This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Quality Driven Frame Rate Optimization for Rate Constrained Video Encoding

Ming-Chen Chien, Ren-Jie Wang, *Student Member, IEEE*, Chien-Hsun Chiu, and Pao-Chi Chang, *Member, IEEE*

***Abstract*—Video quality under rate constraint is mainly controlled by the frame rate and the quantization parameter. This work proposes a mechanism to obtain the optimal frame rate that maximizes video quality under rate constraint. Based on an objective metric of video quality that can reflect subjective quality, this work first proposes a video quality—frame rate—rate constraint model. Second, the relationship between model parameters and video characteristics is formulized. Finally, this work proposes an efficient frame rate optimization mechanism. Experimental results show that the optimal frame rate estimated by our mechanism is identical to the actual optimal frame rate under most bit rate constraints for both training sequences and new test sequences. In addition, the quality loss caused by the estimation error is generally limited within 0.8 dB in our experiments.**

***Index Terms*—Optimal frame rate, rate constraint, video coding.**

I. Introduction

APPLICATIONS of real-time video encoding and transmission, such as video conferencing and video broadcasting, have become increasingly more popular. The allowable bit rate per second R_S of video for these applications is generally constrained by network bandwidth. Therefore, effectively adjusting the encoding parameters to maximize video quality is a critical challenge. Video quality is composed of spatial quality and temporal quality. Spatial quality is dominated by the bit rate per frame R_F [1]. Temporal quality is dominated by the frame rate fr. For real-time video encoding, the encoding buffer should be small for low delay. R_F multiplied by fr is roughly equal to the rate constraint R_S; that is, R_F can be expressed by R_S/fr. Therefore, video quality under rate constraints is actually controlled by frame rate; hence, obtaining the optimal frame rate is critical.

To evaluate video quality, a proper assessment metric is necessary [2], [3]. PSNR is widely used to assess video quality. However, it can only objectively assess the spatial quality but not the temporal quality [4]. Several studies have investigated the impact of frame rate on perceptual video quality [5]–[8]. Wang *et al.* [5] studied the optimal frame rate over a wide range of bandwidth using subjective quality evaluation. The work in

Manuscript received June 27, 2011; revised September 10, 2011; accepted November 28, 2011. This material is based upon work supported in part by the National Science Council, Taiwan, under Grant NSC100-2221-E-008-090.

The authors are with the Department of Communication Engineering, National Central University, Jhong-Li 320, Taiwan. (e-mail: chienm@vaplab.ce.ncu.edu.tw; rjwang@vaplab.ce.ncu.edu.tw; chchiu@vaplab.ce.ncu.edu.tw; pcchang@ce.ncu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBC.2011.2182550

TABLE I
Options for Modeling

Sequences	Bus, Container, Flower , Foreman, Highway, Harbour, Mother Daughter, Soccer
Resolution	QCIF and CIF
Sequence Type	IPPP
Possible Frame Rates	30, 15, 10, 7.5, 6, 5
Number of Frames	150
Software	JM 16.2

[6] performed a double stimulus subjective evaluation to determine preferred frame rates at a fixed bit rate for low bit rate video. In [7], McCarthy *et al.* found that high spatial quality is more preferable than high frame rate for small screens. Chen and Thropp [8] conducted a comprehensive survey of the effects of different frame rates on human performance and summarized them in the areas of psychomotor performance, perceptual performance, behavioral effects, and subjective perception. However, no quality metric for video assessment were derived in these works.

The works in [9]–[13] proposed quality metrics to assess video quality. Lu *et al.* [9] proposed a logarithmic function of the frame rate to model the impact of frame rate dropping on perceptual video quality. The metric proposed in [10] is able to accurately estimate the perceived temporal degradation introduced by both consistent and inconsistent frame dropping. The work in [11] examined the jerkiness and jitter effects caused by different levels of strength, duration and distribution of the temporal impairment. However, these metrics did not consider spatial distortion, which is controlled by the bit rate per frame R_F. The work in [12] proposed a quality metric which emulated human visual perception based on block-fidelity, content richness fidelity, spatial-textural, color, and temporal masking. This model involves sophisticated processes to extract content components from video sequences and hence may be not applicable for practical application. The work in [13] proposed a quality metric as a function of Mean Square Error (MSE) and sequence edge strength. The proposed metric has better correlation with subjective quality compared to popular metrics. However, the relationship between MSE, frame rate, R_F has not been derived.

