

Video Texture and Motion Based Modeling of Rate Variability-Distortion (VD) Curves (Extended Version)

Geert Van der Auwera, Martin Reisslein, and Lina J. Karam

Abstract

We examine and model the bit rate variability-distortion (VD) curve of MPEG-4 Part 2 VBR video encodings, which has applications for optimal statistical multiplexing of low bit rate streaming video. We show that the concave VD curve shape at high compression ratios or large quantization scales, is influenced by both the texture and the motion information. We use linear and quadratic models for the texture and motion bits statistics and devise accurate VD curve models. The model parameters are obtained from statistics that are estimated from two encodings. This work extends our previous work on piecewise models of the VD curve. The texture and motion based VD model based on two encodings has comparable accuracy to a piecewise model based on three encodings. In addition, the texture and motion based VD model provides fundamental insights into how texture and motion affects traffic variability.

Index Terms

Statistical multiplexing, variable bit rate, low bit rate, rate variability-distortion, VD curve, MPEG-4 Part 2, video content, video quality, video traffic, video streaming.

I. INTRODUCTION

A recent study [1] has documented the concave shape of the rate variability-distortion (VD) curve of open-loop or variable bit rate (VBR) MPEG-4 Part 2 [2] encoded video and proposed a crude model for the VD curve. The VD curve is the coefficient of variation (standard deviation normalized by the mean) of the frame size (in bits) as a function of the quantization scale or distortion metric. In [1], the VD curve has been approximated by a piecewise model. The model is based on up to four encoding samples corresponding to four quantization scales. The piecewise model approximation with only two samples of the encodings does not capture the VD curve behavior. In addition, this piecewise model does not provide insights into the properties of the encoded video frames that result in the concave curve shape.

In this study, we build on [1] by examining the underlying effects leading to the concave VD curve and by developing and validating significantly a refined VD curve model for low bit rate encoded video. More specifically, we develop and validate quadratic models for the mean and variance of the texture information (bits) in a frame as a function of the quantization scale. We also develop linear models for the mean and variance of the motion information (bits) in a frame, thus extending [3] where a quadratic

Preliminary results from this study appear in: "Video Texture and Motion Based Modeling of Rate Variability-Distortion (VD) Curves of I-, P-, and B-Frames," *Proc. of IEEE Conf. on Multimedia & Expo*, Toronto, Canada, July 2006.

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.

G. Van der Auwera, L.J. Karam, 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: {geert.vanderauwera, karam, reisslein}@asu.edu, web: <http://www.fulton.asu.edu/~mre>.

model was used for the rate-distortion function of the entire frame size. We combine the models of the texture and motion bits in an overall model of the VD curve. We demonstrate that both the texture and motion bits make significant contributions to the overall concave shape of the VD curve. We find that given the frame size contributions for encodings with only two different quantization scales, this novel model accurately predicts the VD curve across a wide range of quantization scales. This is a significant improvement over the crude model in [1] where encodings for three or more different quantization scales were required.

The VD curve has important implications for statistical multiplexing, as the highest statistical multiplexing gain is typically achieved at the peak of the VD curve [1]. Thus, the refined VD curve model proposed in this paper not only provides fundamental insights into how texture and motion in an encoded video contribute to its traffic variability, but also provides a practical method for estimating the rate variability from only two sample encodings. An application scenario is the streaming of pre-encoded videos through a statistically multiplexed low bit rate network channel. In this scenario, the videos are first encoded with two different quantization scales in order to estimate the VD curves' model parameters and to obtain an estimate of the quantization scales corresponding to the peaks of these VD curves. In the second step, the videos are encoded with their respective peak quantization scales and stored on a streaming video server. If all videos are encoded at their peak rate variability, optimal statistical multiplexing is achieved while streaming them.

This report is organized as follows. In Section II, we discuss related work and position our research on the VD curve. Metric definitions and experimental setup is covered in Section III. The VD curve modeling is introduced in Sections IV-A and IV-B. The quadratic and linear models for the texture and motion bits statistics of video frames are developed and validated in Sections IV-C and IV-D. In Section IV-E, the VD curve models for the I-, P-, and B-frames are validated. The combined VD curve model for all frames is proposed and validated in Section IV-F. The sensitivity of choosing the two quantization scales is analyzed in Section V. The report ends with our conclusions followed by the appendices. Appendix A contains texture and motion bits statistics along with quadratic and linear models. Appendix B illustrates the VD curve modeling with actual VD curves and their models. Appendix C contains the graphs accompanying the sensitivity analysis.

II. RELATED WORK

This study relates most closely to research on the rate-distortion (RD) characteristics of encoded video and research on the analysis and modeling of video traffic. The RD characteristics of encoded video gives the relationship between the (mean) bit rate and the video quality (and encoder quantization scale), see e.g., [4], [5]. RD modeling has been a prolific research area for many years and a wide variety of modeling approaches have been examined, ranging from models primarily based on the source video statistics, e.g., [6], to models primarily based on interpolating sample points of the RD curve, e.g., [7], [8], as well as models incorporating analytical and empirical insights, e.g., [9]–[12]. The RD models are important for rate control in video coding, e.g., [13], [14], and the allocation of mean bit rates to video streams for network transport [15], [16]. Our study differs from the research on the RD characteristics in that we examine the relationship between the variability of the bit rate and the video quality (and

quantization scale). That is, our focus is on the second order statistic of the video traffic, whereas the existing RD studies have focused on the first order statistic of the video traffic.

Research on the statistical analysis of video traffic and video traffic models has also received significant interest, see for instance [17]–[29]. This research focuses on analyzing the statistics of the traffic, which includes the bit rate variability, for a video encoding with a fixed quality level or quantization scale). In contrast, in this paper we examine the bit rate variability of the video as *a function of the quality level/quantization scale*. A detailed understanding of this bit rate variability-quality level relationship is important for video transport mechanisms, for individual streams, e.g., [30]–[32], for multiple streams of a fixed quality level, e.g., [33]–[37], as well as multiple streams with quality adaptation, see e.g., [38]–[42]. Our study complements this literature by uncovering the fundamental relationships between the rate variability produced by the open-loop encoder and the different quantization scales/quality levels.

III. METRIC DEFINITIONS AND EVALUATION SETUP

The rate variability-distortion (VD) curve relates the bit rate variability of an encoded video sequence to its average picture quality. More specifically, the VD curve relates the coefficient of variation CoV of encoded frame sizes (#bits) to the average picture quality (e.g., peak signal-to-noise ratio or PSNR) or the quantization scale q . The CoV_q of the encoded frame sizes of a given video sequence for a particular quantization scale q is defined as

$$CoV_q = \frac{\sigma_q}{\bar{R}_q} \quad (1)$$

with σ_q denoting the standard deviation and \bar{R}_q the mean of the frame sizes (in bits) of the encoded video sequence. The VD curve has a characteristic concave or “hump” shape as depicted in Figure 1, separately for I-, P-, B-frames, and combined for all frame types for a segment of the movie *Star Wars V: The Empire Strikes Back* (1000 QCIF frames). Due to the concave shape, there exists a quantization scale q at which the bit rate variability reaches a maximum. In Figure 1, the abscissa contains the q scale and

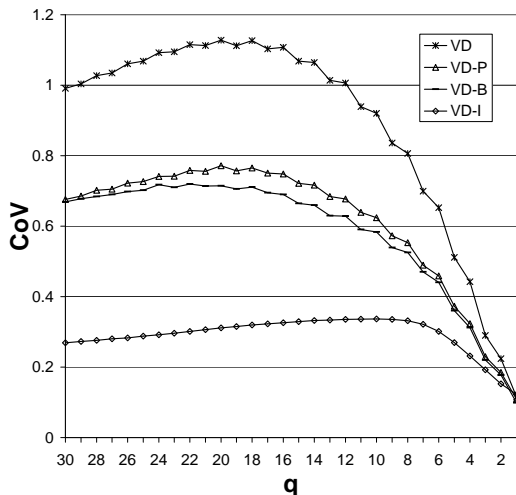


Fig. 1. Concave VD curves for the movie segment from *Star Wars V*, encoded with MPEG-4.

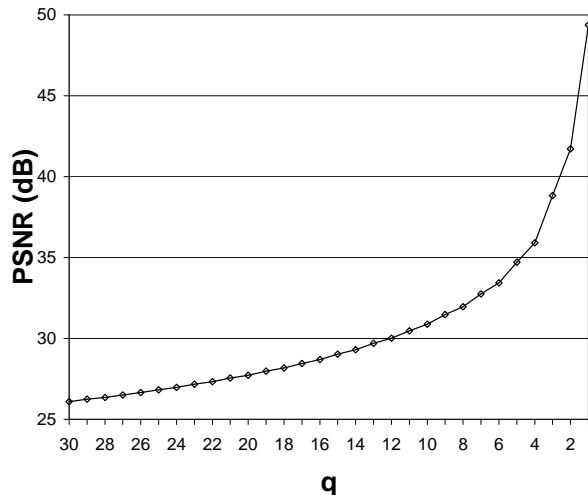


Fig. 2. Average picture quality of the movie segment from *Star Wars V* as a function of q .

decreases from value 30 to 1. This corresponds to the increasing average picture quality as is illustrated in Figure 2, where the average PSNR is plotted as a function of quantization scale q . The PSNR is an objective measure of the quality of a reconstructed video frame $R(x, y)$ with respect to the uncompressed video frame $F(x, y)$. The larger the difference between $R(x, y)$ and $F(x, y)$, or equivalently, the lower the quality of $R(x, y)$, the lower the PSNR. The PSNR is expressed in decibels (dB) to accommodate the logarithmic sensitivity of the human visual system. The PSNR is typically obtained for the luminance video frame and in case of a $N_x \times N_y$ frame consisting of 8-bit pixel values, it is computed as a function of the mean squared error (MSE) as follows:

$$MSE = \frac{1}{N_x \cdot N_y} \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} [F(x, y) - R(x, y)]^2, \quad (2)$$

$$PSNR = 10 \cdot \log \frac{255^2}{MSE}. \quad (3)$$

Other objective video quality metrics exist that correlate better with the quality perception of a human observer, e.g. VQM [43], [44]. These metrics can be applied to the reconstructed video sequences and a transformation curve between q and the metric values can be plotted, as is done for the PSNR in Figure 2. Whatever quality metric is applied, it will have a monotonically increasing or decreasing trend from low quality for $q = 30$ to high quality for $q = 1$. Hence, the concave curve shape of the VD curve will be preserved, although it can be stretched or compressed. The following analysis and conclusions will therefore also be valid for other quality metrics.

A. Scenes

In addition to the *Star Wars V* segment, we employ the same videos as in [1], namely scenes from two movies, *Star Wars IV*, *The Terminator*, and a *Football* game, which are summarized in Table I. All videos are in QCIF format (176×144 pixels). We used the publicly available scene detection software [45] for the determination of the scene boundaries. Following [46], the scenes have been classified 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 video encoding, we use the same video codec as in [1], namely the *Microsoft v2.3.0* reference implementation of MPEG-4 Part 2 [2] in the simple profile. The group-of-pictures (GoP) structure that is used in this work consists of twelve frames, namely one I-frame, three P-frames, and eight B-frames: *IBBPBBPBBPBB*.

IV. VD CURVE MODEL

A. Texture vs. Motion Bits

In this paper, we apply a new modeling approach to the VD curve. We start from the insight that encoded video frames have as constituents texture, motion, and syntax bits. In case of the P- and B-frames, the texture bits represent the encoded prediction error information that remains after the often imperfect temporal prediction step in the video encoder. The motion bits represent side information (motion vectors) about the temporal prediction step and are required to reconstruct the P- and B-frames. In case of I-frames, the texture bits represent the actual encoded frame content. No motion bits are required to

TABLE I
OVERVIEW OF SCENES CLASSIFIED IN 5 MOTION CLASSES.

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

reconstruct the I-frames. The syntax bits assign a meaning to parts of the bit stream, e.g. to distinguish between texture and motion bits.

At low compression ratios or equivalently for small quantization scales, the motion and syntax bits are negligible compared to the bits required to encode the texture. The higher the compression ratio or the larger the quantization scale (e.g., approaching 30 for MPEG-4) the more the texture information is reduced, while the motion and syntax bits vary over a much smaller range. *For large quantization scales the number of texture bits is comparable to the motion bits and therefore the motion information plays a significant role in the bit rate variability.* In other words, the concave VD curve shape at high compression ratios is influenced by the texture information but also by the motion information. We still consider the syntax bits negligible compared to the texture and motion bits, since they are more than an order of magnitude smaller.

This principle is illustrated in Figure 3 where the average P- and B-frame sizes are depicted along with the average texture and motion bits for the *Star Wars V* segment. We observe that for $q = 16$, the P texture information is about twice the size of the motion information and decreases exponentially. For $q = 24$ the average P texture bits are even smaller than the P motion bits. For the B-frames this point is reached for $q = 26$. The average I-frame size curve is not shown in this figure, because it is more than three times larger than the average P-frame size curve. The I-frame sizes as a function of q have an exponential curve shape as well.

B. Texture and Motion VD Curve Model

We first define the notation that is used in this section and in the following sections. Let $\bar{R}_{q,t}$ and $\sigma_{q,t}^2$ respectively denote the average and the variance of the number of texture bits, and $\bar{R}_{q,m}$ and $\sigma_{q,m}^2$ respectively denote the average and variance of the number of motion bits. $cov_q(t, m)$ represents the covariance of the texture and motion information. The index q refers to a particular quantization scale for which these statistics are estimated. We observe that the sum of the number of texture and motion bits approximately equals the total frame size, i.e., when we ignore the syntax bits: $R_q = R_{q,t} + R_{q,m}$.

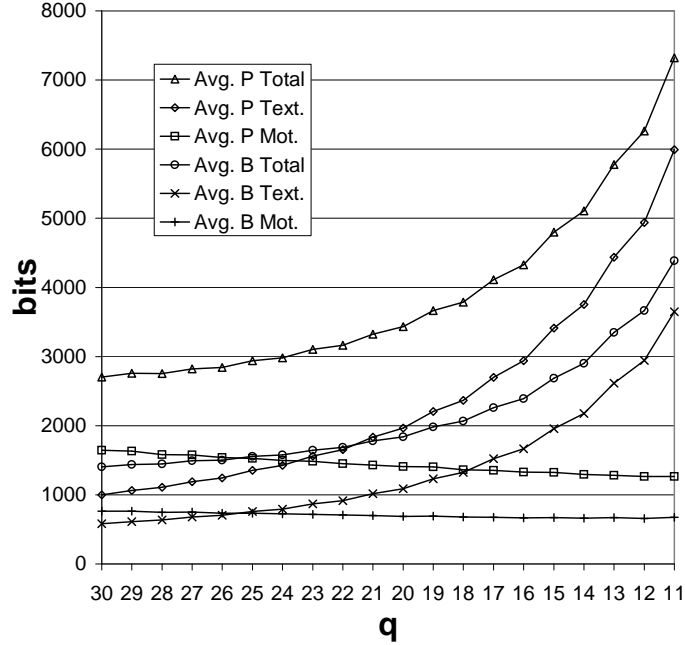


Fig. 3. Average P- and B-frame texture, motion bits and total frame sizes.

With this approximation we can express the CoV_q for the P- and B-frames as:

$$CoV_q^{(P,B)} = \frac{\sigma_q}{R_q} = \frac{\sqrt{\sigma_{q,t}^2 + \sigma_{q,m}^2 + 2 \cdot cov_q(t, m)}}{\bar{R}_{q,t} + \bar{R}_{q,m}}. \quad (4)$$

For small q values, the motion bits are negligible compared to the texture bits and therefore equation (4) reduces to:

$$CoV_{small\ q}^{(P,B)} = CoV_q^{(I)} = \frac{\sigma_{q,t}}{\bar{R}_{q,t}}. \quad (5)$$

Formula (5) is applicable to the I-frames as well, since no motion information is present.

In the following sections, we will formulate models for each of the statistical constituents in equations (4) and (5). We also show that the complex concave VD curve shape is the combined result of simpler linear and quadratic curves.

C. Quadratic Models for Texture Bits Statistics and Covariance

In [3], a quadratic rate-distortion model is devised for rate control. The rate control algorithm based on this model was adopted as part of MPEG-4 VM5.0. The model is formulated with a and b as the model parameters:

$$R_q = a \cdot q^{-1} + b \cdot q^{-2}. \quad (6)$$

This quadratic model is obtained after expanding the rate-distortion function $R(D) = \ln(\frac{1}{\alpha \cdot D})$, $0 < D < \frac{1}{\alpha}$, into a Taylor series and retaining the most important contributing terms. A Laplacian distribution with standard deviation $\sigma = \frac{\sqrt{2}}{\alpha}$ (defines parameter α) for the source model is assumed and the distortion measure $D = |x - \hat{x}|$ is used. Instead of the absolute difference distortion measure D , the quantization scale q is chosen [3] for the quadratic model in equation (6) because of its computational simplicity.

In this paper, we will employ (6) to model the average number of texture bits $\bar{R}_{q,t}$ in (4), since the average texture bits as a function of q represent a rate-distortion curve. We also show that (6) accurately models $\sigma_{q,t}^2$ and is adequate for modeling $cov_q(t, m)$. In the following, we illustrate the modeling of the *Star Wars V* segment.

In Figure 4, the actual average number of texture bits $\bar{R}_{q,t}^{(P)}$ for the P-frames are depicted together with the quadratic model. In Figure 5, the model for the average number of B-frames texture bits $\bar{R}_{q,t}^{(B)}$ is depicted along with the actual $\bar{R}_{q,t}^{(B)}$ curve. In Figure 6, $\bar{R}_{q,t}^{(I)}$ and its model are shown. Figures 7-8-9-10 depict the actual variances $\sigma_{q,t}^2$ and the actual covariances $cov_q(t, m)$ for the P-, B-frames and their corresponding quadratic models. In Figure 11, $\sigma_{q,t}^{2(I)}$ and its quadratic model are shown.

Appendix A contains five modeling examples that are selected from Table I, i.e., one scene from each motion class. All model parameters are obtained from the statistics (average, (co)variance) corresponding to the encodings with quantization scales $q = 10$ and $q = 30$. These two q values are chosen because they include the ‘‘hump’’ of the VD curve and result in a reasonably accurate prediction of the VD curves. Second, the modeling accuracy in the $q \leq 10$ is less of a concern to the application of low bit rate video streaming, although we will illustrate in Section IV-F that the VD curve modeling extends towards small q or high bit rates as well. In Section V, we analyze the optimal choice for the two q values that we have chosen.

We have calculated RMSE values as a measure of the modeling error in the quantization scale interval $q \in [10, 30]$ across all scenes from Table I and the *Star Wars V* segment. The RMSE values are enumerated in Table II. Since the absolute values of the modeled means and variances vary widely, we assess the modeling accuracy with the relative RMSE (RRMSE) values, i.e., the RMSE values divided by the average of the actual data that is modeled on the quantizer scale interval $q \in [10, 30]$. The RMSE values are computed based on the difference between the quadratic models and the actual values of the average texture bits, texture variances, and covariances (motion, texture) of the I-, P-, and B-frames. We enumerate the RRMSE (%) values in Table III. We conclude that the quadratic models match the average and variance statistics curves for the texture bits well, while adequately approximating the covariance curves of the P- and B-frames.

Next, we explain a method for estimating the model parameters a and b . Let X_1 and X_2 represent $\bar{R}_{q,t}$, $\sigma_{q,t}^2$, or $cov_q(t, m)$ corresponding to two quantization scales q_1 and q_2 . The quadratic model parameters a and b from equation (6) are obtained by solving the following system of equations:

$$X_1 = a \cdot q_1^{-1} + b \cdot q_1^{-2} \quad (7)$$

$$X_2 = a \cdot q_2^{-1} + b \cdot q_2^{-2}. \quad (8)$$

The solution to these equations is:

$$a = \frac{q_1^2 \cdot X_1 - q_2^2 \cdot X_2}{q_1 - q_2} \quad (9)$$

$$b = \frac{q_1^2 \cdot q_2 \cdot X_1 - q_2^2 \cdot q_1 \cdot X_2}{q_2 - q_1}. \quad (10)$$

TABLE II
RMSE VALUES FOR LINEAR AND QUADRATIC MODELS.

Video	Scene	Mot. Act.	$\overline{R}_t^{(I)}$	$\sigma_t^{2(I)}$	$\overline{R}_t^{(P)}$	$\overline{R}_m^{(P)}$ (lin.)	$\sigma_t^{2(P)}$	$\sigma_m^{2(P)}$ (lin.)
S.W. V	N/A		242.99	1383866.14	86.80	25.47	228430.62	26404.04
Footb.	298	I	524.14	508888.57	138.34	111.24	90921.64	20248.79
S.W. IV	274	I	62.40	19429.69	109.43	50.59	22325.25	28691.83
Term.	384	I	157.69	27978.20	112.29	71.47	128959.17	27232.71
Mean		I	248.08	185432.15	120.02	77.77	80735.35	25391.11
Footb.	299	II	421.58	2267019.87	115.44	27.53	310131.21	44243.25
S.W. IV	117	II	518.13	16556.77	68.73	14.60	6284.56	7406.55
Term.	462	II	58.24	78753.84	104.83	15.99	98740.58	6952.03
Mean		II	332.65	787443.49	96.33	19.37	138385.45	19533.94
Footb.	557	III	119.43	48063.88	37.91	14.65	33372.99	26115.25
S.W. IV	115	III	191.37	427432.87	42.08	30.26	119411.07	20734.88
Term.	628	III	77.48	167409.71	127.85	24.47	33587.83	12027.97
Mean		III	129.43	214302.15	69.28	23.13	62123.96	19626.04
Footb.	184	IV	111.77	183134.43	231.84	17.56	559152.45	37660.59
S.W. IV	165	IV	263.94	130786.39	48.19	42.06	35043.42	41157.33
Term.	262	IV	337.40	389066.82	326.96	44.23	636830.40	76570.13
Mean		IV	237.70	234329.21	202.33	34.61	410342.09	51796.02
Footb.	336	V	212.16	55411.21	461.95	13.92	142944.65	18132.79
S.W. IV	632	V	273.51	33335.59	30.25	50.68	27530.93	41138.43
Term.	441	V	46.12	87904.78	471.23	23.51	41973.00	8304.15
Mean		V	177.26	58883.86	321.14	29.37	70816.20	22525.12
Mean		I-V	225.02	296078.17	161.82	36.85	152480.61	27774.45
Video	Scene	Mot. Act.	$cov^{(P)}$	$\overline{R}_t^{(B)}$	$\overline{R}_m^{(B)}$ (lin.)	$\sigma_t^{2(B)}$	$\sigma_m^{2(B)}$ (lin.)	$cov^{(B)}$
S.W. V	N/A		187787.04	62.69	19.95	287525.95	21196.52	37082.97
Footb.	298	I	68056.11	159.90	41.55	34024.81	1085.32	5590.46
S.W. IV	274	I	6677.40	122.39	29.71	22534.08	12866.61	17822.93
Term.	384	I	23919.88	120.98	21.21	74835.55	8096.08	11102.44
Mean		I	32884.46	134.42	30.82	43798.15	7349.34	11505.28
Footb.	299	II	196742.83	82.71	22.81	244887.05	19032.25	38513.94
S.W. IV	117	II	2386.20	74.68	79.18	8099.15	60240.12	18270.61
Term.	462	II	40874.80	29.87	17.41	27247.60	25266.38	8292.64
Mean		II	80001.28	62.42	39.80	93411.27	34846.25	21692.40
Footb.	557	III	60151.00	68.23	27.81	22035.69	21810.91	14035.22
S.W. IV	115	III	24724.87	48.39	65.14	74608.32	65144.41	9873.49
Term.	628	III	14623.44	35.33	16.39	85310.52	6407.78	4089.54
Mean		III	33166.44	50.65	36.45	60651.51	31121.03	9332.75
Footb.	184	IV	260958.44	102.26	11.84	237911.63	10894.88	80925.19
S.W. IV	165	IV	58275.53	46.52	8.78	25172.82	17400.08	38827.15
Term.	262	IV	131043.19	175.95	17.45	1121663.04	16135.22	147764.09
Mean		IV	150092.39	108.24	12.69	461582.50	14810.06	89172.14
Footb.	336	V	56462.12	349.28	13.87	15015.62	3467.08	4423.94
S.W. IV	632	V	28993.00	27.17	20.29	26198.93	26310.92	23730.88
Term.	441	V	14170.20	272.35	86.16	354588.50	26054.11	180523.90
Mean		V	33208.44	216.26	40.11	131934.35	18610.70	69559.57
Mean		I-V	65870.60	114.40	31.97	158275.55	21347.48	40252.43

TABLE III
RRMSE (%) VALUES FOR LINEAR AND QUADRATIC MODELS.

