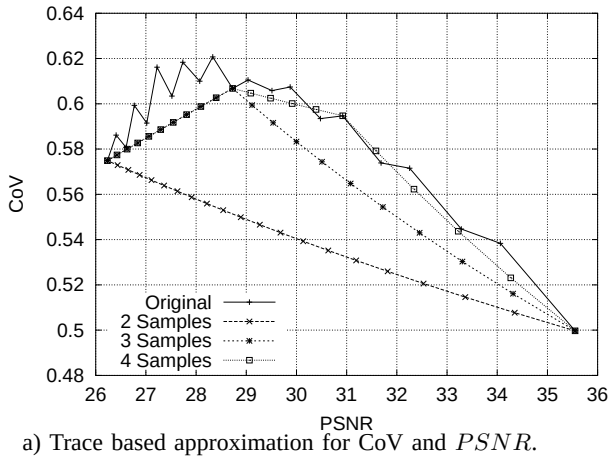


d) Trace based approximation for frame size variation.

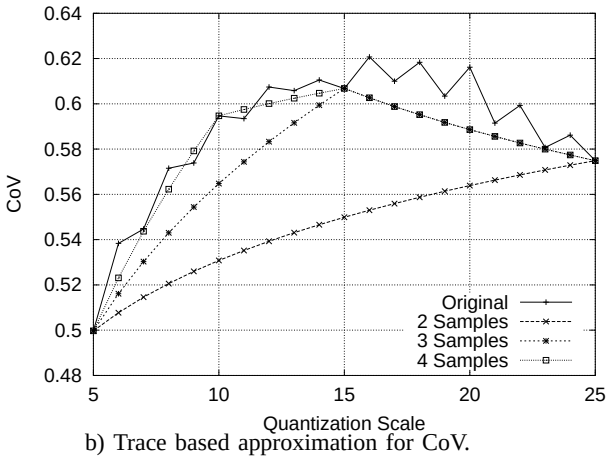


e) Trace based approximation for mean $PSNR$.

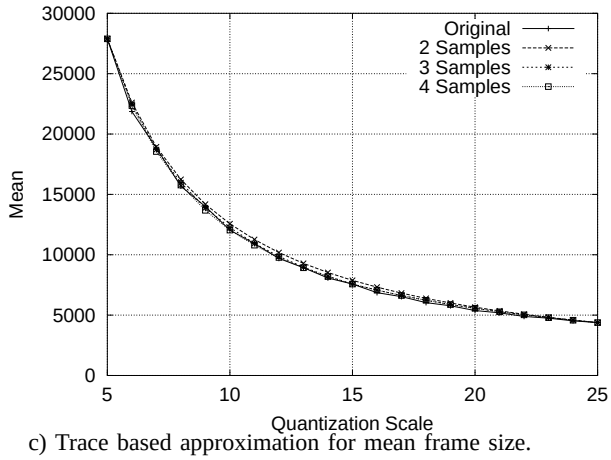
Fig. 58. *Aggregated* level approximation results for scene *628* (MC III) from *The Terminator* .



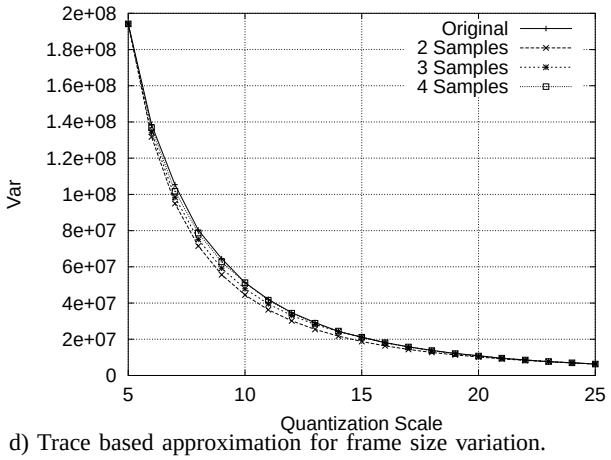
a) Trace based approximation for CoV and $PSNR$.



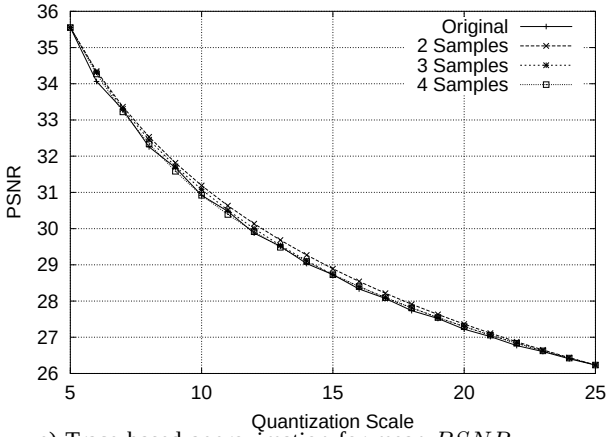
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

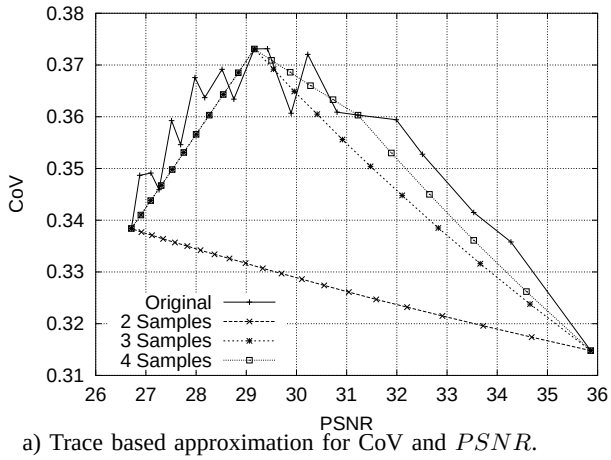


d) Trace based approximation for frame size variation.

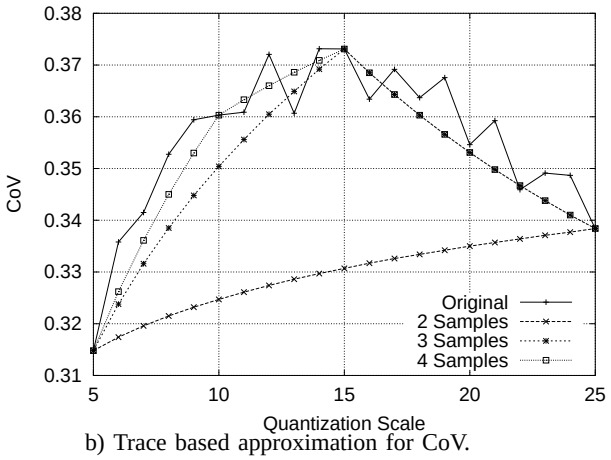


e) Trace based approximation for mean $PSNR$.

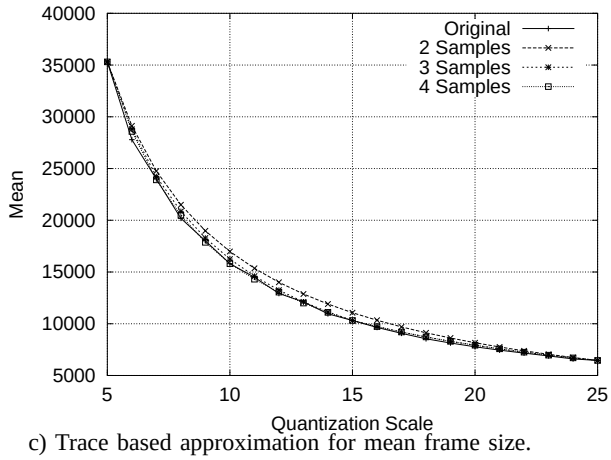
Fig. 59. *Aggregated* level approximation results for scene *262* (MC IV) from *The Terminator* .



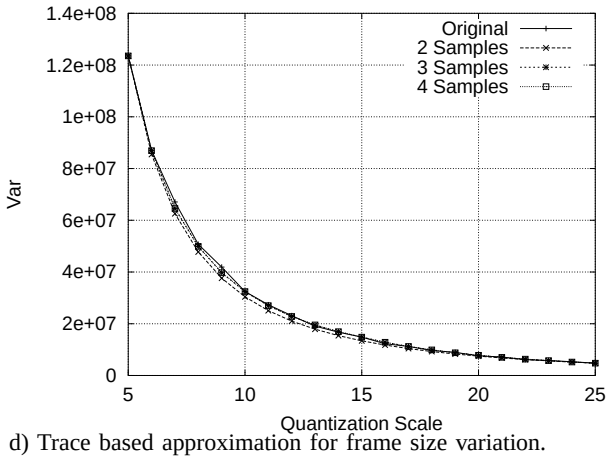
a) Trace based approximation for CoV and $PSNR$.



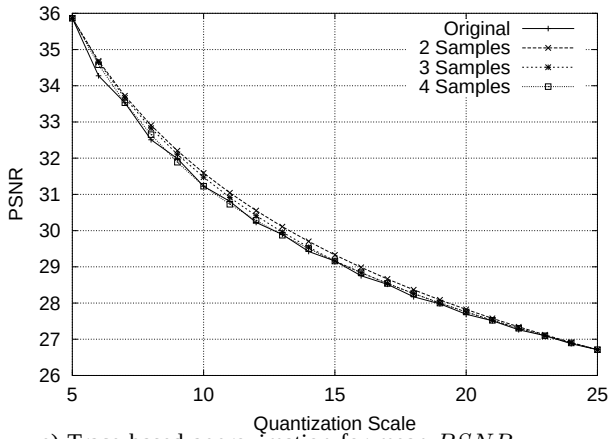
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



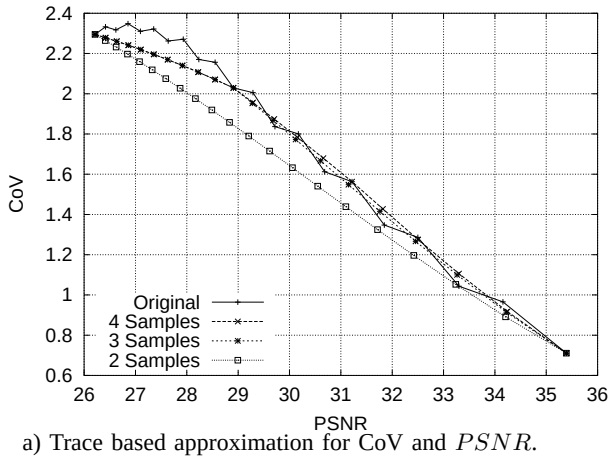
d) Trace based approximation for frame size variation.



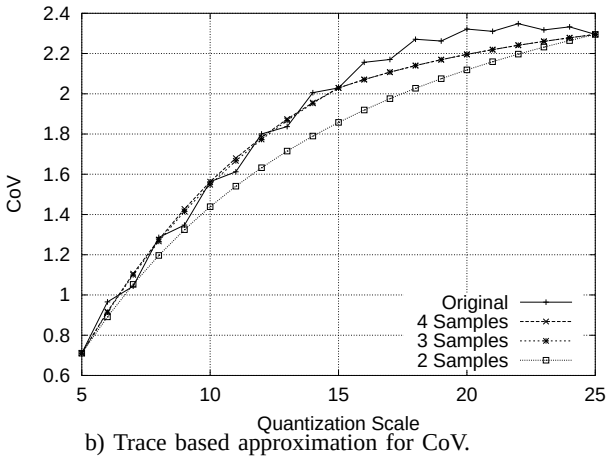
e) Trace based approximation for mean $PSNR$.

Fig. 60. *Aggregated* level approximation results for scene *441* (MC V) from *The Terminator* .

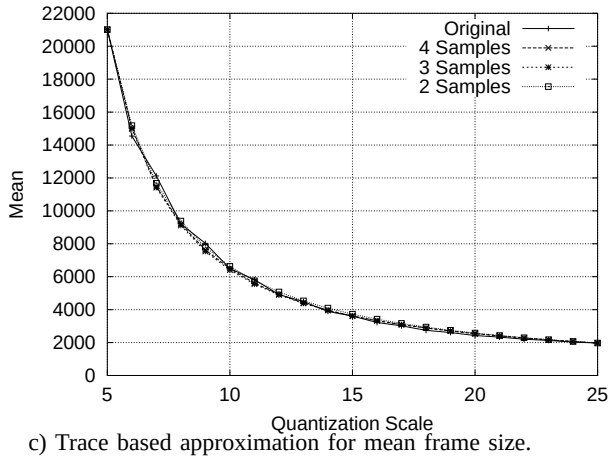
We immediately observe that the approximation with only two samples of the encodings does not capture the VD behavior at all. We furthermore observe that even with four interpolation points, the approximation does not capture the VD behavior for most scenes very well. The fit for the more pronounced hump in the VD curve for the low motion scenes is visibly worse than for the high motion scenes. In addition we note that for the medium to high qualities, the estimation yields a better result and follows the more linear slope closer. From the evaluation of the individual components of the VD curves we observe that for three and four interpolation points the approximations for the individual statistics, namely mean frame size, frame size variance, and mean frame quality ($PSNR$) are close to the original values. We compare the previous results for the *aggregated* level in the following with the results for the *frame* level approximations in Figs. 61, 62, 63, 64, and 65 for scenes from *Football*; in Figs. 66, 67, 68, 69, 70 for scenes from *Star Wars IV*; and in Figs. 71, 72, 73, 74, 75 for scenes from *The Terminator*.



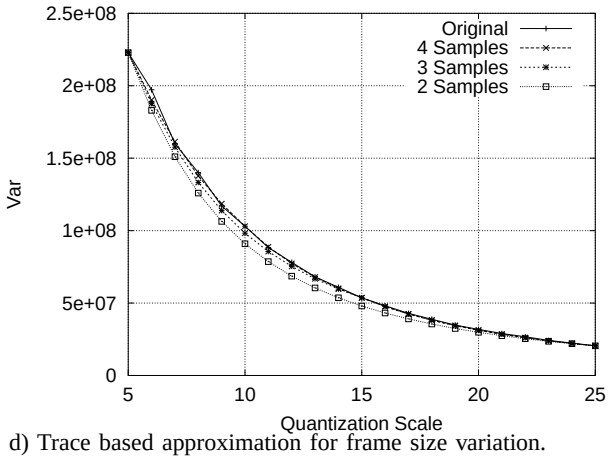
a) Trace based approximation for CoV and $PSNR$.



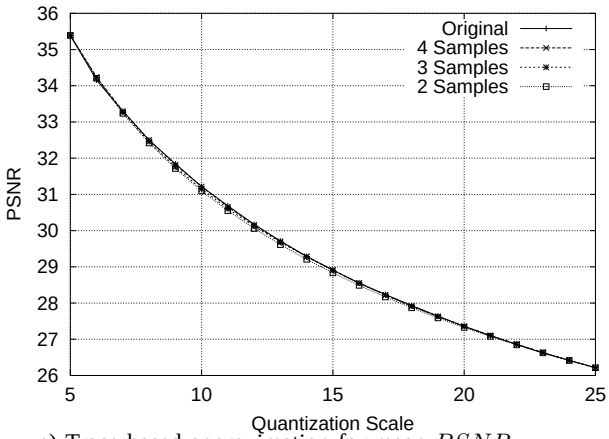
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

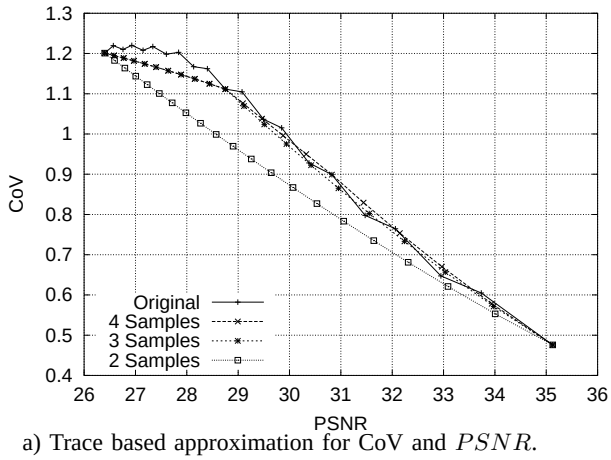


d) Trace based approximation for frame size variation.

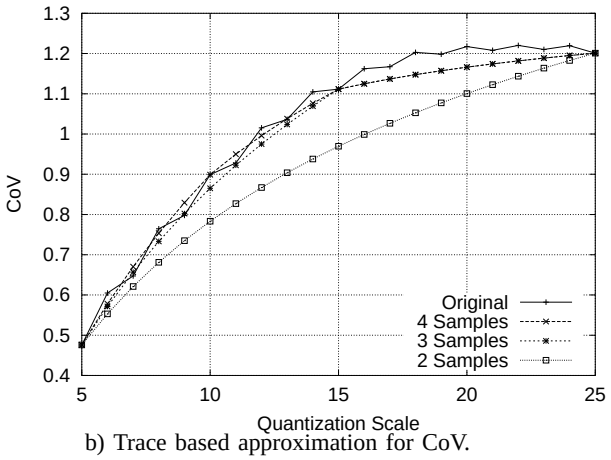


e) Trace based approximation for mean $PSNR$.

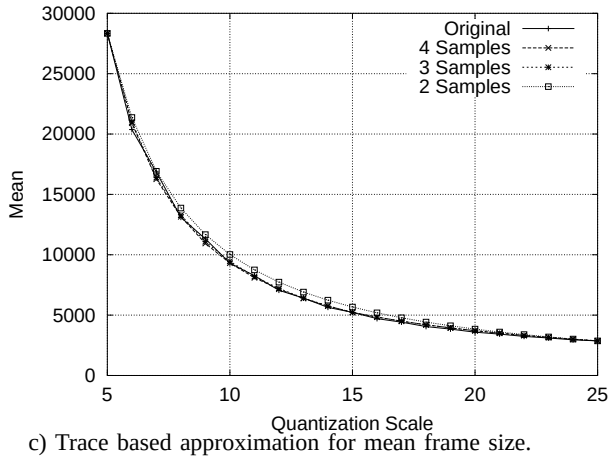
Fig. 61. *Frame* level approximation results for scene *298* (MC I) from *Football* .



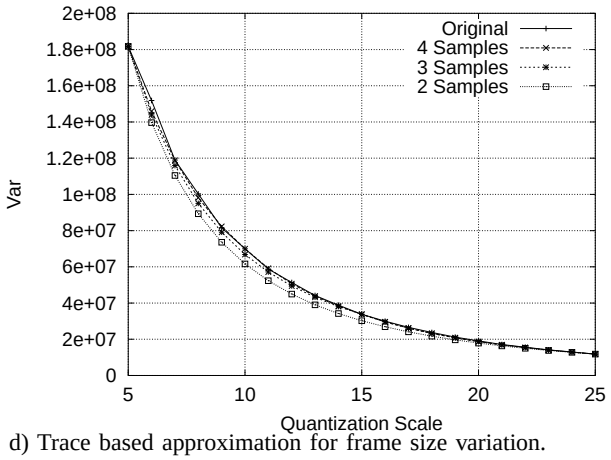
a) Trace based approximation for CoV and $PSNR$.



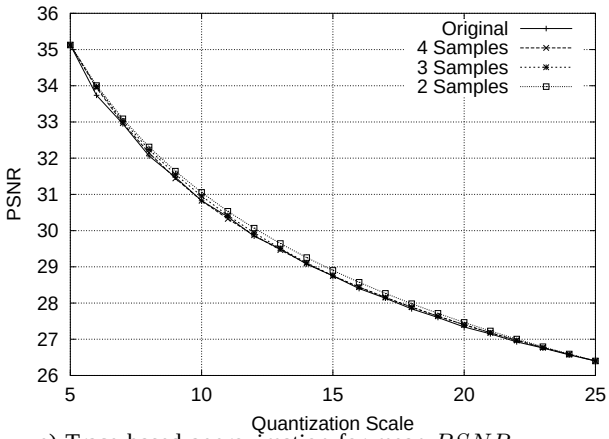
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

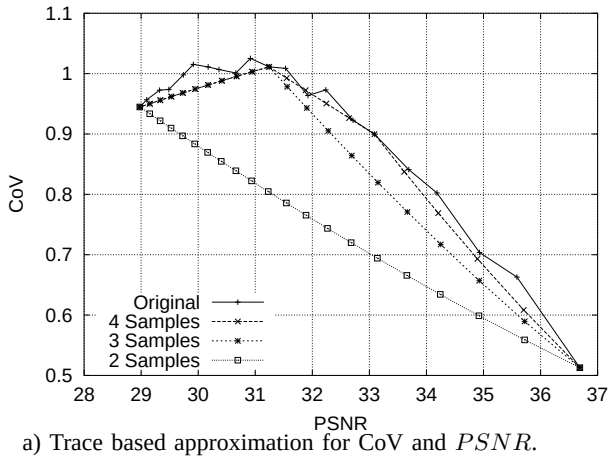


d) Trace based approximation for frame size variation.

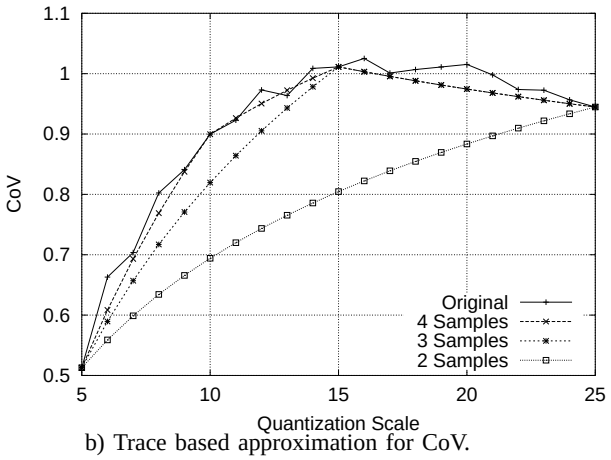


e) Trace based approximation for mean $PSNR$.

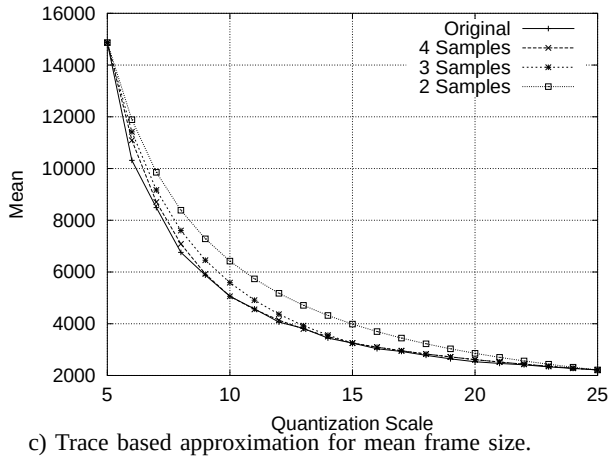
Fig. 62. *Frame* level approximation results for scene *299* (MC II) from *Football* .



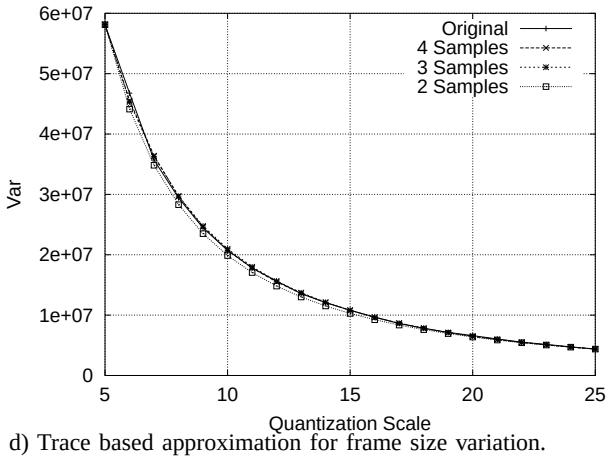
a) Trace based approximation for CoV and $PSNR$.



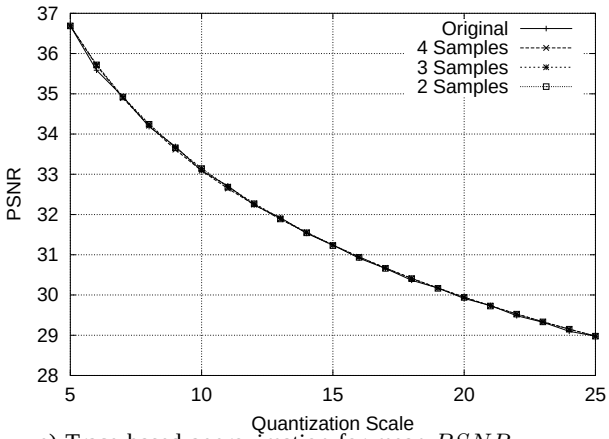
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

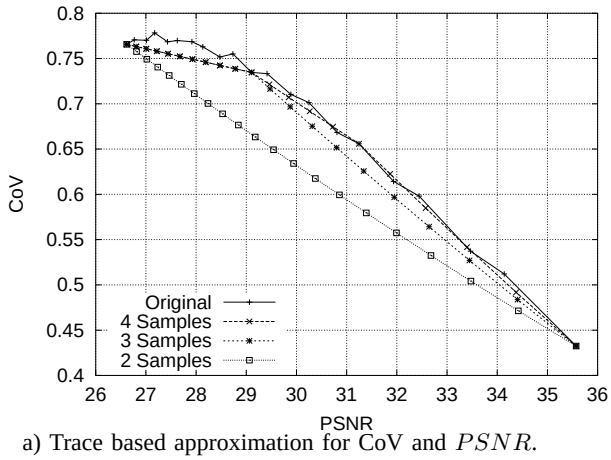


d) Trace based approximation for frame size variation.