Ou *et al.* proposed an accurate metric of video quality [14], [15]. Their metric is formulized as the product of a Spatial Quality Factor (SQF) and a Temporal Correction Factor (TCF). The metric has only two content-dependent parameters, but is with significantly high correlation with Mean Opinion Scores (MOS). In addition, they also derived the two parameters as functions of video characteristics. However, the model depends on Q_0, which is the measured MOS for a video sequence

0018-9316/$31.00 © 2012 IEEE

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

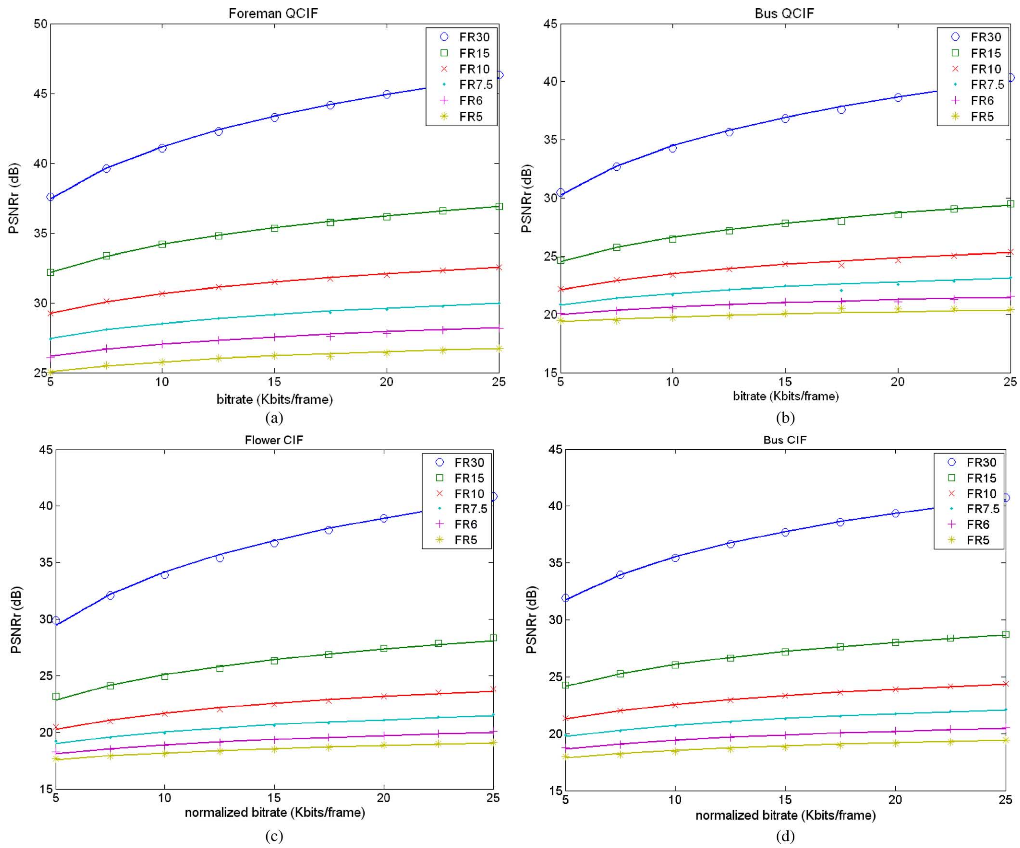


Fig. 1. PSNRr vs. bit rate per frame for fixed frame rates for various sequences. The values of each curve are obtained by the logarithm model (2). (a) Foreman QCIF, $\mathrm{R}^2 = 0.9979, \mathrm{RMSE} = 0.0625$ (b) Bus QCIF, $\mathrm{R}^2 = 0.9815, \mathrm{RMSE} = 0.1151$ (c) Flower CIF, $\mathrm{R}^2 = 0.9848, \mathrm{RMSE} = 0.1577$ (d) Bus CIF, $\mathrm{R}^2 = 0.9925, \mathrm{RMSE} = 0.0762$.