Video	Scene	Mot.	$\overline{R}_t^{(I)}$	$\sigma_t^{2(I)}$	$\overline{R}_t^{(P)}$	$\overline{R}_m^{(P)}$	$\sigma_t^{2(P)}$	$\sigma_m^{2(P)}$	$cov^{(P)}$	$\overline{R}_t^{(B)}$	$\overline{R}_m^{(B)}$	$\sigma_t^{2(B)}$	$\sigma_m^{2(B)}$	$cov^{(B)}$
<i>S.W. V</i>	N/A		1.79	6.41	3.32	1.79	4.63	2.54	11.32	4.12	2.84	18.01	6.76	10.82
<i>Footb.</i>	298	I	2.39	12.88	9.82	30.38	2.87	5.42	7.06	22.46	32.64	32.45	7.43	25.69
<i>S.W. IV</i>	274	I	0.73	9.66	12.87	7.12	12.05	18.92	5.96	21.40	5.53	28.46	10.97	39.14
<i>Term.</i>	384	I	1.54	12.69	11.56	13.83	16.08	10.76	6.68	23.58	7.13	27.23	8.28	11.91
<i>Mean</i>		I	1.55	11.74	11.42	17.11	10.34	11.70	6.56	22.48	15.10	29.38	8.89	25.58
<i>Footb.</i>	299	II	2.36	8.84	3.61	1.75	8.72	4.63	14.13	4.87	3.41	20.07	7.60	12.84
<i>S.W. IV</i>	117	II	12.99	5.82	16.00	2.78	12.96	8.18	8.53	23.40	8.95	25.61	16.03	26.81
<i>Term.</i>	462	II	0.52	12.62	5.59	2.36	7.09	4.96	11.72	3.27	3.13	11.06	11.86	11.75
<i>Mean</i>		II	5.29	9.09	8.40	2.30	9.59	5.92	11.46	10.51	5.16	18.91	11.83	17.13
<i>Footb.</i>	557	III	1.11	4.12	2.12	1.01	4.57	9.62	18.46	7.02	4.34	3.98	15.03	8.68
<i>S.W. IV</i>	115	III	2.15	4.31	2.76	2.91	19.95	6.77	8.27	5.10	8.85	27.25	31.44	8.10
<i>Term.</i>	628	III	0.74	8.42	4.63	1.32	7.75	5.68	10.26	2.24	1.75	19.40	4.62	3.95
<i>Mean</i>		III	1.33	5.62	3.17	1.75	10.76	7.36	12.33	4.78	4.98	16.88	17.03	6.91
<i>Footb.</i>	184	IV	0.69	7.81	4.24	0.68	5.52	2.72	8.35	3.05	1.19	3.48	3.04	6.42
<i>S.W. IV</i>	165	IV	3.35	20.34	2.28	1.62	7.69	5.89	13.93	3.60	0.70	6.34	4.26	12.84
<i>Term.</i>	262	IV	2.54	3.92	6.59	1.84	10.74	10.05	8.47	5.28	1.20	17.42	2.96	13.00
<i>Mean</i>		IV	2.19	10.69	4.37	1.38	7.98	6.22	10.25	3.97	1.03	9.08	3.42	10.75
<i>Footb.</i>	336	V	1.41	6.87	6.48	0.35	8.67	9.15	35.13	6.54	0.75	4.72	8.15	29.28
<i>S.W. IV</i>	632	V	4.94	10.50	1.63	1.46	6.74	24.54	15.40	2.08	0.98	3.81	10.76	10.98
<i>Term.</i>	441	V	0.38	3.34	6.86	0.57	3.61	4.74	10.45	5.18	4.16	5.83	4.98	11.99
<i>Mean</i>		V	2.24	6.90	4.99	0.79	6.34	12.81	20.33	4.60	1.96	4.79	7.97	17.41
<i>Mean</i>		I-V	2.52	8.81	6.47	4.66	9.00	8.80	12.19	9.27	5.65	15.81	9.83	15.56

D. Linear Models for Motion Bits Statistics

We observe in Figure 3 that the average number of motion bits follows a linear trend as a function of q . Hence, we propose a linear model with c and d as the model parameters:

$$\overline{R}_{q,m} = c \cdot q + d. \quad (11)$$

The model parameters can be estimated easily by solving a system of two equations as before for the quadratic model:

$$X_1 = c \cdot q_1 + d \quad (12)$$

$$X_2 = c \cdot q_2 + d. \quad (13)$$

The solution to these equations is:

$$c = \frac{X_1 - X_2}{q_1 - q_2} \quad (14)$$

$$d = \frac{q_1 \cdot X_2 - q_2 \cdot X_1}{q_1 - q_2}. \quad (15)$$

Figure 12 depicts the actual $\overline{R}_{q,m}^{(P)}$ and the linear model for the *Star Wars V* video segment. Also the actual $\overline{R}_{q,m}^{(B)}$ curve is shown together with its linear model. We observe that the linear model accurately fits the average curves. As for the quadratic models, the model parameters are obtained from the motion bits statistics (average, variance) corresponding to the encodings with quantization scales $q = 10$ and $q = 30$.

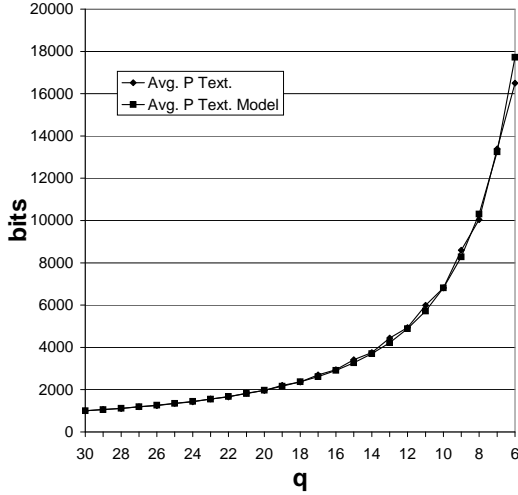


Fig. 4. Comparison of actual average number of texture bits per P-frame with corresponding quadratic model.

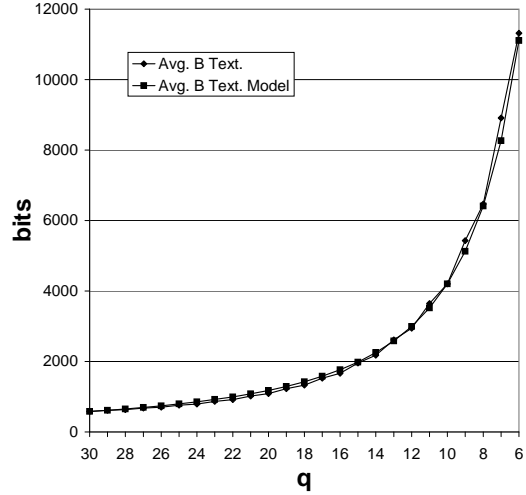


Fig. 5. Comparison of actual average number of texture bits per B-frame with corresponding quadratic model.

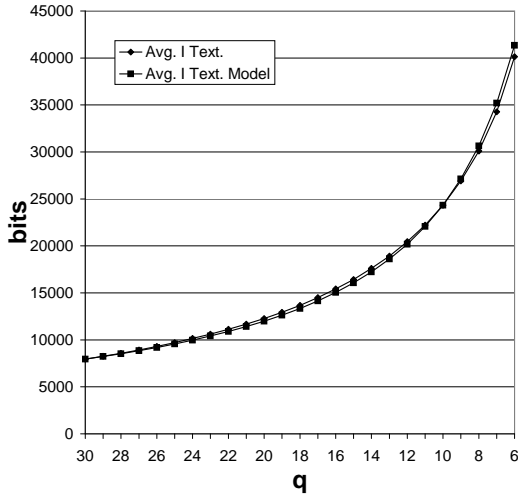


Fig. 6. Comparison of actual average number of texture bits per I-frame with corresponding quadratic model.

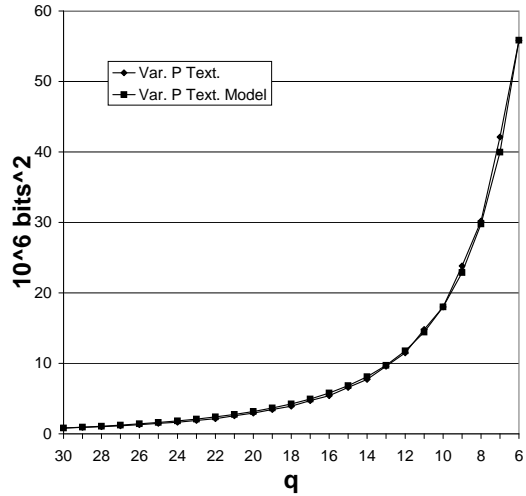


Fig. 7. Comparison of actual variance of P-frame texture bits with corresponding quadratic model.

The last constituent of equation (4) to be modeled is the variance of the motion bits, $\sigma_{q,m}^2$. Figure 13 depicts $\sigma_{q,m}^2$ for the P- and B-frames. The linear model is also the most appropriate in this case. Appendix A illustrates these linear models for five selected scenes, one for each motion class.

We have calculated the RMSE values for the linear motion bits statistics models and this for all scenes from Table I and for the *Star Wars V* segment. The RMSE values are enumerated in Table II for the P- and B-frames. As for the quadratic models, we also computed the RRMSE values (%) and enumerate them in Table III. We empirically conclude that the linear models fit the average and the variance of the number of motion bits well.

Now that we have developed the models for the texture and motion bits statistics, we need to assemble the VD curve model. We first illustrate this with the VD curve models for the I-, P-, and B-frames separately in the next Section IV-E. Second, the overall VD curve model, or “combined VD curve model”, is derived

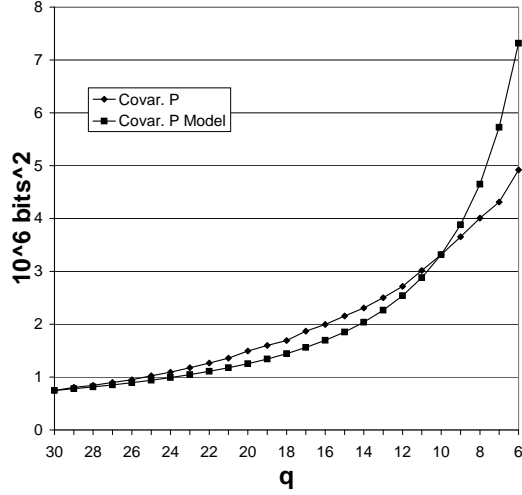


Fig. 8. Comparison of actual covariance of P-frame bits (texture, motion) with corresponding quadratic model.

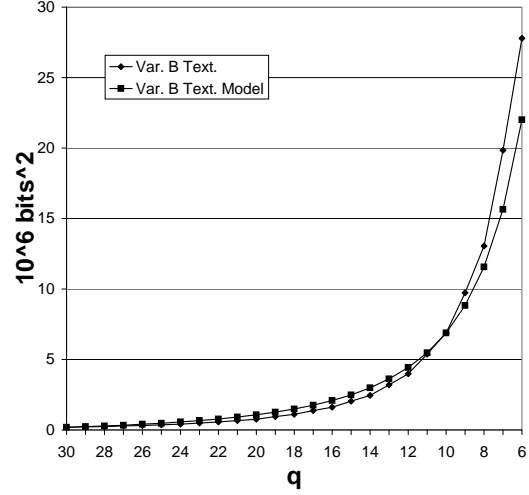


Fig. 9. Comparison of actual variance of B-frame texture bits with corresponding quadratic model.

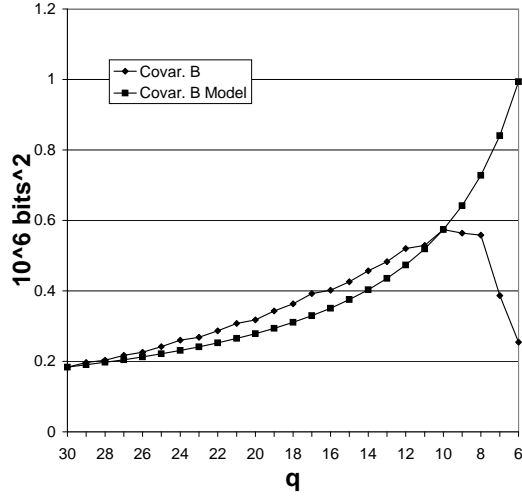


Fig. 10. Comparison of actual covariance of B-frame bits (texture, motion) with corresponding quadratic model.

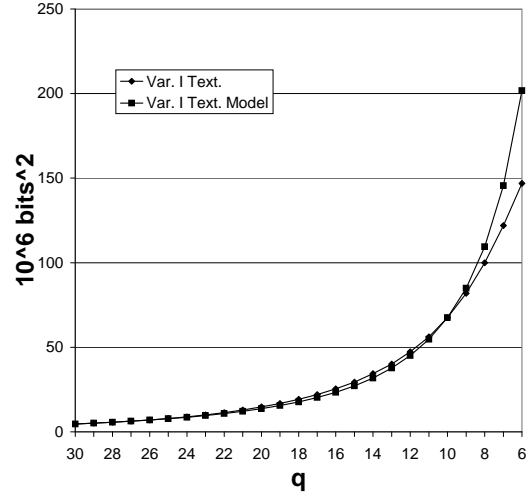


Fig. 11. Comparison of actual variance of I-frame texture bits with corresponding quadratic model.

and validated in Section IV-F.

E. VD Curve Model for I-, P-, and B-Frames

In the previous sections, we developed models for each of the constituents of the VD curves as given by (4) and (5). The complete VD curve models for the P- and B-frames can be reformulated as a function of q and ten model parameters, two parameters for each linear and quadratic model:

$$\sigma_{q,t}^2 = a_1/q + b_1/q^2 \quad (16)$$

$$\text{cov}_q(t, m) = a_2/q + b_2/q^2 \quad (17)$$

$$\overline{R}_{q,t} = a_3/q + b_3/q^2 \quad (18)$$

$$\sigma_{q,m}^2 = c_1 \cdot q + d_1 \quad (19)$$

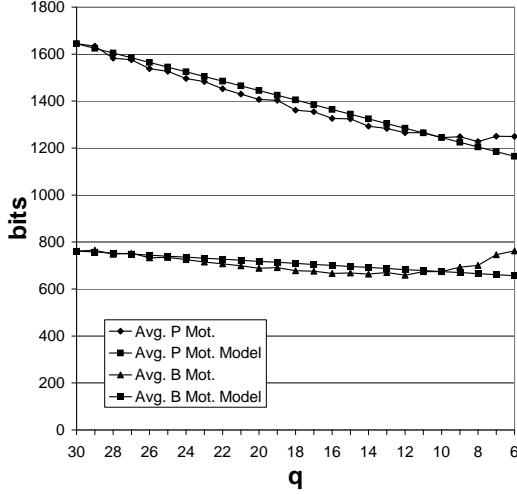


Fig. 12. Comparison of actual average number of P, B motion bits with corresponding linear models.

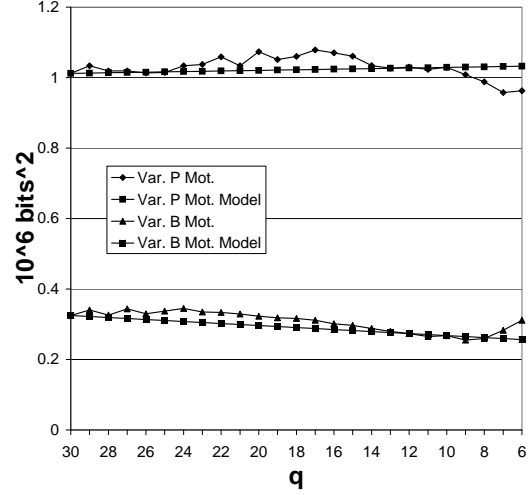


Fig. 13. Comparison of actual variances of P, B motion bits with corresponding linear models.

$$\bar{R}_{q,m} = c_2 \cdot q + d_2 \quad (20)$$

$$CoV_q^{(P,B)} = \frac{\sqrt{a_1/q + b_1/q^2 + c_1 \cdot q + d_1 + 2 \cdot (a_2/q + b_2/q^2)}}{a_3/q + b_3/q^2 + c_2 \cdot q + d_2}. \quad (21)$$

Analogously, the VD curve model for small q values and for the I-frames includes four model parameters and is given by:

$$\sigma_{q,t}^2 = a_1/q + b_1/q^2 \quad (22)$$

$$\bar{R}_{q,t} = a_2/q + b_2/q^2 \quad (23)$$

$$CoV_{small\ q}^{(P,B)} = CoV_q^{(I)} = \frac{a_1/q + b_1/q^2}{a_2/q + b_2/q^2}. \quad (24)$$

From equations (21) and (24) it is clear that the *complex concave VD curve shape is the combined result of simpler linear and quadratic curves for the averages, variances, and covariances of the texture and motion bits.*

In Figures 14 and 15, the actual VD curves for the P- and B-frames from the *Star Wars V* video segment are depicted. The VD models estimated from encoding settings $q = 10$, $q = 30$ and employing equation (21) are also shown. The models match the original curves well for $10 \leq q \leq 30$ and capture the concave VD curve shape. The VD-P and VD-B models are also an accurate representation for small q or equivalently the highest qualities. In Figure 16, the actual VD curve and model for the I-frames are depicted. The model matches the VD curve well for $10 \leq q \leq 30$. However, the VD-I model becomes inaccurate for $q < 10$. Appendix B illustrates the VD-I, VD-P, and VD-B curve models for all video scenes.

Table IV enumerates the RMSE values of the VD-I, VD-P, and VD-B models corresponding to all video scenes in Table I and the *Star Wars V* segment, computed for the quantization scale range above. Table V contains the RRMSE values. We empirically conclude that the VD curve models for the I-, P-, and B-frames offer a good prediction of the actual VD curves.

TABLE IV
RMSE VALUES FOR VD CURVE MODELS OF I-, P-, AND B-FRAMES.

Video	Scene	Mot. Act.	VD-I	VD-P	VD-B
<i>S.W. V</i>	N/A		0.0463	0.0481	0.0558
<i>Footb.</i>	298	I	0.0459	0.2255	0.0885
<i>S.W. IV</i>	274	I	0.0455	0.0583	0.0557
<i>Term.</i>	384	I	0.0457	0.0870	0.0745
<i>Mean</i>		I	0.0457	0.1236	0.0729
<i>Footb.</i>	299	II	0.0462	0.0486	0.0517
<i>S.W. IV</i>	117	II	0.0529	0.0512	0.0838
<i>Term.</i>	462	II	0.0463	0.0495	0.0619
<i>Mean</i>		II	0.0485	0.0498	0.0658
<i>Footb.</i>	557	III	0.0460	0.0484	0.0567
<i>S.W. IV</i>	115	III	0.0492	0.0496	0.0562
<i>Term.</i>	628	III	0.0460	0.0476	0.0589
<i>Mean</i>		III	0.0471	0.0485	0.0573
<i>Footb.</i>	184	IV	0.0458	0.0459	0.0495
<i>S.W. IV</i>	165	IV	0.0467	0.0489	0.0490
<i>Term.</i>	262	IV	0.0462	0.0569	0.0809
<i>Mean</i>		IV	0.0462	0.0505	0.0598
<i>Footb.</i>	336	V	0.0458	0.0457	0.0459
<i>S.W. IV</i>	632	V	0.0479	0.0471	0.0476
<i>Term.</i>	441	V	0.0458	0.0462	0.0548
<i>Mean</i>		V	0.0465	0.0463	0.0494
<i>Mean</i>		I-V	0.0468	0.0638	0.0610

TABLE V
RRMSE (%) VALUES FOR VD CURVE MODELS OF I-, P-, AND B-FRAMES.

Video	Scene	Mot. Act.	VD-I	VD-P	VD-B
<i>S.W. V</i>	N/A		1.07	1.47	3.45
<i>Footb.</i>	298	I	4.72	14.70	11.17
<i>S.W. IV</i>	274	I	3.63	5.63	4.82
<i>Term.</i>	384	I	8.80	7.04	5.24
<i>Mean</i>		I	4.18	9.12	7.08
<i>Footb.</i>	299	II	1.61	2.07	2.94
<i>S.W. IV</i>	117	II	8.88	3.50	8.58
<i>Term.</i>	462	II	8.08	2.37	6.18
<i>Mean</i>		II	6.19	2.65	5.90
<i>Footb.</i>	557	III	2.79	3.10	4.16
<i>S.W. IV</i>	115	III	3.40	2.85	4.89
<i>Term.</i>	628	III	3.05	4.76	8.59
<i>Mean</i>		III	3.08	3.57	5.88
<i>Footb.</i>	184	IV	3.68	0.76	2.02
<i>S.W. IV</i>	165	IV	7.38	4.23	2.66
<i>Term.</i>	262	IV	2.34	6.27	9.63
<i>Mean</i>		IV	4.47	3.75	4.77
<i>Footb.</i>	336	V	3.65	2.76	5.40
<i>S.W. IV</i>	632	V	7.63	4.57	2.25
<i>Term.</i>	441	V	3.42	4.55	5.68
<i>Mean</i>		V	4.90	3.96	4.44
<i>Mean</i>		I-V	4.56	4.61	5.61

In Figures 17 and 18, the “small q ” models (5) for the P- and B-frames from the *Star Wars V* segment are depicted (models use $q = 10$ and $q = 30$). They are adequate approximations for the respective VD curves in the small q range. From this, it can also be seen that in the small q range the motion statistics and the covariance do not influence the VD curve shape or equivalently there is no influence on the bit rate variability. On the other hand, in the range $10 \leq q \leq 30$, the statistics of the motion bits need to be modeled, as well as the covariance, otherwise the curve based on the texture statistics alone deviates strongly from the actual VD curve.

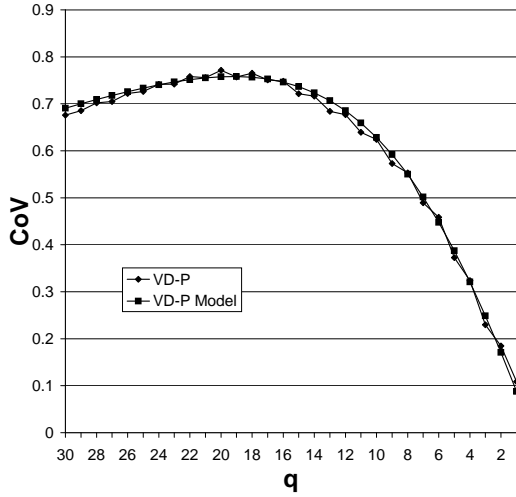


Fig. 14. Comparison of actual VD-P curve with corresponding model.