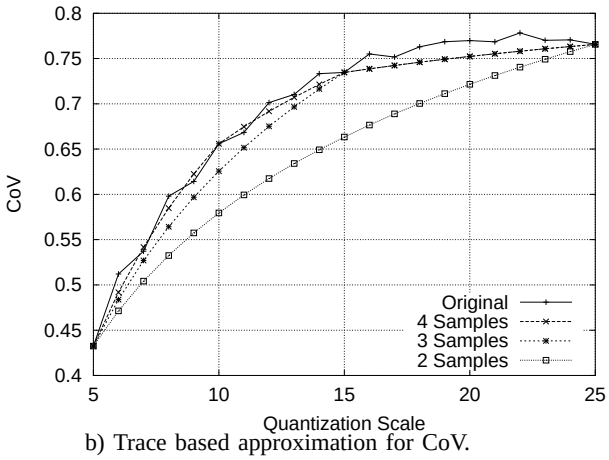


e) Trace based approximation for mean $PSNR$.

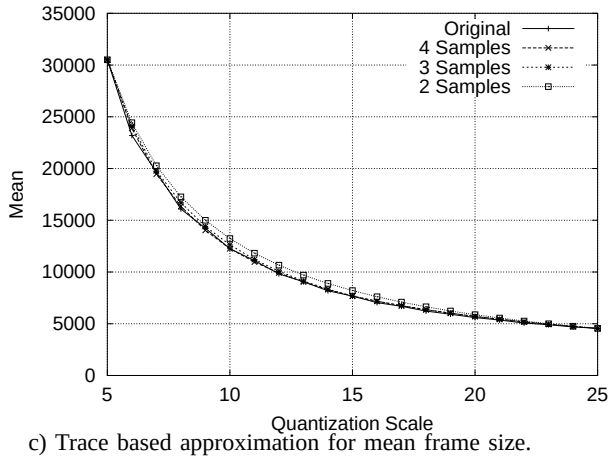
Fig. 63. *Frame* level approximation results for scene *557* (MC III) from *Football* .



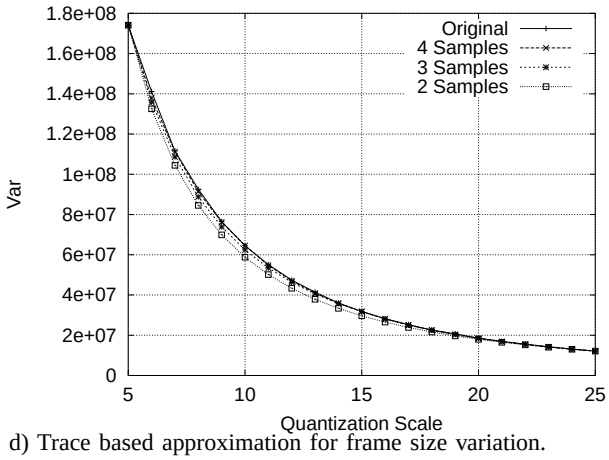
a) Trace based approximation for CoV and $PSNR$.



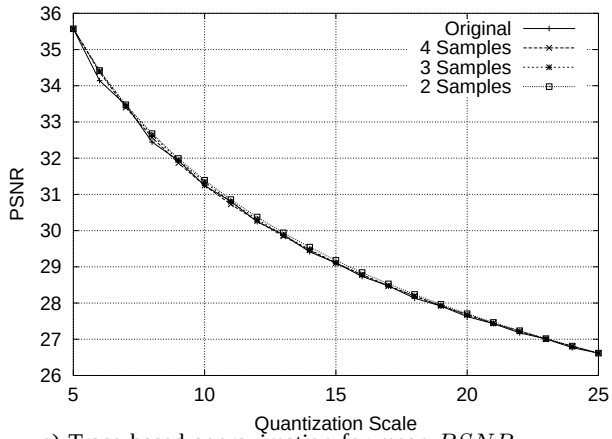
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

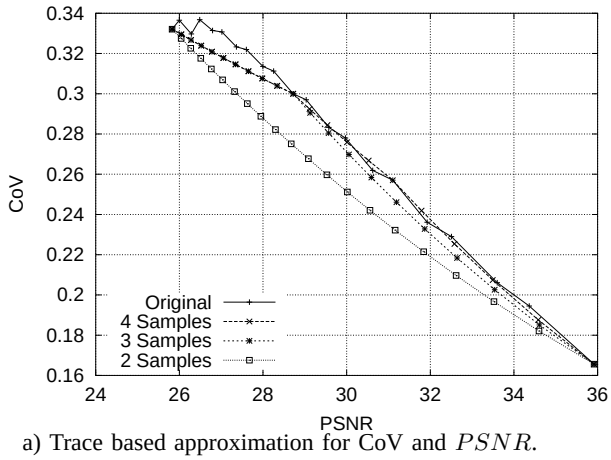


d) Trace based approximation for frame size variation.

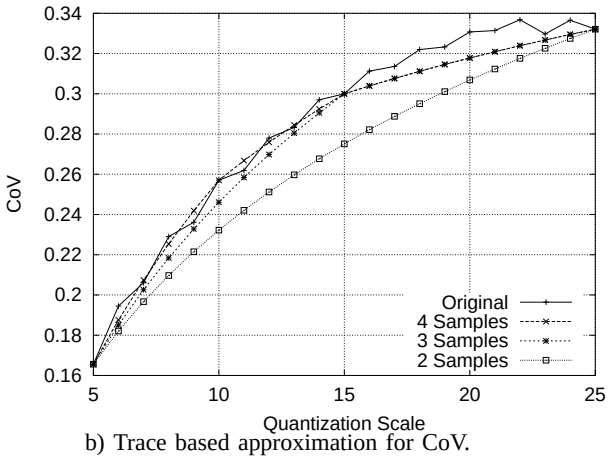


e) Trace based approximation for mean $PSNR$.

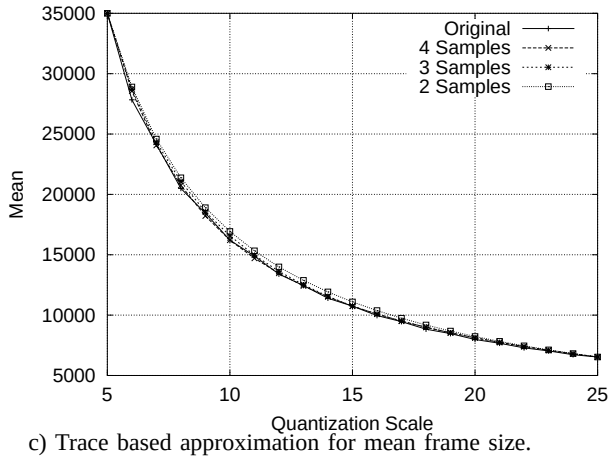
Fig. 64. *Frame* level approximation results for scene *184* (MC IV) from *Football* .



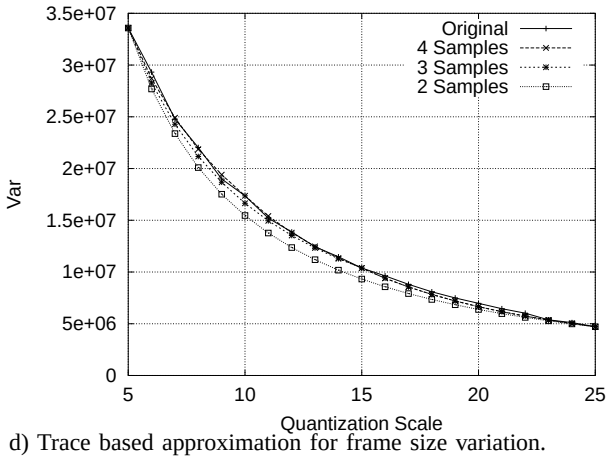
a) Trace based approximation for CoV and $PSNR$.



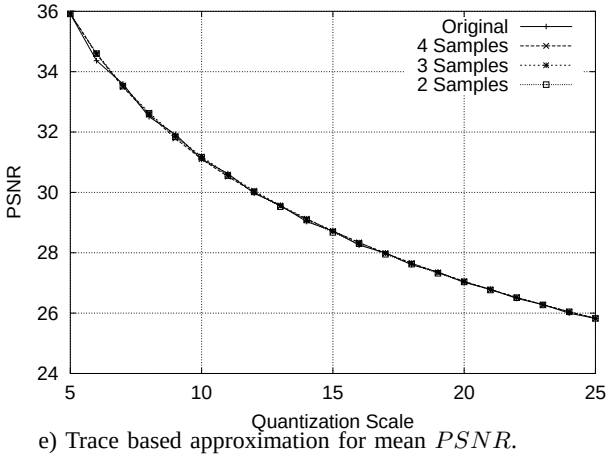
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

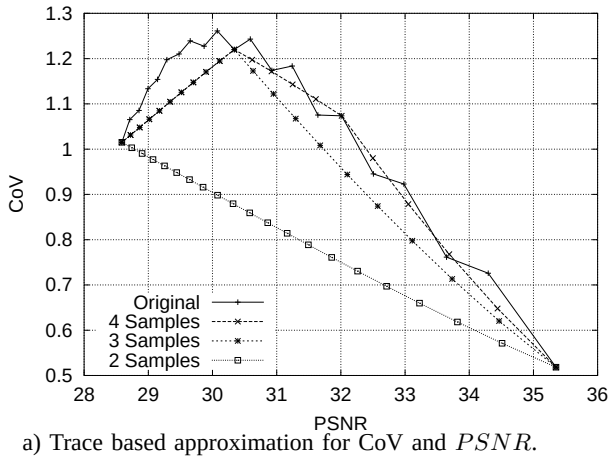


d) Trace based approximation for frame size variation.

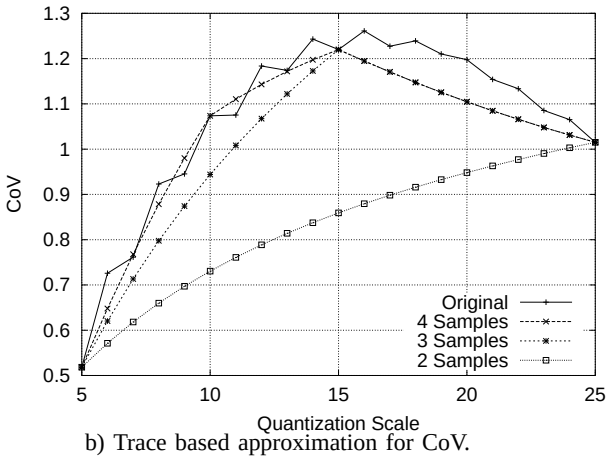


e) Trace based approximation for mean $PSNR$.

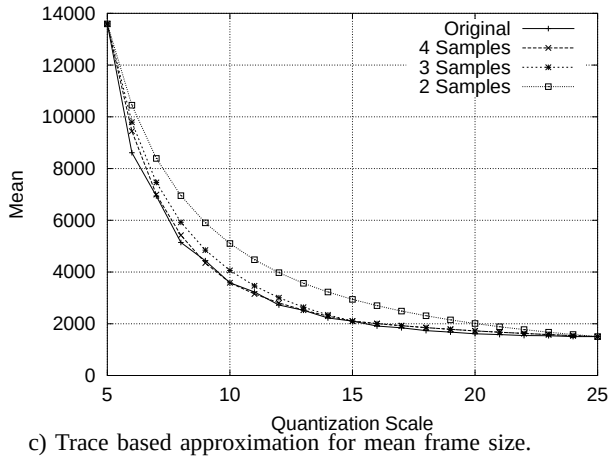
Fig. 65. *Frame* level approximation results for scene *336* (MC V) from *Football* .



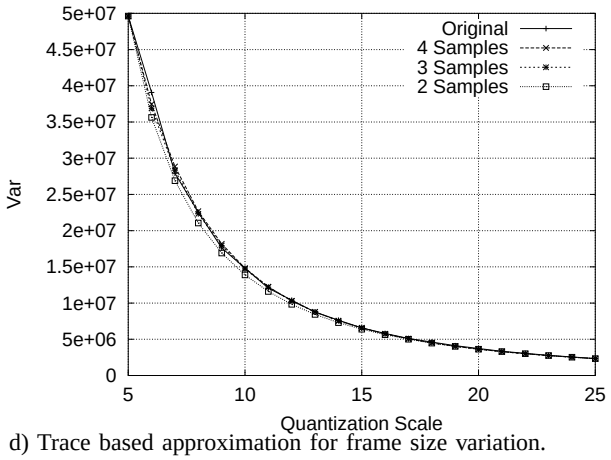
a) Trace based approximation for CoV and $PSNR$.



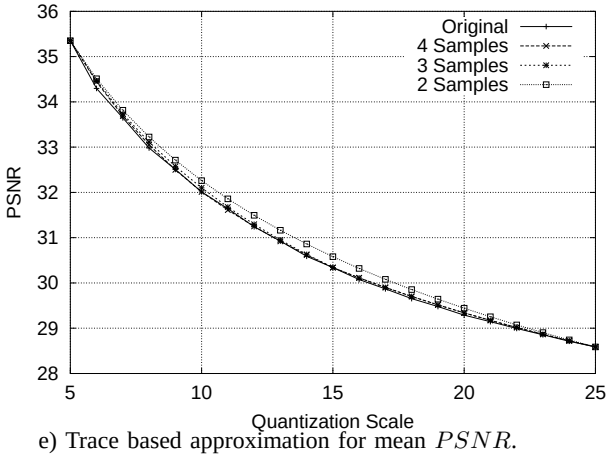
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

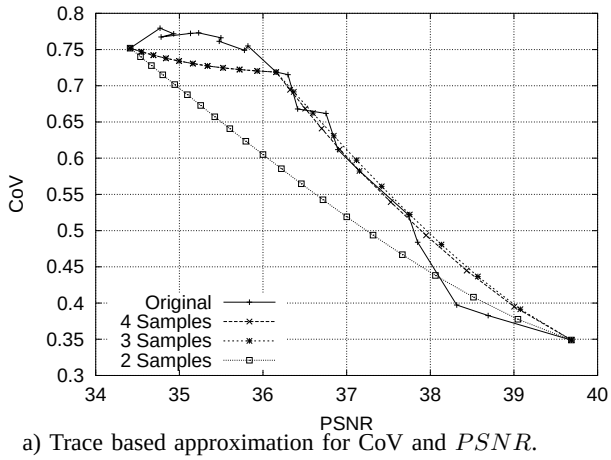


d) Trace based approximation for frame size variation.

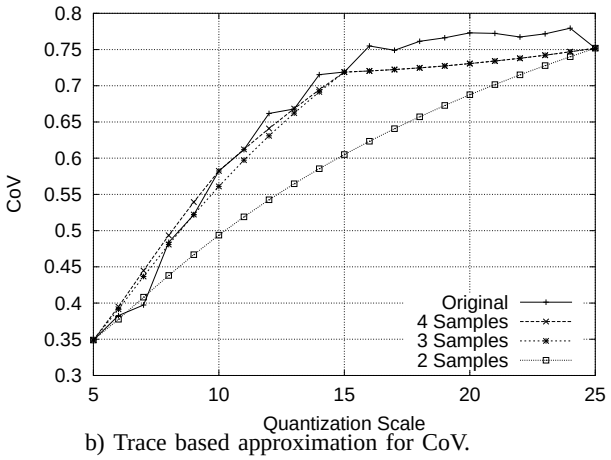


e) Trace based approximation for mean $PSNR$.

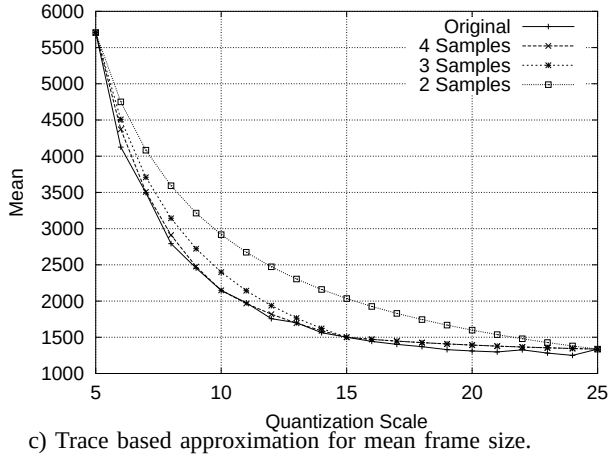
Fig. 66. *Frame* level approximation results for scene *274* (MC I) from *Star Wars IV* .



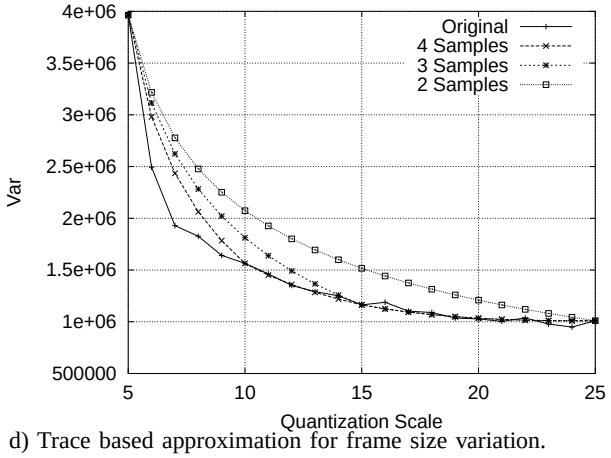
a) Trace based approximation for CoV and $PSNR$.



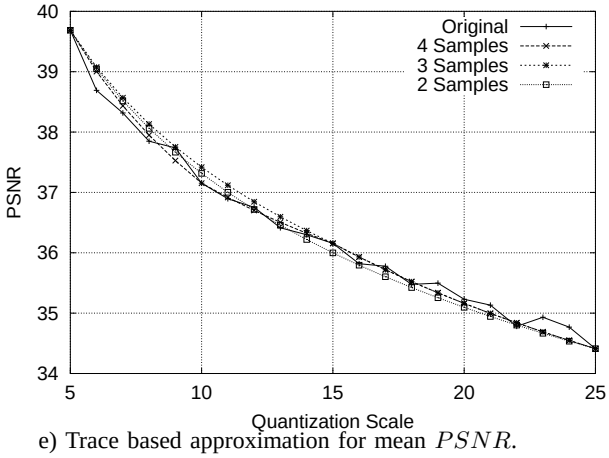
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

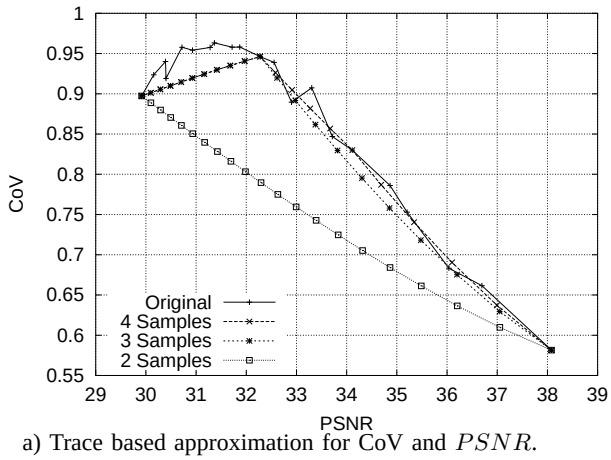


d) Trace based approximation for frame size variation.

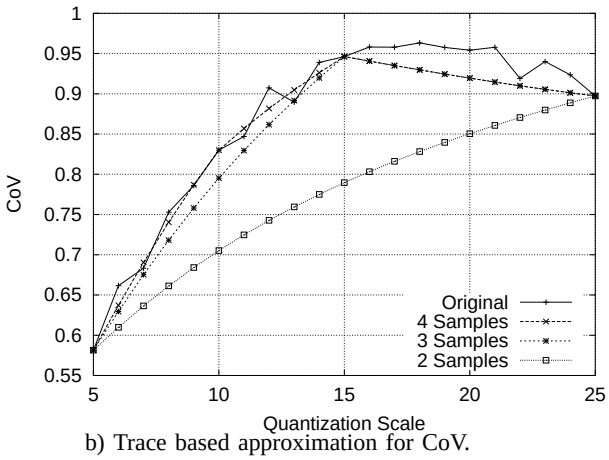


e) Trace based approximation for mean $PSNR$.

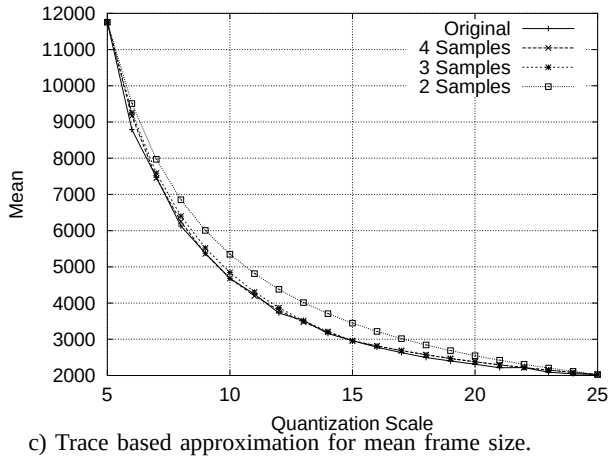
Fig. 67. *Frame* level approximation results for scene *117* (MC II) from *Star Wars IV* .



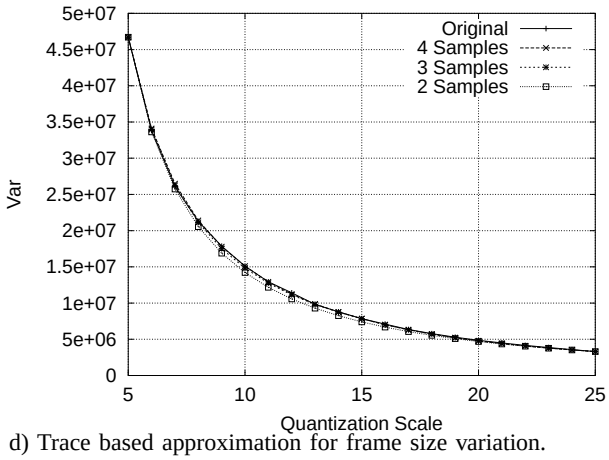
a) Trace based approximation for CoV and $PSNR$.



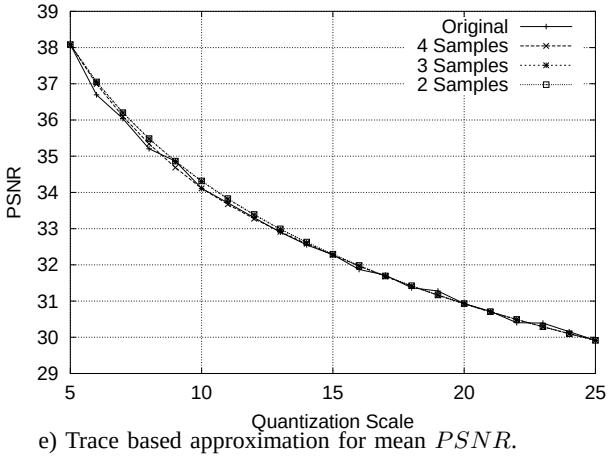
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

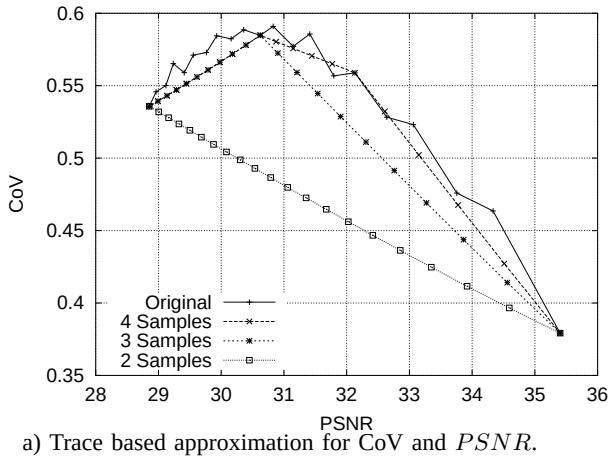


d) Trace based approximation for frame size variation.

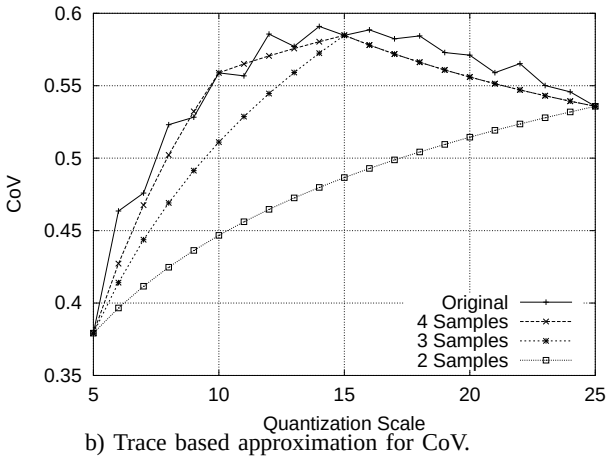


e) Trace based approximation for mean $PSNR$.

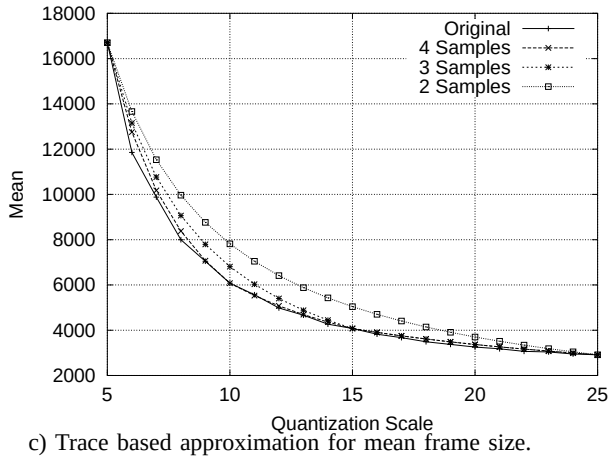
Fig. 68. *Frame* level approximation results for scene *115* (MC III) from *Star Wars IV* .