decoded at the lowest QP and the highest frame rate. Q_0 is content-dependent and cannot be obtained automatically by the processor of the video encoder because it is a score given by human beings [16]. Feghali *et al.* proposed the Quality Metric (QM) to assess the video quality [17]. QM is formulized as a function of frame rate and $PSNR_r$, the average PSNR of each frame with the skipped frame being reconstructed by repeating the previous frame. QM is objective, proportional to subject quality, and has been used to assess video quality in numerous studies [11], [18], [19]. However, $PSNR_r$ depends on R_F, frame rate, and hence should be further modeled. Regarding to up and down conversions of frame rate, sophisticate interpolation methods instead of simple repeating were proposed in the literatures [20]–[22].

In this work, we build a new video Quality—Frame rate—Rate constraint (Q-F-R) model based on QM. We model $PSNR_r$ as a concise function of R_F and frame rate. Furthermore, we determine three factors to represent sequence characteristics and formulized the relationship between model parameters and these three factors. Using this new model, the optimal frame rate of all the training sequences and other sequences can be obtained.

This paper is organized as follows. Section II proposes the new Q-F-R model and formulizes the relationship between model parameters and sequence characteristics. Section III proposes our mechanism of frame rate optimization during video encoding. Section IV presents experimental results. Finally, Section V draws conclusions.

II. The Proposed Video Quality-Frame Rate-Rate Constraint Model

The video quality metric QM is defined as [17]

$$QM = PSNR_r + c_1 m^{c_2}(30 - fr) \tag{1}$$

where $PSNR_r$ represents the average PSNR of each frame in the sequence. The skipped frame is reconstructed by repeating the previous frame. If the frame rate is lower than 30, the distortion of the skipped frame greatly degrades the value of $PSNR_r$. The lower the frame rate is, the higher the degradation is. Therefore, $PSNR_r$ can reflect the temporal quality dominated by the frame rate. Parameter m is the normalized value of the top 25 % of the largest motion vectors in the sequence. c_1 and c_2 are constants. Because the latter term of QM in (1) is effectively modeled as a function of frame rate, the modeling work can be reduced to model $PSNR_r$ as a function of frame rate and rate constraint.

The modeling environment is set as shown in Table I. According to our extensive experiments as Fig. 1 shows, $PSNR_r$ vs. bit rate per frame R_F for each fixed frame rate is a logarithm function for both QCIF and CIF sequences. Similar to [23], in order to build a concurrent Q-F-R model for both QCIF and CIF sequences, the values of R_F for CIF resolution are normalized as one quarter of total bits per frame so that the curves of $PSNR_r$ vs. R_F for CIF resolution can be similar to QCIF resolution. Consequently, the model parameters of CIF resolution can be close to that of QCIF resolution. The measurement criteria R^2 (Coefficient of determination) and RMSE (Root Mean