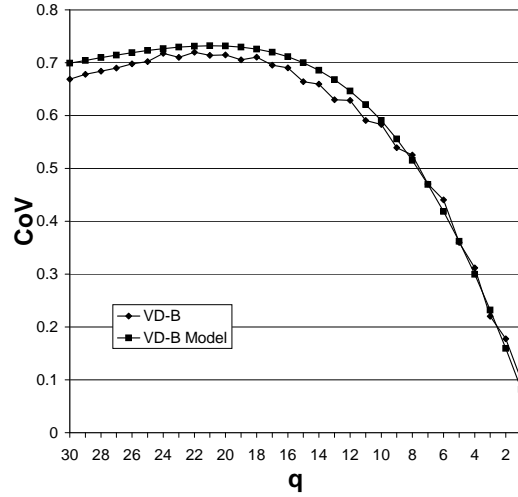


Fig. 15. Comparison of actual VD-B curve with corresponding model.

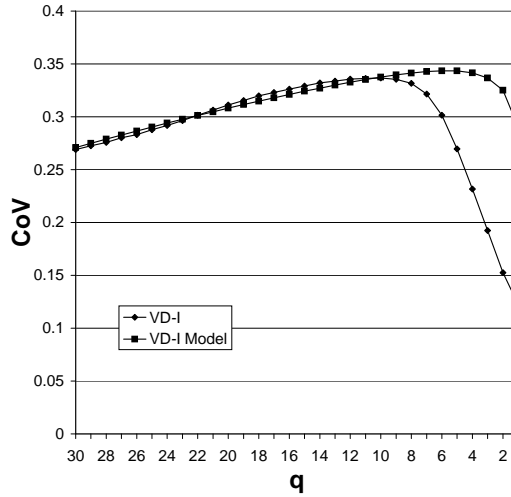
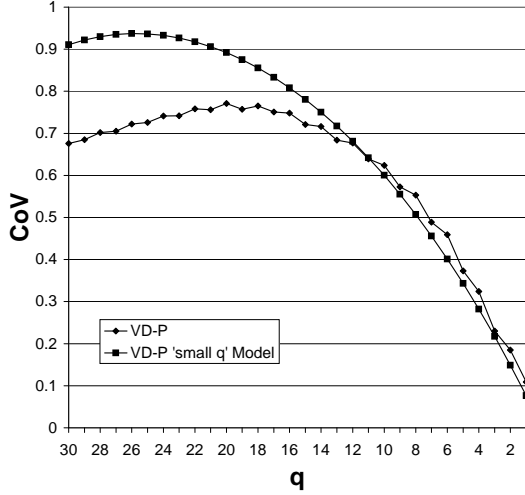
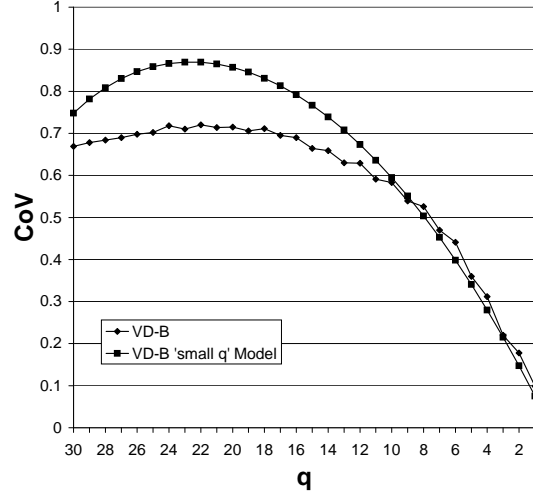


Fig. 16. Comparison of actual VD-I curve with corresponding model.

F. Combined VD Curve Model

The last step in the modeling is combining all linear and quadratic models of the statistical parameters for the I-, P-, and B-frames into a single VD curve model. The frame size variabilities expressed by the

Fig. 17. Actual VD-P curve and the “small q ” model.Fig. 18. Actual VD-B curve and the “small q ” model.

VD-I, VD-P, and VD-B models will each contribute to the overall variability of the entire frame sequence. Intuition tells us that the relative contributions of the I-, P-, and B-frames to the overall variability depend on the group-of-pictures (GoP) structure. Furthermore, it is mathematically obvious that the variance is not a simple linear combination of the frame type variances. However, the overall average is equal to the linear combination of the frame type averages as we show next.

The GoP structure that is used in this work consists of twelve frames and is given by: *IBBPBBPBBPBB*, i.e., one I-frame, three P-frames and eight B-frames. Denote the I-, P-, and B-frame fractions by: $a^{(I)}$, $a^{(P)}$, and $a^{(B)}$. The fractions for our chosen GoP structure are: $a^{(I)} = \frac{1}{12}$, $a^{(P)} = \frac{1}{4}$, and $a^{(B)} = \frac{2}{3}$. In the following, we assume there are an integer number N of GoPs in the sequence so that the total number of frames equals $12 \times N$. The overall average of frame sizes for a particular quantization scale q , \bar{R}_q , can therefore be expressed as follows, with $X_{q,i}$ denoting individual frame sizes:

$$\bar{R}_q = \frac{1}{12N} \left(\sum_{i=1}^N X_{q,i}^{(I)} + \sum_{i=1}^{3N} X_{q,i}^{(P)} + \sum_{i=1}^{8N} X_{q,i}^{(B)} \right) \quad (25)$$

$$= \frac{1}{12N} \left(N \cdot \bar{R}_q^{(I)} + 3N \cdot \bar{R}_q^{(P)} + 8N \cdot \bar{R}_q^{(B)} \right) \quad (26)$$

$$= \frac{1}{12} \cdot \bar{R}_q^{(I)} + \frac{1}{4} \cdot \bar{R}_q^{(P)} + \frac{2}{3} \cdot \bar{R}_q^{(B)}. \quad (27)$$

Hence, the overall average of frame sizes is the weighted sum of the averages per frame type (I-P-B) with weighting coefficients equal to the fractions of each frame type in the GoP structure. Furthermore, we have $\bar{R}_q^{(I,P,B)} = \bar{R}_{q,t}^{(I,P,B)} + \bar{R}_{q,m}^{(I,P,B)}$, with each texture and motion average modeled by the linear and quadratic models that we presented in Sections IV-C and IV-D.

To calculate the overall variance, we follow a similar approach. The overall variance for a particular quantization scale q , σ_q^2 , is given by:

$$\sigma_q^2 = \frac{1}{12N} \left[\sum_{i=1}^N \left(X_{q,i}^{(I)} - \bar{R}_q \right)^2 + \sum_{i=1}^{3N} \left(X_{q,i}^{(P)} - \bar{R}_q \right)^2 + \sum_{i=1}^{8N} \left(X_{q,i}^{(B)} - \bar{R}_q \right)^2 \right] \quad (28)$$

$$\begin{aligned}
&= \frac{1}{12N} \sum_{i=1}^N \left(\left\{ X_{q,i}^{(I)} - \bar{R}_q^{(I)} \right\} + \left\{ \bar{R}_q^{(I)} - \bar{R}_q \right\} \right)^2 + \frac{1}{12N} \sum_{i=1}^{3N} \left(\left\{ X_{q,i}^{(P)} - \bar{R}_q^{(P)} \right\} + \left\{ \bar{R}_q^{(P)} - \bar{R}_q \right\} \right)^2 \\
&\quad + \frac{1}{12N} \sum_{i=1}^{8N} \left(\left\{ X_{q,i}^{(B)} - \bar{R}_q^{(B)} \right\} + \left\{ \bar{R}_q^{(B)} - \bar{R}_q \right\} \right)^2 \tag{29}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{12N} \left[\sum_{i=1}^N \left(X_{q,i}^{(I)} - \bar{R}_q^{(I)} \right)^2 + \sum_{i=1}^{3N} \left(X_{q,i}^{(P)} - \bar{R}_q^{(P)} \right)^2 + \sum_{i=1}^{8N} \left(X_{q,i}^{(B)} - \bar{R}_q^{(B)} \right)^2 \right] \\
&\quad + \frac{1}{6N} \left[\left(\bar{R}_q^{(I)} - \bar{R}_q \right) \sum_{i=1}^N \left(X_{q,i}^{(I)} - \bar{R}_q^{(I)} \right) \right] + \frac{1}{6N} \left[\left(\bar{R}_q^{(P)} - \bar{R}_q \right) \sum_{i=1}^{3N} \left(X_{q,i}^{(P)} - \bar{R}_q^{(P)} \right) \right] \\
&\quad + \frac{1}{6N} \left[\left(\bar{R}_q^{(B)} - \bar{R}_q \right) \sum_{i=1}^{8N} \left(X_{q,i}^{(B)} - \bar{R}_q^{(B)} \right) \right] \\
&\quad + \frac{1}{12} \left[\left(\bar{R}_q^{(I)} - \bar{R}_q \right)^2 + 3 \cdot \left(\bar{R}_q^{(P)} - \bar{R}_q \right)^2 + 8 \cdot \left(\bar{R}_q^{(B)} - \bar{R}_q \right)^2 \right]. \tag{30}
\end{aligned}$$

The $\frac{1}{12N}$ terms can clearly be rewritten as the variances of the three frame types (I-P-B). The $\frac{1}{6N}$ terms each sum over the individual frame sizes about the mean for each frame type, which is the first central moment. Hence, all $\frac{1}{6N}$ terms equal zero. The last terms with coefficient $\frac{1}{12}$ need some expansion to simplify them. This results in the following formula for the overall variance of the frame sizes:

$$\sigma_q^2 = \frac{1}{12} \cdot \sigma_q^{2(I)} + \frac{1}{4} \cdot \sigma_q^{2(P)} + \frac{2}{3} \cdot \sigma_q^{2(B)} + \frac{1}{12} \cdot \left(\bar{R}_q^{(I)} \right)^2 + \frac{1}{4} \cdot \left(\bar{R}_q^{(P)} \right)^2 + \frac{2}{3} \cdot \left(\bar{R}_q^{(B)} \right)^2 - \bar{R}_q^2. \tag{31}$$

We split each of the variance and average terms into texture and motion contributions that we modeled with linear and quadratic models in sections IV-C and IV-D. We can easily generalize equations (27) and (31) for arbitrary GoP structures by adjusting the coefficients to represent the fractions of each frame type in the GoP. The general formula for the VD curve model that combines the variabilities of all frame types, or “combined VD curve model”, is therefore:

$$CoV_q = \sqrt{\frac{a^{(I)} \sigma_q^{2(I)} + a^{(P)} \sigma_q^{2(P)} + a^{(B)} \sigma_q^{2(B)} + a^{(I)} \left(\bar{R}_q^{(I)} \right)^2 + a^{(P)} \left(\bar{R}_q^{(P)} \right)^2 + a^{(B)} \left(\bar{R}_q^{(B)} \right)^2}{\left[a^{(I)} \bar{R}_q^{(I)} + a^{(P)} \bar{R}_q^{(P)} + a^{(B)} \bar{R}_q^{(B)} \right]^2} - 1. \tag{32}$$

We can rewrite this expression as a function of the coefficients of variation of the I-, P-, and B-frames as follows:

$$CoV_q = \sqrt{\frac{a^{(I)} \bar{R}_q^{(I) 2} CoV_q^{(I) 2} + 1 + a^{(P)} \bar{R}_q^{(P) 2} CoV_q^{(P) 2} + 1 + a^{(B)} \bar{R}_q^{(B) 2} CoV_q^{(B) 2} + 1}{\left[a^{(I)} \bar{R}_q^{(I)} + a^{(P)} \bar{R}_q^{(P)} + a^{(B)} \bar{R}_q^{(B)} \right]^2} - 1. \tag{33}$$

Figure 19 depicts the combined VD model for the *Star Wars V* video segment. The model accurately approximates the actual VD curve of this sequence. In Appendix B, we present the combined VD curve models for all video scenes from Table I. The combined models approximate the actual VD curves well in the quantization scale range $10 \leq q \leq 30$ and fairly well in the small quantization scale range, corresponding to high frame qualities or high bit rates. The modeling error of the I-frames statistics in this small q range influences the combined VD curve model in some instances, such as for *Football* scene 557, *Star Wars IV* scene 117, and *Terminator* scene 462. However this modeling error is acceptable in case of low bit rate video streaming.

In Tables VI and VII, we compare the combined VD curve model with the piecewise approximation model applied in [1] for all scenes from Table I. We used the RMSE and RRMSE to compare both approaches. For the piecewise modeling approach, we present the RMSE and RRMSE values for approximations obtained from respectively 2, 3, and 4 quantization scales. The RMSE and RRMSE values are in all cases significantly lower for the combined VD model compared to the piecewise model employing two quantization scales. For ten out of the fifteen video scenes, the RMSE and RRMSE values of the combined VD curve model are smaller than the RMSE and RRMSE values of the piecewise model obtained from three quantization scales, and for about half of the scenes the combined VD curve model outperforms the piecewise RMSE and RRMSE values obtained from four scales. We conclude from the overall mean that the approximation performance of the combined VD curve model in the quantization scale interval $10 \leq q \leq 30$, based on quantization scales $q = 10$ and $q = 30$, is comparable to the piecewise model employing three quantization scales.

We wish to point out that the combined VD curve model derived in this section is independent from the linear and quadratic models employed for the texture and motion bits statistics. Other models for these statistics can be plugged into the combined VD curve model to further improve the modeling accuracy. This is a possible direction for future research.

TABLE VI

COMPARISON BETWEEN PIECEWISE APPROXIMATION MODEL FROM [1] WITH 2, 3, 4 SAMPLES AND COMBINED VD MODEL (RMSE VALUES).

Video	Scene	Mot. Act.	Piecew.-2	Piecew.-3	Piecew.-4	Comb. VD
<i>Footb.</i>	298	I	0.1448	0.0680	0.0690	0.1335
<i>S.W. IV</i>	274	I	0.2711	0.0758	0.0526	0.0721
<i>Term.</i>	384	I	0.2620	0.0613	0.0501	0.0786
<i>Mean</i>		I	0.2260	0.0684	0.0572	0.0947
<i>Footb.</i>	299	II	0.1068	0.0307	0.0288	0.0287
<i>S.W. IV</i>	117	II	0.0827	0.0264	0.0263	0.0189
<i>Term.</i>	462	II	0.1122	0.0383	0.0230	0.0362
<i>Mean</i>		II	0.1006	0.0318	0.0260	0.0279
<i>Footb.</i>	557	III	0.1534	0.0439	0.0216	0.0178
<i>S.W. IV</i>	115	III	0.1098	0.0267	0.0216	0.0173
<i>Term.</i>	628	III	0.0819	0.0286	0.0161	0.0156
<i>Mean</i>		III	0.1150	0.0330	0.0198	0.0169
<i>Footb.</i>	184	IV	0.0573	0.0179	0.0120	0.0114
<i>S.W. IV</i>	165	IV	0.0789	0.0264	0.0132	0.0091
<i>Term.</i>	262	IV	0.0535	0.0186	0.0114	0.0341
<i>Mean</i>		IV	0.0632	0.0210	0.0122	0.0182
<i>Footb.</i>	336	V	0.0204	0.0077	0.0067	0.0097
<i>S.W. IV</i>	632	V	0.0461	0.0226	0.0068	0.0110
<i>Term.</i>	441	V	0.0314	0.0072	0.0052	0.0170
<i>Mean</i>		V	0.0326	0.0125	0.0062	0.0125
<i>Mean</i>		I-V	0.1075	0.0333	0.0243	0.0341

V. SENSITIVITY OF VD CURVE MODEL

In this section, we study the optimal choice for the two quantization scales, q_1 and q_2 , that are used for encoding the video sequence and subsequently for estimating the averages, variances, and covariances of the number of texture and motion bits. The linear and quadratic models that we have developed, could strongly depend on the choice of q_1 and q_2 . For example, if q_1 and q_2 are not spread out over a significant portion of the quantization scale range $1 \leq q \leq 30$, then it is likely that a modeling error will be substantial for quantization scales that are located far from q_1 and q_2 . Furthermore, the ‘‘hump’’ shape is influenced by both the motion and the texture statistics, as observed earlier. Therefore, the motion statistics need to

TABLE VII
COMPARISON BETWEEN PIECEWISE APPROXIMATION MODEL FROM [1] WITH 2, 3, 4 SAMPLES AND COMBINED VD MODEL
(RRMSE (%) VALUES).

Video	Scene	Mot. Act.	Piecew.-2	Piecew.-3	Piecew.-4	Comb. VD
<i>Footb.</i>	298	I	7.80	3.66	3.72	6.28
<i>S.W. IV</i>	274	I	25.38	7.09	4.93	6.54
<i>Term.</i>	384	I	18.57	4.35	3.55	5.29
<i>Mean</i>		I	17.25	5.03	4.06	6.04
<i>Footb.</i>	299	II	10.48	3.01	2.82	2.52
<i>S.W. IV</i>	117	II	12.64	4.04	4.02	2.58
<i>Term.</i>	462	II	10.91	3.72	2.23	3.28
<i>Mean</i>		II	11.34	3.59	3.02	2.79
<i>Footb.</i>	557	III	16.77	4.80	2.37	1.86
<i>S.W. IV</i>	115	III	12.63	3.07	2.49	1.89
<i>Term.</i>	628	III	13.14	4.59	2.58	2.47
<i>Mean</i>		III	14.18	4.15	2.48	2.07
<i>Footb.</i>	184	IV	8.26	2.59	1.72	1.53
<i>S.W. IV</i>	165	IV	14.44	4.82	2.41	1.64
<i>Term.</i>	262	IV	9.10	3.17	1.93	5.80
<i>Mean</i>		IV	10.60	3.53	2.02	2.99
<i>Footb.</i>	336	V	7.15	2.72	2.34	3.07
<i>S.W. IV</i>	632	V	12.70	6.23	1.86	3.28
<i>Term.</i>	441	V	8.83	2.03	1.48	4.78
<i>Mean</i>		V	9.56	3.66	1.89	3.71
<i>Mean</i>		I-V	12.59	3.99	2.70	3.52

be modeled accurately by choosing at least one q value in the higher compression ratio region (large q). Otherwise, if q_1 and q_2 would be selected in the low compression ratio region, then we would estimate motion statistics in a region where the texture information dominates the motion information, resulting in a good approximation of the motion bits statistics in a region where motion information is insignificant for determining the VD curve shape. From these considerations, we can already formulate two high-level recommendations for the choice of q_1 and q_2 : (i) there should be a reasonable interval size $\Delta q = |q_1 - q_2|$, and (ii) at least one q should be chosen in the high compression ratio region where the “hump” is typically situated.

An important unknown is the possibility that the choice of q_1 and q_2 depends on the motion activity in the video scene, since the “hump” amplitude is influenced by the motion activity [1]. We define a scene as a collection of consecutive video frames contained in between two significant video frame content changes, such as director cuts (i.e., the abrupt change of frame content between two consecutive frames) or fades (i.e., the dissolving of content into other content). It seems reasonable to examine whether there could be better choices for q_1 and q_2 than 10 and 30 (which we used in the preceding section), depending on the motion activity, and resulting in more accurate concave VD curve models.

We use the Microsoft v2.3.0 implementation of MPEG-4 (simple profile) for the encoding of the fifteen QCIF scenes from Table I. In our sensitivity analysis of the choices for q_1 and q_2 , we examine interval sizes $\Delta q \in \{20, 30\}$, and choices for q_1 and q_2 from the range $1 \leq q \leq 30$, as enumerated in Table VIII. These four possible combinations cover the entire quantization scale interval. The encodings are followed by estimating the texture and motion statistics, and subsequently by calculating the VD curve models for the P-, B-, and I-frames. We have omitted the interval size $\Delta q = 10$, since our experiments indicate that the VD curve models for $\Delta q = 10$ are poor approximations of the actual VD curves for most video scenes. Appendix C groups the results according to motion class and video scene into Figures 28 to 42. Each figure contains the actual VD curves for the P-, B-, and I-frames. Alongside every VD curve, four models are plotted corresponding to the values for q_1 and q_2 in Table VIII. Visual inspection quickly

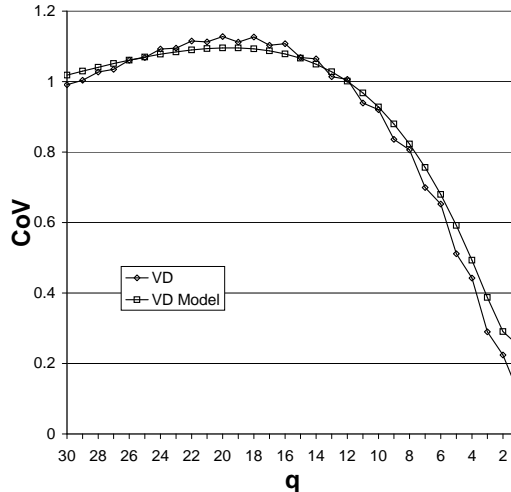


Fig. 19. VD curve for all frames and the combined VD model for the *Star Wars V* fragment.

TABLE VIII
QUANTIZATION SCALE INTERVALS Δq AND POSSIBLE CHOICES FOR q_1 AND q_2 .

Δq	q_1	q_2
20	1	20
	5	25
	10	30
30	1	30

reveals that not all choices for q_1 and q_2 result in models that are good approximations of the actual VD curves. Overall, the choice $q_1 = 1$ leads to models that are poor approximations of the actual VD curves, e.g., in Figures 32 and 35. For $q_1 = 1$ and $q_2 = 20$, the model seems unbounded to the actual curve for large q values as is the case, e.g., in Figures 28 and 30. If $q_1 = 1$ and $q_2 = 30$, the models are tied to the extreme q values and result in intermediate values that are often far from the target values such as in Figures 33 and 39. This leaves $q_1 = 5$ and $q_2 = 25$, or, $q_1 = 10$ and $q_2 = 30$. Overall, both choices result overall in better approximations than the two previous choices.

We have computed the RMSE for each quantization scale choice from Table VIII and for each video scene from Table I. The RMSE is computed for the entire range of the quantization scale. Table IX enumerates the average, maximum, and maximum standard deviation of the RMSE values from each of the five motion classes. We observe that the quantization scale choices $q_1 = 10$ and $q_2 = 30$ has overall the lowest average RMSE, lowest maximum RMSE, and lowest maximum standard deviation of all RMSE values.

From this sensitivity analysis we can recommend that the best choice for the two quantization scales is $q_1 = 10$ and $q_2 = 30$. Our experiments have revealed no evidence that the motion class has a significant impact on this choice.

VI. CONCLUSION

We have modeled the bit rate variability-distortion (VD) curve for low bit rate video streaming. The VD curve has a typical concave or “hump” shape for MPEG-4 VBR encodings as previously observed in [1]. We have refined the coefficient of variation (CoV) into the CoV of texture and motion information, since we have found that the motion information is an important constituent at medium to high compression ratios. Furthermore, the texture and motion information statistics are different functions of quantization

TABLE IX
SENSITIVITY ANALYSIS OF QUANTIZATION SCALES USED FOR MODEL PARAMETER ESTIMATION (RMSE VALUES).