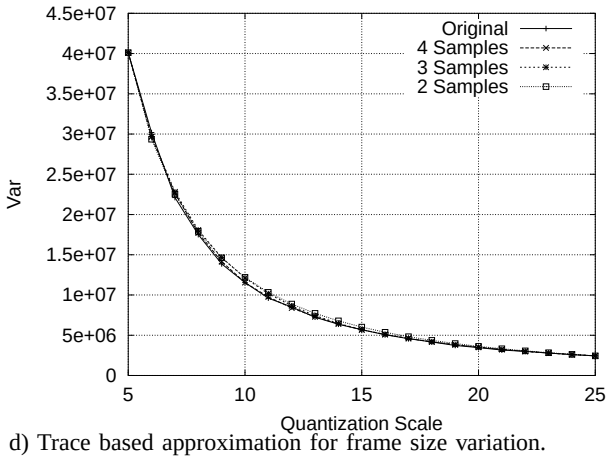
a) Trace based approximation for CoV and $PSNR$.



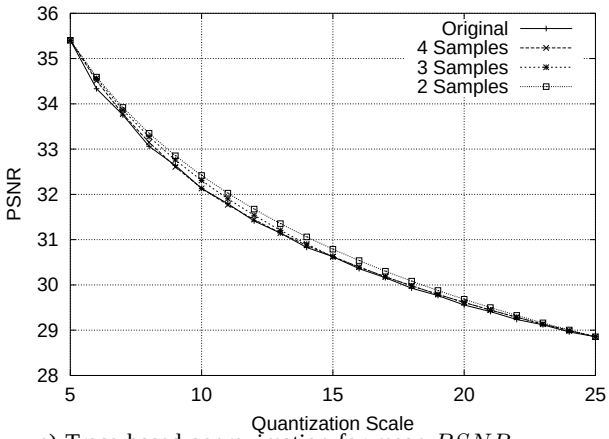
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

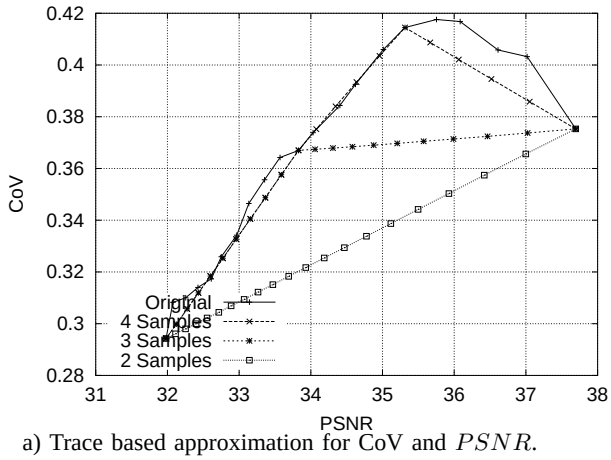


d) Trace based approximation for frame size variation.

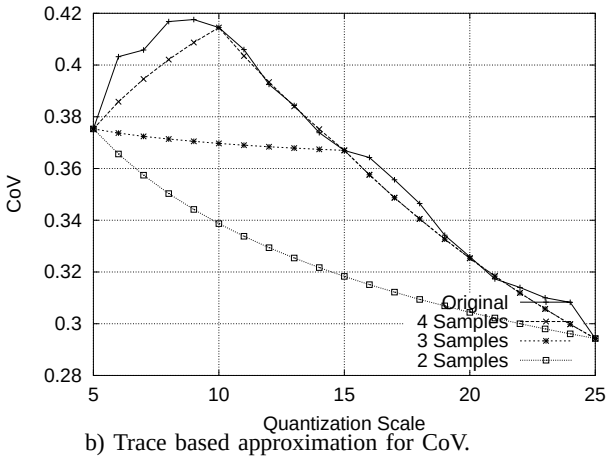


e) Trace based approximation for mean $PSNR$.

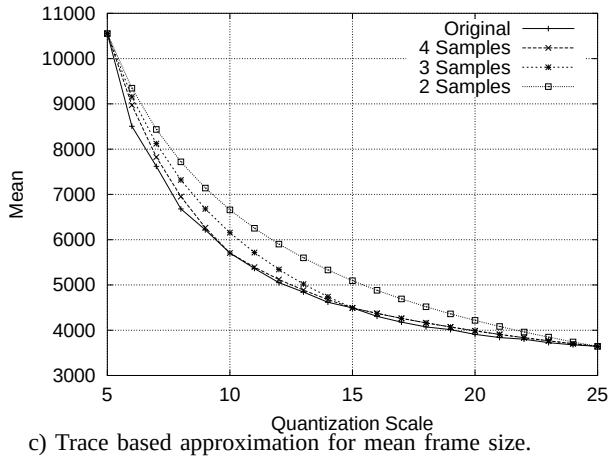
Fig. 69. *Frame* level approximation results for scene *165* (MC IV) from *Star Wars IV* .



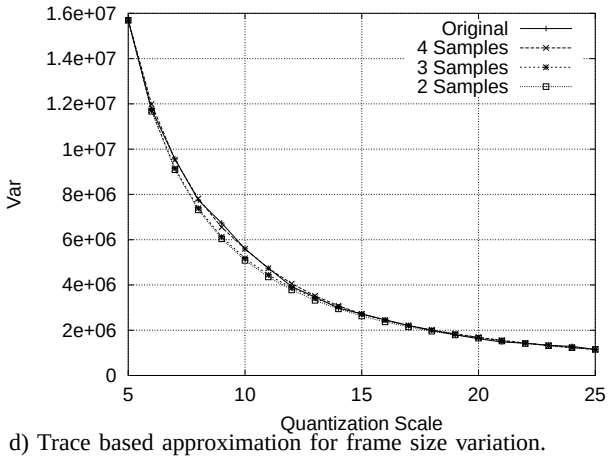
a) Trace based approximation for CoV and $PSNR$.



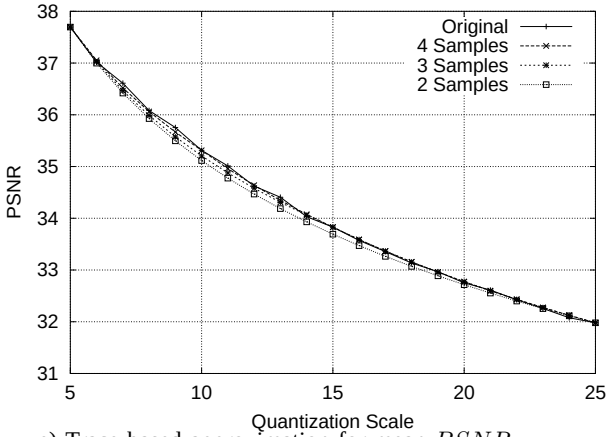
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

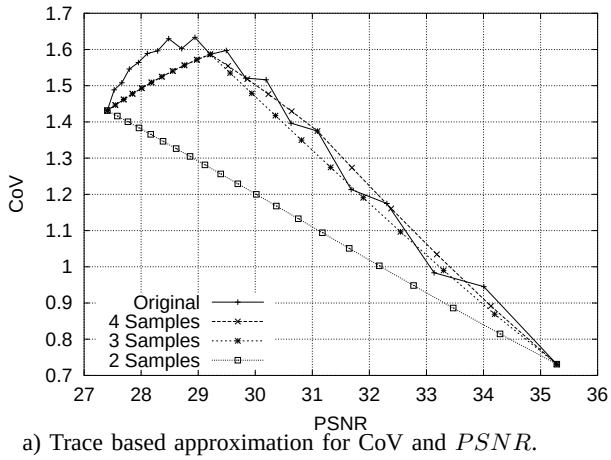


d) Trace based approximation for frame size variation.

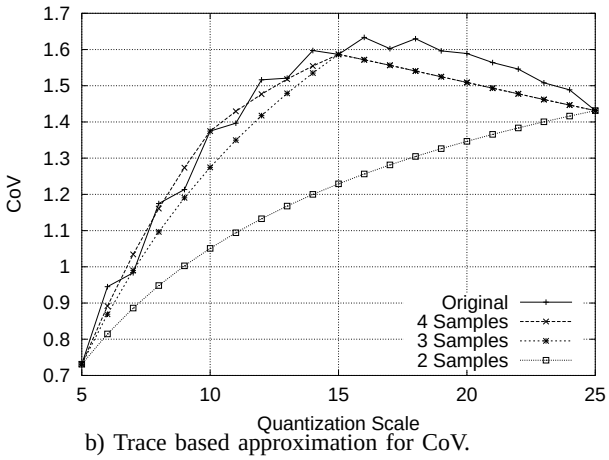


e) Trace based approximation for mean $PSNR$.

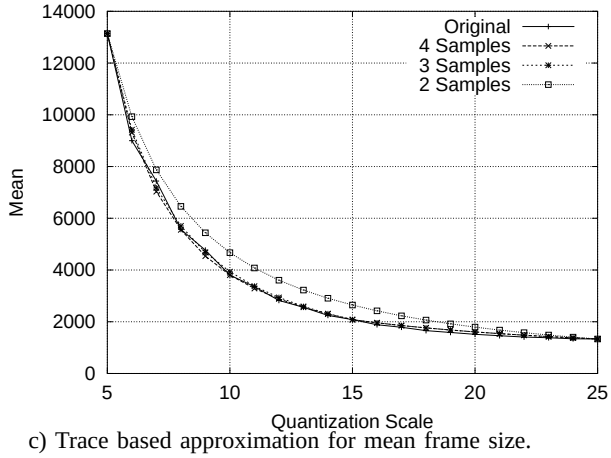
Fig. 70. *Frame* level approximation results for scene *632* (MC V) from *Star Wars IV* .



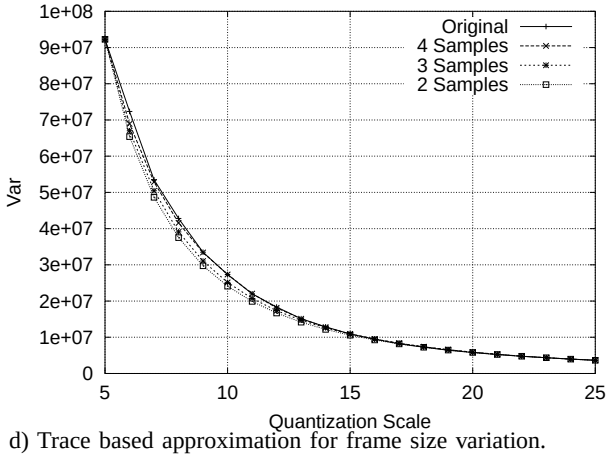
a) Trace based approximation for CoV and $PSNR$.



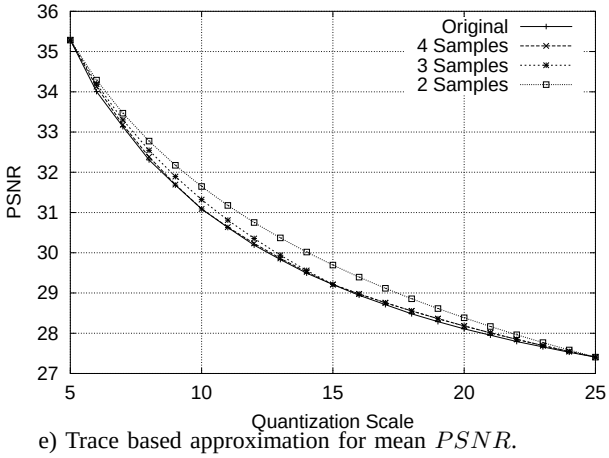
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

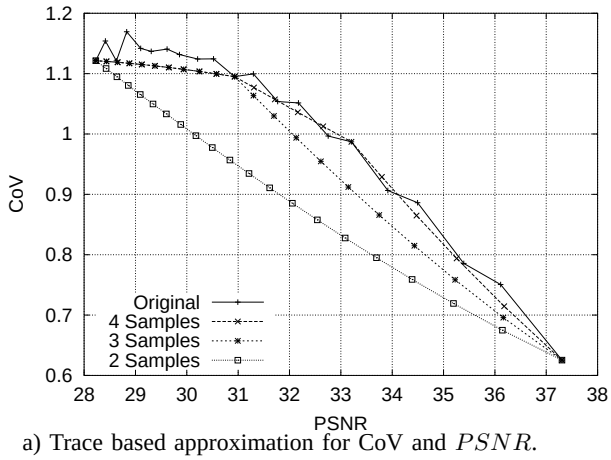


d) Trace based approximation for frame size variation.

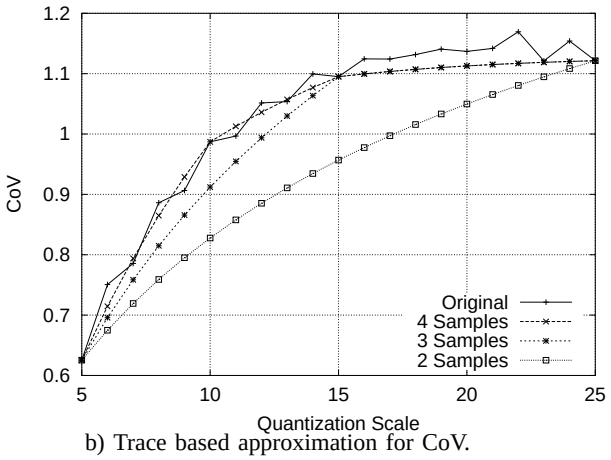


e) Trace based approximation for mean $PSNR$.

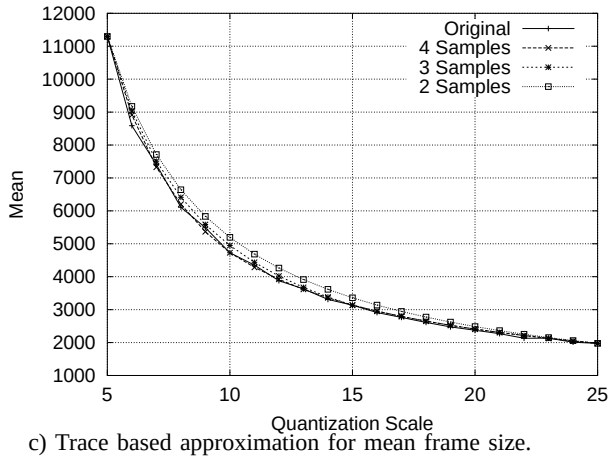
Fig. 71. *Frame* level approximation results for scene *384* (MC I) from *The Terminator* .



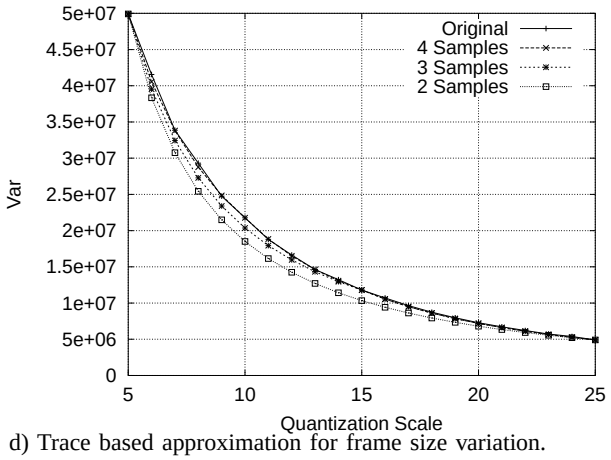
a) Trace based approximation for CoV and $PSNR$.



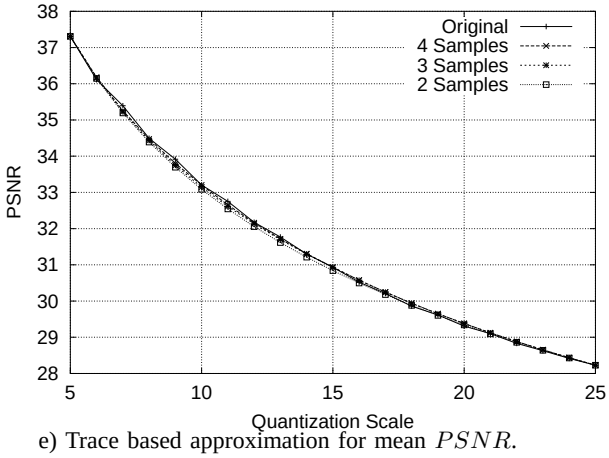
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

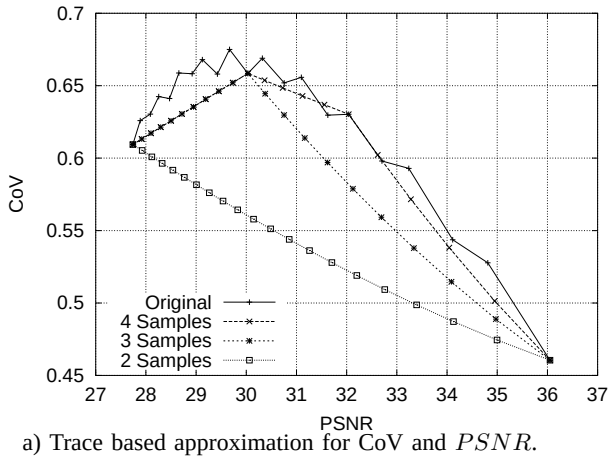


d) Trace based approximation for frame size variation.

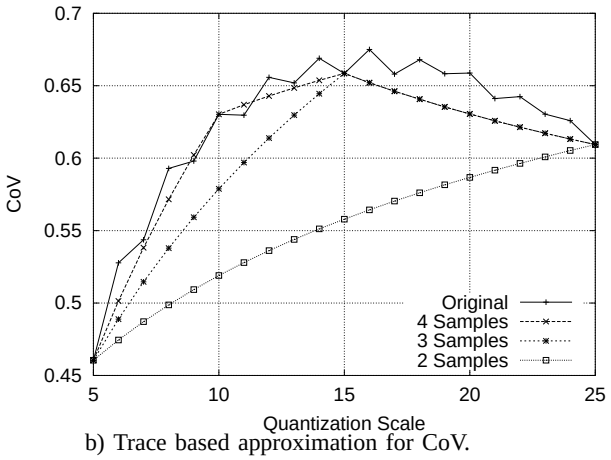


e) Trace based approximation for mean $PSNR$.

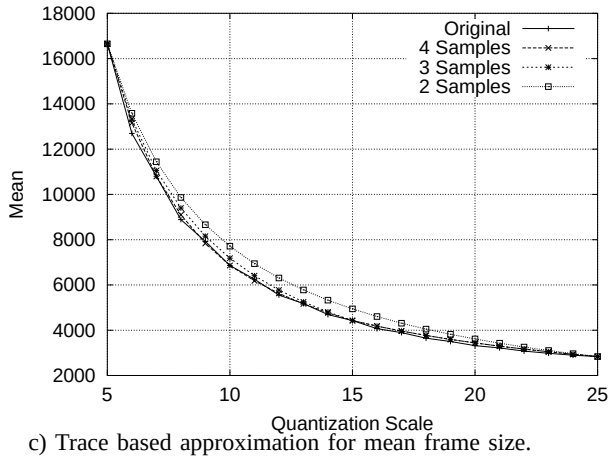
Fig. 72. *Frame* level approximation results for scene *462* (MC II) from *The Terminator* .



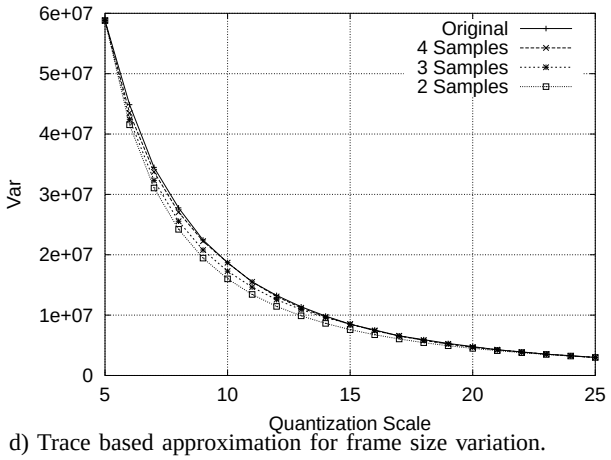
a) Trace based approximation for CoV and $PSNR$.



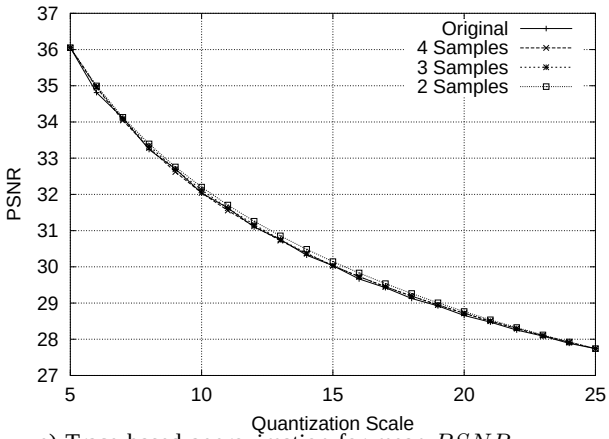
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

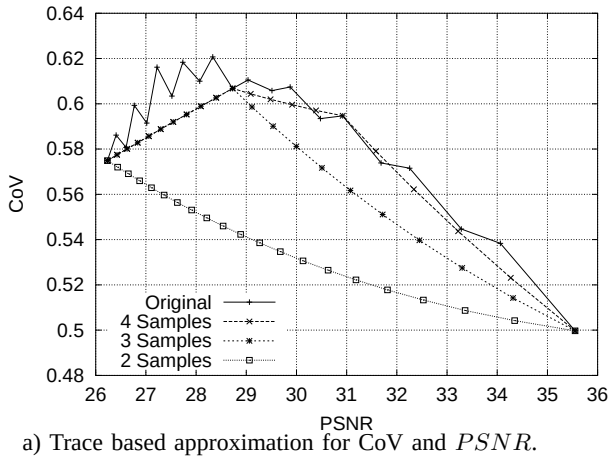


d) Trace based approximation for frame size variation.

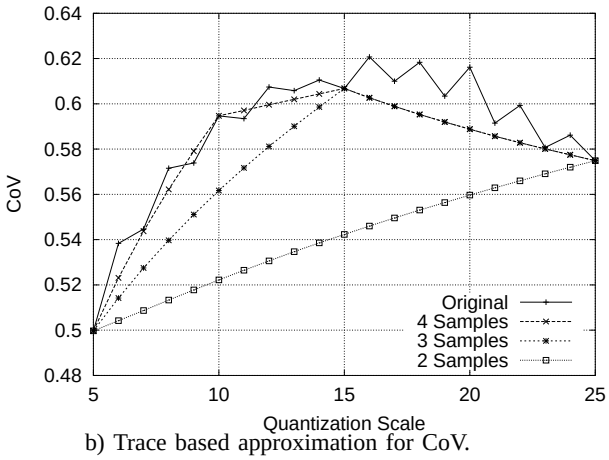


e) Trace based approximation for mean $PSNR$.

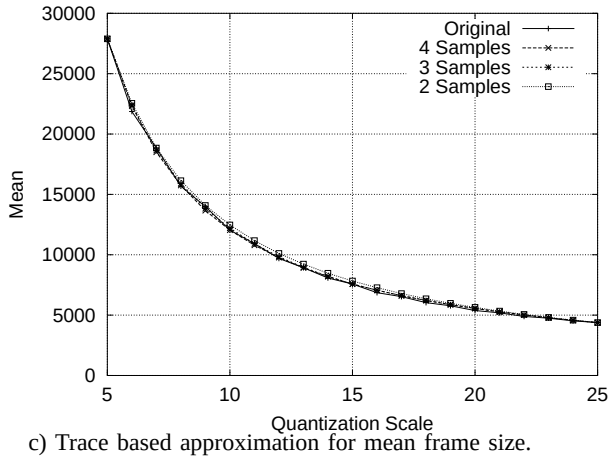
Fig. 73. *Frame* level approximation results for scene *628* (MC III) from *The Terminator* .



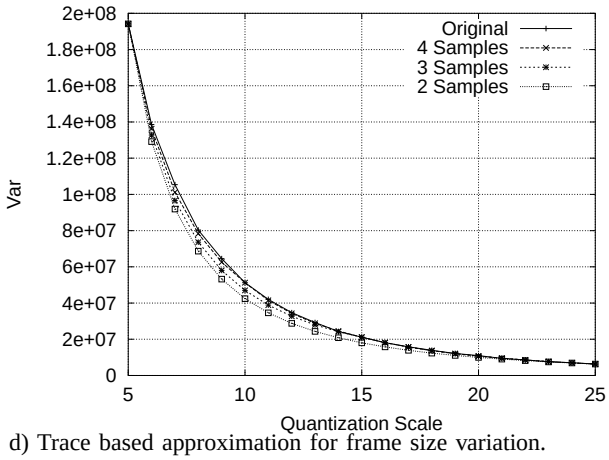
a) Trace based approximation for CoV and $PSNR$.