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

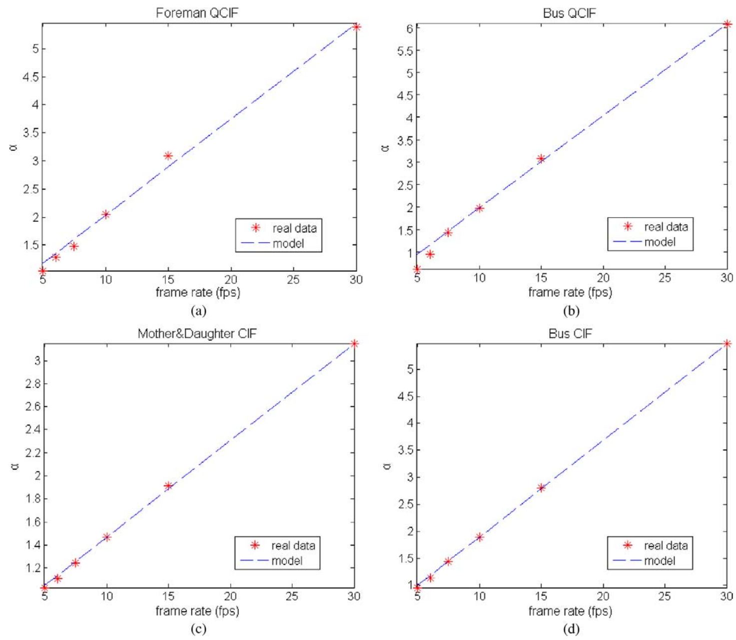


Fig. 2. Model parameter α vs. frame rate fr for various sequences. (a) Foreman QCIF, $\mathrm{R}^2 = 0.9934, \mathrm{RMSE} = 0.1712$ (b) Bus QCIF, $\mathrm{R}^2 = 0.9995, \mathrm{RMSE} = 0.05468$ (c) Mother & Daughter CIF, $\mathrm{R}^2 = 0.9996, \mathrm{RMSE} = 0.02137$ (d) Bus CIF, $\mathrm{R}^2 = 0.9999, \mathrm{RMSE} = 0.007703$.

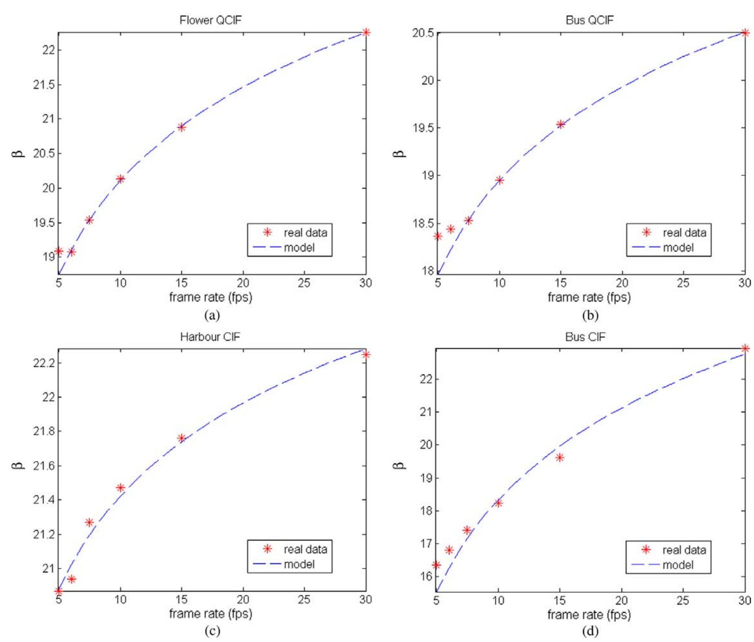


Fig. 3. Model parameter β vs. frame rate fr for various sequences. (a) Flower QCIF, $\mathrm{R}^2 = 0.9989, \mathrm{RMSE} = 0.06245$ (b) Bus QCIF, $\mathrm{R}^2 = 0.9998, \mathrm{RMSE} = 0.01453$ (c) Harbor CIF, $\mathrm{R}^2 = 0.9883, \mathrm{RMSE} = 0.06302$ (d) Bus CIF, $\mathrm{R}^2 = 0.9878, \mathrm{RMSE} = 0.3294$.