Average RMSE	$q_1 = 1, q_2 = 30$	$q_1 = 1, q_2 = 20$	$q_1 = 10, q_2 = 30$	$q_1 = 5, q_2 = 25$
I (Motion Class)	0.0733	0.1299	0.0479	0.2318
II	0.1543	0.1750	0.0740	0.0705
III	0.0877	0.0746	0.0925	0.0709
IV	0.0449	0.0367	0.0321	0.0307
V	0.0347	0.0266	0.0399	0.0400
I-V	0.0790	0.0886	0.0573	0.0888
Max. RMSE	$q_1 = 1, q_2 = 30$	$q_1 = 1, q_2 = 20$	$q_1 = 10, q_2 = 30$	$q_1 = 5, q_2 = 25$
I	0.1887	0.3254	0.1738	1.4355
II	0.5355	0.5982	0.2910	0.2306
III	0.2278	0.1919	0.4539	0.2317
IV	0.0770	0.0690	0.0770	0.0740
V	0.1315	0.0954	0.1232	0.1858
I-V	0.5355	0.5982	0.4539	1.4355
Max. $\sigma(RMSE)$	$q_1 = 1, q_2 = 30$	$q_1 = 1, q_2 = 20$	$q_1 = 10, q_2 = 30$	$q_1 = 5, q_2 = 25$
I	0.0657	0.1203	0.0516	0.4567
II	0.1847	0.2107	0.0921	0.0665
III	0.0765	0.0624	0.1400	0.0690
IV	0.0204	0.0177	0.0221	0.0213
V	0.0378	0.0278	0.0419	0.0576
I-V	0.1847	0.2107	0.1400	0.4567

scale q and therefore splitting the frame sizes into texture and motion bits is warranted. The VD curve's concave shape is the combined result of simpler linear and quadratic functions of the statistical texture and motion constituents.

We modeled the average and variance of the texture bits using a quadratic model based on the rate-distortion model employed for rate control in MPEG-4 [3]. We also modeled the covariance of the texture and motion bits by this quadratic model. We modeled the average and the variance of the motion bits by a linear model. We have thus extended [3] where a quadratic model was used for the rate-distortion function of the entire frame size. Overall, the models result in good predictions of the actual curves and are obtained from only two samples of the statistical parameters for texture and motion bits (average, variance, covariance).

Our sensitivity analysis suggests that a good choice for the two q values is $q_1 = 10$ and $q_2 = 30$. The texture and motion based VD curve model improves upon the piecewise modeling in [1] and results in a good prediction of the actual VD curves. The texture and motion based VD curve model for two quantization scales has a modeling accuracy that is comparable to the piecewise model for three quantization scales.

A future research direction is adapting the VD curve modeling approach derived in this report to the new H.264/AVC [47] video encoder.

REFERENCES

- [1] P. Seeling and M. Reisslein, "The rate variability-distortion (VD) curve of encoded video and its impact on statistical multiplexing," *IEEE Transactions on Broadcasting*, vol. 51, no. 4, pp. 473–492, Dec. 2005.
- [2] ISO/IEC JTC 1/SC 29/WG 11 N2802, "Information technology—generic coding of audio-visual objects—part 2: Visual, final proposed draft amendment 1," Geneva, July 1999.
- [3] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 1, Feb. 1997.
- [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.

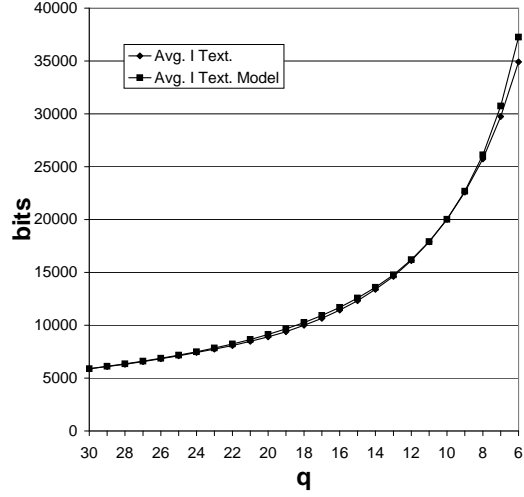
- [6] H.-M. Hang and J.-J. Chen, "Source model for transform video coder and its application —part I: Fundamental theory," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 287–298, Apr. 1997.
- [7] C. Cai, R. Ding, and S. K. Mitra, "New bit-rate control models for 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] M. Rezaei, M. Gabbouj, and S. Wenger, "Analyzed rate distortion model in standard video codecs for rate control," in *Proceedings of IEEE Workshop on Signal Processing Systems Design and Implementation*, Nov. 2005, pp. 550–555.
- [11] J. Sun, W. Gao, D. Zhao, and Q. Huang, "Statistical model, analysis and approximation of rate-distortion function in MPEG-4 FGS videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 4, pp. 535–539, Apr. 2006.
- [12] Z. Zhang, G. Liu, H. Li, and Y. Li, "A novel PDE-based rate-distortion model for rate control," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 11, pp. 1354–1364, Nov. 2005.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] A. Alheraish, S. Alshebeili, and T. Alamri, "A GACS modeling approach for MPEG broadcast video," *IEEE Transactions on Broadcasting*, vol. 50, no. 2, pp. 132–141, June 2004.
- [18] 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.
- [19] 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.
- [20] M. Dai and D. Loguinov, "Analysis and modeling of MPEG-4 and H.264 multi-layer video traffic," in *Proceedings of IEEE Infocom*, Mar. 2005, pp. 2257–2267.
- [21] M. Ghanbari and D. E. Pearson, "Components of bit-rate variation in videoconference signals," *Electronics Letters*, vol. 25, no. 4, pp. 285–286, 1989.
- [22] 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.
- [23] X.-D. Huang, Y.-H. Zhou, and R.-F. Zhang, "A multiscale model for MPEG-4 varied bit rate video traffic," *IEEE Transactions on Broadcasting*, vol. 50, no. 3, pp. 323–334, Sept. 2004.
- [24] 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.
- [25] C. H. Liew, C. K. Kodikara, and A. M. Kondoz, "MPEG-encoded variable bit-rate video traffic modelling," *IEE Proceedings Communications*, vol. 152, no. 5, pp. 749–756, Oct. 2005.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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.

- [35] 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.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] 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.
- [41] 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.
- [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] "American national standard for telecommunications—digital transport of one-way video signals—parameters for objective performance assessment," ANSI T1.801.03-2003, American National Standards Institute.
- [44] "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Phase II," Aug. 2003, available from <http://www.vqeg.org>.
- [45] "Automatic Movie Content Analysis (MoCA), Shot Boundary Detection tool," available from <http://www.informatik.uni-mannheim.de/pi4/projects/MoCA>.
- [46] S. Jeannin and A. Divakaran, "MPEG-7 visual motion descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 720–724, June 2001.
- [47] T. Wiegand, G. Sullivan, G. Bjontegaard, 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.

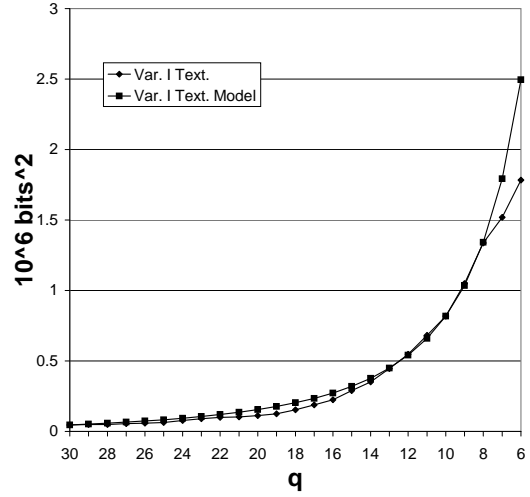
APPENDIX A: STATISTICS MODELING

We present numerous figures depicting statistics and models for a number of scenes from the *Football*, *Starwars IV* and *Terminator* sequences. For each motion class one representative scene is selected.

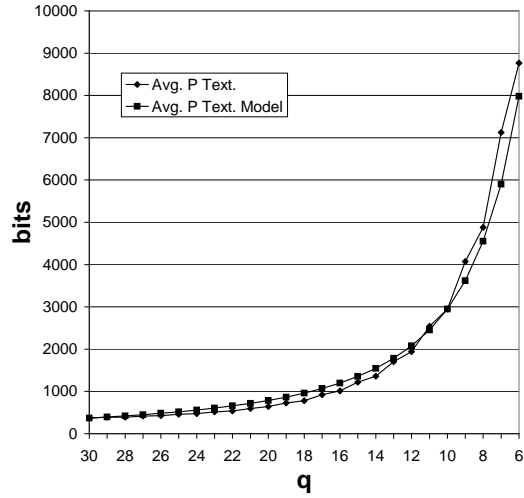
Terminator Scene 384, Motion Class I



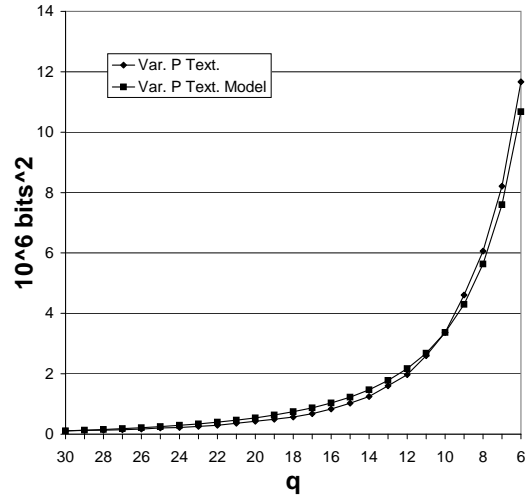
a) I-frames texture averages.



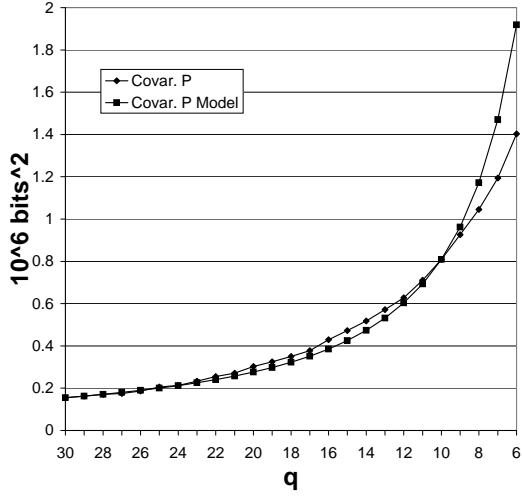
b) I-frames texture variances.



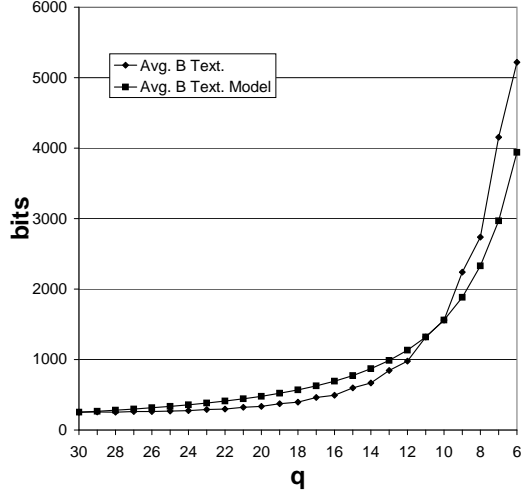
c) P-frames texture averages.



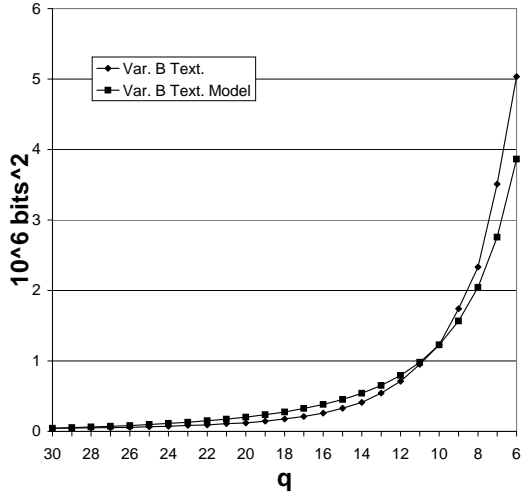
d) P-frames texture variances.



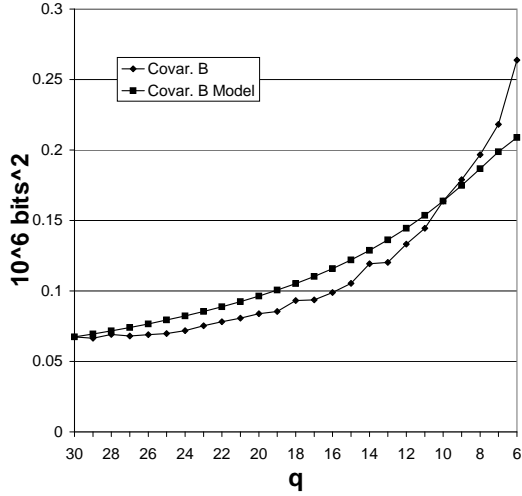
e) P-frames covariances.



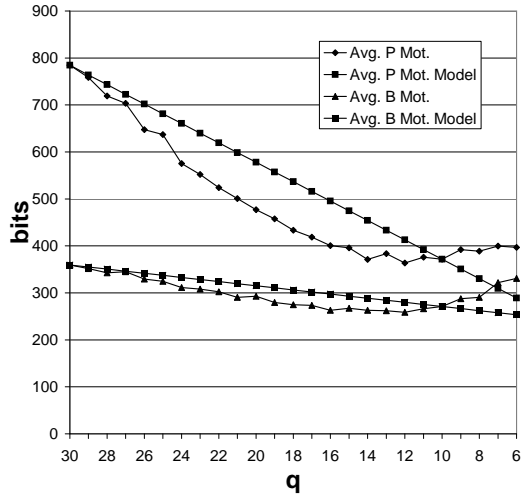
f) B-frames texture averages.



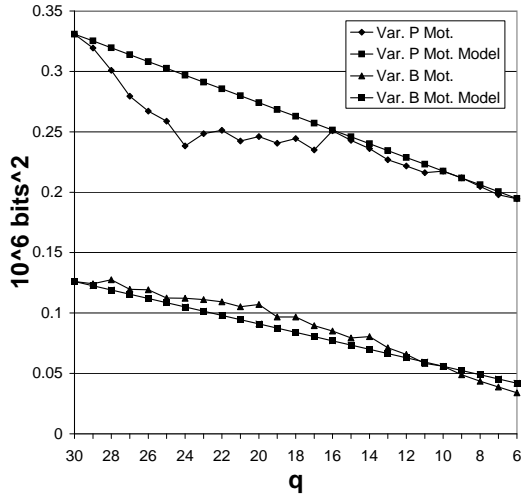
g) B-frames texture variances.



h) B-frames covariances.



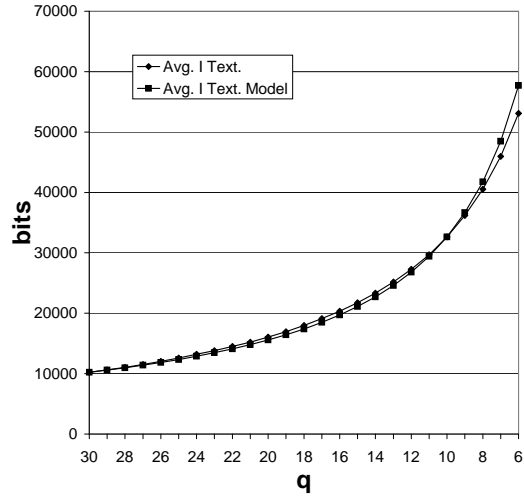
i) P- and B-frames motion averages.



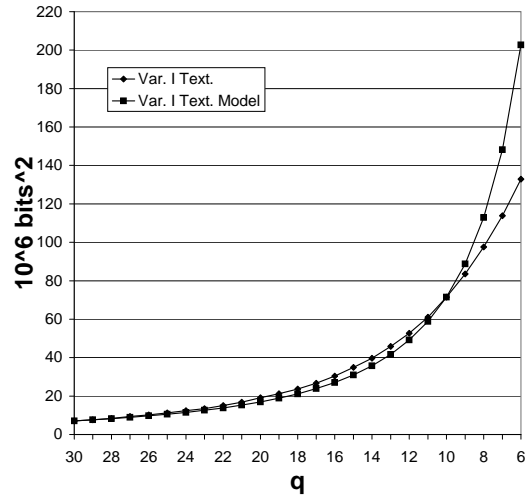
j) P- and B-frames motion variances.

Fig. 20. Statistics and models for MPEG-4 coded Terminator scene 384, motion class I.

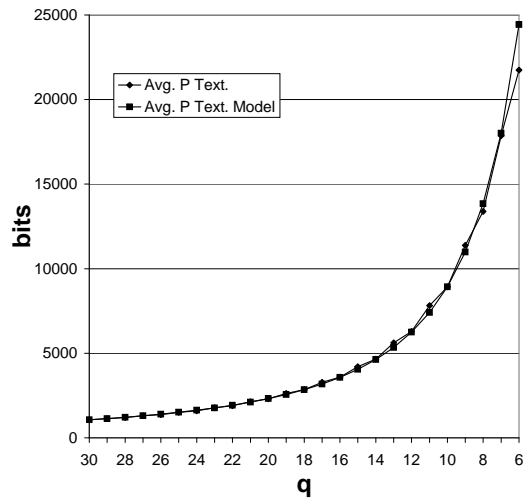
Football Scene 299, Motion Class II



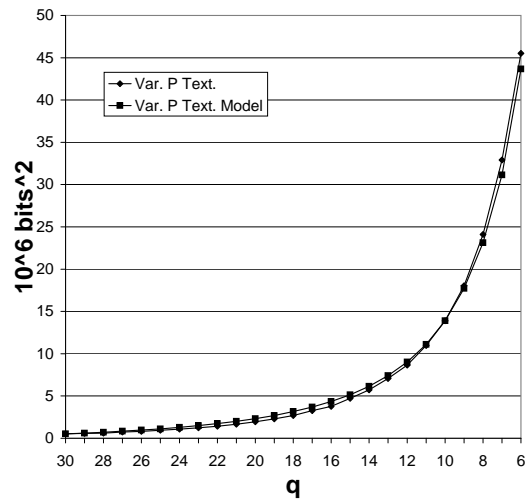
a) I-frames texture averages.



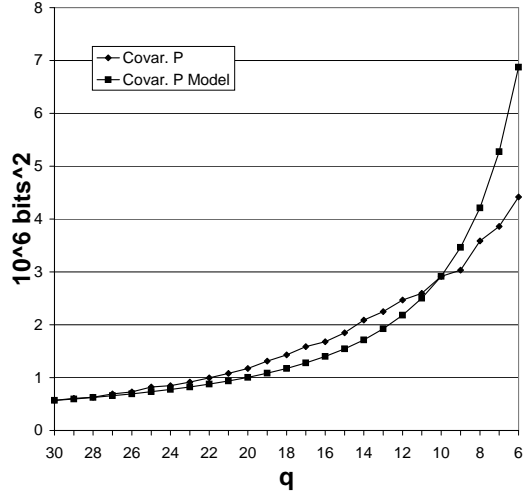
b) I-frames texture variances.



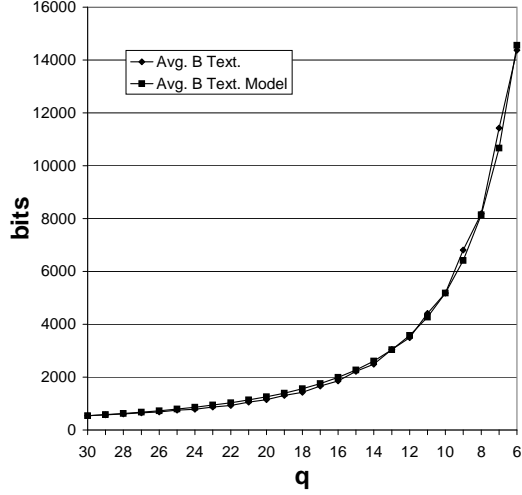
c) P-frames texture averages.



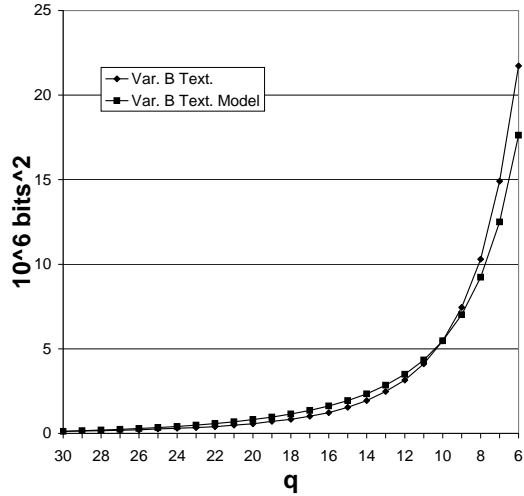
d) P-frames texture variances.



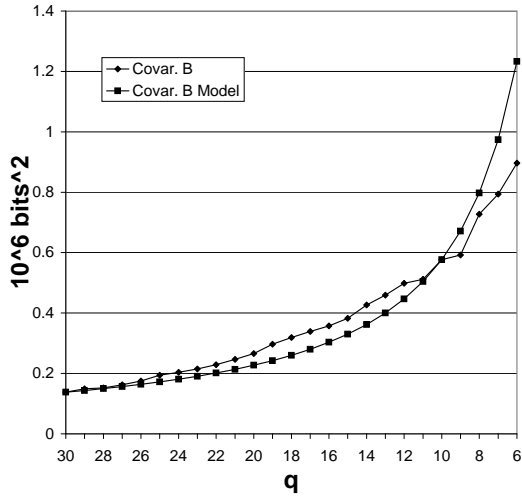
e) P-frames covariances.



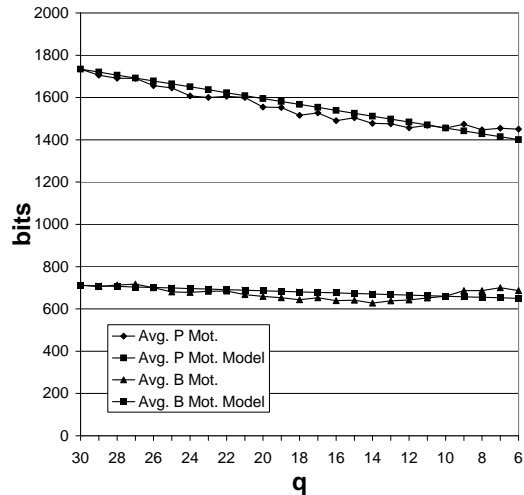
f) B-frames texture averages.



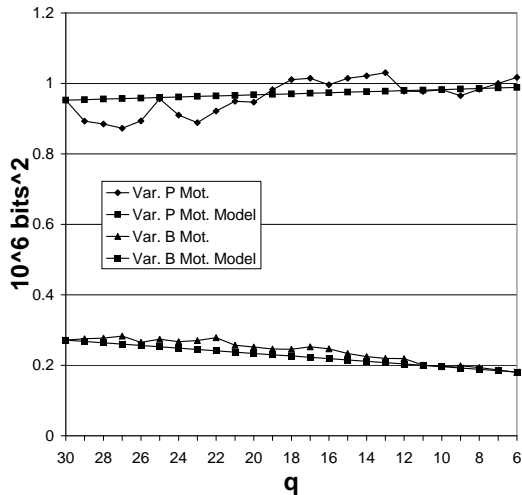
g) B-frames texture variances.



h) B-frames covariances.