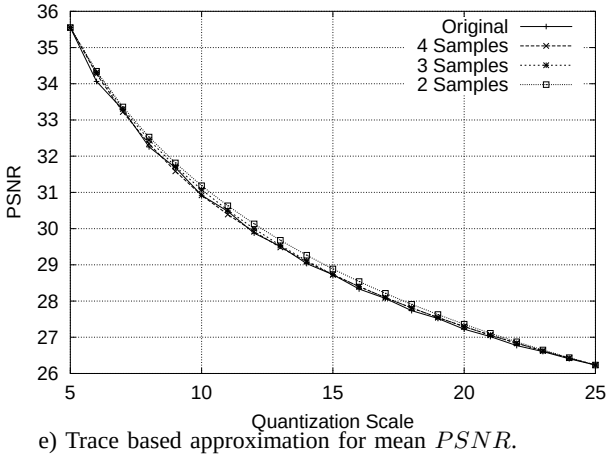
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.

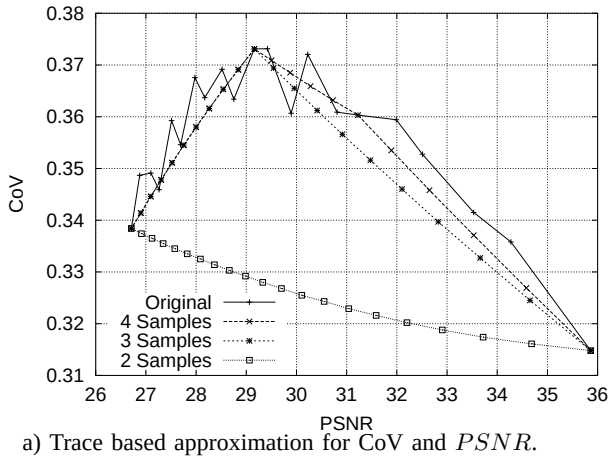


d) Trace based approximation for frame size variation.

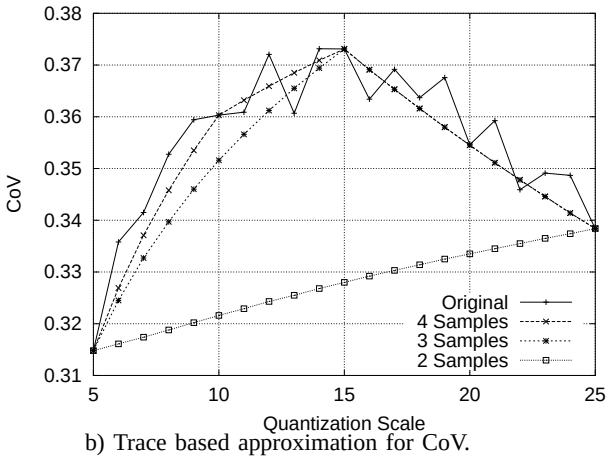


e) Trace based approximation for mean $PSNR$.

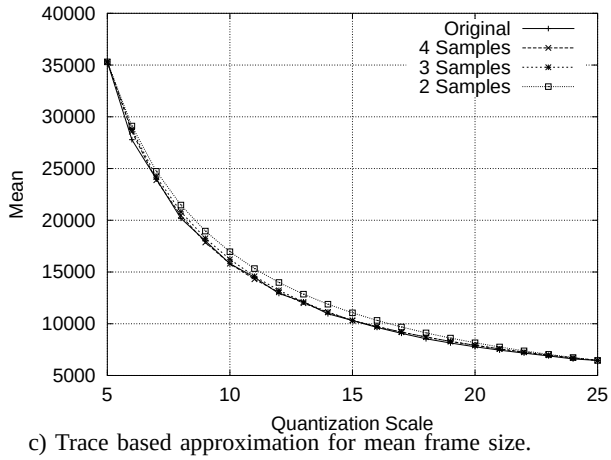
Fig. 74. *Frame* level approximation results for scene *262* (MC IV) from *The Terminator* .



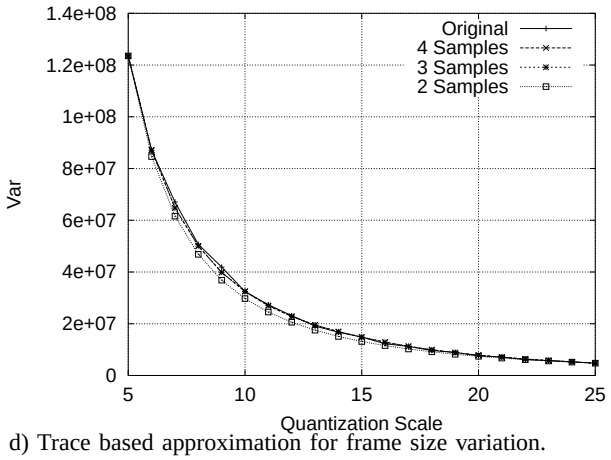
a) Trace based approximation for CoV and $PSNR$.



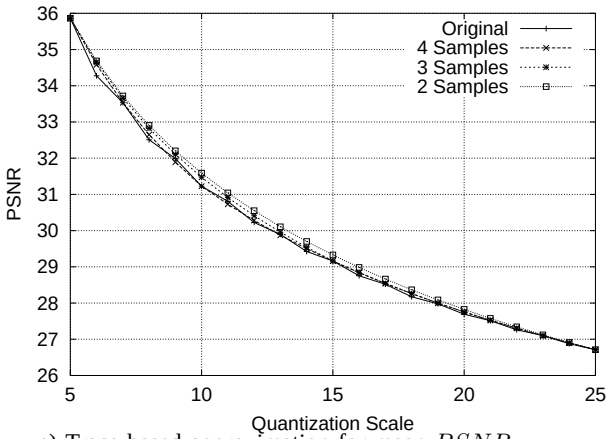
b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



d) Trace based approximation for frame size variation.



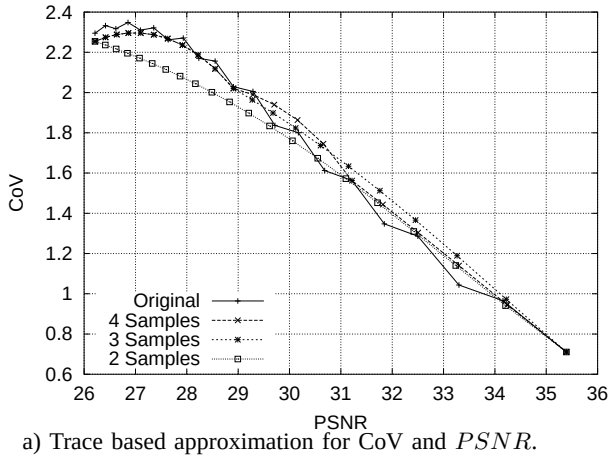
e) Trace based approximation for mean $PSNR$.

Fig. 75. *Frame* level approximation results for scene *441* (MC V) from *The Terminator* .

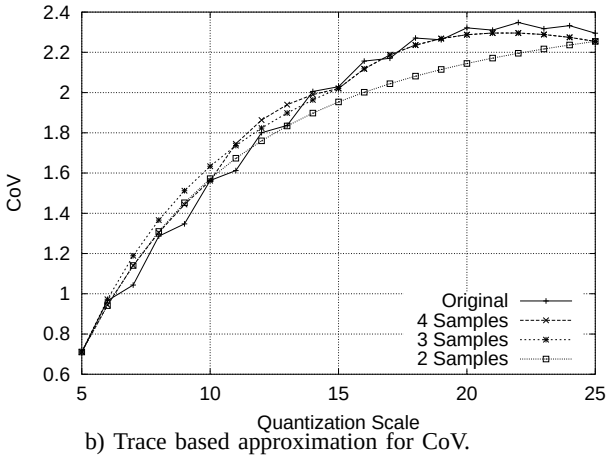
We observe that for most scenes the differences between the *aggregated* and the *frame* level approximations is only minor. In both cases, the fit between the estimated VD curve and the real VD curve is better for the region of higher PSNR values, as previously observed on the *aggregated* level. We also observe that the frame level approximation gives approximately the same performance as the aggregate level approximation. We furthermore observe that the individual statistics (i.e., frame size mean and standard deviation, and PSNR quality) are approximated very well with both aggregate and frame level approximation for a small number of samples. The approximation of the typically highly non-linear VD curve appears to pose a particular challenge.

In Figs. 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, and 90 we plot the approximation at the *macroblock* level. In addition we calculate the coefficient of correlation between the approximated and the real statistics as in Eqn. 10 and the mean squared error (MSE) for all three approximation levels. The results are given in Tables VI, VII, VIII, IX, X, XI, XII, XIII, XIV, XV, XVI, XVII, XVIII, XIX, and XX.

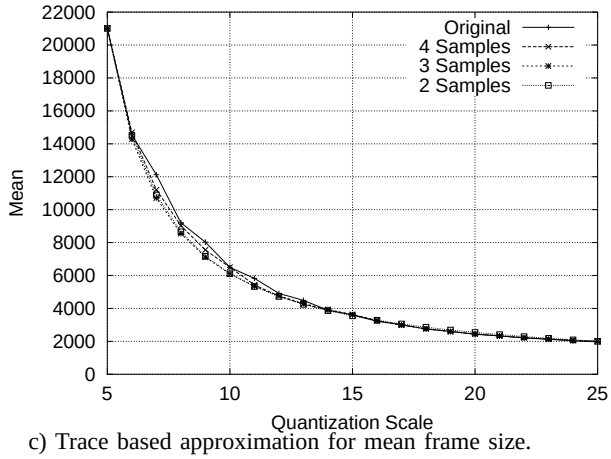
Similar to the two other levels, the VD approximation with two interpolation points gives a poor fit. We observe however, that with three interpolation points the fit becomes fairly good, and becomes even closer with four interpolation points. We conclude that the *macroblock* level yields the best approximation for a given number of sample points.



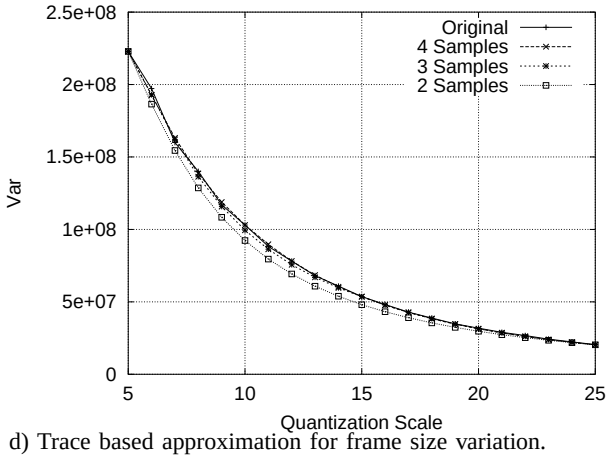
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



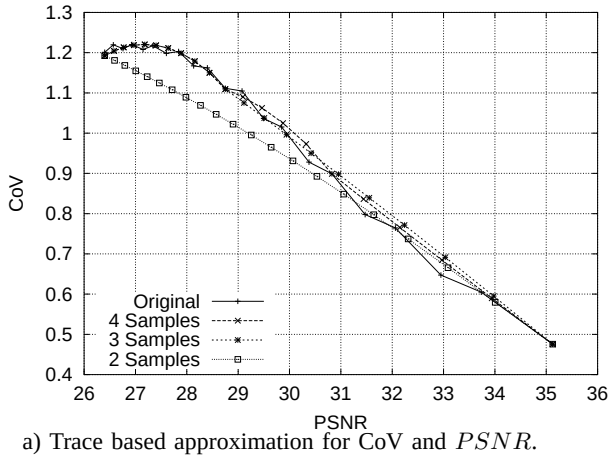
d) Trace based approximation for frame size variation.

Fig. 76. *Macroblock* level approximation results for scene *298* (MC I) from *Football* .

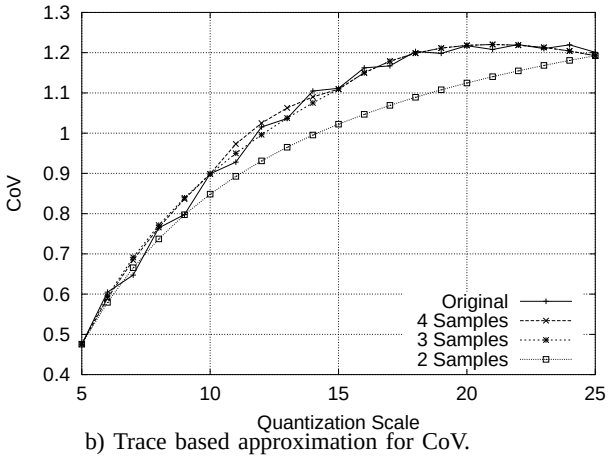
TABLE VI

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *298* WITH GOP PATTERN 7.

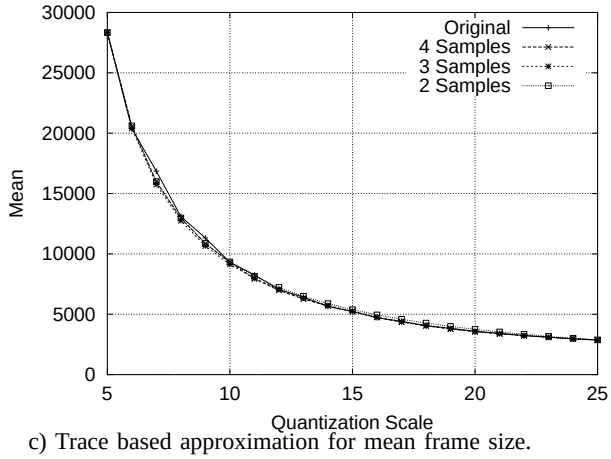
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.9078	0.9893	0.9888	0.1151	0.0121	0.0063
	σ^2	0.9786	0.9917	0.9974	2.2057e+014	7.1976e+013	1.6691e+013
	$\overline{X}$	0.9939	0.9977	0.9980	481702.6713	104399.7190	54121.5672
	$PSNR$	0.9997	0.9998	0.9999	0.0054	0.0015	0.0005
Frame	CoV	0.9821	0.9909	0.9899	0.0210	0.0046	0.0048
	σ^2	0.9953	0.9989	0.9993	5.4204e+013	9.4109e+012	3.0180e+012
	$\overline{X}$	0.9983	0.9978	0.9981	46065.0784	53460.3440	44732.8947
	$PSNR$	0.9996	0.9998	0.9999	0.0055	0.0015	0.0005
MB	CoV	0.9849	0.9885	0.9901	0.0113	0.0045	0.0031
	σ^2	0.9966	0.9994	0.9994	3.9252e+013	3.1585e+012	1.9164e+012
	$\overline{X}$	0.9953	0.9949	0.9977	147613.0230	180179.8715	65281.8491
	$PSNR$	0.9996	0.9998	0.9999	0.0055	0.0015	0.0005



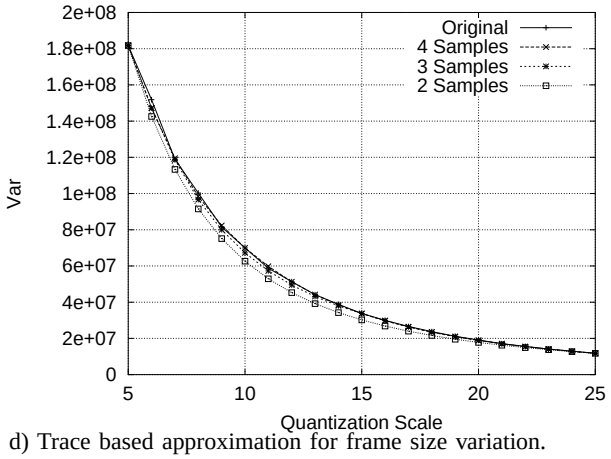
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



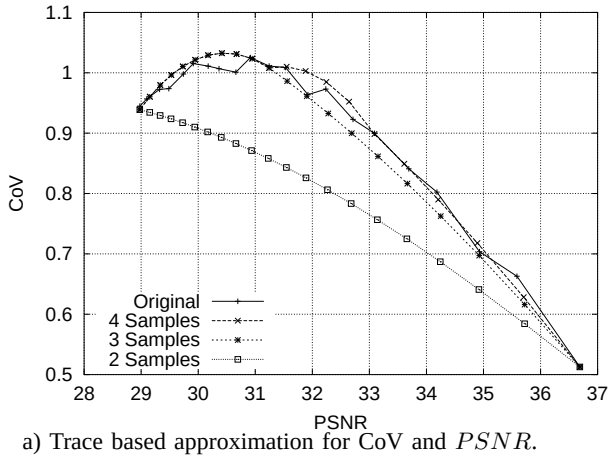
d) Trace based approximation for frame size variation.

Fig. 77. *Macroblock* level approximation results for scene *299* (MC II) from *Football* .

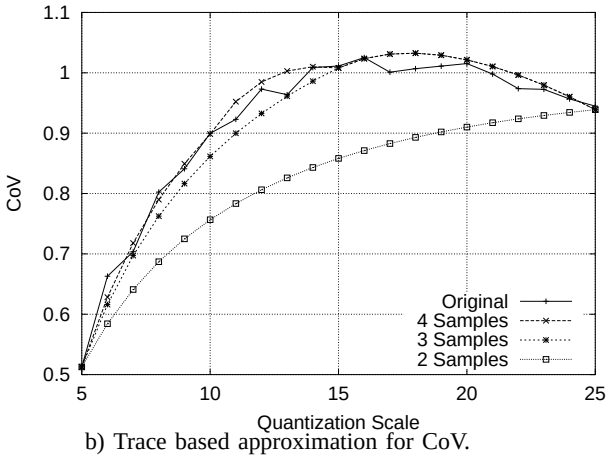
TABLE VII

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *299* WITH GOP PATTERN 7.

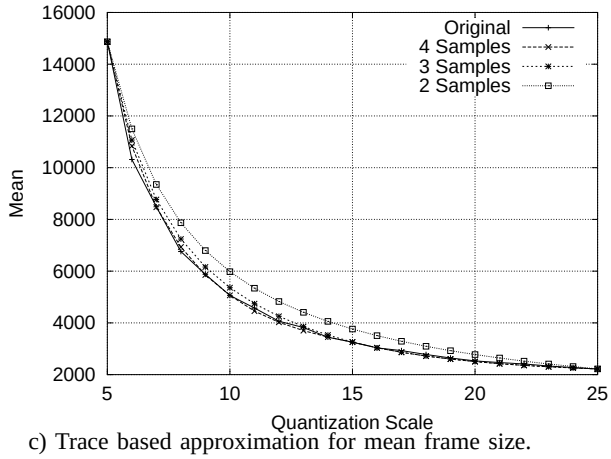
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.8929	0.9894	0.9914	0.0261	0.0024	0.0010
	σ^2	0.9864	0.9946	0.9981	8.5423e+013	2.8918e+013	7.6157e+012
	$\overline{X}$	0.9960	0.9989	0.9990	605441.5634	95010.7475	48050.0164
	$PSNR$	0.9992	0.9996	0.9996	0.0249	0.0071	0.0033
Frame	CoV	0.9520	0.9944	0.9920	0.0114	0.0009	0.0008
	σ^2	0.9955	0.9989	0.9994	3.0656e+013	5.9759e+012	1.9733e+012
	$\overline{X}$	0.9984	0.9992	0.9990	211002.9508	39973.5182	43279.8265
	$PSNR$	0.9993	0.9996	0.9996	0.0236	0.0069	0.0033
MB	CoV	0.9794	0.9946	0.9944	0.0049	0.0003	0.0003
	σ^2	0.9970	0.9995	0.9995	2.1273e+013	2.5206e+012	1.3134e+012
	$\overline{X}$	0.9987	0.9984	0.9989	64086.3566	91442.7848	57144.5720
	$PSNR$	0.9993	0.9996	0.9996	0.0236	0.0069	0.0033



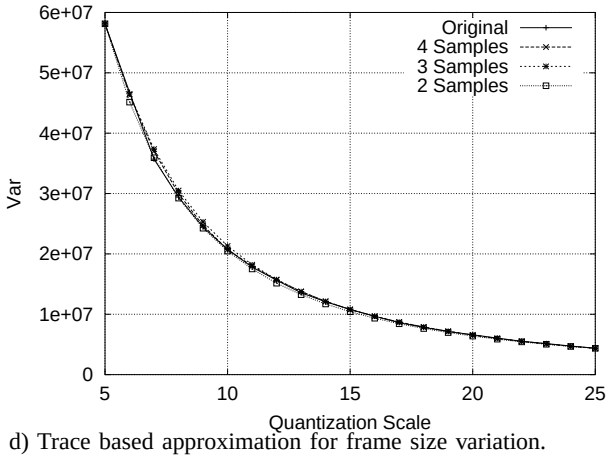
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



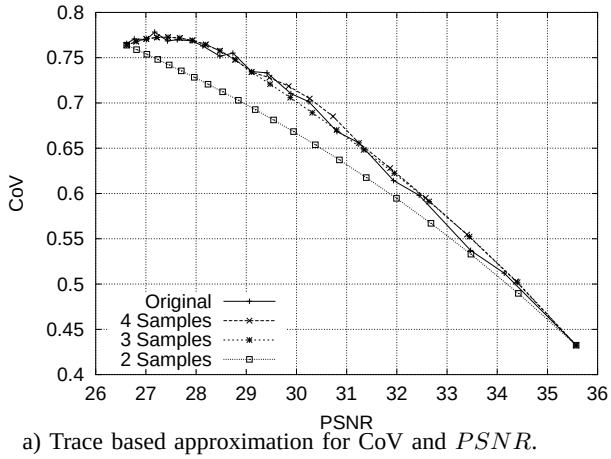
d) Trace based approximation for frame size variation.

Fig. 78. *Macroblock* level approximation results for scene *557* (MC III) from *Football* .

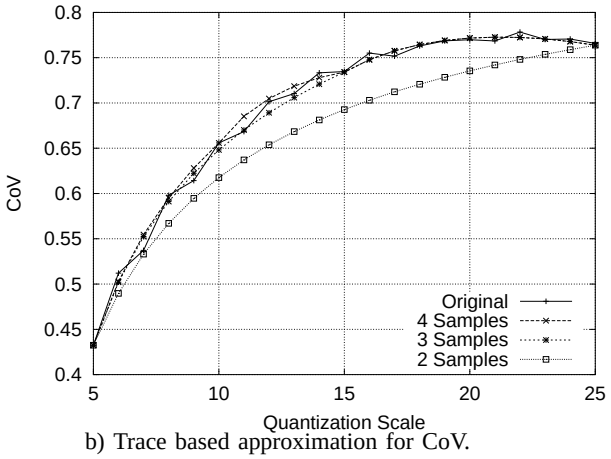
TABLE VIII

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *557* WITH GOP PATTERN 7.

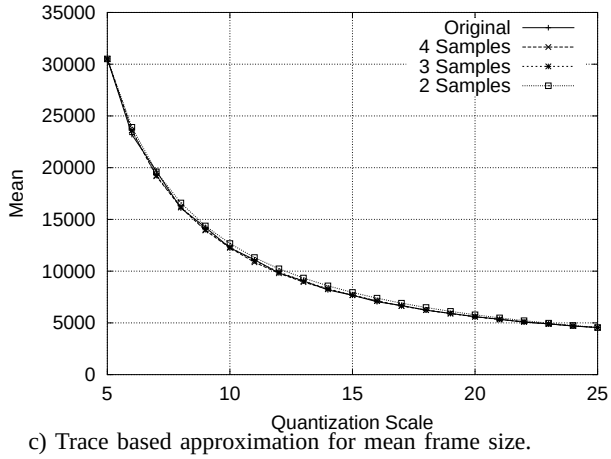
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.6538	0.9370	0.9825	0.0301	0.0036	0.0007
	σ^2	0.9974	0.9987	0.9990	1.6126e+012	5.5411e+011	3.1518e+011
	$\overline{X}$	0.9734	0.9910	0.9970	942154.5169	215401.4534	51754.4875
	$PSNR$	0.9997	0.9997	0.9996	0.0017	0.0019	0.0018
Frame	CoV	0.7209	0.9653	0.9863	0.0235	0.0019	0.0005
	σ^2	0.9988	0.9994	0.9995	6.3659e+011	1.3570e+011	1.0603e+011
	$\overline{X}$	0.9773	0.9934	0.9977	795969.9326	154592.9675	37575.2991
	$PSNR$	0.9997	0.9997	0.9996	0.0017	0.0019	0.0018
MB	CoV	0.8592	0.9797	0.9893	0.0124	0.0005	0.0003
	σ^2	0.9995	0.9991	0.9995	1.8763e+011	2.5344e+011	1.1353e+011
	$\overline{X}$	0.9891	0.9972	0.9987	373890.7352	53796.4218	16632.0050
	$PSNR$	0.9997	0.9997	0.9996	0.0017	0.0019	0.0018



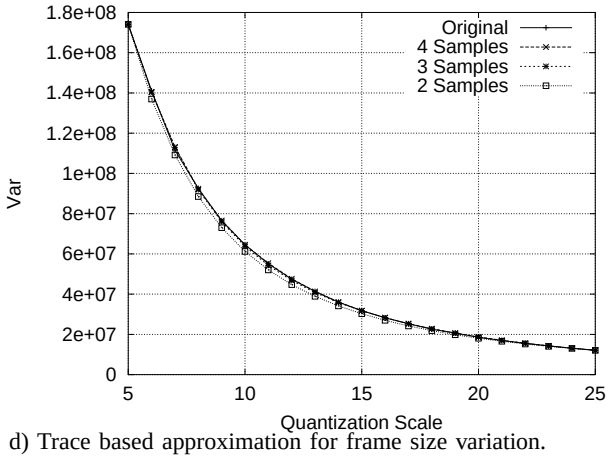
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



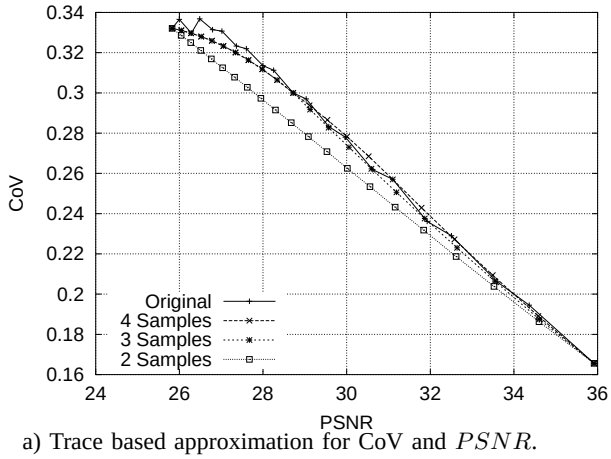
d) Trace based approximation for frame size variation.