Square Error) show that the error of the modeling is very limited. The same phenomenon is observed for the other sequences. We first model the relationship as

$$PSNR_r = \alpha \ln R_F + \beta \tag{2}$$

where α and β are model parameters. We perform curve fitting on each curve of $PSNR_r$ vs. bit rate for each specific frame rate and obtain the value of α and β for each curve. As we measure the relationship between α and frame rate as shown in Fig. 2, we find that their relationship is linear and model it as

$$\alpha = a_1 fr + a_2 \tag{3}$$

where a_1 and a_2 are sequence dependent parameters. As we measure the relationship between β and frame rate as shown in Fig. 3, we find that their relationship is logarithmic and can be modeled as

$$\beta = b_1 \ln fr + b_2 \tag{4}$$

where b_1 and b_2 are sequence dependent parameters. Placing (3) and (4) into (2), we derive the QM—frame rate—bit rate model for both QCIF and CIF sequences as

$$QM = (a_1 fr + a_2) \ln \frac{R_S}{fr} + b_1 \ln fr + b_2 + c_1 m^{c_2} (30 - fr) \tag{5}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

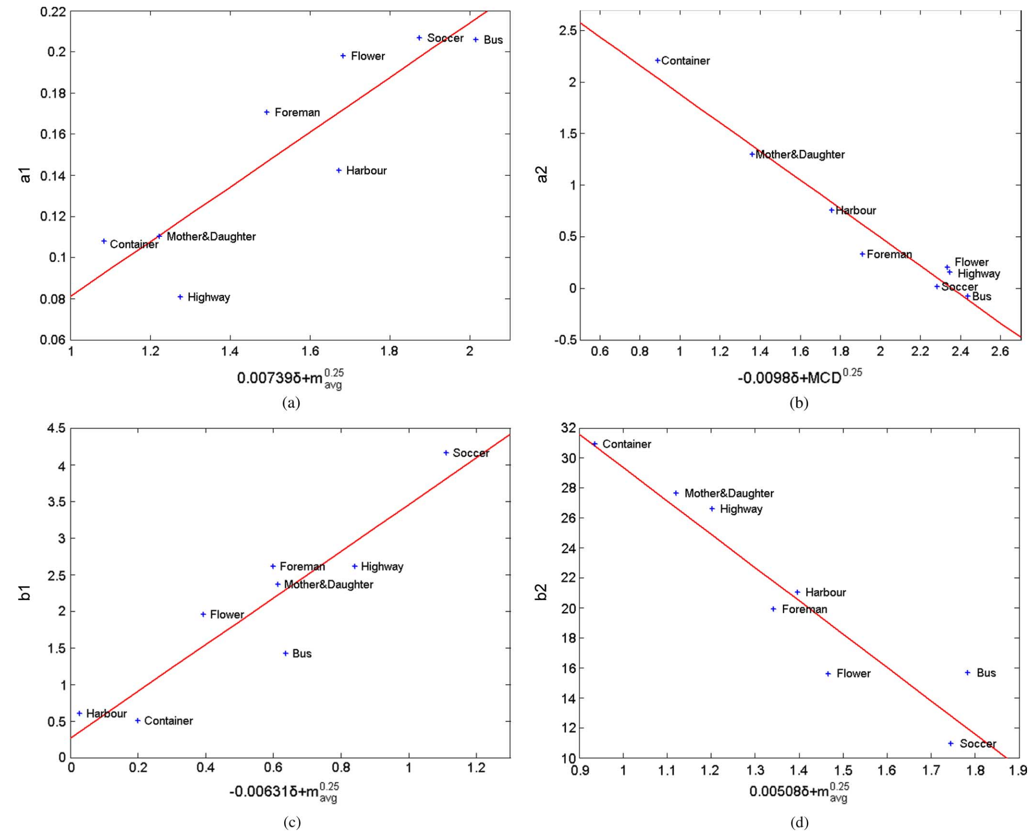


Fig. 4. Model parameter vs. weighted sum of characteristic parameters for QCIF training sequences, (a) model parameter a1, $\mathrm{R}^2 = 0.7752, \mathrm{RMSE} = 0.0254$ (b) a2, $\mathrm{R}^2 = 0.9584, \mathrm{RMSE} = 0.1734$ (c) b1, $\mathrm{R}^2 = 0.8397, \mathrm{RMSE} = 0.5186$, (d) b2, $\mathrm{R}^2 = 0.8905, \mathrm{RMSE} = 2.469$.