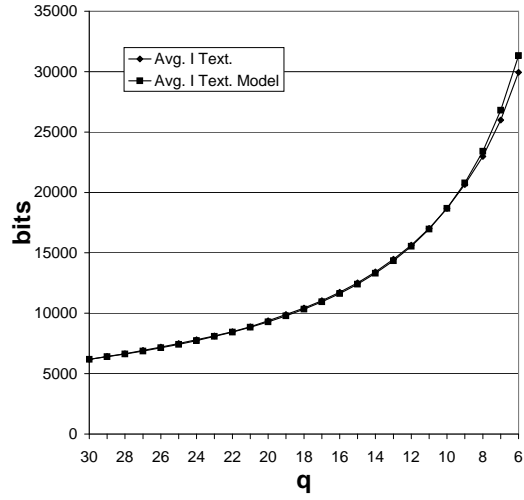
i) P- and B-frames motion averages.



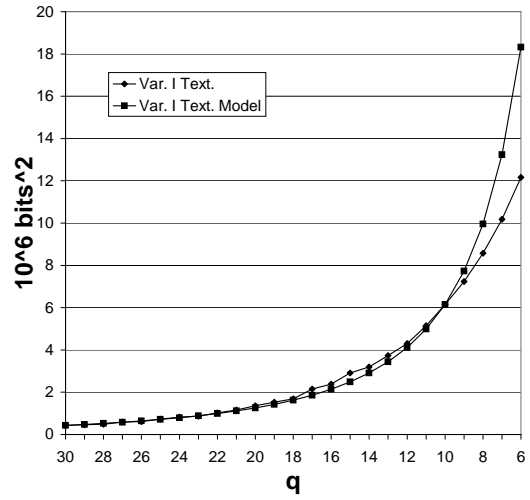
j) P- and B-frames motion variances.

Fig. 21. Statistics and models for MPEG-4 coded *Football* scene 299, motion class II.

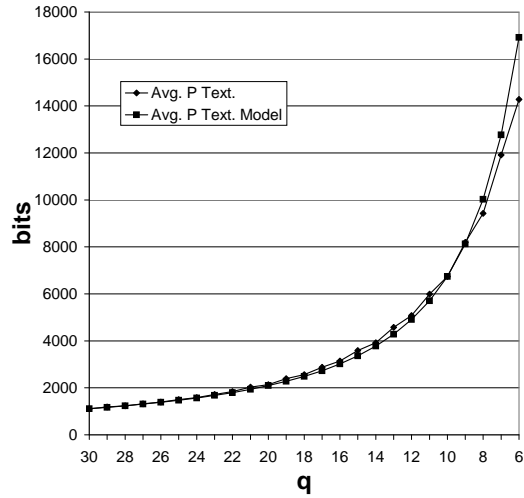
Terminator Scene 628, Motion Class III



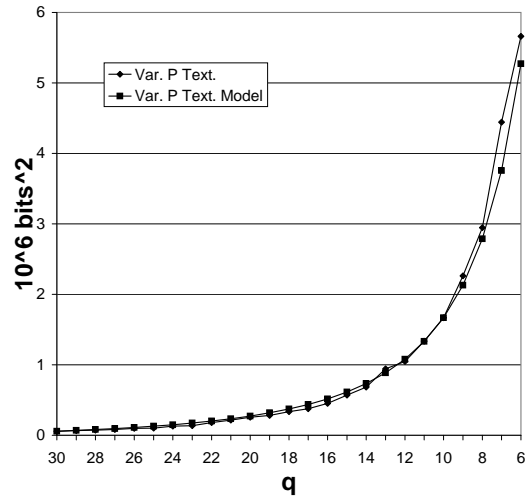
a) I-frames texture averages.



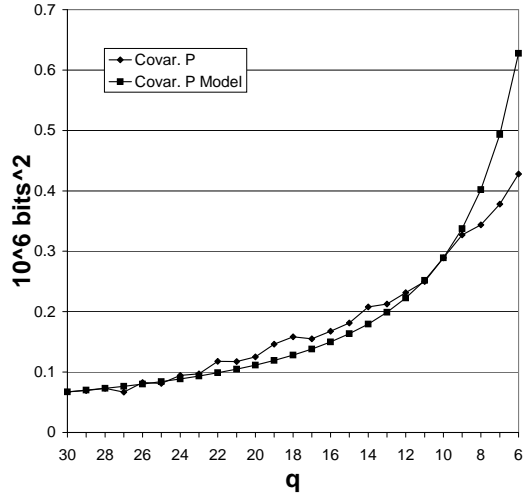
b) I-frames texture variances.



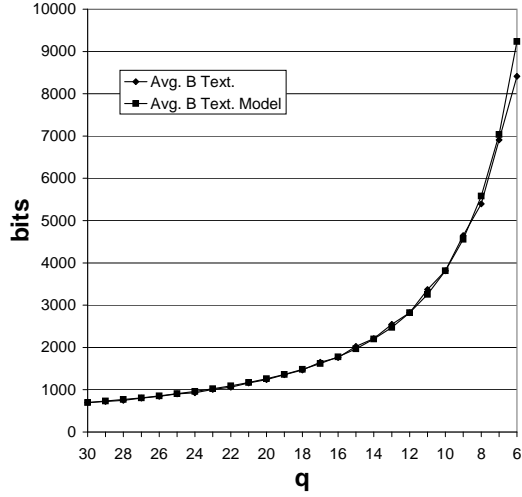
c) P-frames texture averages.



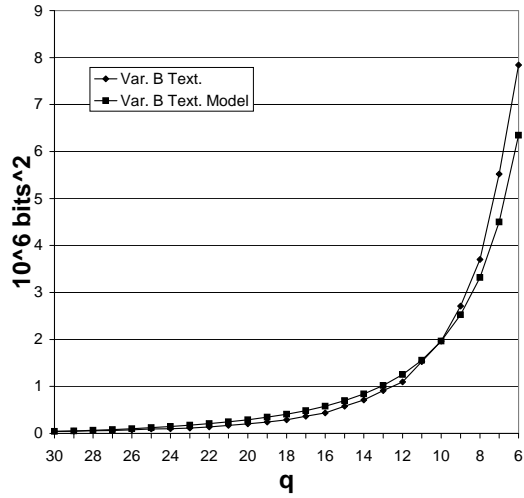
d) P-frames texture variances.



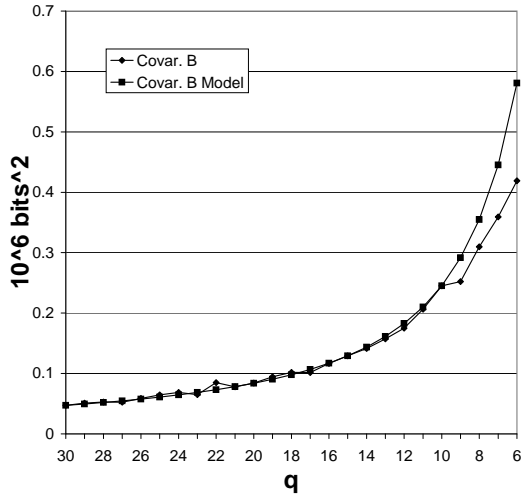
e) P-frames covariances.



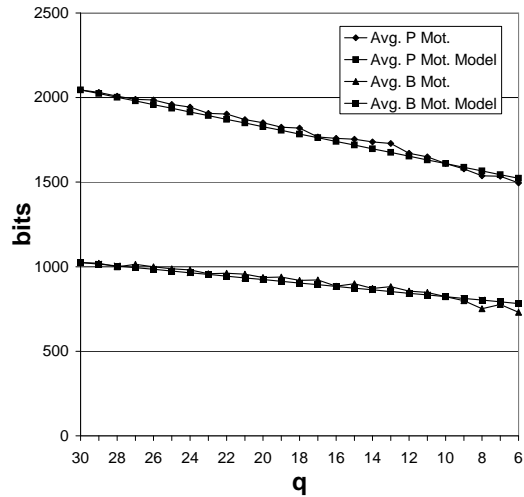
f) B-frames texture averages.



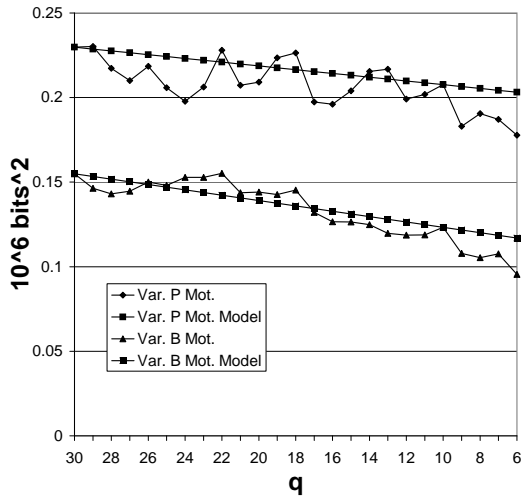
g) B-frames texture variances.



h) B-frames covariances.



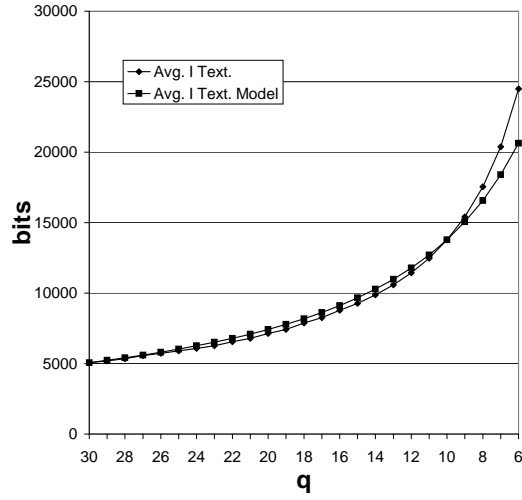
i) P- and B-frames motion averages.



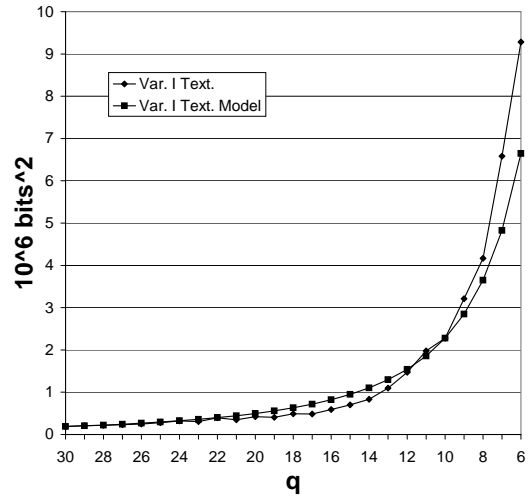
j) P- and B-frames motion variances.

Fig. 22. Statistics and models for MPEG-4 coded Terminator scene 628, motion class III.

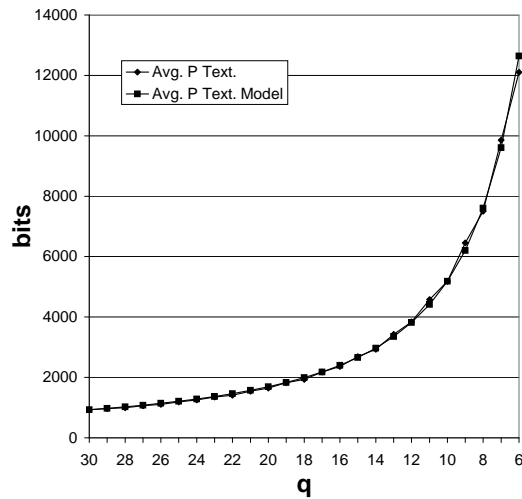
Star Wars IV Scene 165, Motion Class IV



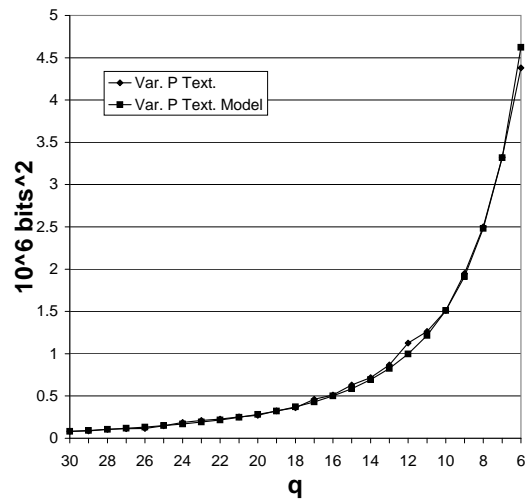
a) I-frames texture averages.



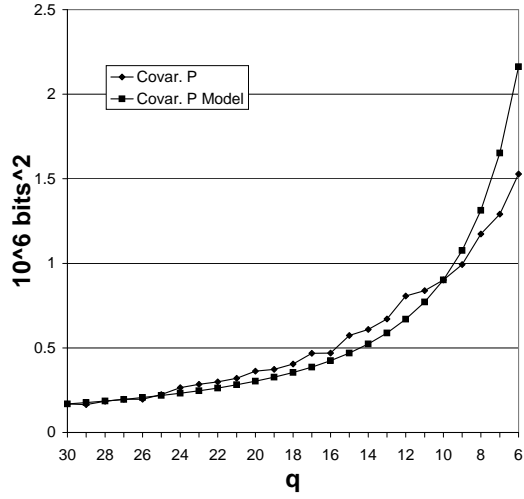
b) I-frames texture variances.



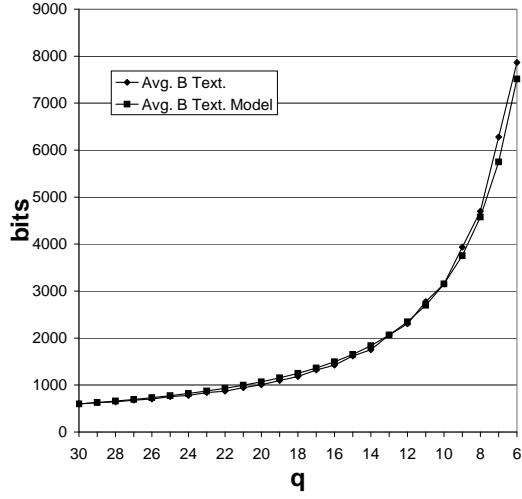
c) P-frames texture averages.



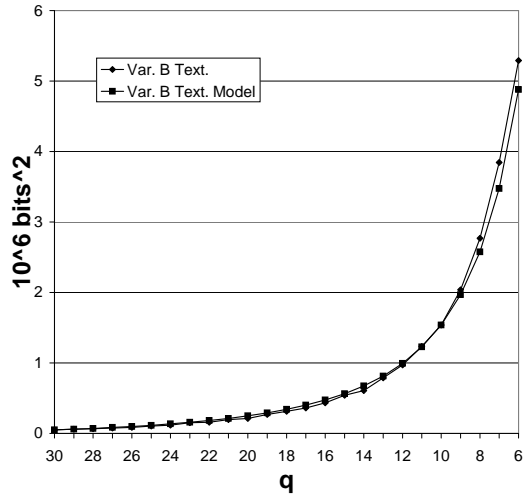
d) P-frames texture variances.



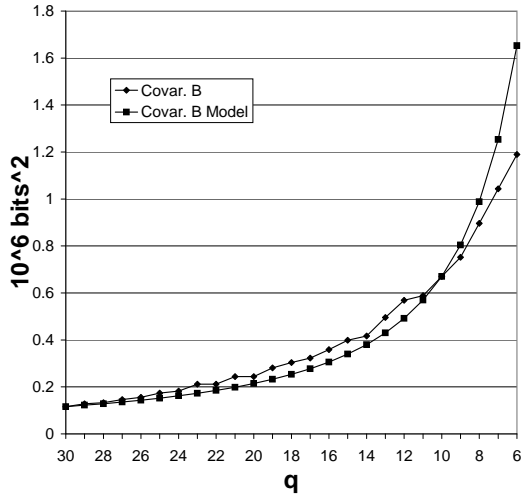
e) P-frames covariances.



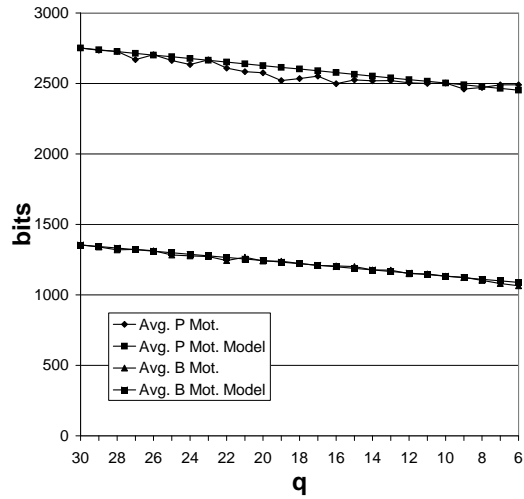
f) B-frames texture averages.



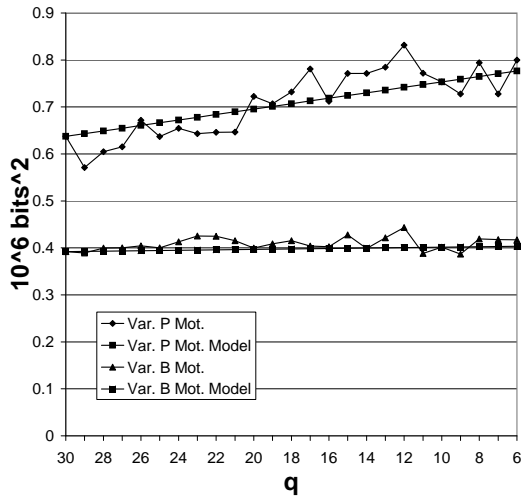
g) B-frames texture variances.



h) B-frames covariances.



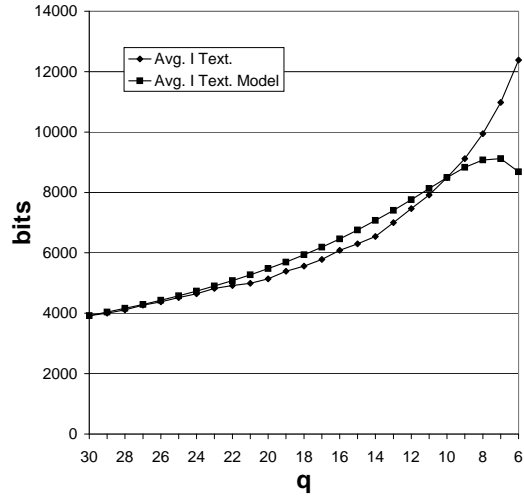
i) P- and B-frames motion averages.



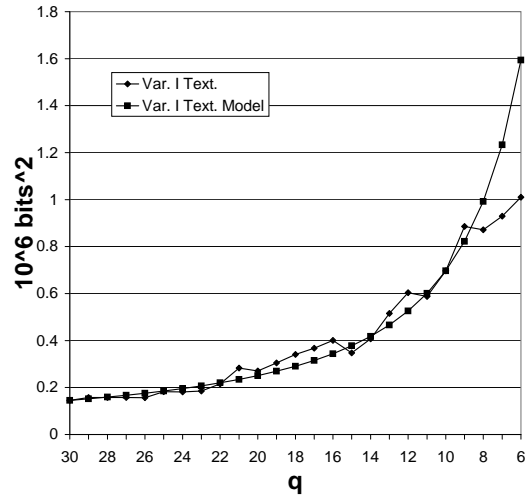
j) P- and B-frames motion variances.

Fig. 23. Statistics and models for MPEG-4 coded *Star Wars IV* scene 165, motion class IV.

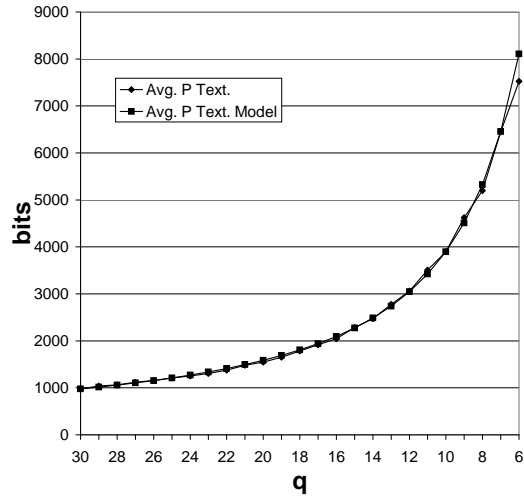
Star Wars IV Scene 632, Motion Class V



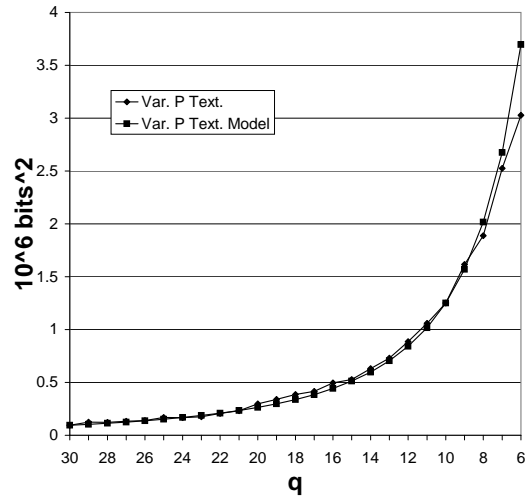
a) I-frames texture averages.



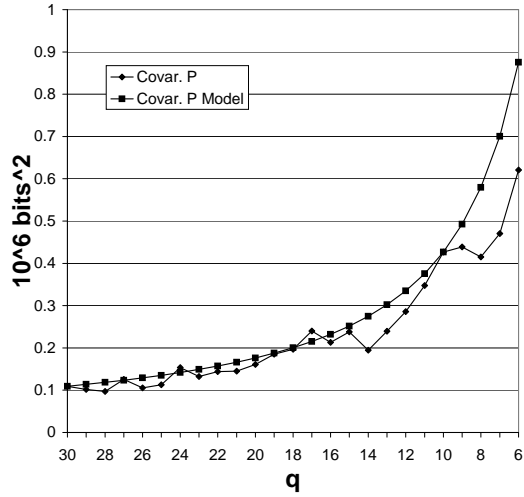
b) I-frames texture variances.



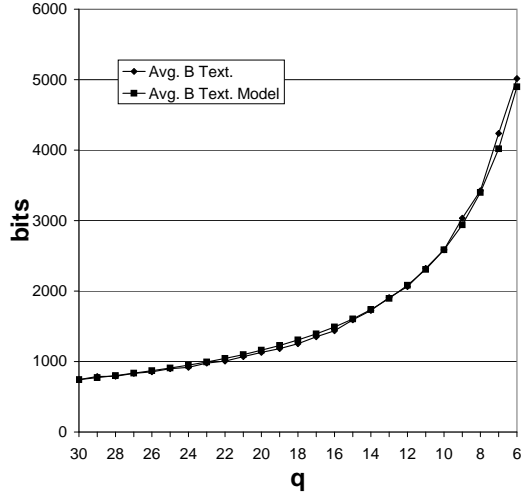
c) P-frames texture averages.



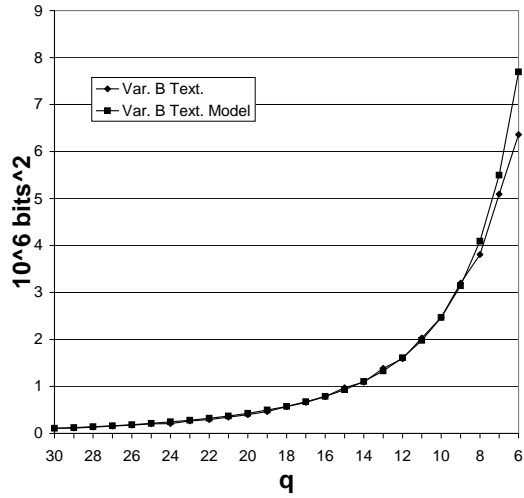
d) P-frames texture variances.