Fig. 79. *Macroblock* level approximation results for scene *184* (MC IV) from *Football* .

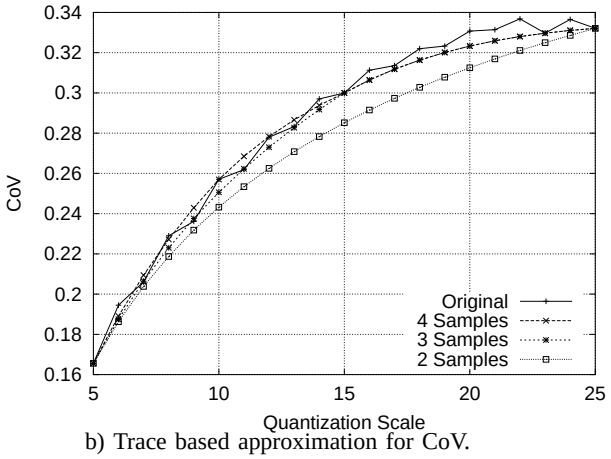
TABLE IX

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *184* WITH GOP PATTERN 7.

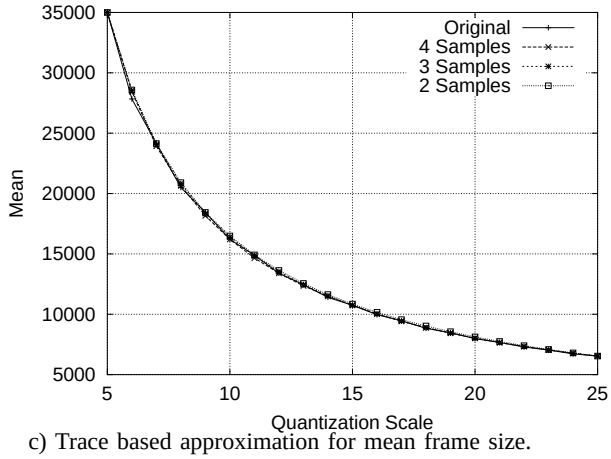
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.8798	0.9793	0.9911	0.0058	0.0008	0.0002
	σ^2	0.9926	0.9966	0.9990	4.0099e+013	1.5463e+013	3.8768e+012
	$\overline{X}$	0.9967	0.9989	0.9993	566049.5371	123830.0066	44081.8381
	$PSNR$	0.9993	0.9994	0.9993	0.0121	0.0068	0.0051
Frame	CoV	0.9311	0.9915	0.9924	0.0033	0.0003	0.0001
	σ^2	0.9971	0.9994	0.9998	1.6103e+013	3.1519e+012	6.1213e+011
	$\overline{X}$	0.9977	0.9992	0.9993	388284.7900	74254.9970	34981.5910
	$PSNR$	0.9993	0.9994	0.9993	0.0115	0.0066	0.0050
MB	CoV	0.9773	0.9953	0.9947	0.0011	0.0000	0.0001
	σ^2	0.9992	0.9999	0.9999	4.5641e+012	2.4823e+011	1.4959e+011
	$\overline{X}$	0.9993	0.9995	0.9994	86311.6645	22459.4637	26337.0295
	$PSNR$	0.9993	0.9994	0.9993	0.0115	0.0066	0.0050



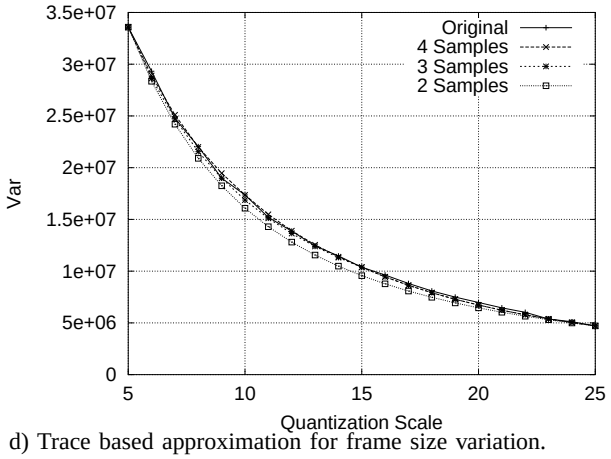
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



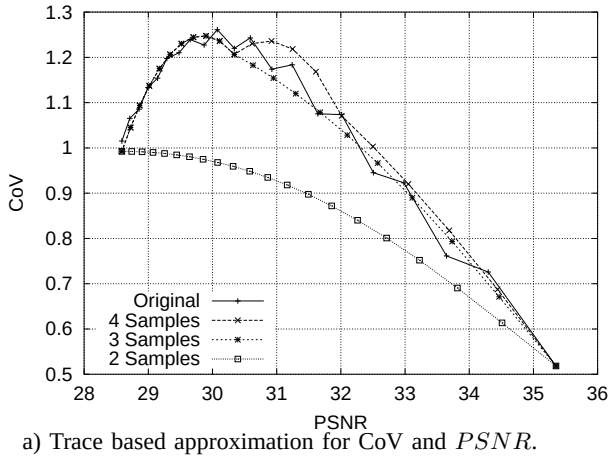
d) Trace based approximation for frame size variation.

Fig. 80. *Macroblock* level approximation results for scene *336* (MC V) from *Football* .

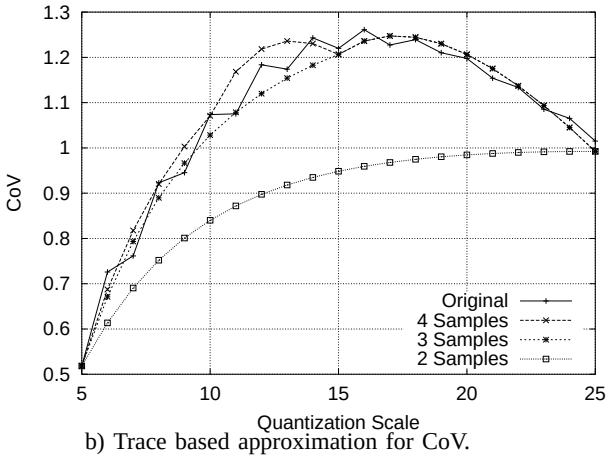
TABLE X

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *336* WITH GOP PATTERN 7.

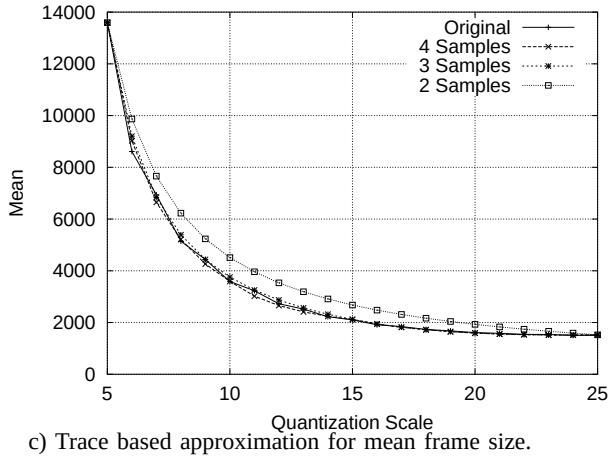
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.9487	0.9919	0.9910	0.0007	0.0001	0.0000
	σ^2	0.9873	0.9970	0.9989	3.1275e+012	5.8393e+011	1.0965e+011
	$\overline{X}$	0.9989	0.9994	0.9994	246785.5237	80152.4159	39786.7247
	$PSNR$	0.9994	0.9994	0.9993	0.0051	0.0059	0.0055
Frame	CoV	0.9711	0.9937	0.9912	0.0004	0.0001	0.0000
	σ^2	0.9953	0.9991	0.9993	1.2627e+012	1.6619e+011	5.6977e+010
	$\overline{X}$	0.9990	0.9994	0.9994	213749.5419	70513.2477	37900.8983
	$PSNR$	0.9994	0.9994	0.9993	0.0051	0.0059	0.0055
MB	CoV	0.9876	0.9963	0.9943	0.0002	0.0000	0.0000
	σ^2	0.9980	0.9996	0.9994	5.5510e+011	6.4481e+010	4.3498e+010
	$\overline{X}$	0.9995	0.9996	0.9995	50923.9406	33257.4683	28794.1210
	$PSNR$	0.9994	0.9994	0.9993	0.0051	0.0059	0.0055



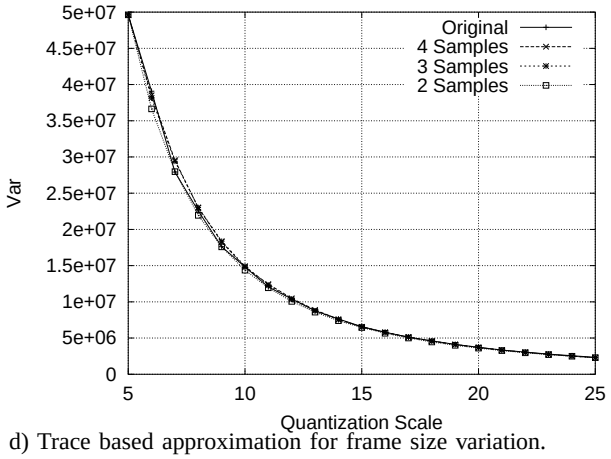
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



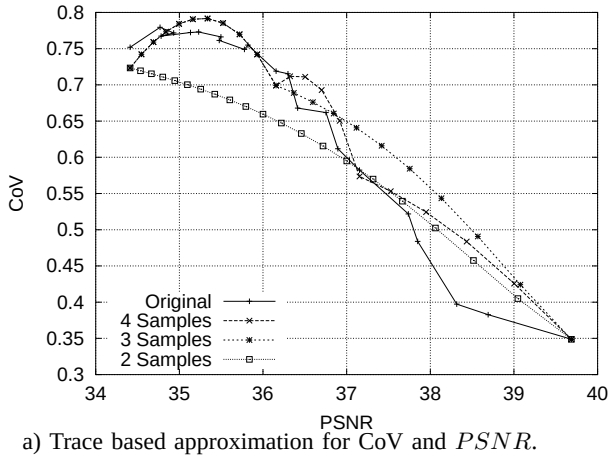
d) Trace based approximation for frame size variation.

Fig. 81. *Macroblock* level approximation results for scene *274* (MC I) from *Star Wars IV* .

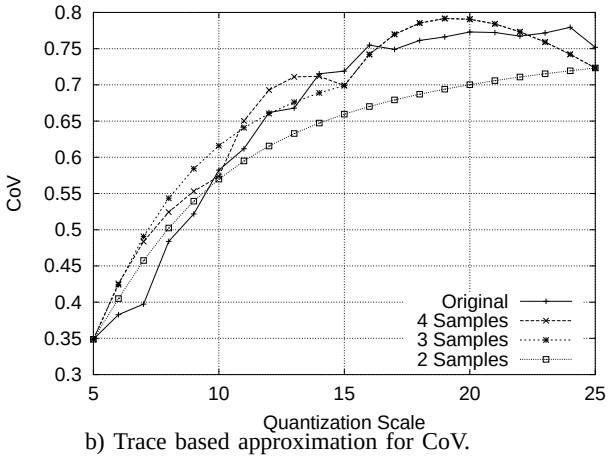
TABLE XI

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *274* WITH GOP PATTERN 7.

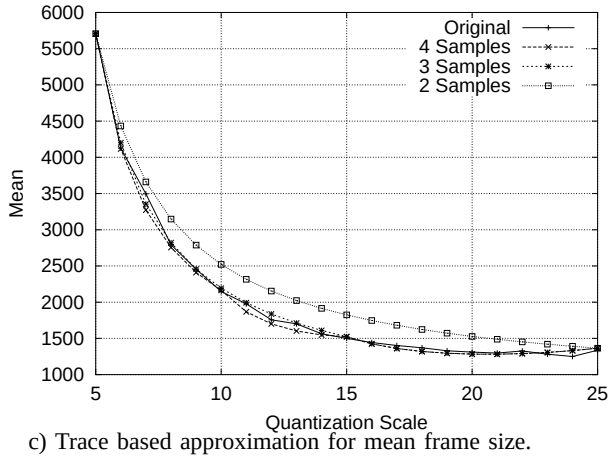
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.4848	0.9314	0.9576	0.0901	0.0098	0.0033
	σ^2	0.9959	0.9970	0.9980	1.5017e+012	9.1510e+011	4.5321e+011
	$\overline{X}$	0.9615	0.9899	0.9958	1220250.9867	207923.3227	57131.8500
	$PSNR$	0.9981	0.9997	0.9997	0.0359	0.0039	0.0019
Frame	CoV	0.5755	0.9624	0.9593	0.0735	0.0057	0.0028
	σ^2	0.9974	0.9987	0.9988	8.3190e+011	2.4998e+011	1.8991e+011
	$\overline{X}$	0.9681	0.9930	0.9967	1002019.4664	138130.3653	41558.3004
	$PSNR$	0.9981	0.9997	0.9997	0.0357	0.0039	0.0019
MB	CoV	0.8196	0.9786	0.9725	0.0414	0.0009	0.0012
	σ^2	0.9986	0.9988	0.9990	3.2765e+011	1.9081e+011	1.8106e+011
	$\overline{X}$	0.9879	0.9980	0.9982	395256.9521	23209.2210	17175.2000
	$PSNR$	0.9981	0.9997	0.9997	0.0357	0.0039	0.0019



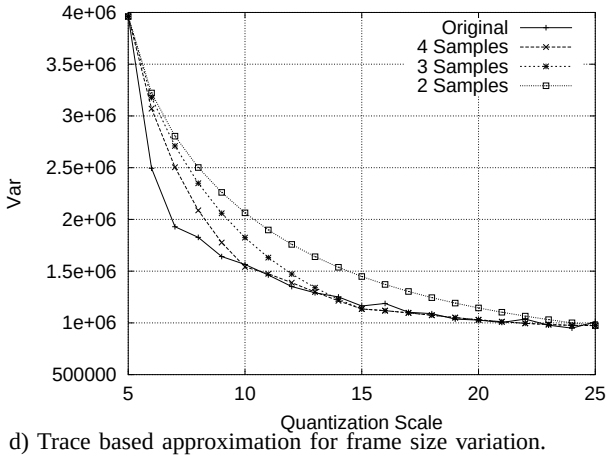
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



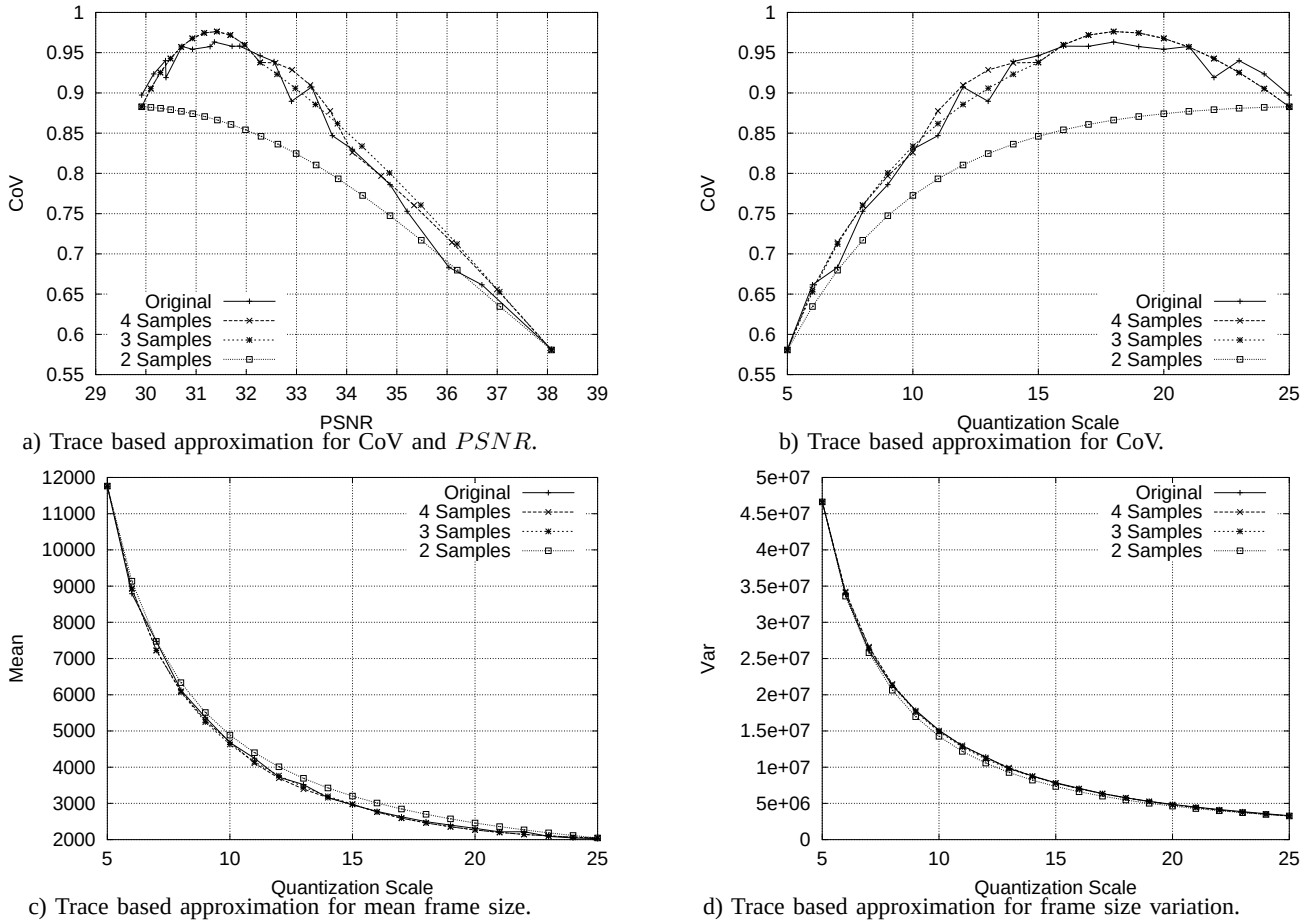
d) Trace based approximation for frame size variation.

Fig. 82. *Macroblock* level approximation results for scene *117* (MC II) from *Star Wars IV* .

TABLE XII

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *117* WITH GOP PATTERN 7.

Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.8837	0.9892	0.9900	0.0086	0.0007	0.0006
	σ^2	0.8716	0.9237	0.9521	2.4377e+011	8.4185e+010	4.0639e+010
	$\overline{X}$	0.9368	0.9889	0.9969	353409.5823	39689.7312	8060.1276
	$PSNR$	0.9936	0.9928	0.9938	0.0267	0.0306	0.0169
Frame	CoV	0.9115	0.9899	0.9880	0.0068	0.0007	0.0007
	σ^2	0.9113	0.9375	0.9648	1.6416e+011	6.4205e+010	2.7618e+010
	$\overline{X}$	0.9540	0.9924	0.9975	258184.9568	26529.8308	5737.2708
	$PSNR$	0.9936	0.9929	0.9938	0.0267	0.0303	0.0169
MB	CoV	0.9752	0.9587	0.9643	0.0029	0.0012	0.0010
	σ^2	0.9037	0.9236	0.9552	1.5142e+011	7.8475e+010	3.6313e+010
	$\overline{X}$	0.9905	0.9980	0.9972	73110.5820	2575.3346	4554.2534
	$PSNR$	0.9936	0.9929	0.9938	0.0267	0.0303	0.0169



a) Trace based approximation for CoV and $PSNR$.

b) Trace based approximation for CoV.

c) Trace based approximation for mean frame size.

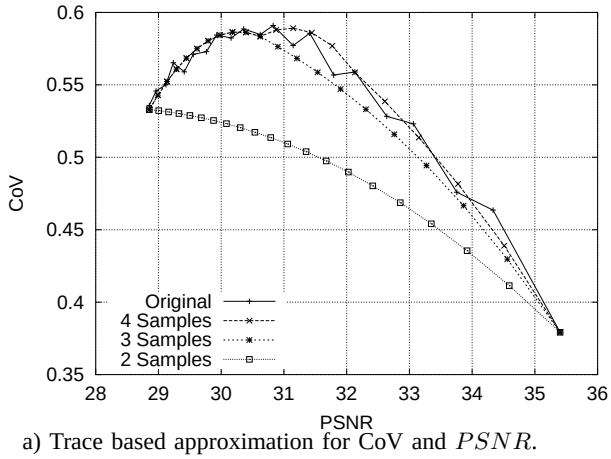
d) Trace based approximation for frame size variation.

Fig. 83. *Macroblock* level approximation results for scene *115* (MC III) from *Star Wars IV* .

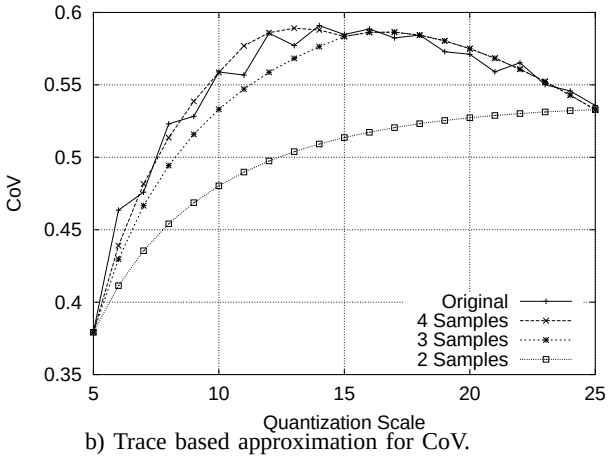
TABLE XIII

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *115* WITH GOP PATTERN 7.

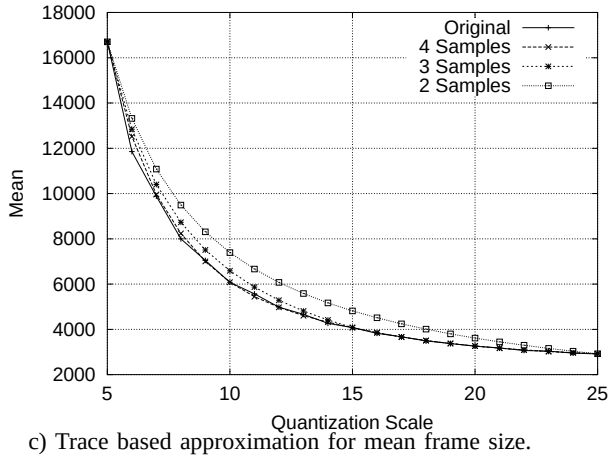
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.7688	0.9812	0.9795	0.0132	0.0008	0.0005
	σ^2	0.9996	0.9995	0.9996	7.9676e+010	1.0035e+011	8.8578e+010
	$\overline{X}$	0.9882	0.9978	0.9987	312308.0059	35193.5517	13160.4002
	$PSNR$	0.9984	0.9985	0.9985	0.0178	0.0163	0.0094
Frame	CoV	0.7916	0.9825	0.9796	0.0121	0.0007	0.0005
	σ^2	0.9993	0.9999	0.9999	2.3450e+011	2.0469e+010	1.6506e+010
	$\overline{X}$	0.9923	0.9986	0.9989	205885.3443	19556.6460	9467.4826
	$PSNR$	0.9984	0.9985	0.9985	0.0173	0.0161	0.0093
MB	CoV	0.9456	0.9825	0.9803	0.0048	0.0002	0.0003
	σ^2	0.9994	0.9999	0.9998	2.3388e+011	2.2076e+010	2.7925e+010
	$\overline{X}$	0.9987	0.9991	0.9991	34332.0665	6652.9219	6574.7703
	$PSNR$	0.9984	0.9985	0.9985	0.0173	0.0161	0.0093



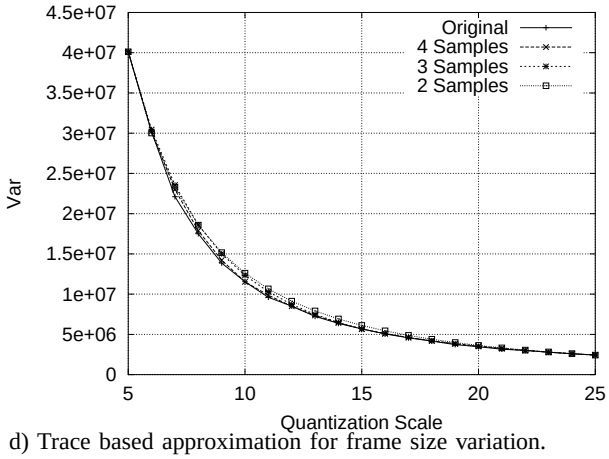
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



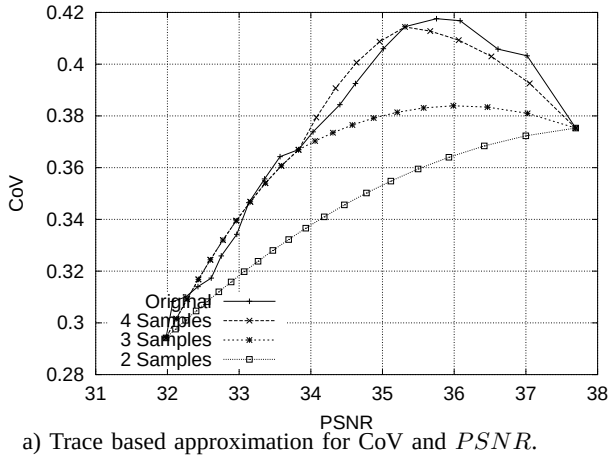
d) Trace based approximation for frame size variation.