Here the value of R_S as well as R_F for CIF size should be normalized by dividing by 4. Because the values of a_1, a_2, b_1, and b_2 depend on sequence characteristics, their relationship with sequence characteristics must be derived. In general, the characteristics of a video sequence include the temporal and spatial characteristics. The temporal characteristics can be measured by the motion prediction and the prediction error. The spatial characteristics can be measured by the edge strength that can represent the texture. We use average motion m_{avg} defined as (6) to represent the motion of a video sequence.

$$m_{avg} = \frac{1}{XYT}\sum_{t=1}^{T}\sum_{y=1}^{Y}\sum_{x=1}^{X}\sqrt{mvx_{x,y,t}^2 + mvy_{x,y,t}^2} \quad (6)$$

where X, Y, and T are the number of Macro Blocks (MBs) in a row, the number of MBs in a column, and the number of frames, respectively. (mv_x, mv_y) represents the motion vector in 16x16 block size.

The motion prediction error is represented by MCD, defined as

$$\begin{aligned} MCD &= \frac{1}{XYT}\sum_{t=1}^{T}\sum_{y=1}^{Y}\sum_{x=1}^{X} \times (F(x,y,t) \\ &\quad - F(x - mvx_{x,y,t}, y - mvy_{x,y,t}, t-1))^2 \end{aligned} \quad (7)$$

Note that $F(x - mvx_{x,y,t}, y - mvy_{x,y,t}, t-1)$ is the pixel value in the previously original frame with motion compensation. Although the residual signal, which refers to the previously reconstructed frame, can easily obtained in H.264, it is not suitable to represent the video complexity because it depends on the bit rate. The block size of motion compensation for MCD is 16×16.

The Sobel filter defined as (8) is widely used to obtain the edge strength that can represent the spatial complexity of a video sequence [13].

$$\delta = \frac{1}{XYT}\sum_{t=1}^{T}\sum_{y=1}^{Y}\sum_{x=1}^{X}(|G_h(x,y,t)| + |G_v(x,y,t)|) \quad (8)$$

where $G_h(x,y,t)$ and $G_v(x,y,t)$ are horizontal and vertical edge strengths, respectively. They are defined as (9) and (10).

$$G_h = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \otimes F(x,y,t) \quad (9)$$

$$G_v = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \otimes F(x,y,t) \quad (10)$$

where $F(x,y,t)$ represents the pixel value located at (x,y) in frame t, and $\otimes$ denotes 2D convolution. The values of characteristic parameters m_{avg}, δ and MCD of the training sequences are listed in Table II.

To analyze the relationship between the model parameters (a_1, a_2, b_1, b_2) and video characteristics ($\mathrm{m_{avg}}$, δ, MCD), we perform extensive curve fitting. We try all possible combinations and consider both the conciseness and accuracy. We discover that a_1 is approximately linear to $(0.00739\delta + m_{avg}^{0.25})$ for