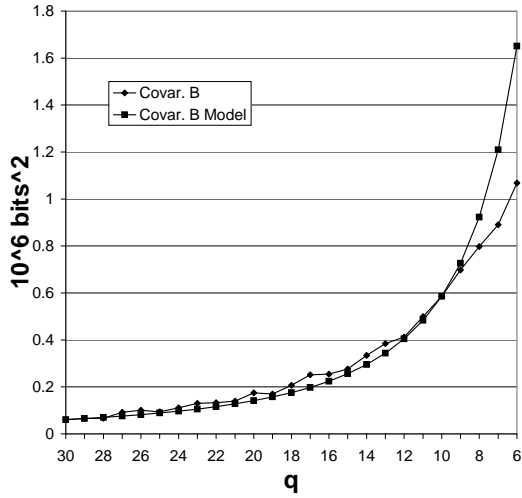
e) P-frames covariances.



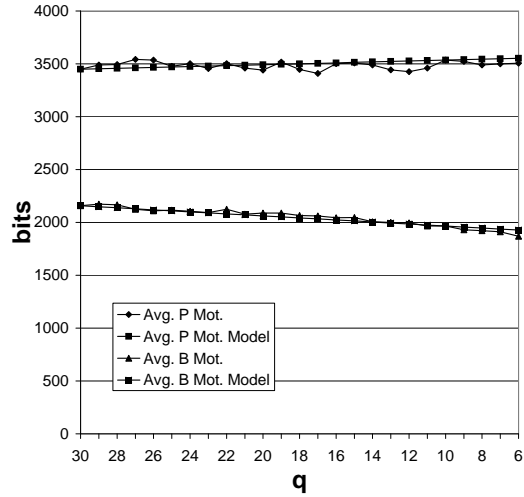
f) B-frames texture averages.



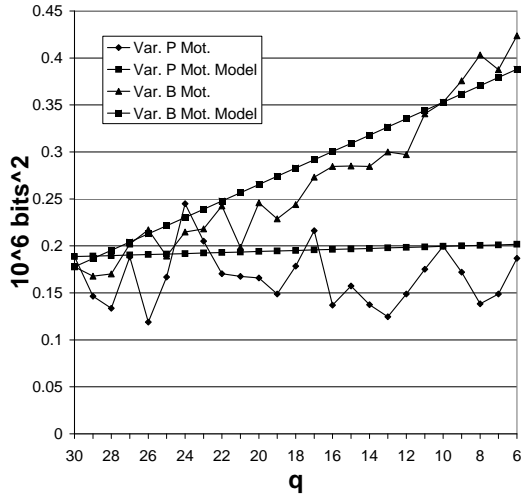
g) B-frames texture variances.



h) B-frames covariances.



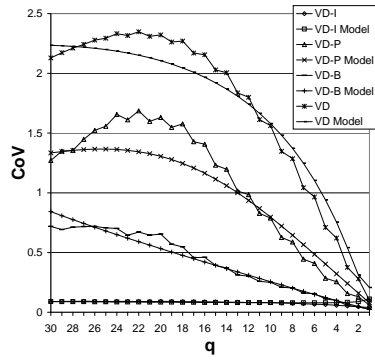
i) P- and B-frames motion averages.



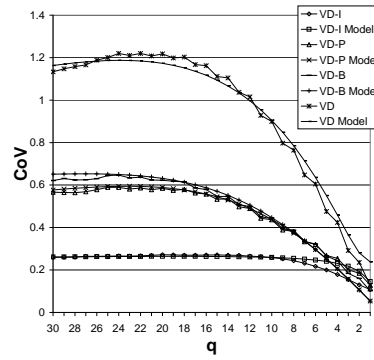
j) P- and B-frames motion variances.

Fig. 24. Statistics and models for MPEG-4 coded *Star Wars IV* scene 632, motion class V.

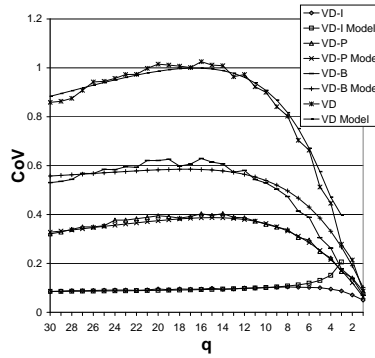
APPENDIX B: VD CURVE MODELS



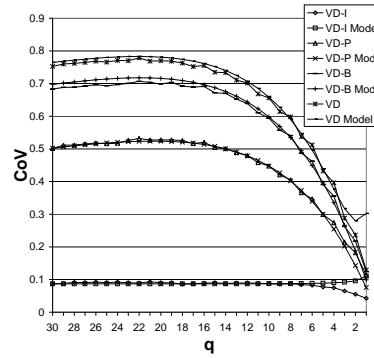
a) Scene 298, motion class I.



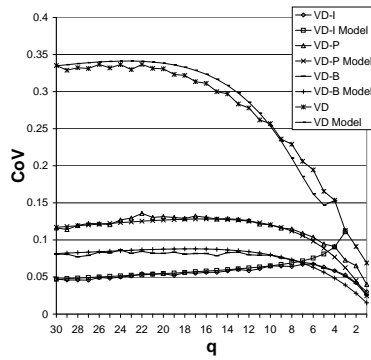
b) Scene 299, motion class II.



c) Scene 557, motion class III.

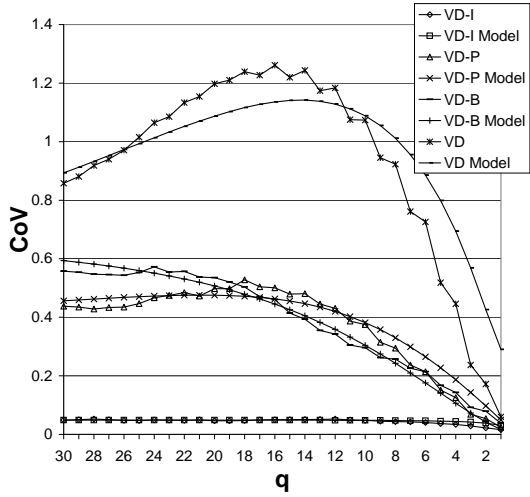


d) Scene 184, motion class IV.

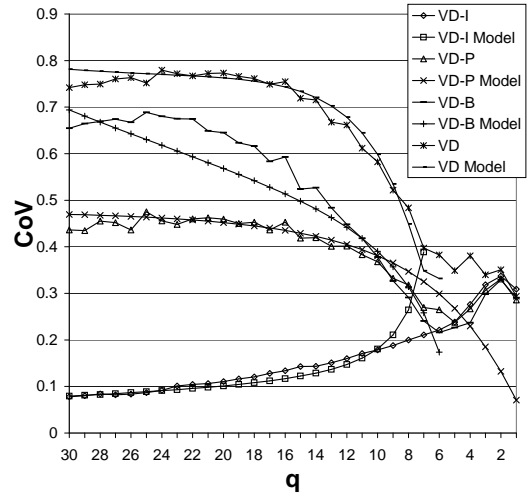


e) Scene 336, motion class V.

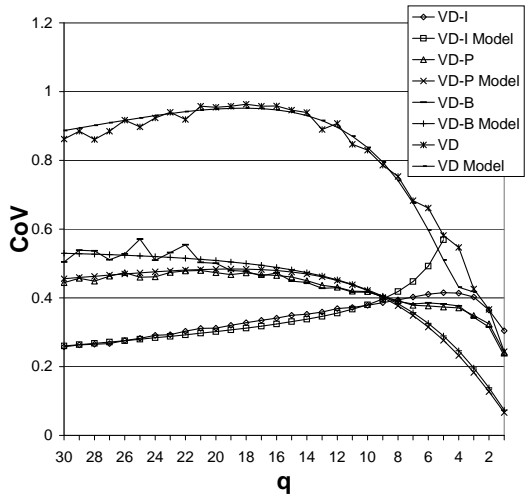
Fig. 25. VD curves and models for MPEG-4 coded *Football* scenes.



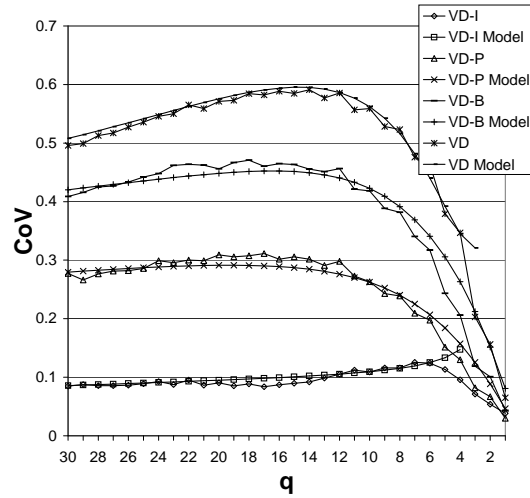
a) Scene 274, motion class I.



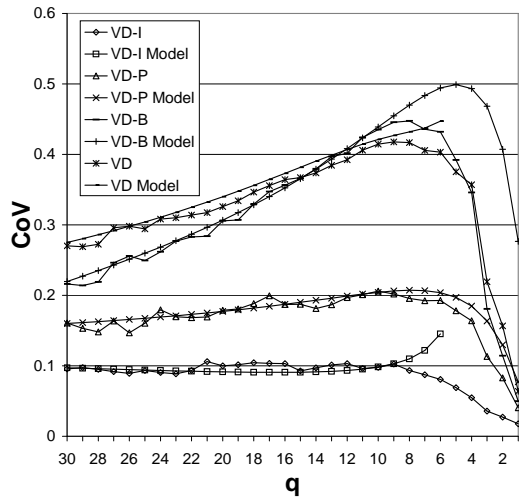
b) Scene 117, motion class II.



c) Scene 115, motion class III.

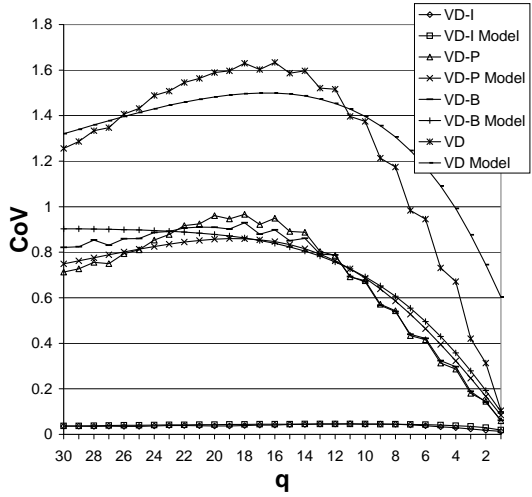


d) Scene 165, motion class IV.

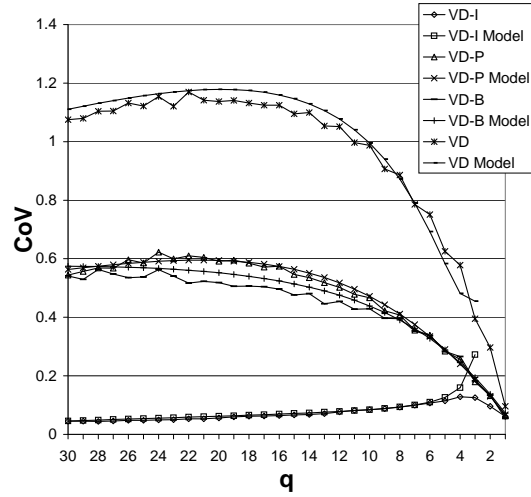


e) Scene 632, motion class V.

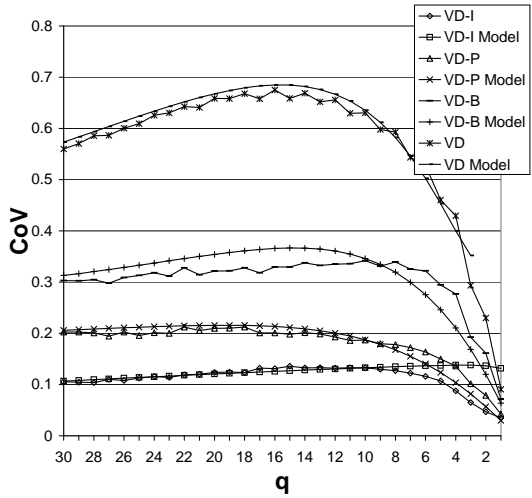
Fig. 26. VD curves and models for MPEG-4 coded *Star Wars IV* scenes.



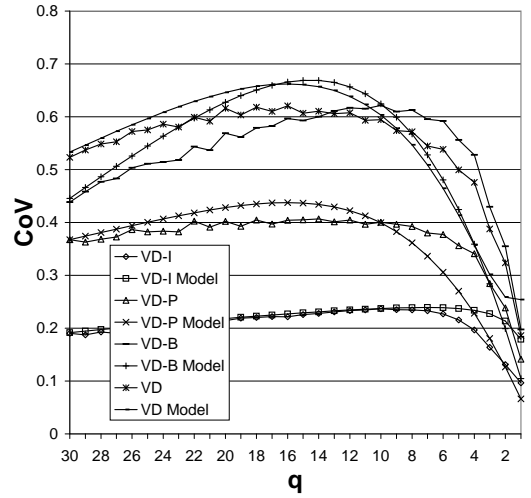
a) Scene 384, motion class I.



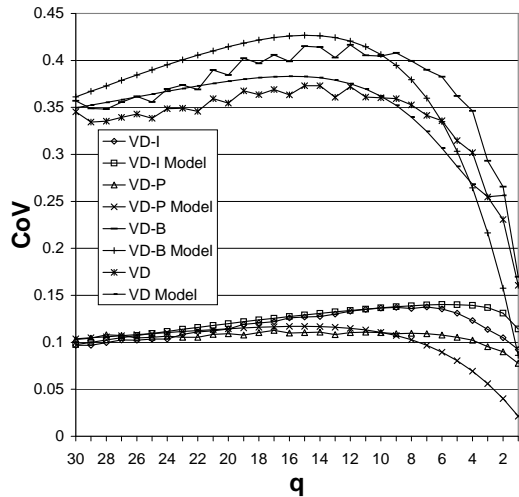
b) Scene 462, motion class II.



c) Scene 628, motion class III.



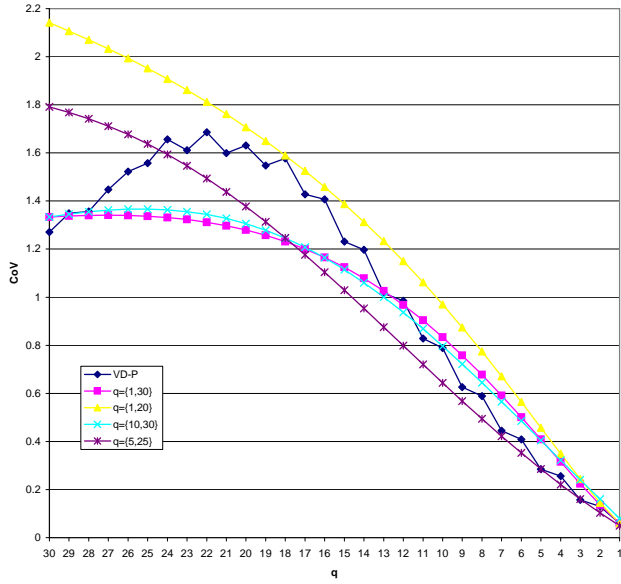
d) Scene 262, motion class IV.



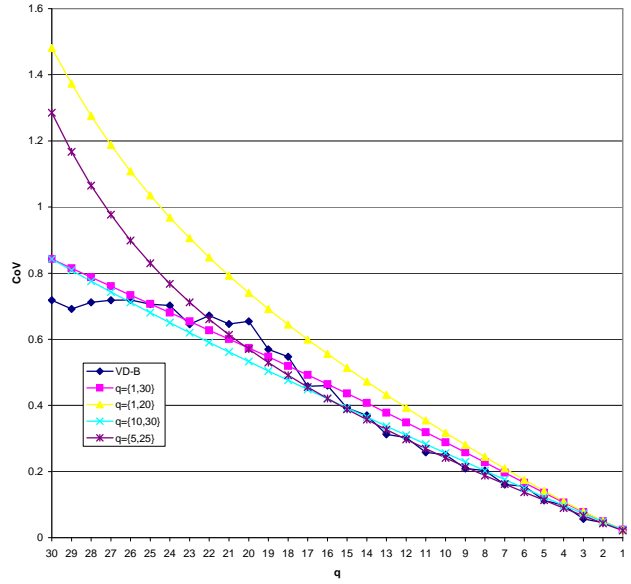
e) Scene 441, motion class V.

Fig. 27. VD curves and models for MPEG-4 coded Terminator scenes.

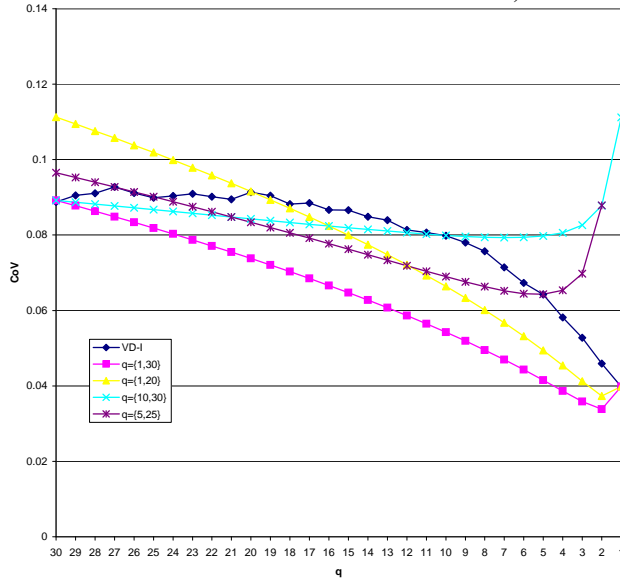
APPENDIX C: SENSITIVITY ANALYSIS



a) VD-P curve and models.

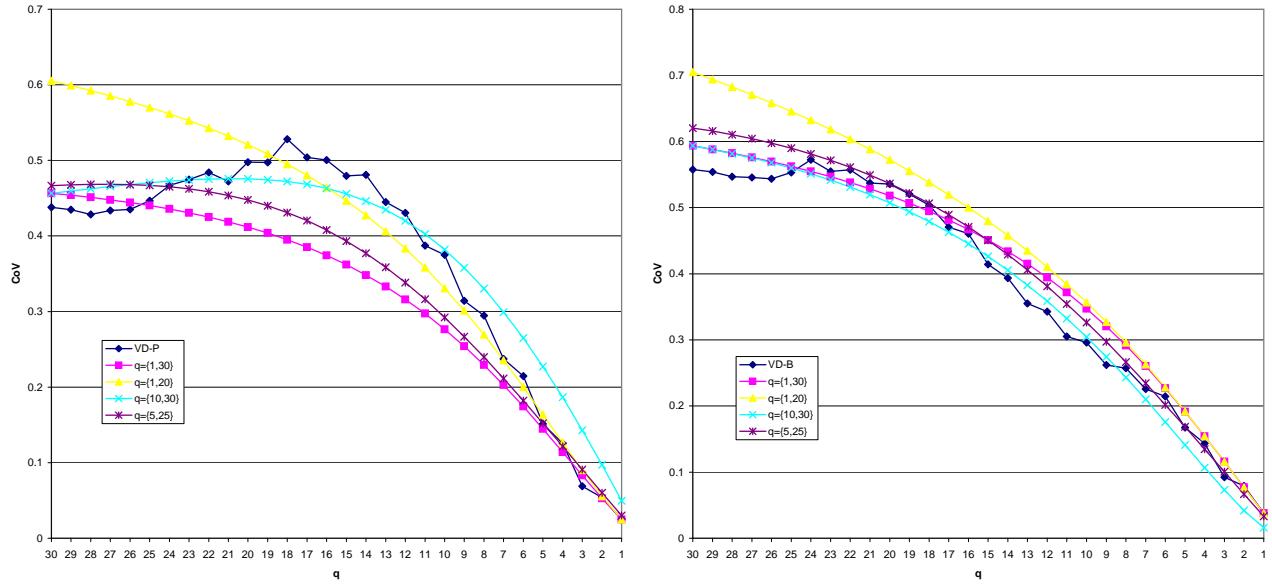


b) VD-B curve and models.



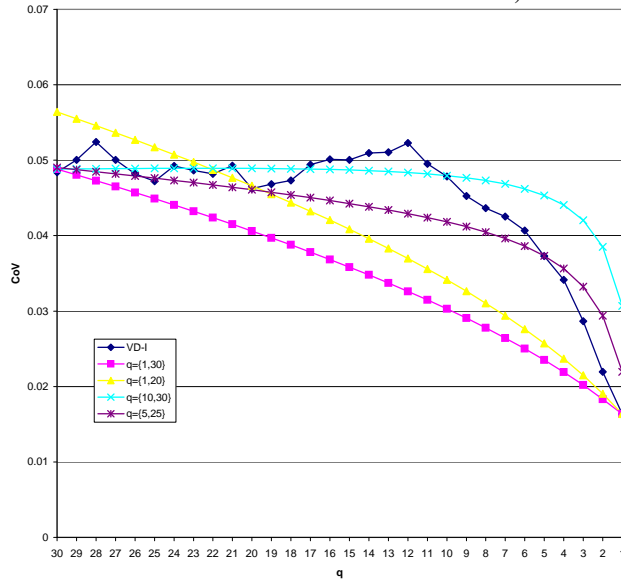
c) VD-I curve and models.

Fig. 28. VD curves and models for MPEG-4 coded *Football* scene 298, belonging to motion class I.



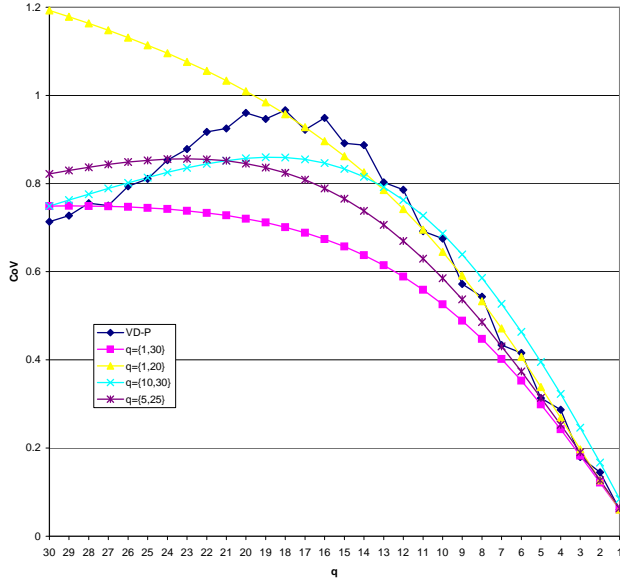
a) VD-P curve and models.

b) VD-B curve and models.