Fig. 84. *Macroblock* level approximation results for scene *165* (MC IV) from *Star Wars IV* .

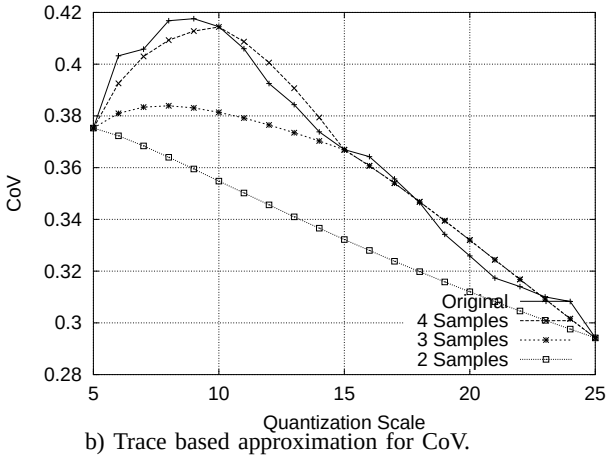
TABLE XIV

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *165* WITH GOP PATTERN 7.

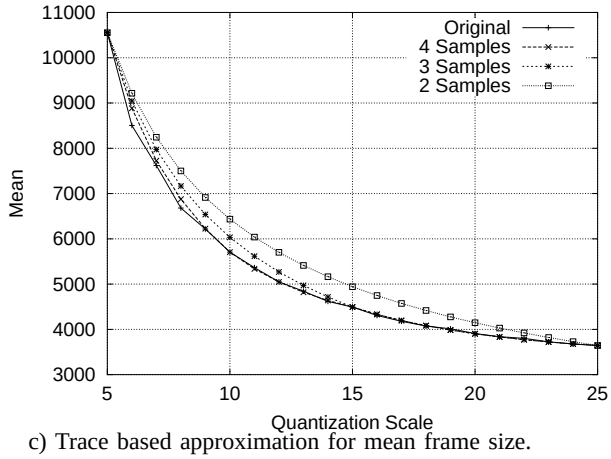
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.4541	0.8635	0.9571	0.0070	0.0011	0.0002
	σ^2	0.9992	0.9994	0.9994	1.1809e+011	7.7145e+010	8.7895e+010
	$\overline{X}$	0.9707	0.9909	0.9971	1336948.0082	284345.3317	61978.5827
	$PSNR$	0.9982	0.9992	0.9995	0.0324	0.0097	0.0027
Frame	CoV	0.5032	0.9099	0.9643	0.0062	0.0007	0.0002
	σ^2	0.9989	0.9991	0.9997	1.6882e+011	1.1682e+011	3.4538e+010
	$\overline{X}$	0.9725	0.9919	0.9974	1251289.0775	249873.2863	54061.5554
	$PSNR$	0.9982	0.9992	0.9995	0.0322	0.0097	0.0027
MB	CoV	0.7475	0.9520	0.9755	0.0032	0.0002	0.0001
	σ^2	0.9980	0.9983	0.9996	3.9319e+011	2.8365e+011	5.5770e+010
	$\overline{X}$	0.9839	0.9956	0.9985	727088.4480	118462.1985	25328.8165
	$PSNR$	0.9982	0.9992	0.9995	0.0322	0.0097	0.0027



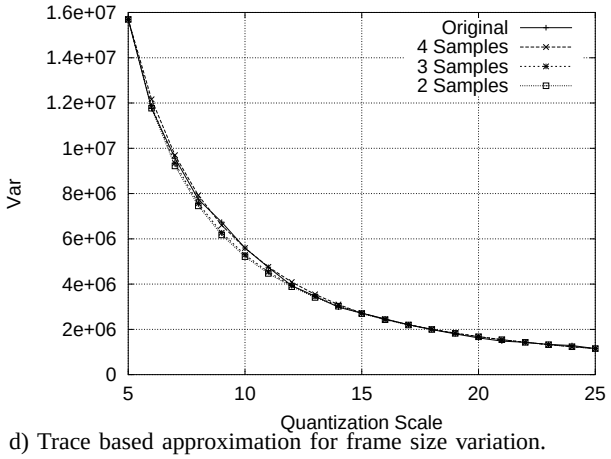
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



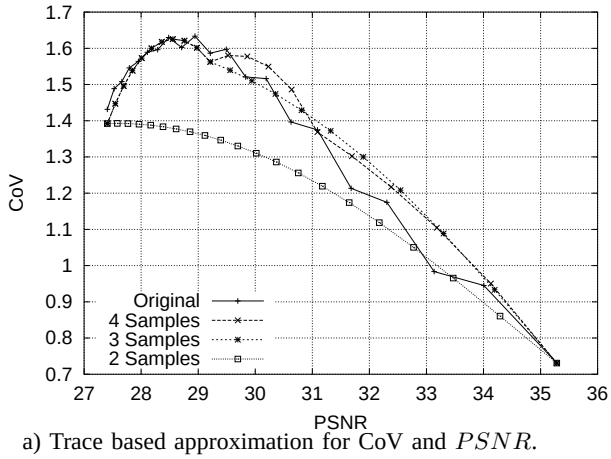
d) Trace based approximation for frame size variation.

Fig. 85. *Macroblock* level approximation results for scene *632* (MC V) from *Star Wars IV* .

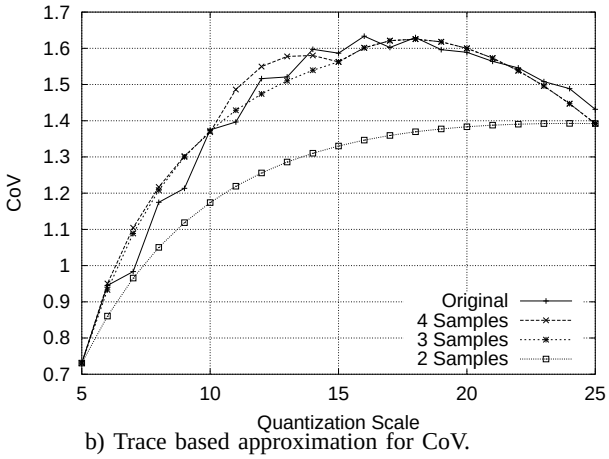
TABLE XV

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *632* WITH GOP PATTERN 7.

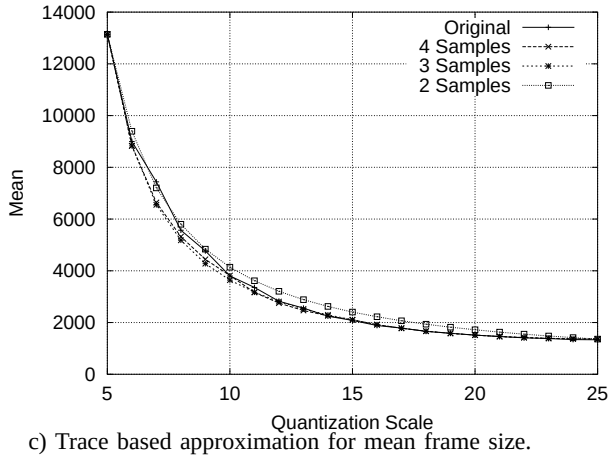
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.6888	0.8913	0.9828	0.0022	0.0005	0.0001
	σ^2	0.9974	0.9980	0.9996	6.5688e+010	4.2216e+010	5.6635e+009
	$\overline{X}$	0.9710	0.9911	0.9972	408099.5896	86594.2989	19148.0985
	$PSNR$	0.9979	0.9989	0.9996	0.0167	0.0051	0.0015
Frame	CoV	0.6971	0.8855	0.9831	0.0021	0.0005	0.0000
	σ^2	0.9974	0.9979	0.9996	6.6320e+010	4.6078e+010	6.3330e+009
	$\overline{X}$	0.9719	0.9911	0.9973	394595.8025	85313.6359	18352.9577
	$PSNR$	0.9978	0.9989	0.9996	0.0170	0.0052	0.0015
MB	CoV	0.8594	0.9364	0.9843	0.0012	0.0003	0.0000
	σ^2	0.9983	0.9987	0.9994	3.6286e+010	2.2167e+010	1.3666e+010
	$\overline{X}$	0.9828	0.9943	0.9982	235489.4584	47641.9119	9435.5042
	$PSNR$	0.9978	0.9989	0.9996	0.0170	0.0052	0.0015



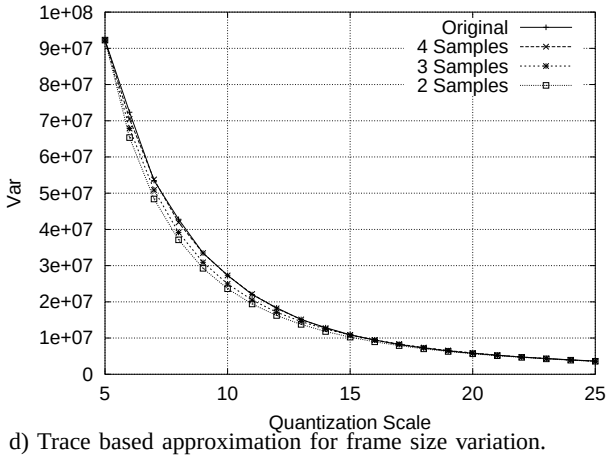
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



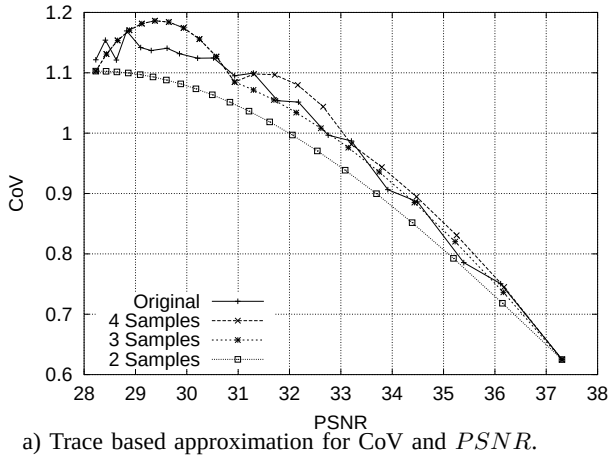
d) Trace based approximation for frame size variation.

Fig. 86. *Macroblock* level approximation results for scene *384* (MC I) from *The Terminator* .

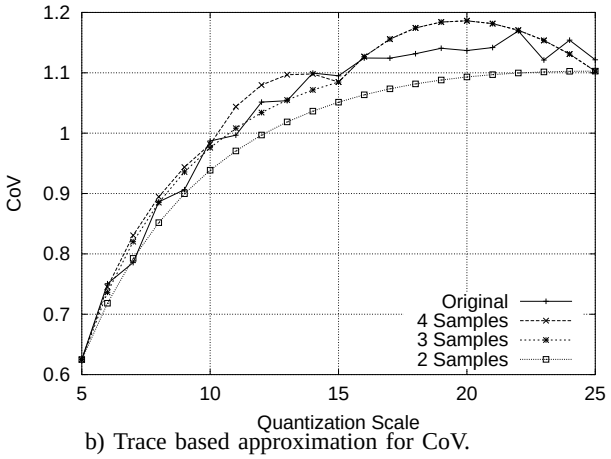
TABLE XVI

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *384* WITH GOP PATTERN 7.

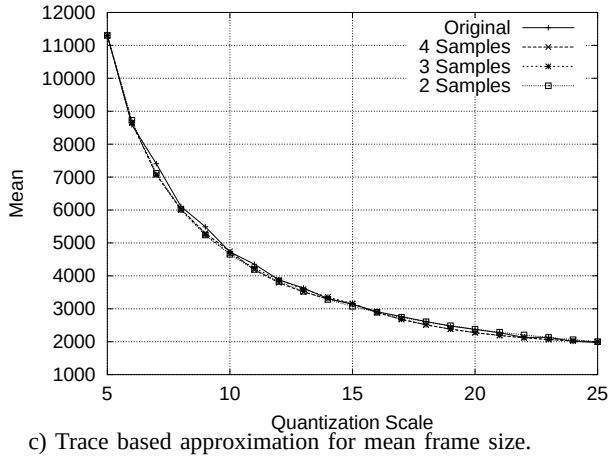
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.6515	0.9629	0.9757	0.1002	0.0091	0.0032
	σ^2	0.9914	0.9937	0.9981	1.0797e+013	7.4575e+012	1.9101e+012
	$\overline{X}$	0.9858	0.9977	0.9981	476980.0219	42191.9574	19863.8921
	$PSNR$	0.9935	0.9991	0.9998	0.1508	0.0154	0.0026
Frame	CoV	0.7608	0.9844	0.9764	0.0686	0.0038	0.0025
	σ^2	0.9949	0.9974	0.9993	6.3397e+012	3.0954e+012	6.5636e+011
	$\overline{X}$	0.9911	0.9985	0.9980	299785.4444	17328.2805	18236.9374
	$PSNR$	0.9935	0.9991	0.9998	0.1504	0.0154	0.0026
MB	CoV	0.9203	0.9773	0.9718	0.0359	0.0015	0.0021
	σ^2	0.9944	0.9978	0.9997	7.3306e+012	2.7035e+012	2.0951e+011
	$\overline{X}$	0.9975	0.9952	0.9966	66032.9372	60897.6148	42339.3178
	$PSNR$	0.9935	0.9991	0.9998	0.1504	0.0154	0.0026



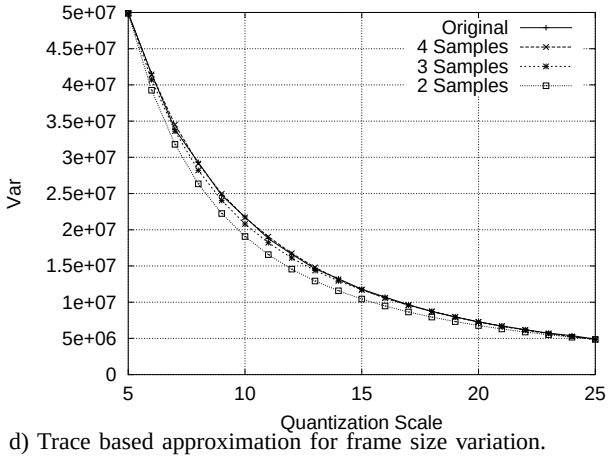
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



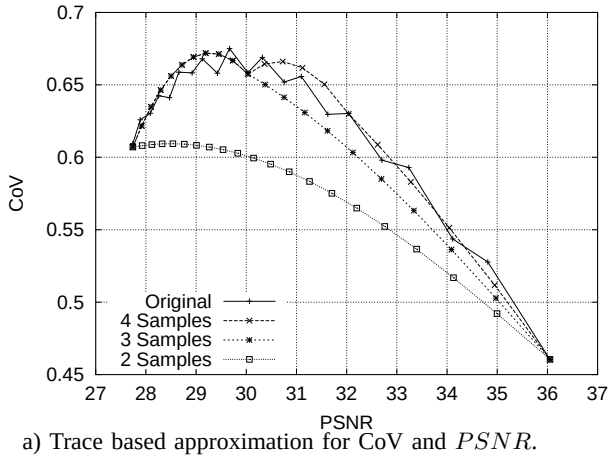
d) Trace based approximation for frame size variation.

Fig. 87. *Macroblock* level approximation results for scene *462* (MC II) from *The Terminator* .

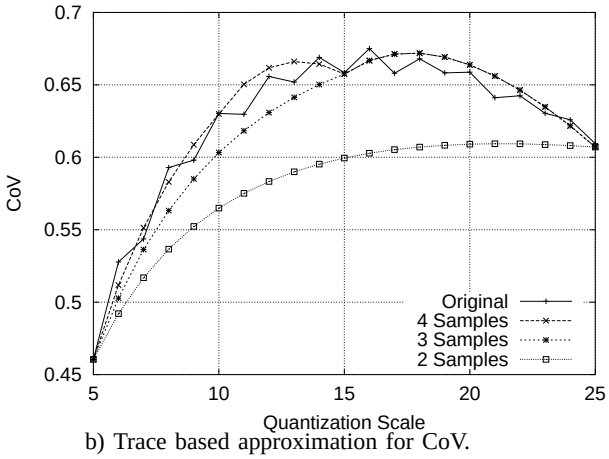
TABLE XVII

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *462* WITH GOP PATTERN 7.

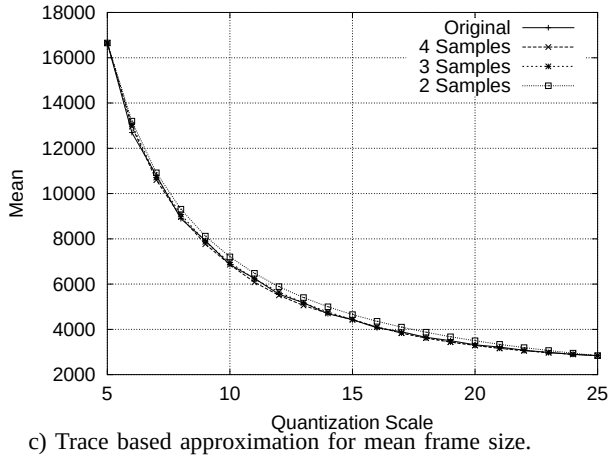
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.8515	0.9711	0.9844	0.0170	0.0023	0.0006
	σ^2	0.9914	0.9973	0.9994	4.1375e+012	1.0445e+012	1.5292e+011
	$\overline{X}$	0.9939	0.9976	0.9987	136010.1819	35774.7426	10386.4111
	$PSNR$	0.9992	0.9994	0.9996	0.0103	0.0055	0.0034
Frame	CoV	0.8896	0.9811	0.9856	0.0126	0.0015	0.0005
	σ^2	0.9918	0.9981	0.9997	3.9325e+012	7.4172e+011	7.1325e+010
	$\overline{X}$	0.9962	0.9985	0.9989	81502.6656	19671.8690	7631.6823
	$PSNR$	0.9992	0.9994	0.9996	0.0104	0.0055	0.0034
MB	CoV	0.9887	0.9753	0.9755	0.0019	0.0007	0.0009
	σ^2	0.9952	0.9994	0.9998	2.5147e+012	2.0404e+011	3.2223e+010
	$\overline{X}$	0.9984	0.9985	0.9987	10894.1578	12484.3441	12304.3418
	$PSNR$	0.9992	0.9994	0.9996	0.0104	0.0055	0.0034



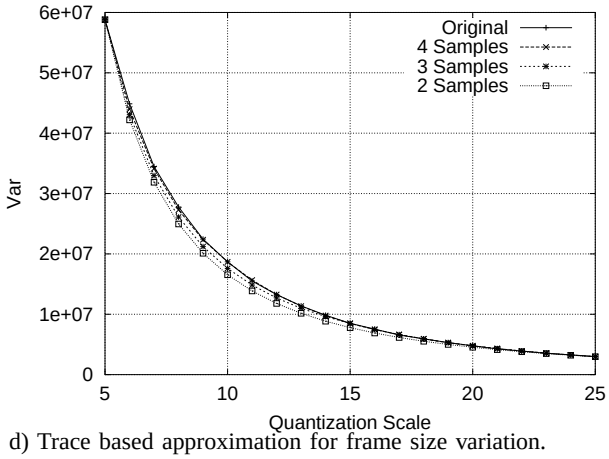
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



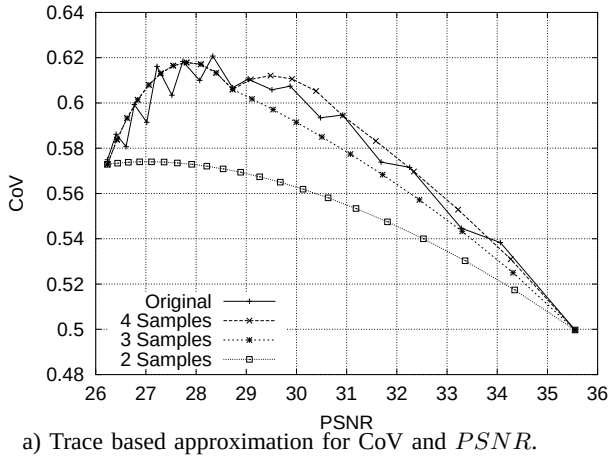
d) Trace based approximation for frame size variation.

Fig. 88. *Macroblock* level approximation results for scene *628* (MC III) from *The Terminator* .

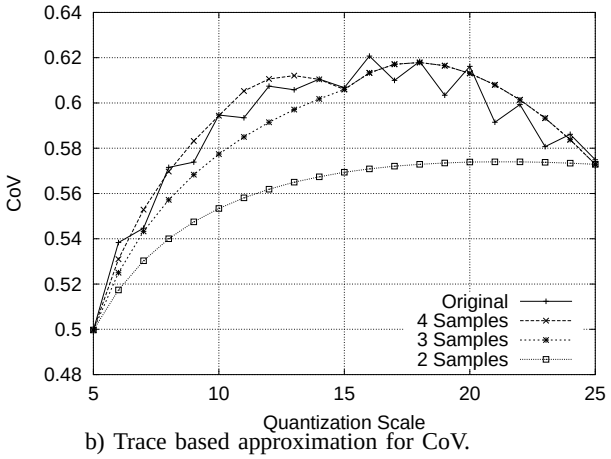
TABLE XVIII

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *628* WITH GOP PATTERN 7.

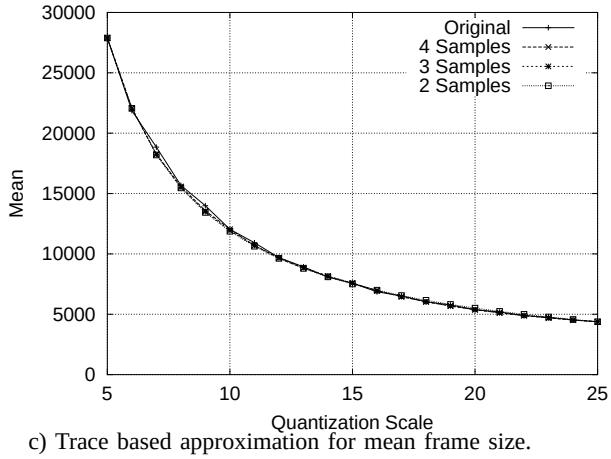
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.5226	0.9175	0.9573	0.0067	0.0009	0.0003
	σ^2	0.9957	0.9982	0.9996	2.4137e+012	8.8814e+011	1.5177e+011
	$\overline{X}$	0.9935	0.9983	0.9990	347590.5979	60033.3799	19477.8726
	$PSNR$	0.9994	0.9997	0.9996	0.0126	0.0034	0.0028
Frame	CoV	0.5237	0.9222	0.9579	0.0067	0.0008	0.0003
	σ^2	0.9944	0.9979	0.9997	3.1022e+012	1.0348e+012	1.4641e+011
	$\overline{X}$	0.9943	0.9986	0.9991	305425.7669	49897.9149	17041.2591
	$PSNR$	0.9994	0.9997	0.9996	0.0124	0.0033	0.0028
MB	CoV	0.8485	0.9495	0.9761	0.0025	0.0002	0.0001
	σ^2	0.9966	0.9988	0.9998	1.9420e+012	5.7141e+011	5.6415e+010
	$\overline{X}$	0.9988	0.9994	0.9993	59036.7299	10033.8056	9594.3275
	$PSNR$	0.9994	0.9997	0.9996	0.0124	0.0033	0.0028



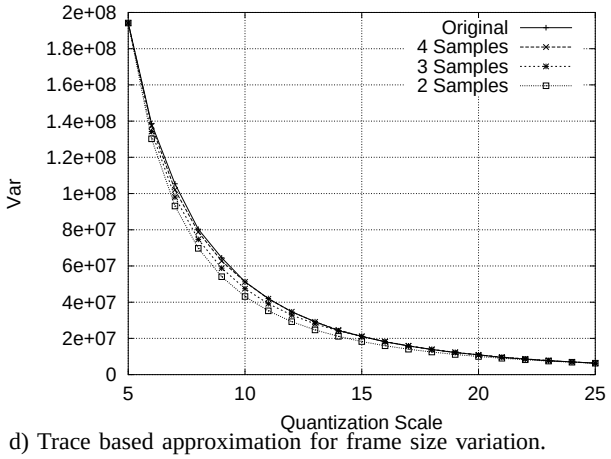
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



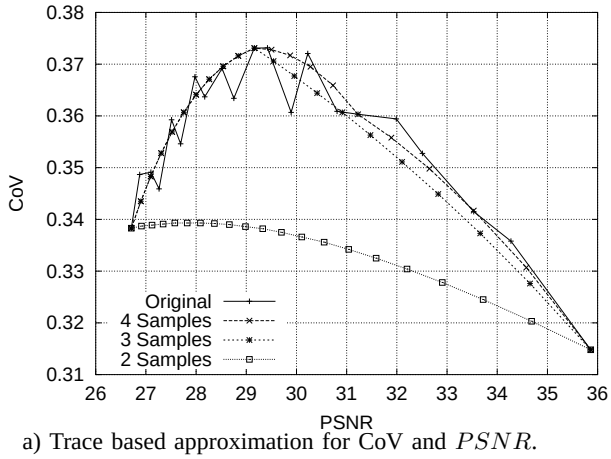
d) Trace based approximation for frame size variation.

Fig. 89. *Macroblock* level approximation results for scene *262* (MC IV) from *The Terminator* .

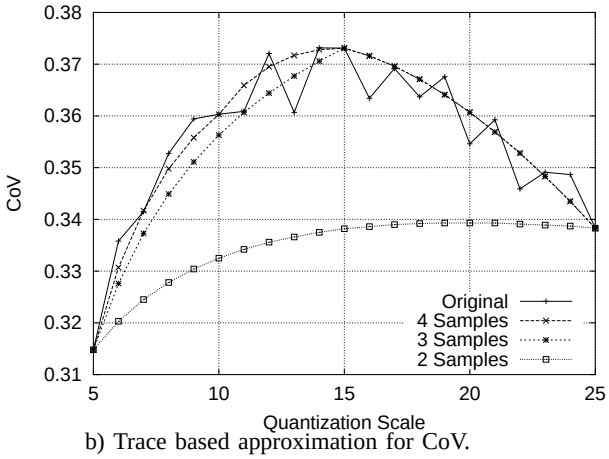
TABLE XIX

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *262* WITH GOP PATTERN 7.

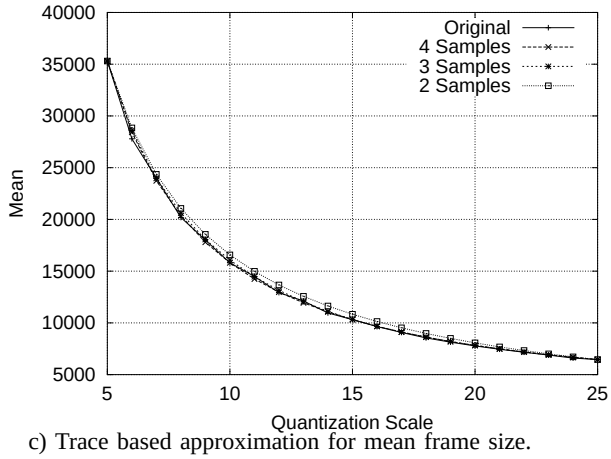
Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.4553	0.8981	0.9146	0.0023	0.0003	0.0001
	σ^2	0.9961	0.9985	0.9997	2.1264e+013	6.7437e+012	1.1015e+012
	$\overline{X}$	0.9991	0.9994	0.9993	122974.0110	34925.3184	30662.3918
	$PSNR$	0.9990	0.9995	0.9994	0.0289	0.0084	0.0048
Frame	CoV	0.3423	0.8816	0.9162	0.0029	0.0003	0.0001
	σ^2	0.9933	0.9975	0.9996	3.6336e+013	1.1088e+013	1.5110e+012
	$\overline{X}$	0.9992	0.9994	0.9993	91628.5594	29865.2514	29670.0293
	$PSNR$	0.9991	0.9995	0.9994	0.0281	0.0083	0.0048
MB	CoV	0.7809	0.9136	0.9518	0.0011	0.0001	0.0001
	σ^2	0.9944	0.9982	0.9998	3.0184e+013	7.8525e+012	8.7530e+011
	$\overline{X}$	0.9990	0.9992	0.9993	49500.4273	38951.5499	31724.8877
	$PSNR$	0.9991	0.9995	0.9994	0.0281	0.0083	0.0048



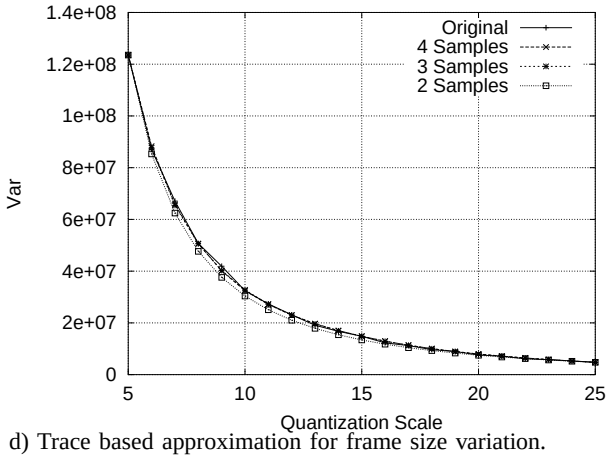
a) Trace based approximation for CoV and $PSNR$.



b) Trace based approximation for CoV.



c) Trace based approximation for mean frame size.



d) Trace based approximation for frame size variation.

Fig. 90. *Macroblock* level approximation results for scene *441* (MC V) from *The Terminator* .

TABLE XX

COEFFICIENT OF CORRELATION AND MEAN SQUARED ERROR (MSE) FOR SCENE *441* WITH GOP PATTERN 7.

Level	Metric	Coefficient of Correlation Interpolation Points			MSE Interpolation Points		
		2	3	4	2	3	4
Trace	CoV	0.1410	0.8525	0.8956	0.0009	0.0001	0.0000
	σ^2	0.9983	0.9995	0.9995	3.3767e+012	6.1688e+011	5.9996e+011
	$\overline{X}$	0.9976	0.9993	0.9994	576562.9865	98889.5527	46267.2788
	$PSNR$	0.9983	0.9989	0.9990	0.0483	0.0203	0.0078
Frame	CoV	0.0625	0.8633	0.9049	0.0010	0.0001	0.0000
	σ^2	0.9975	0.9995	0.9995	5.1874e+012	5.1890e+011	5.2947e+011
	$\overline{X}$	0.9977	0.9993	0.9994	555474.3411	90799.7775	43743.4701
	$PSNR$	0.9983	0.9989	0.9991	0.0477	0.0201	0.0077
MB	CoV	0.4725	0.8913	0.9191	0.0006	0.0000	0.0000
	σ^2	0.9982	0.9996	0.9996	3.4889e+012	3.5556e+011	3.7193e+011
	$\overline{X}$	0.9988	0.9995	0.9995	264669.8110	47232.9464	31524.1059
	$PSNR$	0.9983	0.9989	0.9991	0.0477	0.0201	0.0077

VI. VD Implications for Statistical Multiplexing

In this section we provide a methodology for assessing the statistical multiplexing gain from the VD characteristics and examine the implications of the rate variability-distortion (VD) characteristics of encoded video on statistical multiplexing, which is a key element in many video traffic management schemes.

First we augment the frame size notation defined in Section III-A from an individual video stream to multiple different streams. We let $j,\ j = 1, \ldots, J$, index the ongoing streams. We let $X_n^q(j)$ denote the size (in bit) of video frame n of stream j encoded with quantization scale q and assume that all streams have the same number of frames N. In order to fix ideas in describing our methodology we adopt the video traffic model with random phase shifts defined in [57]. In this model the frame size is modelled by a steady state random variable, whereby the distribution of the random variable is given by the histogram of the frame sizes. In our context we let $X^q(j)$ be a random variable denoting the frame size of stream j encoded with quantization scale q. The distribution of $X^q(j)$ is given by

$$\pi_j^q(x) = P(X^q(j) = x) = \frac{1}{N} \sum_{n=1}^{N} 1_{(X_n^q(j)=x)}, \tag{18}$$

where $1_{(A)}$ denotes the indicator function, which is 1 if A is true and 0 otherwise. We let T denote the frame period (display time) of a given video frame in seconds. In order not to obscure our main points we consider an elementary frame based real-time video streaming scenario, where each individual video frame is transmitted at the constant bit rate $X_n^q(j)/T$ during one frame period of length T and the streams are statistically multiplexed onto a bufferless link, as illustrated in Figure 91. Loss occurs at

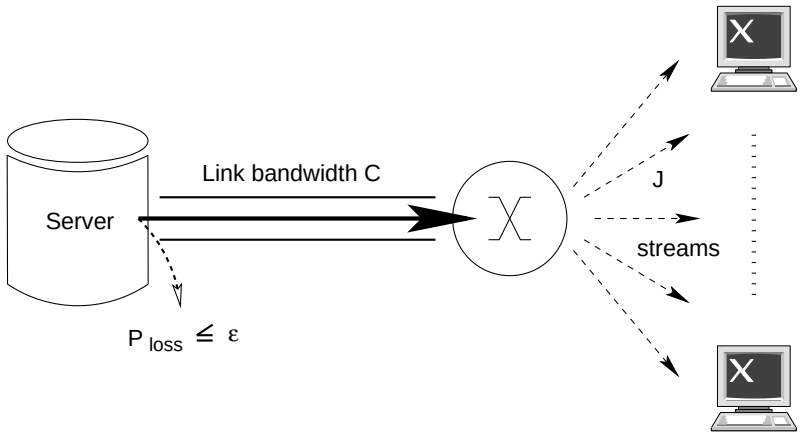


Fig. 91. Bufferless statistical multiplexing model

the link whenever the aggregated traffic from the ongoing video streams $\sum_{j=1}^{J} X^q(j)/T$ exceeds the link bandwidth C. We measure the loss in terms of the long run fraction of frame periods during which loss occurs, i.e., we define the loss probability as

$$P_{\text{loss}} = P\left(\frac{1}{T}\sum_{j=1}^{J} X^q(j) > C\right) \tag{19}$$

and require that the loss probability be less than some minuscule ϵ.

In order to determine whether a set of J streams can be supported without violating the statistical quality of service requirement that the loss probability be less than ϵ we need to determine the loss probability. This could be done using discrete event simulation, which however tends to be computationally very demanding. We outline two alternative approaches—a central limit based approach and a large deviations based approach—which require relatively little computational effort and give accurate results. The Central Limit approach models the aggregate traffic load as a Normal random variable with mean $\sum_{j=1}^{J} \bar{X}_q(j)$ and variance $\sum_{j=1}^{J} \sigma_q^2(j)$ and approximates the loss probability by the probability that a Normal random variable with the specified mean and variance exceeds the link capacity $C \cdot T$ in one frame period. Note that for each video stream j, the mean $\bar{X}_q(j)$ and the variance $\sigma_q^2(j)$ of the frame size as a function of the quantization scale q can be obtained from the traditional rate-distortion curve in conjunction with the rate variability-distortion curve of the video encoding. Also note that by using piece-wise approximation models, see Section V, a few sample encodings of each of the multiplexed videos are sufficient to obtain the frame size mean and variance for the range of quantization scales q and thus to assess the loss probability at a range of quantization scales. The outlined Central Limit based approach is computationally very simple and requires only the rate-distortion and rate variability curves of the encoding, however, it tends to slightly underestimate the loss probability especially in the range of very small loss probabilities, as we will study numerically in Section VII-A.

The large deviations based approach is very accurate for the entire range of loss probabilities, but requires a more specific characterization of the frame sizes. In particular, the large deviations approach requires the logarithmic moment generation function of the frame size $\mu_{X^q(j)}(s) = \ln E[e^{s \cdot X^q(j)}]$, where s denotes the real valued transform variable. For each of the multiplexed video streams the function $\mu_{X^q(j)}(s)$ can be explicitly expressed in terms the histogram (18), as

$$\mu_{X^q(j)}(s) = \ln \sum_{x=x_{\min}}^{x_{\max}} \pi_j^q(x) \cdot e^{sx}, \tag{20}$$

where $x_{\min}$ and $x_{\max}$ denote the smallest and largest frame size, respectively. The individual logarithmic moment generating functions are then used to compute the logarithmic moment generating function of the aggregate traffic as $\mu_{X^q}(s) = \sum_{j=1}^{J} \mu_{X^q(j)}(s)$, which in turn is employed to compute the large deviations estimate of the loss probability as

$$P_{\text{loss}} = \frac{1}{s^* \sqrt{2\pi \mu''_{X^q}(s^*)}} e^{-s^* CT + \mu_{X^q}(s^*)}. \tag{21}$$

In (21), s^* is the unique solution to $\mu'_{X^q}(s) = CT$, and the prime denotes derivative with respect to s. This computation of the large deviations estimate can be computationally demanding due to the direct computation of the logarithmic moment generating function for the different s values from the histogram. In addition, when assessing the stream admissibility and utility for a range of quantization scales q for different streams j, the direct computation requires the histogram of the frame sizes for each quantization scale q for each stream j.

The computation of the large deviations estimate can be made computationally more efficient as we describe next. Following the general series expansion technique developed in [58, p. 206] we can expand

the logarithmic moment generating function of stream j encoded with quantization scale q as

$$\mu_{X^q(j)}(s) = \sum_{k=1}^{K} c_k^q(j) \cdot s^k. \tag{22}$$

The series expansion coefficients can be expressed as

$$c_k^q = a_k^q(j) - \frac{1}{k} \sum_{i=1}^{k-1} i \cdot a_{k-i}(j) \cdot c_i^q(j), \tag{23}$$

where

$$a_i^q(j) = \frac{1}{i!} \sum_x \pi_j^q(x) \cdot x^i. \tag{24}$$

We found in our numerical work that a relatively small number of K = 15 coefficients is sufficient for a good series approximation, leading to a significantly reduced computational effort when computing the moment generating function for different s for a given (fixed) quantization scale q. To make the calculation of the moment generating function for a range of different quantization scales q more efficient, we use the following technique. We obtain the coefficients $c_k^q(j),\ k = 1, \ldots, K$, for a small number of sample encodings for different quantization scales q for each video j. For a given video j and coefficient index k, we then construct a piecewise approximation of the coefficient $c_k^q(j)$ across the full range of quantization scales q, using the approximation techniques described in Section V.

A. Numerical Results

In this section we present numerical examples to illustrate the use of the methodology described in the preceding section as well as the typical characteristics of the multiplexing behavior as a function of the video quality. For the illustrative examples presented here we use the *Star Wars* video sequence, which has been widely used in video multiplexing studies, for all ongoing video streams. We also present the results for our two additional movies, namely *Football* and *The Terminator*. Each stream has its own independent random phase, which models the random start time and user interactivity for the stream. We denote $J_{\max}$ for the maximum number of simultaneous video streams that can be supported by the link while maintaining the loss constraint, which we verify with the LD approach. In Figs. 92, 94, and 96 we plot the maximum number of supported connections $J_{\max}$ as a function of the PSNR quality. The videos are encoded with GoP pattern 7 with the same quantization scale for all frame types and the link capacity is set to $C = 50$ Mbps. We plot the maximum number of supported streams for peak rate allocation ($\epsilon = 0$) and two non-zero ϵ, namely $\epsilon = 10^{-6}$ and $\epsilon = 10^{-4}$, as well as mean rate allocation, which is obtained by dividing the link bandwidth by the average bit rate of the video. We consider both the frame based real-time streaming of the video as well as the GoP based streaming of the video, where the frames are aggregated over a GoP (a = 12) and transmitted at a constant rate over the duration of one GoP. This GoP smoothing reduces the variability of the video traffic at the expense of an increase in the delay of the video streaming by roughly two GoP durations. We observe that the introduction of a small loss probability ϵ results in a significant increase in the maximum number of supported streams over the peak rate allocation. This increase appears to be especially significant for video streams with lower PSNR quality; and we will examine this effect shortly in more detail. We also observe that the increase

in the number of permitted streams for a given increase in the loss requirement appears particularly significant for low PSNR qualities. Increasing the permitted loss probability from $\epsilon = 10^{-7}$ to $\epsilon = 10^{-3}$ for frame based streaming at 31 dB of PSNR video quality, for instance, allows for approximately 50 additional streams. We also observe that GoP streaming allows generally for more streams than frame based streaming, especially for low video qualities.

To further examine the statistical multiplexing effect for different video qualities and different permitted loss probabilities ϵ, we examine the statistical multiplexing gain achieved by allowing for an $\epsilon \geq 0$. The statistical multiplexing gain is defined as

$$g(\epsilon) = \frac{J_{\max}(\epsilon) - J_{\max}(0)}{J_{\max}(0)}, \tag{25}$$

where $J_{\max}(0)$ denotes the maximum number of supported connections for lossless transmission (i.e., peak rate allocation). In Figs. 93, 95, and 97 we plot the multiplexing gain $g(\epsilon)$ as a function of the PSNR quality for frame based streaming with different ϵ. We observe that the multiplexing gain as a function of the PSNR video quality exhibits a "hump", similar to the corresponding VD curves (see Figs. 41, 42, and 43). The explanation for this behavior of the multiplexing gains is as follows. At very low quality and at very high quality, the variability of the video traffic is relatively low (compared to the quality region where the traffic variability peaks). For the lower variability traffic, the peak rate allocation allows for a relatively larger number of streams, i.e., a higher long run average utilization of the link (defined as the sum of the average bit rates of the supported streams divided by the link bandwidth). For the higher variability streams, the utilization is lower. When statistically multiplexing with some non-zero permitted loss probability ϵ, the statistical multiplexing effect (i.e., the effect of temporarily high bit rates in some streams being compensated for by the temporarily low bit rates in other streams) becomes stronger when more streams are multiplexed, i.e., for lower stream quality. As a result, the statistical multiplexing gain is small for high quality streams because relatively few streams can be supported and the utilization with peak rate allocation is already relatively high. The statistical multiplexing gain is the highest in the region where the VD curve peaks since the number of streams with statistical multiplexing is relatively high and the utilization with peak rate allocation is the lowest. For very low quality streams, the number of streams with statistical multiplexing is higher, but so is the number of streams with peak rate allocation, resulting a somewhat smaller statistical multiplexing gain. In summary, we conclude from the results shown here that the highest multiplexing gains are achieved around the peak of the VD curve, i.e., in the region of high variability.

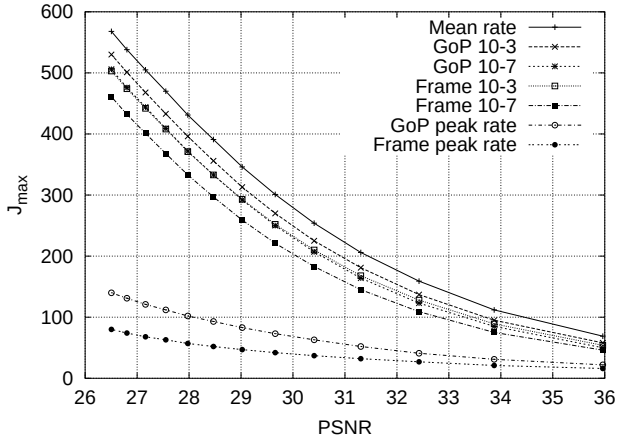


Fig. 92. Maximum number of supported streams $J_{\max}$ as a function of PSNR quality for *Football*, and link bandwidth $C = 50$ Mbps.

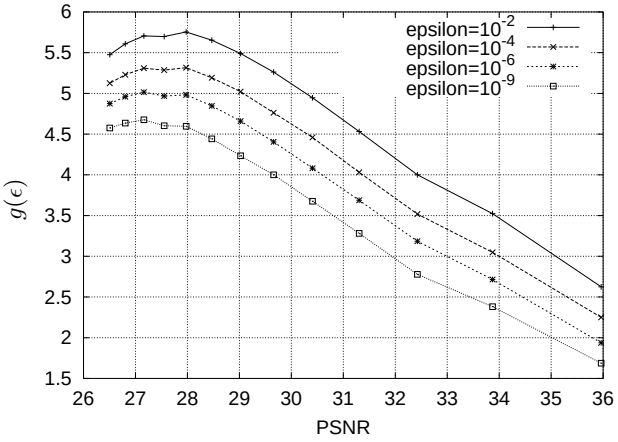


Fig. 93. Multiplexing gain $g(\epsilon)$ as a function of PSNR video quality for *Football* and bandwidth $C = 50$ Mbps.

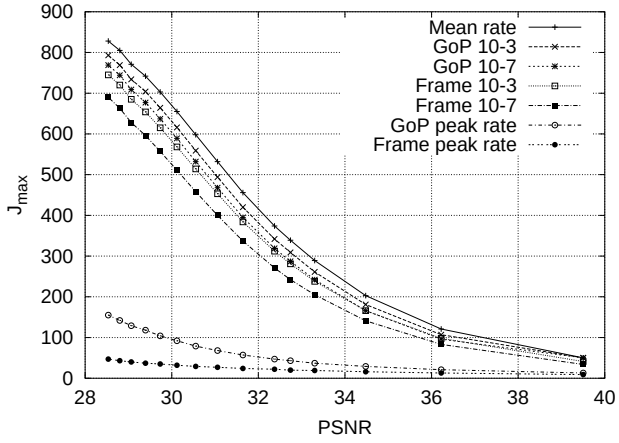


Fig. 94. Maximum number of supported streams $J_{\max}$ as a function of PSNR quality for *Star Wars*, and link bandwidth $C = 50$ Mbps.

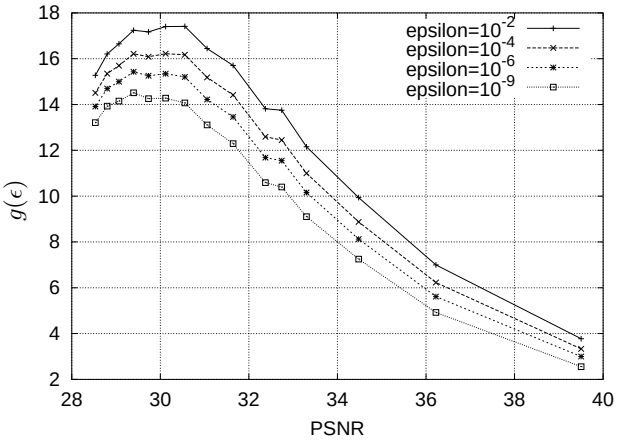


Fig. 95. Multiplexing gain $g(\epsilon)$ as a function of PSNR video quality for *Star Wars* and bandwidth $C = 50$ Mbps.

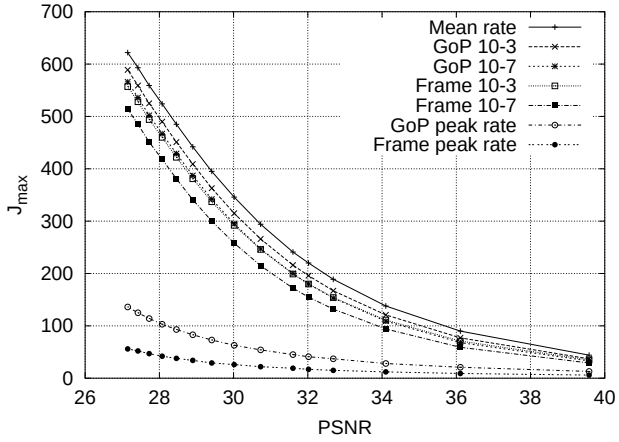


Fig. 96. Maximum number of supported streams $J_{\max}$ as a function of PSNR quality for *The Terminator*, and link bandwidth $C = 50$ Mbps.

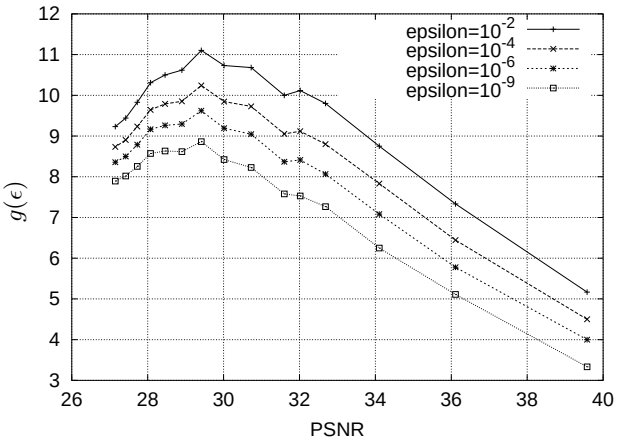


Fig. 97. Multiplexing gain $g(\epsilon)$ as a function of PSNR video quality for *The Terminator* and bandwidth $C = 50$ Mbps.

VII. Network Utility: Assessment Methodology and Typical Characteristics

In this section we provide a methodology for assessing the utility (revenue) earned from the video streaming service and examine its typical characteristics for open-loop encoded video, We let $Q_q(j)$ denote the average quality of video stream j encoded with quantization scale q. Note that if the utility depends only on the number of supported streams (irrespective of their video quality), then the revenue is maximized by streaming the lowest quality $\arg\min_q\{Q_q(j)\}$ for each stream j. On the other hand, if we assume that the utility for the content provider is maximized by the highest multiplexing gain, then it would be preferable to stream the videos with $q_{\max}$, i.e., with the highest variability. These initial observations do not consider the quality of the streamed video.