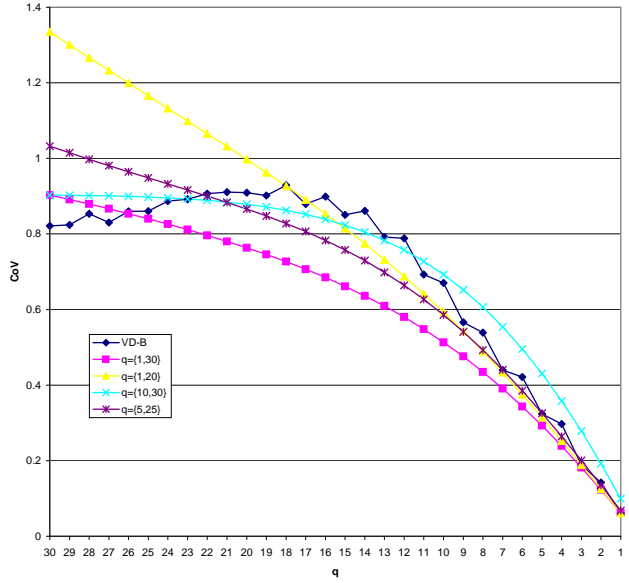


c) VD-I curve and models.

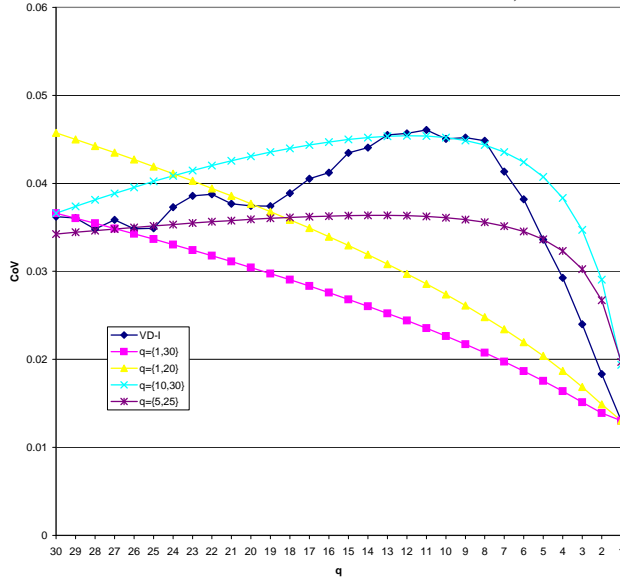
Fig. 29. VD curves and models for MPEG-4 coded *Star Wars IV* scene 274, belonging to motion class I.



a) VD-P curve and models.

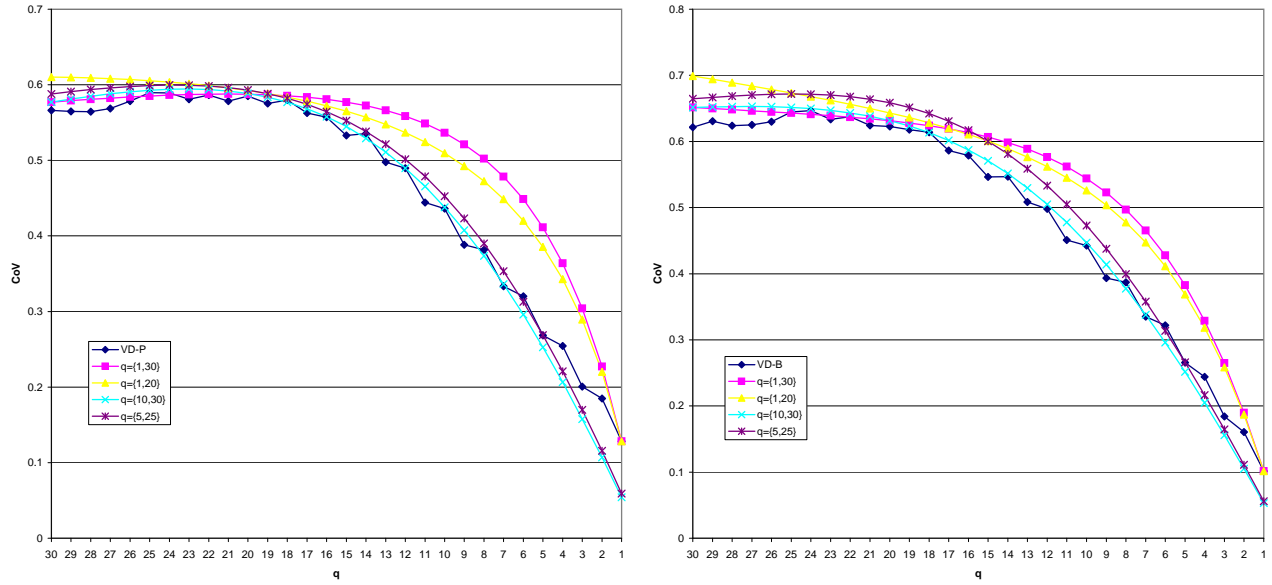


b) VD-B curve and models.



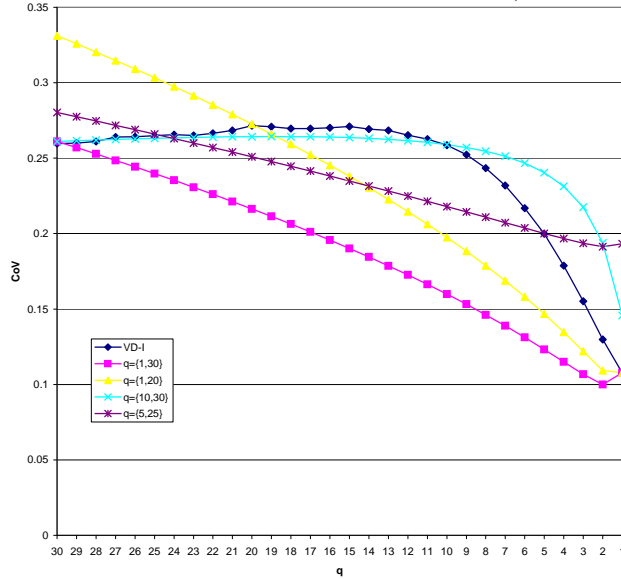
c) VD-I curve and models.

Fig. 30. VD curves and models for MPEG-4 coded Terminator scene 384, belonging to motion class I.



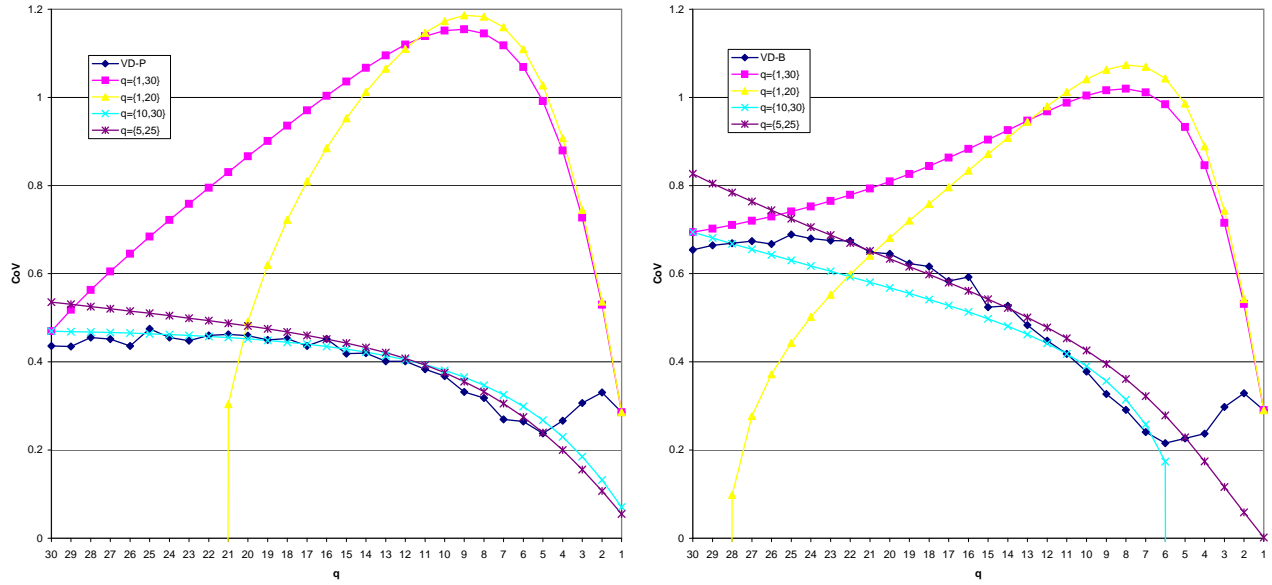
a) VD-P curve and models.

b) VD-B curve and models.



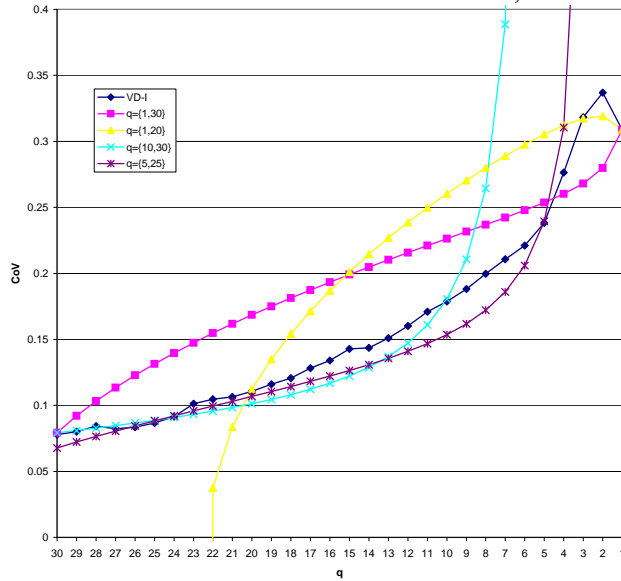
c) VD-I curve and models.

Fig. 31. VD curves and models for MPEG-4 coded *Football* scene 299, belonging to motion class II.



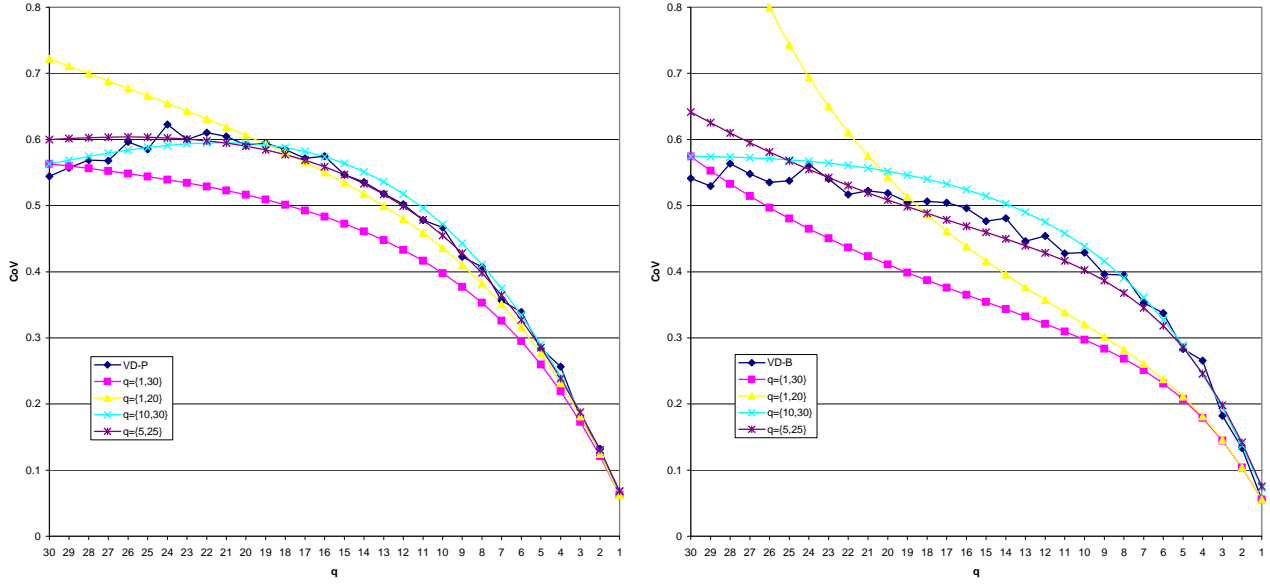
a) VD-P curve and models.

b) VD-B curve and models.



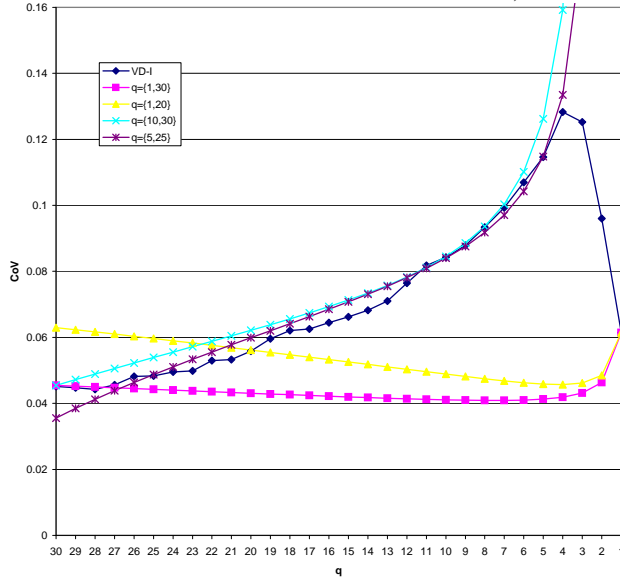
c) VD-I curve and models.

Fig. 32. VD curves and models for MPEG-4 coded *Star Wars IV* scene 117, belonging to motion class II.



a) VD-P curve and models.

b) VD-B curve and models.



c) VD-I curve and models.

Fig. 33. VD curves and models for MPEG-4 coded Terminator scene 462, belonging to motion class II.

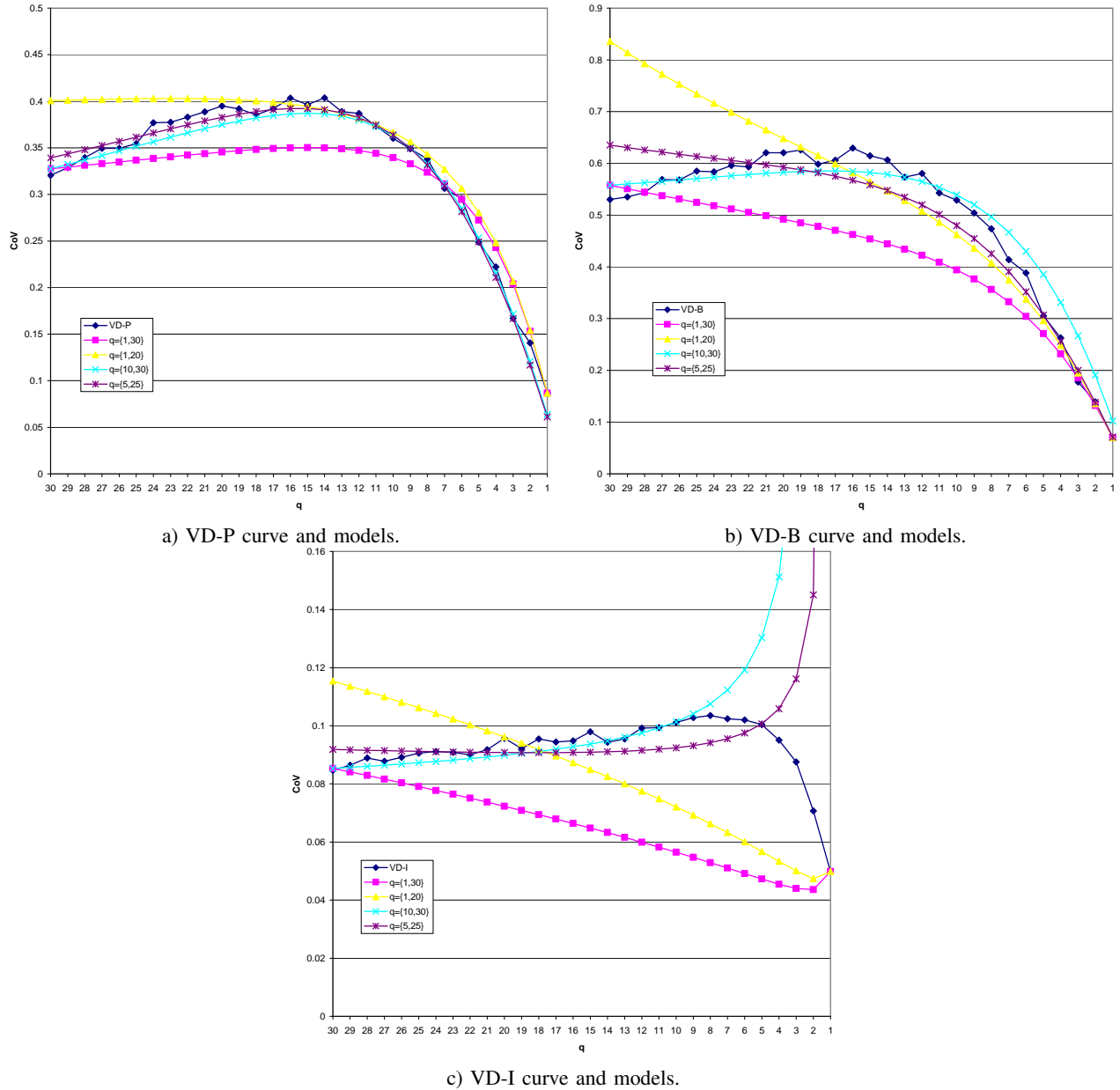
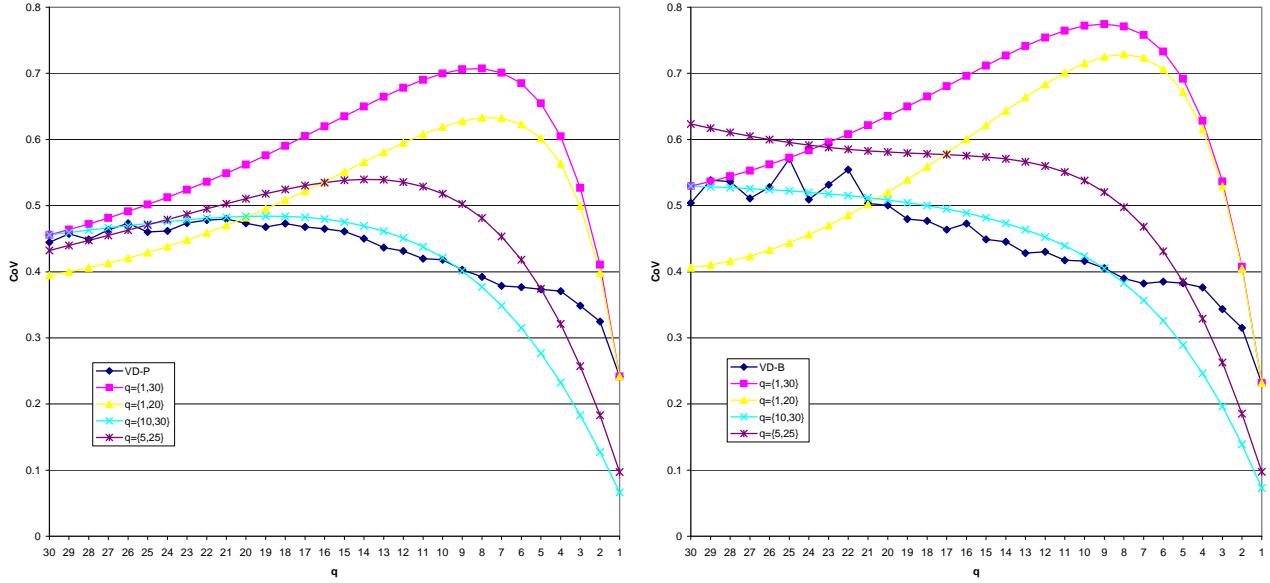
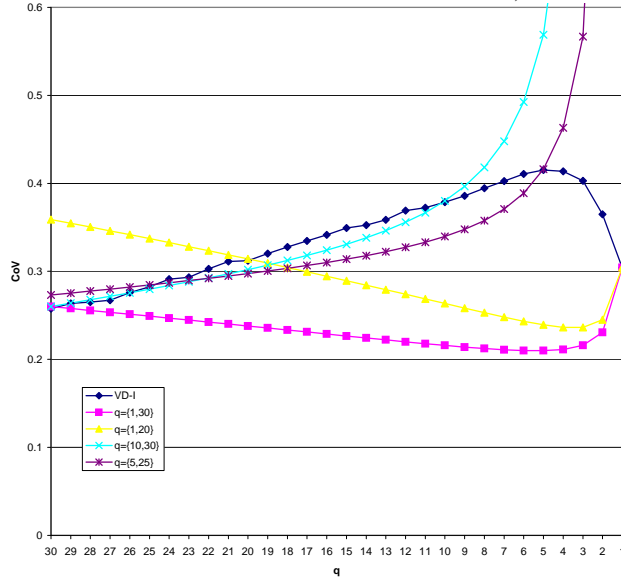


Fig. 34. VD curves and models for MPEG-4 coded *Football* scene 557, belonging to motion class III.



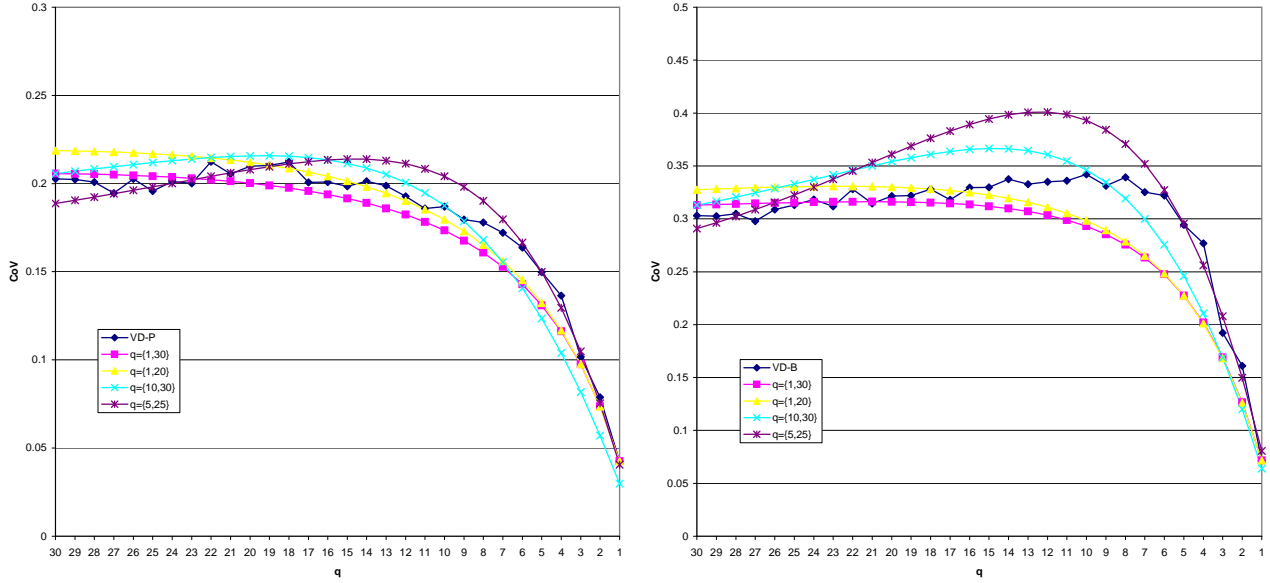
a) VD-P curve and models.

b) VD-B curve and models.