In a more realistic scenario, the revenue from the video streaming is likely determined by the number of supported streams as well as their quality. To capture this effect we adopt utility functions which are widely employed in microeconomics to relate the preferences of consumers (in our case clients) to specific goods or services (in our case the average video qualities $Q_q(j)$). In our context the utility function models the value (utility) that a video stream of a given quality has for a user. A widely employed type of utility function in microeconomics are functions with diminishing marginal utility. With such functions, the marginal increment in the utility of a good for a fixed increment in the quality of the good decreases as the absolute quality level increases. Such functions appear also appropriate in our context because a client currently receiving a low quality video (encoded with a large quantization scale) typically perceives a noticeable increase in quality if the video quality is slightly increased (the quantization scale reduced). On the other hand, for a client receiving a high quality video, a slight further increase in quality is typically barely noticeable [59]. A typical utility function [43], [48] reflecting this behavior is given by

$$U_q(j) = 1 + u(j) \cdot \log_{10}[1 + Q_q(j) - Q_{\min}(j)], \tag{26}$$

where $u(j)$ is a tuning parameter quantifying the value of the quality increases to the user receiving stream j. Note that the minimum quality video (with the largest quantization scale) is assigned a utility of one in this definition. We assume that the minuscule losses at the multiplexer do not deteriorate the video quality; the incorporation of the impact of the losses on the video quality is left as future work, as outlined in Section VIII. The total network utility earned from the video streaming as a function of the vector of quantization scales $\mathbf{q} = (q(1)\ldots, q(J))$, is obtained by summing the utilities for the individual supported streams, i.e.,

$$U_J(\mathbf{q}) = \sum_{j=1}^{J} U_q(j), \tag{27}$$

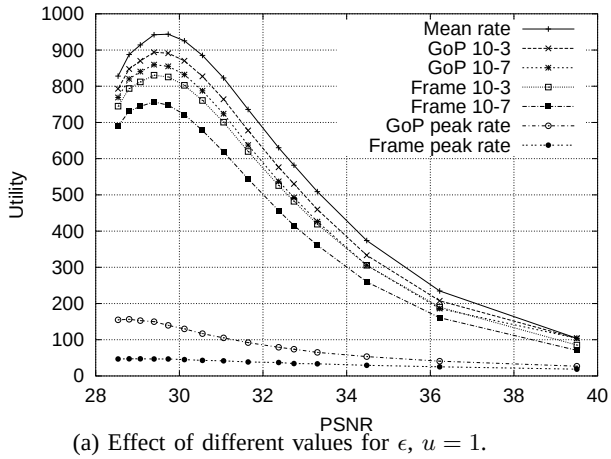
whereby the techniques from Section VI are used to assess whether a set of J streams can be supported for a permissible loss probability. The function $U(\mathbf{q})$ can now be maximized to determine the largest network utility value and the quantization scales attaining that value.

A. Numerical Results

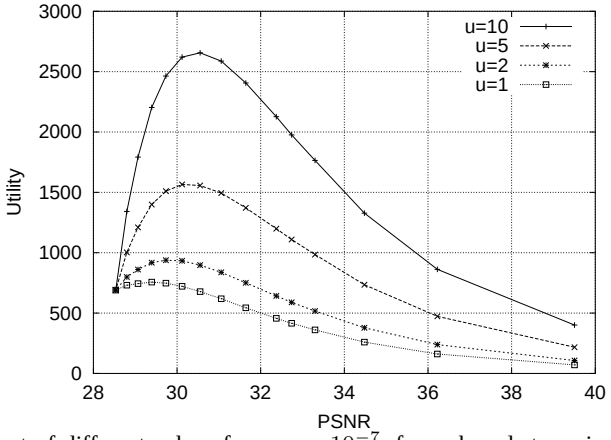
For simplicity of exposition we consider in our illustrative numerical work a homogeneous streaming scenario where all ongoing streams have the same utility parameter $u(j) = u$ and are obtained from the

Football, *Star Wars*, or *Terminator* video with random phase shifts. In this scenario, the overall network utility is obtained by multiplying the utility for a given stream by the number of supported streams, i.e., $U_J(q) = J \cdot U_q(j)$, and we refer to $J_{\max} \cdot U_q(j)$ as the maximum network utility.

In Fig. 98 we plot the maximum network utility as a function of the PSNR video quality for different parameters u in the utility function definition and for different permitted loss probabilities ϵ for *Star Wars*. The $J_{\max}$ values calculated with the large deviations approach and plotted in Figs. 92, 94, and 96 are used. The bandwidth is set to $C = 50$ Mbps. We observe from Fig. 98(a) that the maximum network



(a) Effect of different values for ϵ, $u = 1$.

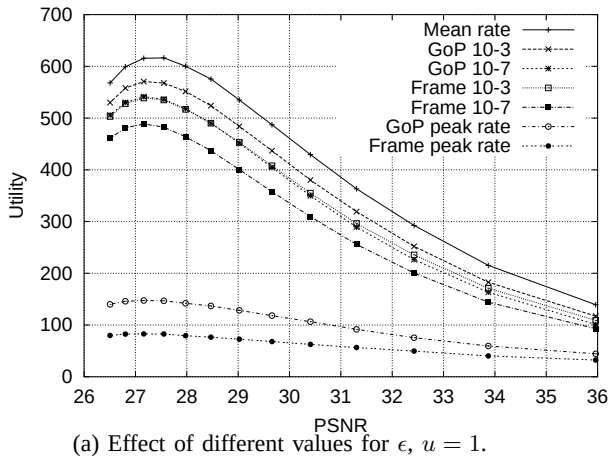


(b) Effect of different values for u, $\epsilon = 10^{-7}$, frame-based streaming.

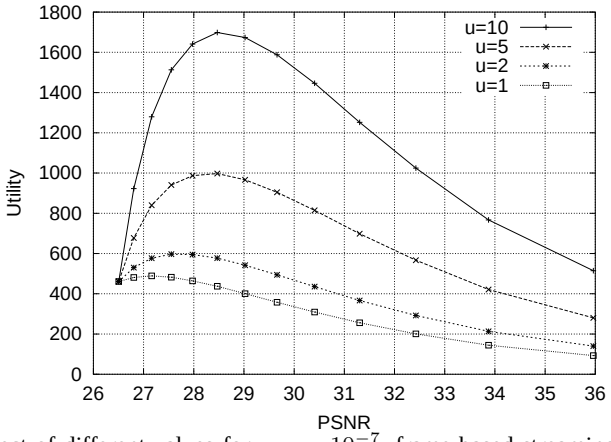
Fig. 98. Maximum network utility as a function of PSNR video quality for *Star Wars*.

utility exhibits a characteristic hump similar to the VD curve and the statistical multiplexing gain. We furthermore observe from Fig. 98(b) that for larger utility parameter u, i.e., when more value is assigned to the video quality, the peak in the maximum utility tends to drift towards higher video qualities. In other words, the largest network utility value is achieved by statistically multiplexing fewer streams, but each stream has a higher quality. Examining Fig. 98(a) more closely, we observe that higher permitted loss probabilities and GoP smoothing result in significant increases in the maximum utility, except for very high PSNR qualities. For a permissible loss probability of $\epsilon = 10^{-7}$, for instance, GoP based streaming gives approximately 13% higher utility than frame-based streaming around the peak of the maximum utility curve. Similar results are obtained for the movies *Football* and *Terminator*, as illustrated in Figs. 99 and 100, respectively.

The results presented so far have used the large deviations approach to determine the admissibility of a set of streams. In Fig. 101 we plot the maximum network utility obtained when employing the Normal approximation to determine the stream admissability and compare with the large deviations based results for *Star Wars*. We observe that the Normal approximation gives generally fairly accurate results, especially for the larger permitted loss probability $\epsilon = 10^{-3}$. For the smaller loss probability $\epsilon = 10^{-7}$ the Normal approximation slightly overestimates (by typically less than 5%) the achievable utility. We obtain similar outcomes for the movies *Football* and *Terminator* as illustrated in Figs. 102 and 103. Overall, we may conclude that the Normal approximation gives a fairly good assessment of the utility. Note that the Normal approximation requires only the rate-distortion (RD) function and the rate variability-distortion (VD) function of the encoded video, both of which can be obtained with high accuracy by a piecewise

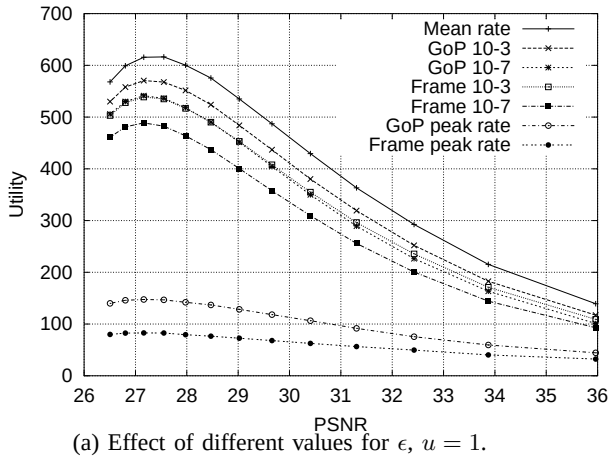


(a) Effect of different values for ϵ, $u = 1$.

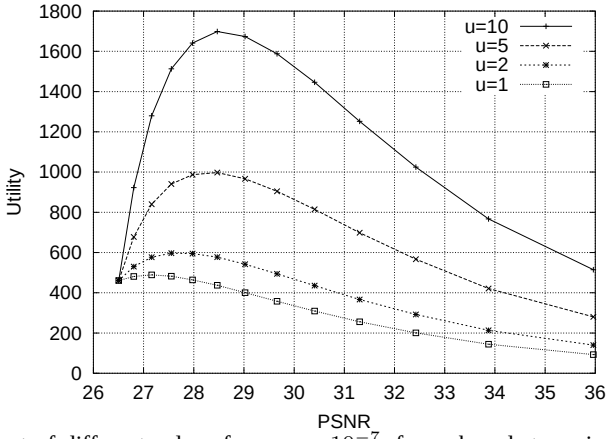


(b) Effect of different values for u, $\epsilon = 10^{-7}$, frame-based streaming.

Fig. 99. Maximum network utility as a function of PSNR video quality for *Football*.



(a) Effect of different values for ϵ, $u = 1$.



(b) Effect of different values for u, $\epsilon = 10^{-7}$, frame-based streaming.

Fig. 100. Maximum network utility as a function of PSNR video quality for *Terminator*.

linear approximation based on a few sample encodings.

VIII. Conclusion

We have examined the relationships between video quality, bit rate variability, and the utility from a streaming service with statistical multiplexing for open-loop encoded video. We have found that the rate variability-distortion (VD) curve of open-loop encoded video exhibits typically a characteristic "hump" behavior and have investigated how this hump behavior is influenced by the different motion levels in the video content, the video encoding parameters and traffic smoothing. We have found that the bit rate variability is the highest and the hump in the VD curve is most pronounced for low motion video scenes. We have also found that larger quantization scales for the predictive frame types in MPEG (compared to the quantization scale for the intra coded frames) and shorter GoP patterns tend to increase the level of bit rate variability. Furthermore, we have observed that traffic smoothing within a scene is highly effective in reducing the bit rate variability within the scene, whereas traffic smoothing over a long video is less effective in reducing the bit rate variability of the video.

We have described a methodology for assessing the admissibility of a set of video streams on a link subject to a statistical quality of service criterion and for assessing the utility (revenue) earned by a service provider when statistically multiplexing video of different quality levels over the link. In summary, the methodology first determines whether a set of streams can be supported while ensuring a small long run loss probability. The utility from the supported streams is then computed by adding the utilities corresponding to the quality levels of the individual streams. Our numerical work for homogeneous streaming scenarios indicates that the statistical multiplexing gain and the utility as a function of the video quality level typically exhibit a characteristic "hump" similar to the VD curve. The peaks in these functions are typically in the vicinity of the quantization scale attaining the peak in the VD curve. Finally, we have demonstrated that the Normal approximation which relies on the first and second moment of the video traffic (as a function of the quality level) is quite accurate in assessing the network utility in the bufferless statistical multiplexing model.

There are many exciting avenues for future work. One avenue is to incorporate the effect of the lost video traffic on the video quality, which becomes important for large loss probabilities. To incorporate the effect of these losses in our utility evaluation, the quality of a stream could be modeled as a function of both the quantization scale q and the loss probability limit ϵ, i.e., as $Q_q^\epsilon(j)$, using for instance the models studied in [23], [24], [59]–[62]. This adjusted quality $Q_q^\epsilon(j)$ can then be employed in the utility evaluation, e.g., by using a utility function such as (26).

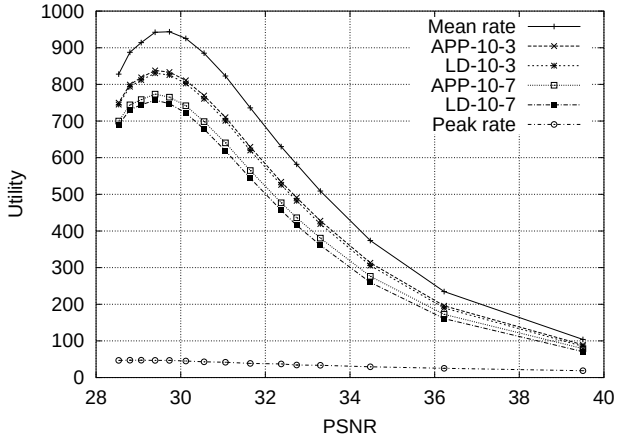


Fig. 101. Maximum utility evaluated with Normal approximation (APP) and Large Deviations (LD) approach as a function of PSNR video quality, frame based streaming, u = 1 for *Star Wars*.

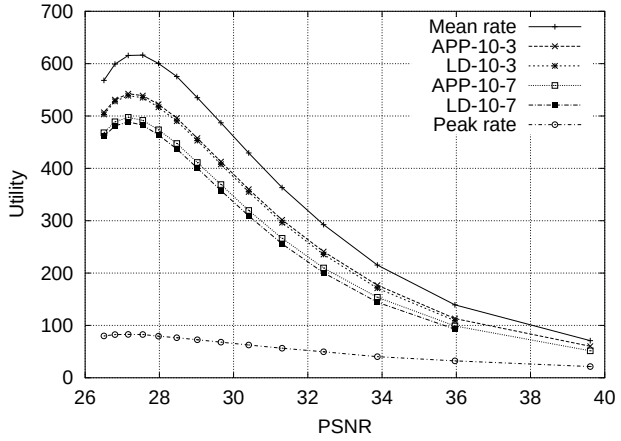


Fig. 102. Maximum utility evaluated with Normal approximation (APP) and Large Deviations (LD) approach as a function of PSNR video quality, frame based streaming, $u = 1$ for *Football*.

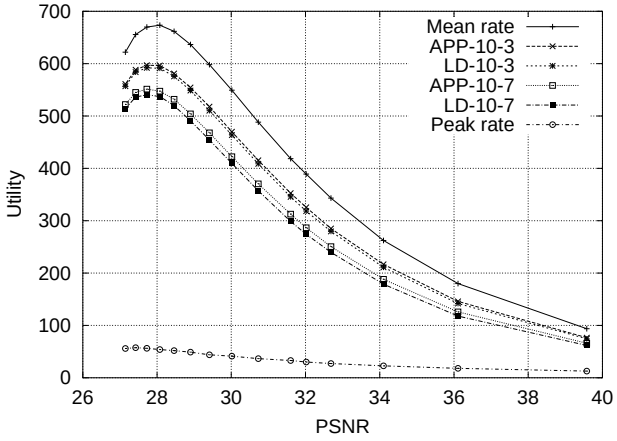


Fig. 103. Maximum utility evaluated with Normal approximation (APP) and Large Deviations (LD) approach as a function of PSNR video quality, frame based streaming, $u = 1$ for *Terminator*.

REFERENCES

[1] T. V. Lakshman, A. Ortega, and A. R. Reibman, "VBR video: Tradeoffs and potentials," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 952–973, May 1998.

[2] A. R. Reibman and M. T. Sun, *Compressed Video over Networks*. Marcel Dekker, New York, 2000.

[3] D. Wu, Y. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming video over the internet: Approaches and directions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 282–300, Mar. 2001.

[4] A. Ortega and K. Ramachandran, "Rate-distortion methods for image and video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[5] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[6] H.-M. Hang and J.-J. Chen, "Source model for transform video coder and its application —part I: Fundamental theory," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 287–298, Apr. 1997.

[7] C. Cai, R. Ding, and S. K. Mitra, "New bit-rate control models for image and video transmission," in *Proc. of the IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions*, vol. 2, June 2002, pp. 958–962.

[8] L.-J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 4, pp. 446–459, Dec. 2001.

[9] Z. He and S. K. Mitra, "A unified rate-distortion analysis framework for transform coding," *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 11, no. 12, pp. 1221–1236, Dec. 2001.

[10] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 1, pp. 12–20, Feb. 1996.

[11] K. N. Ngan, T. Meier, and Z. Chen, "Improved single-video-object rate control for MPEG-4," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 5, pp. 385–393, May 2003.

[12] M. Dai and D. Loguinov, "Analysis of rate-distortion functions and congestion control in scalable internet video streaming," in *Proc. of the Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, Monterey, CA, June 2003, pp. 60–69.

[13] X. Lu, R. O. Morando, and M. ElZarki, "Understanding video quality and its use in feedback control," in *Proceedings of International Packet Video Workshop 2002*, Pittsburgh, PA, 2002.

[14] N. Ansari, H. Liu, Y. Q. Shi, and H. Zhao, "On modeling MPEG video traffics," *IEEE Transactions on Broadcasting*, vol. 48, no. 4, pp. 337–347, Dec. 2002.

[15] K. Chandra and A. R. Reibman, "Modeling one– and two–layer variable bit rate video," *IEEE/ACM Transactions on Networking*, vol. 7, no. 3, pp. 398–413, June 1999.

[16] M. Ghanbari and D. E. Pearson, "Components of bit-rate variation in videoconference signals," *Electronics Letters*, vol. 25, no. 4, pp. 285–286, 1989.

[17] D. P. Heyman and T. V. Lakshman, "Source models for VBR broadcast video traffic," *IEEE/ACM Transactions on Networking*, vol. 4, pp. 40–48, Jan. 1996.

[18] M. M. Krunz and A. M. Makowski, "Modeling video traffic using $M/G/\infty$ input processes: A compromise between markovian and LRD models," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 733–748, June 1998.

[19] A. Lombaedo, G. Schembra, and G. Morabito, "Traffic specifications for the transmission of stored MPEG video on the internet," *IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 5–17, Mar. 2001.

[20] B. Melamed and D. E. Pendarakis, "Modelling full-length VBR video using markov-renewal-modulated TES models," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 600–611, June 1998.

[21] A. Robinson, D. Pearson, and J. Xiong, "The influence of scene content on bit-rate variations in ATM video," in *Proceedings of IEE Colloquium on Coding for Packet Video and Speech Transmission*, Nov. 1992, pp. 7/1–7/5.

[22] U. K. Sarkar, S. Ramakrishnan, and D. Sarkar, "Modeling full-length video using markov-modulated gamma-based framework," *IEEE/ACM Transactions on Networking*, vol. 11, no. 4, pp. 638–649, Aug. 2003.

[23] R. Aravind, M. Civanlar, and A. Reibman, "Packet loss resilience of MPEG–2 scalable video coding algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 5, pp. 426–435, Oct. 1996.

[24] P. Frossard and O. Verscheure, "Joint source/FEC rate selection for quality-optimal MPEG–2 video delivery," *IEEE Transactions on Image Processing*, vol. 10, no. 12, pp. 1815–1825, Dec. 2001.

[25] W. Luo and M. ElZarki, "Quality control for VBR video over ATM networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 6, pp. 1029–1039, Aug. 1997.

[26] O. Verscheure, P. Frossard, and M. Hamdi, "User-oriented QoS analysis in MPEG-2 video delivery," *Journal of Real-Time Imaging*, vol. 5, no. 5, pp. 305–314, Oct. 1999.

[27] Y. Yang and S. S. Hernami, "Rate control for VBR video over ATM: simplification and implementation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 9, pp. 1045–1058, Sept. 2001.

[28] S. Bakiras and V. O. K. Li, "Maximizing the number of users in an interactive video-on-demand system," *IEEE Transactions on Broadcasting*, vol. 48, no. 4, pp. 281–292, Dec. 2002.

[29] I. Dalgic and F. A. Tobagi, "Performance evaluation of ATM networks carrying constant and variable bit-rate video traffic," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 6, pp. 1115–1131, Aug. 1997.

[30] N. Duffield, K. Ramakrisnan, and A. Reibman, "Issues of quality and multiplexing when smoothing rate adaptive video," *IEEE Transactions on Multimedia*, vol. 1, no. 4, pp. 352–364, Dec. 1999.
[31] W.-C. Feng and J. Rexford, "A comparison of bandwidth smoothing techniques for the transmission of prerecorded compressed video," in *Proceedings of IEEE Infocom*, Kobe, Japan, Apr. 1997, pp. 58–67.
[32] M. Grossglauser, S. Keshav, and D. Tse, "RCBR: A simple and efficient service for multiple time–scale traffic," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 741–755, 1997.
[33] M. Krunz, W. Zhao, and I. Matta, "Scheduling and bandwidth allocation for distribution of archived video in VoD systems," *Telecommunication Systems*, vol. 9, no. 3/4, pp. 335–355, Sept. 1998.
[34] S. C. Liew and C.-Y. Tse, "Video aggregation: adapting video traffic for transport over broadband networks by integrating data compression and statistical multiplexing," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 7, pp. 1455–1471, Sept. 1996.
[35] D. Wrege, E. Knightly, H. Zhang, and J. Liebeherr, "Deterministic delay bounds for VBR video in packet–switching networks: Fundamental limits and tradeoffs," *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp. 352–362, June 1996.
[36] Z. Zhang, J. Kurose, J. Salehi, and D. Towsley, "Smoothing, statistical multiplexing and call admission control for stored video," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 1148–1166, Aug. 1997.
[37] M. Balakrishnan, R. Cohen, E. Fert, and G. Keesman, "Benefits of statistical multiplexing in multi-program broadcasting," in *Proceedings of Int. Broadcasting Convention*, Sept. 1997, pp. 560–565.
[38] M. W. Garrett and M. Vetterli, "Joint source/channel coding of statistically multiplexed real-time services on packet networks," *IEEE/ACM Transactions on Networking*, vol. 1, no. 1, pp. 71–80, 1993.
[39] J. Jordan and A. Bock, "Analysis, modelling and performance prediction of digital video statistical multiplexing," in *Proceedings of Int. Broadcasting Convention*, Sept. 1997, pp. 553–559.
[40] P. Cuenca, A. Garrido, F. Quiles, and L. Orozco-Barbosa, "An efficient protocol architecture for error-resilient MPEG-2 video communications over ATM networks," *IEEE Transactions on Broadcasting*, vol. 45, no. 1, pp. 129–140, Mar. 1999.
[41] R. Kuceren and J. W. Modestino, "A joint source-channel coding approach to network transport of digital video," in *Proceedings of IEEE Infocom*, Mar. 2000, pp. 717–726.
[42] G.-M. Muntean, P. Perry, and L. Murphy, "A new adaptive multimedia streaming system for all-IP multi-service networks," *IEEE Transactions on Broadcasting*, vol. 50, no. 1, pp. 1–10, Mar. 2004.
[43] P. Bocheck, A. T. Campbell, S.-F. Chang, and R. R.-F. Liao, "Utility-based network adaptation for MPEG-4 systems," in *Proceedings of NOSSDAV*, June 1999.
[44] J. Shin, J. Kim, and C.-C. Kuo, "Quality–of–service mapping mechanisms for packet video in differentiated services network," *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 219–231, June 2001.
[45] K. Yanori, H. Akimaru, and M. Finley, "Optimum design of videoconferencing systems over ATM networks," in *Proceedings of the ICCT 2000*, vol. 2, Aug. 2000, pp. 1213–1219.
[46] E. Fulp, M. Ott, D. Reininger, and D. Reeve, "Paying for QoS: An optimal distributed algorithm for pricing network ressources," in *Proceedings of the Sixth International Workshop on Quality of Service*, Napa, CA, US, May 1998, pp. 75–84.
[47] R.-F. Liao and A. Campbell, "A utiliy–based approach for quantitative adaption in wireless packet networks," *Wireless Networks*, vol. 7, no. 5, pp. 541–551, 2001.
[48] A.E.Luna, L. Kondi, and A. Katsaggelos, "Maximizing user utility in video streaming applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 2, pp. 141–148, Feb. 2003.
[49] S.-F. Chang, T. Sikora, and A. Puri, "Overview of the MPEG-7 standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 688–695, June 2001.
[50] M. Roach, J. Mason, and M. Pawlewski, "Video genre classification using dynamics," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Salt Lake City, UT, May 2001, pp. 1557–1560.
[51] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the h.264/avc video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, July 2003.
[52] S.-T. Hsiang and J. W. Woods, "Invertible three-dimensional analysis/synthesis system for video coding with half-pixel-accurate motion compensation," *Proceedings of the SPIE Conference on Visual Communications and image processing*, 1999.
[53] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991.
[54] A. M. Law and W. D. Kelton, *Simulation, Modeling and Analysis*, 2nd ed. McGraw-Hill, 2000.
[55] S. Winkler, "Vision models and quality metrics for image processing applications," Ph.D. dissertation, EPFL, Switzerland, 2000.
[56] A. Rohaly, J. Libert, P. Corriveau, and A. Webster, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000.
[57] M. Reisslein and K. Ross, "Call admission for prerecorded sources with packet loss," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 6, pp. 1167–1180, Aug. 1997.
[58] J. Y. Hui, *Switching and Traffic Theory for Integrated Broadband Networks*. Boston, MA: Kluwer, 1990.
[59] C. J. van den Branden Lambrecht and O. Verscheure, "Perceptual quality measure using a spatio-temporal model of the human visual system," in *Proceedings of SPIE, Vol. 2668*, San Jose, CA, Jan. 1996, pp. 450–461.

[60] O. Verscheure, P. Frossard, and M. Hamdi, “Joint impact of MPEG-2 encoding rate and ATM cell losses on video quality,” in *Proceedings of IEEE Globecom*, Nov. 1998, pp. 71–76.

[61] O. Verscheure and C. J. van den Branden Lambrecht, “Adaptive quantization using a perceptual visibility predictor,” in *Proceedings of IEEE Int. Conf. on Image Processing*, Santa Barbara, CA, Oct. 1997, pp. 298–301.

[62] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, “Analysis of video transmission over lossy channels,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, Aug. 2000.