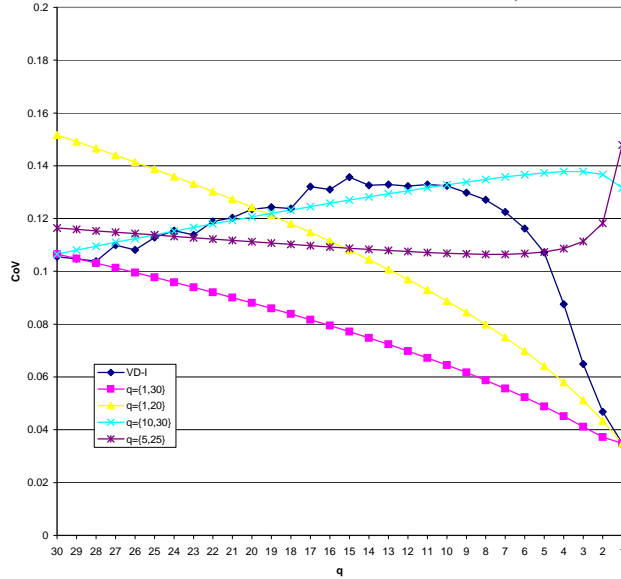
c) VD-I curve and models.

Fig. 35. VD curves and models for MPEG-4 coded *Star Wars IV* scene 115, belonging to motion class III.



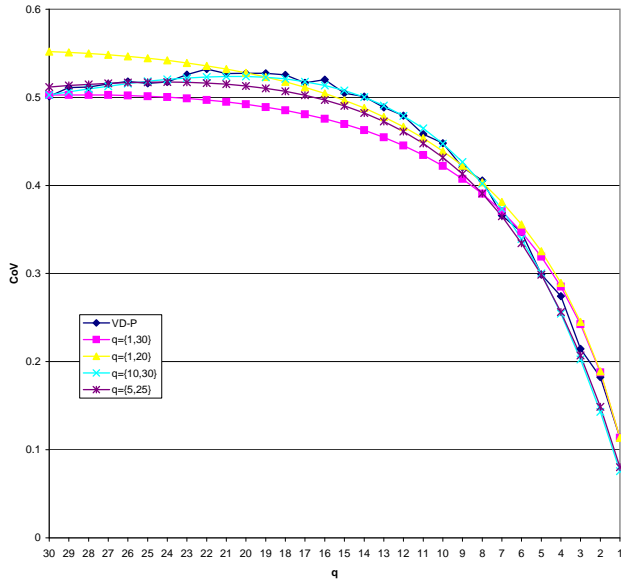
a) VD-P curve and models.

b) VD-B curve and models.

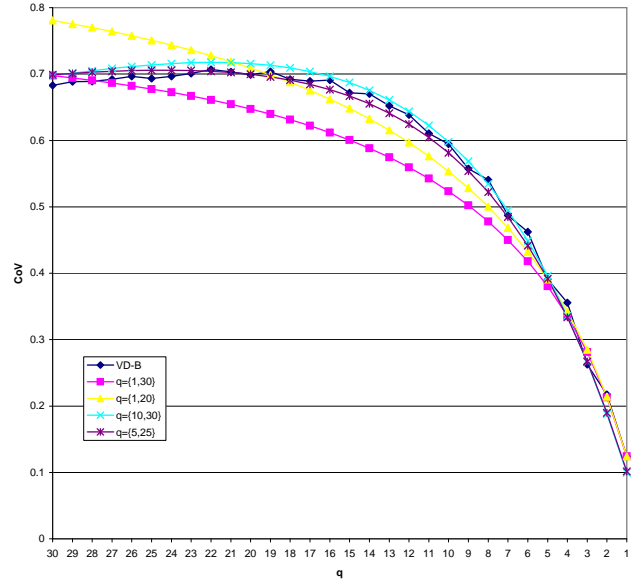


c) VD-I curve and models.

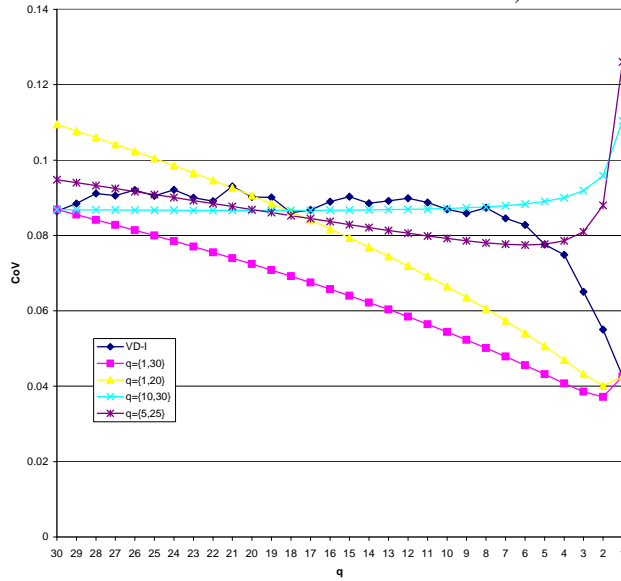
Fig. 36. VD curves and models for MPEG-4 coded Terminator scene 628, belonging to motion class III.



a) VD-P curve and models.

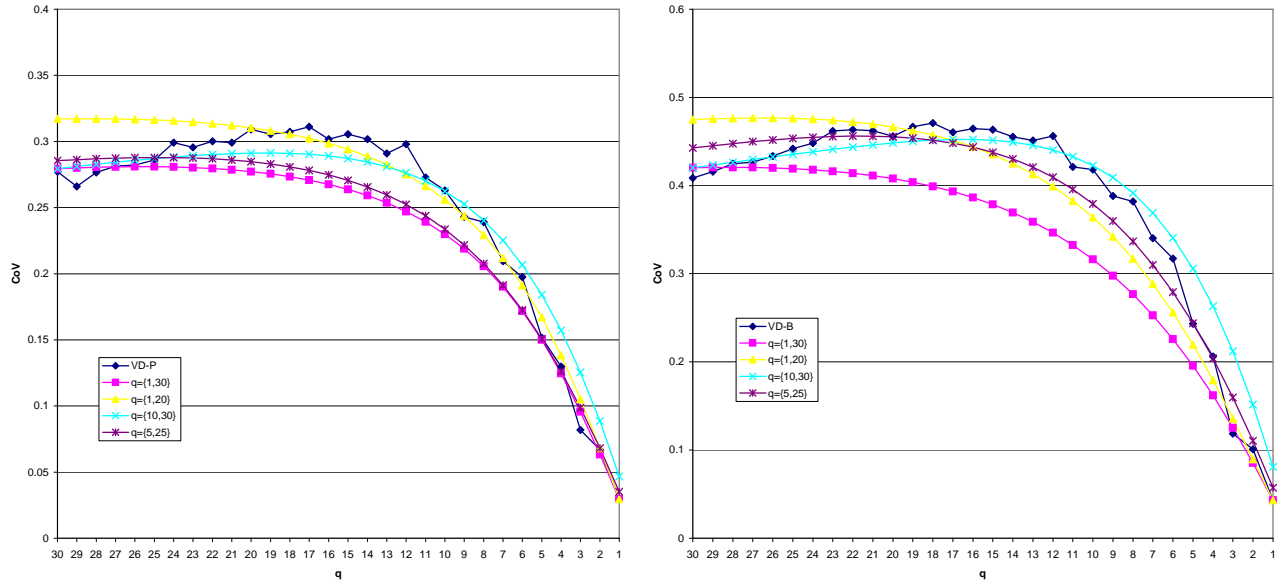


b) VD-B curve and models.



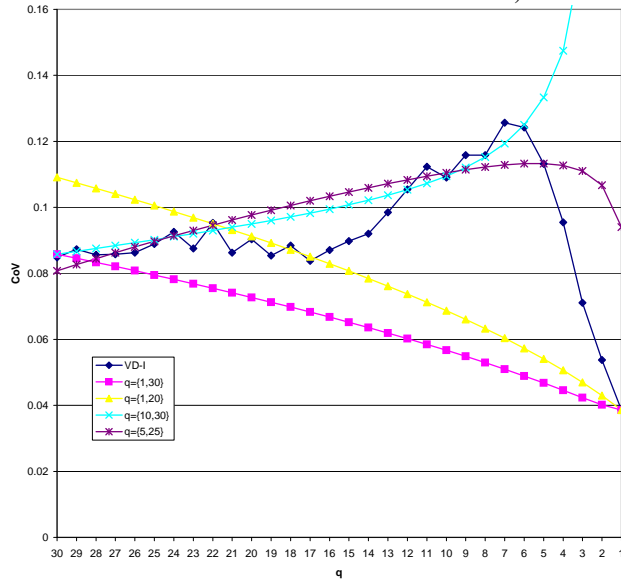
c) VD-I curve and models.

Fig. 37. VD curves and models for MPEG-4 coded *Football* scene 184, belonging to motion class IV.



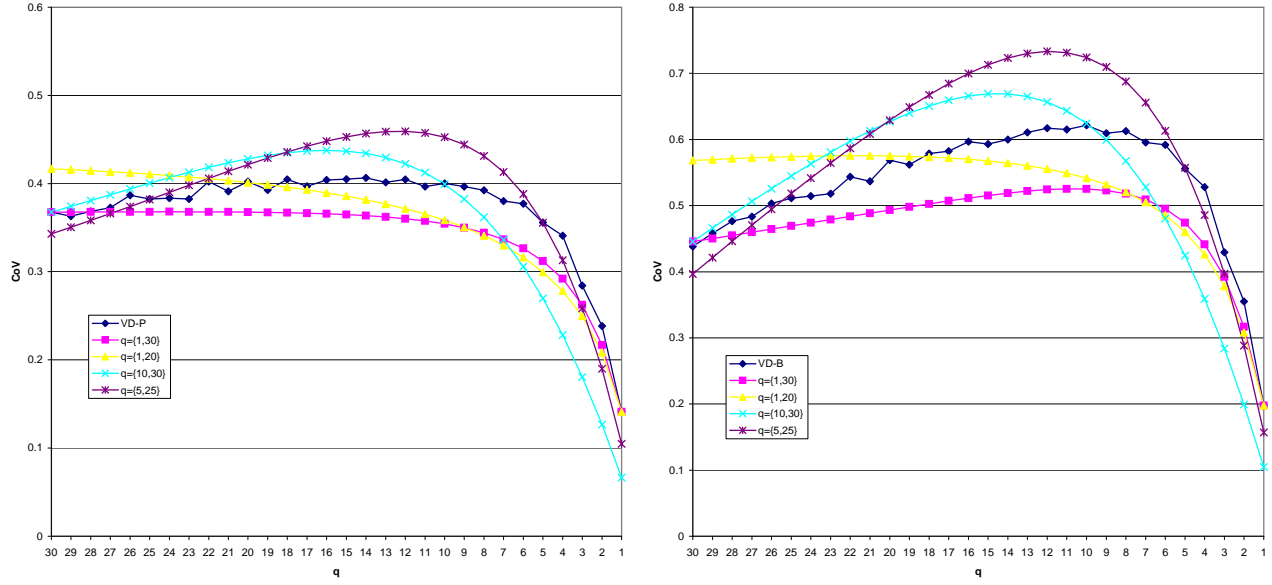
a) VD-P curve and models.

b) VD-B curve and models.



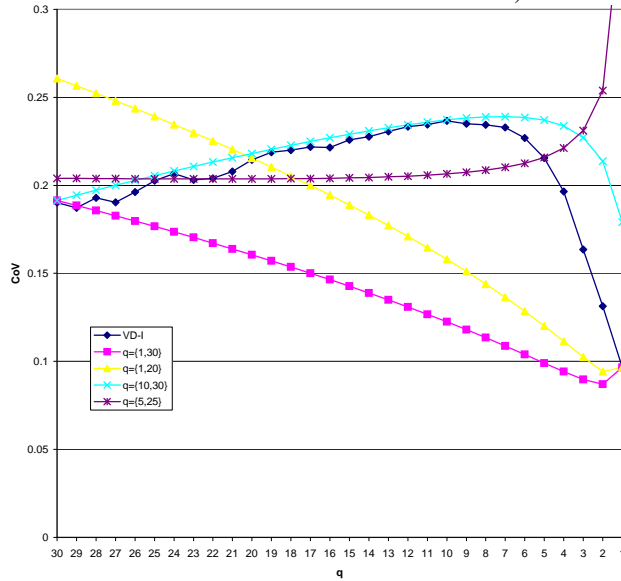
c) VD-I curve and models.

Fig. 38. VD curves and models for MPEG-4 coded *Star Wars IV* scene 165, belonging to motion class IV.



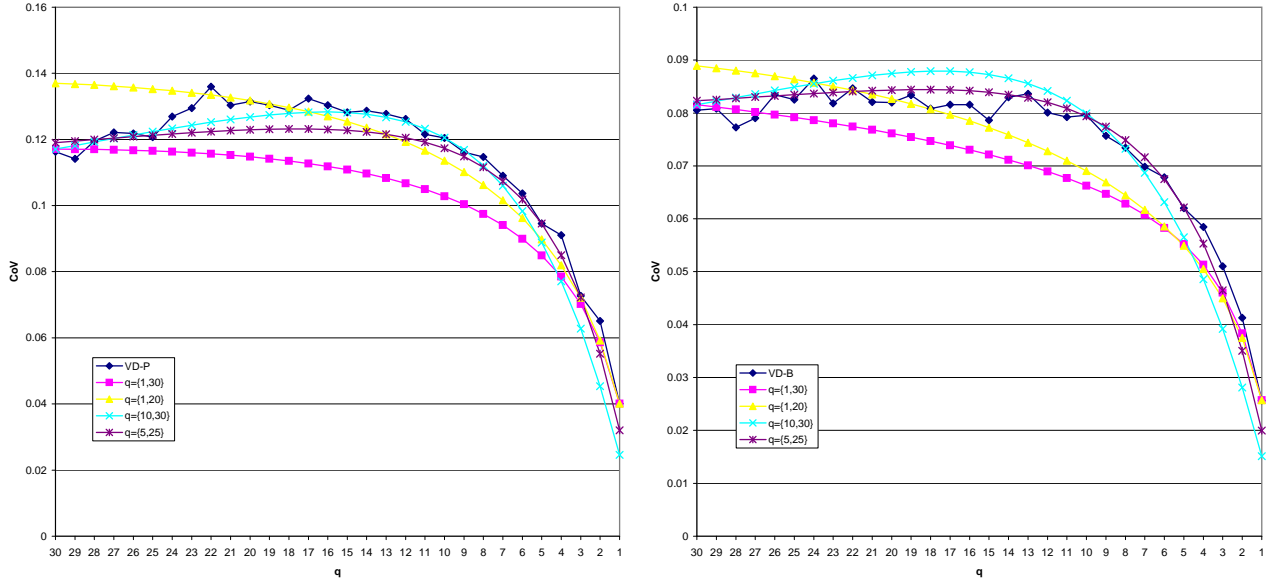
a) VD-P curve and models.

b) VD-B curve and models.



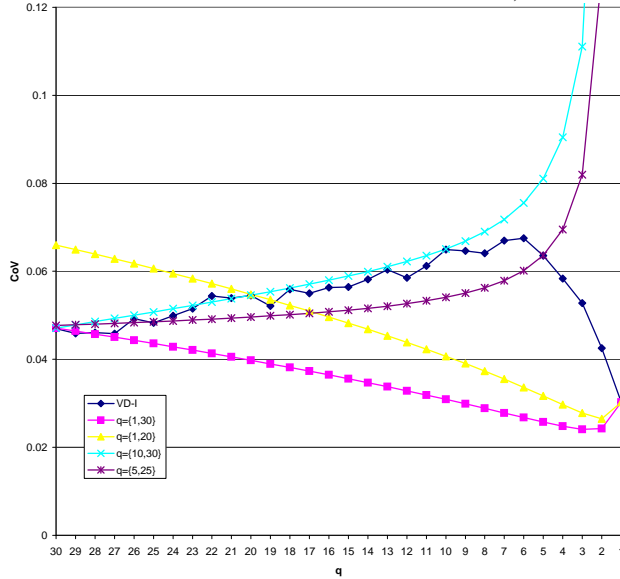
c) VD-I curve and models.

Fig. 39. VD curves and models for MPEG-4 coded Terminator scene 262, belonging to motion class IV.



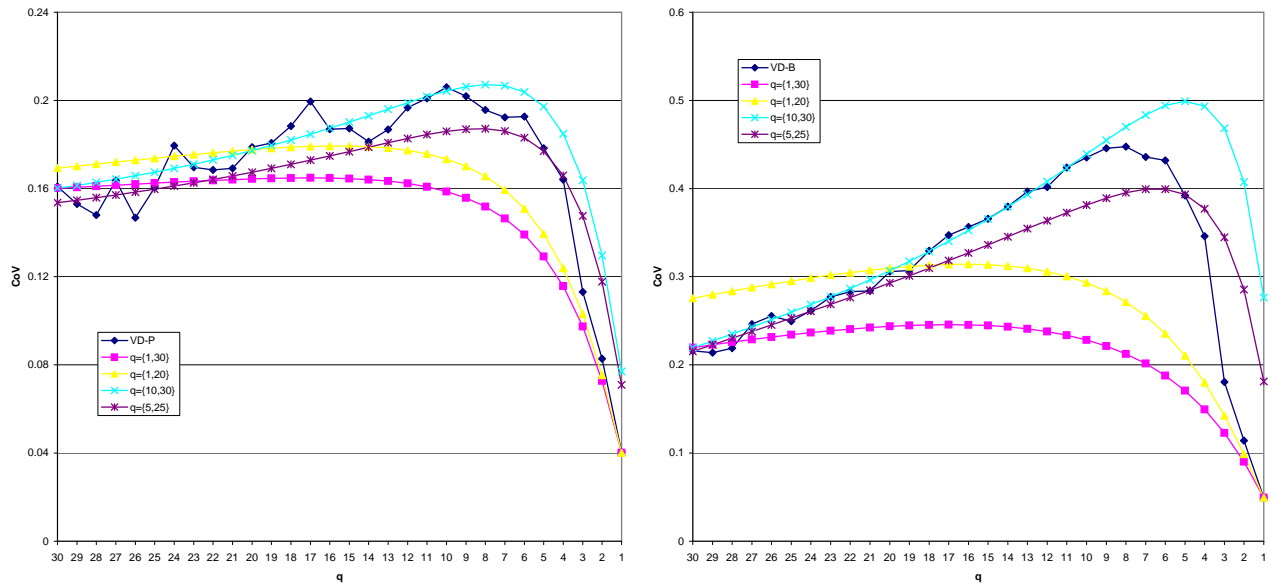
a) VD-P curve and models.

b) VD-B curve and models.



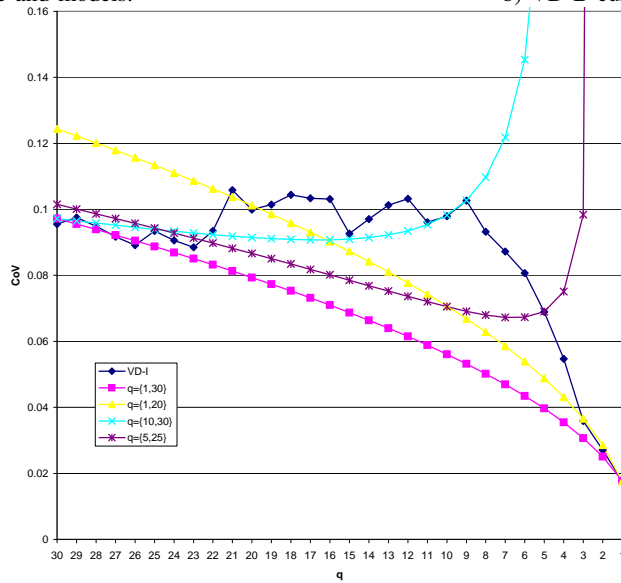
c) VD-I curve and models.

Fig. 40. VD curves and models for MPEG-4 coded *Football* scene 336, belonging to motion class V.



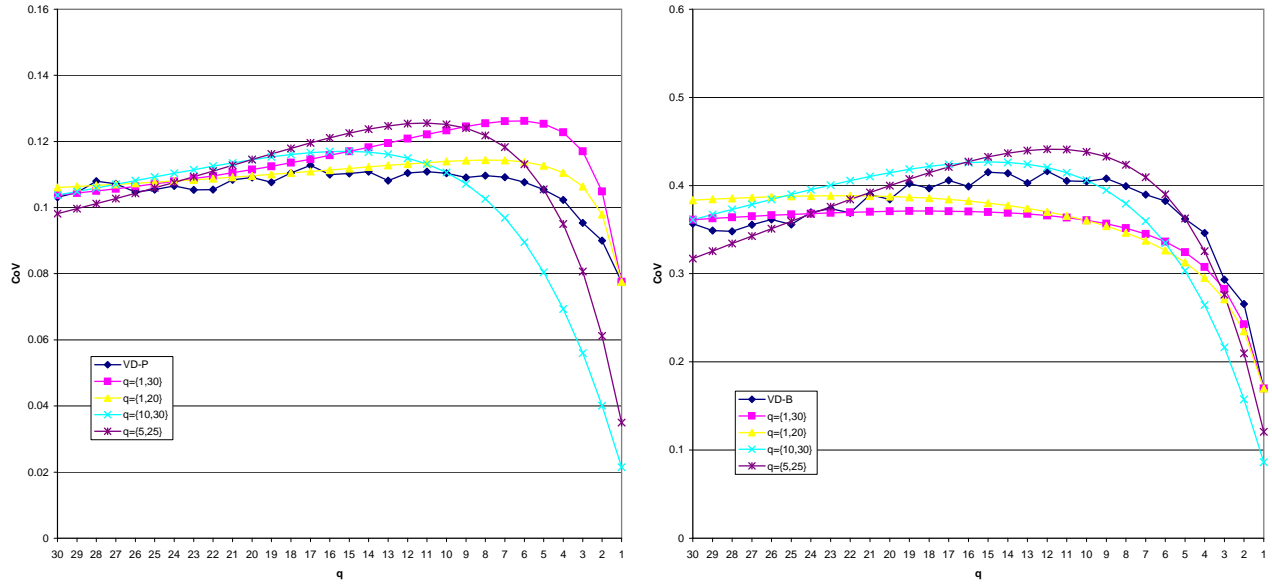
a) VD-P curve and models.

b) VD-B curve and models.



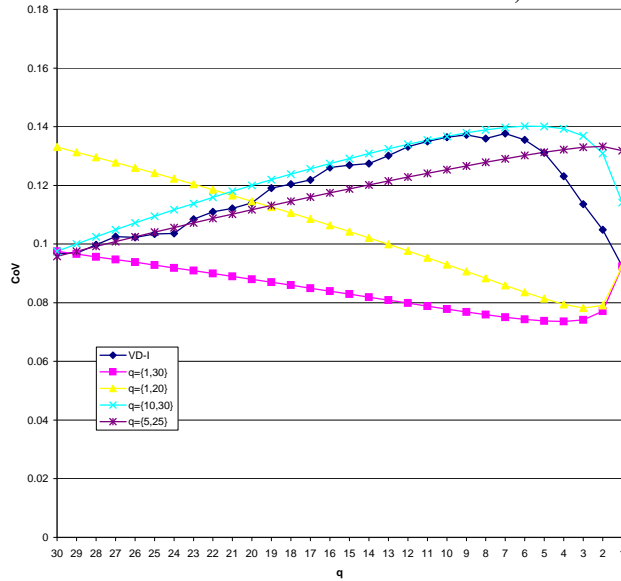
c) VD-I curve and models.

Fig. 41. VD curves and models for MPEG-4 coded *Star Wars IV* scene 632, belonging to motion class V.



a) VD-P curve and models.

b) VD-B curve and models.



c) VD-I curve and models.

Fig. 42. VD curves and models for MPEG-4 coded Terminator scene 441, belonging to motion class V.