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

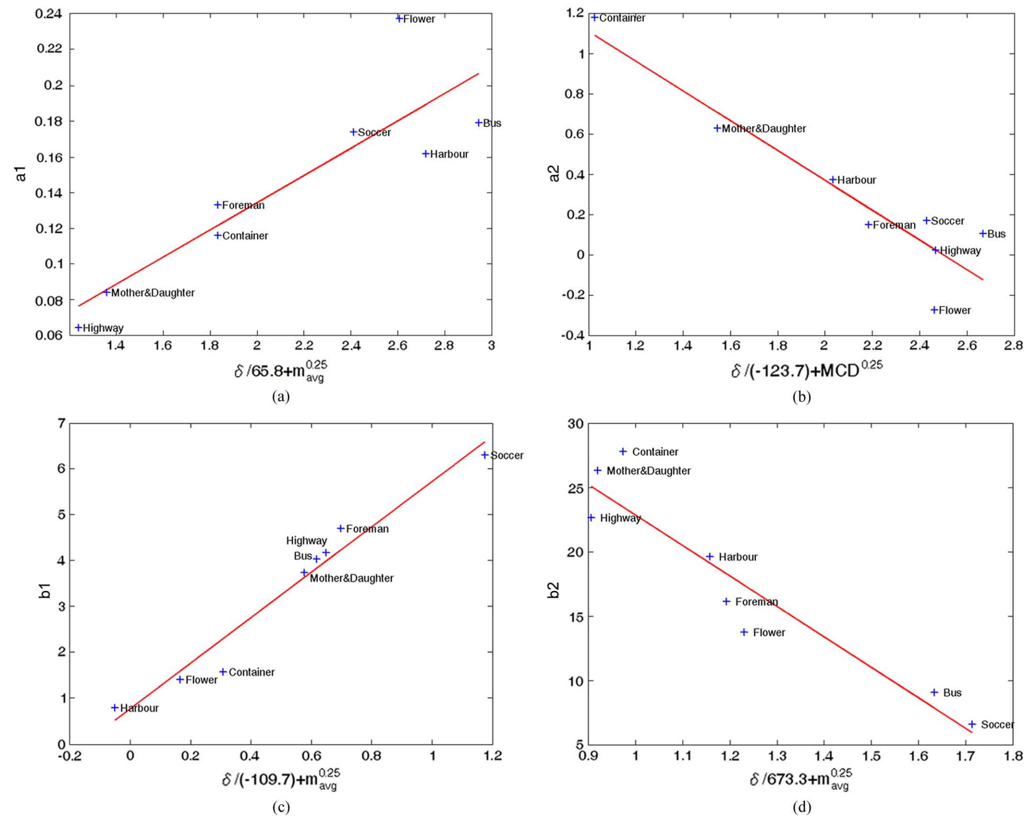


Fig. 5. Model parameter vs. weighted sum of characteristic parameters for CIF training sequences, (a) model parameter a1, $\mathrm{R}^2 = 0.7696, \mathrm{RMSE} = 0.0252$(b) a2, $\mathrm{R}^2 = 0.8688, \mathrm{RMSE} = 0.1496$ (c) b1, $\mathrm{R}^2 = 0.9593, \mathrm{RMSE} = 0.3598$, (d) b2, $\mathrm{R}^2 = 0.8886, \mathrm{RMSE} = 2.4201$.

TABLE II
VIDEO SEQUENCES FOR THE SIMULATION

Sequence Parameters	m_{avg}	δ	MCD
Bus	2.613	100.71	140.44
Container	0.136	64.52	5.256
Flower	0.95	94.24	115.03
Foreman	1.043	65.11	43.01
Harbour	0.381	120.07	76.47
Highway	1.174	31.7	50.41
Mother Daughter	0.64	44.53	10.61
Soccer	4.566	55.7	65.15

QCIF sequences as Fig. 4(a) shows. With linear regression, we derive the relationship between the model parameters a_1 and video characteristics as (11). With the same process, we derive (12)–(14) for QCIF sequences.

$$\begin{aligned} a_1 &= 0.133\left(0.00739\delta + m_{avg}^{0.25}\right) - 0.0519 \\ &= 0.133 m_{avg}^{0.25} + 0.000983\delta - 0.0519 \qquad (11) \\ a_2 &= 0.0138\delta - 1.387 MCD^{0.25} + 3.269 \qquad (12) \\ b_1 &= 3.187 m_{avg}^{0.25} - 0.0201\delta + 0.2669 \qquad (13) \\ b_2 &= -22.24 m_{avg}^{0.25} - 0.113\delta + 51.61 \qquad (14) \end{aligned}$$

Similarly, for CIF sequences (see Fig. 5), the relationship between model parameter and sequence characteristics can be found as (15)–(18).

$$\begin{aligned} a_1 &= 0.0766\left(0.0152\delta + m_{avg}^{0.25}\right) - 0.0187 \\ &= 0.0766 m_{avg}^{0.25} + 0.0012\delta - 0.0187 \qquad (15) \\ a_2 &= 0.006\delta - 0.739 MCD^{0.25} + 1.849 \qquad (16) \\ b_1 &= 4.971 m_{avg}^{0.25} - 0.0453\delta + 0.759 \qquad (17) \\ b_2 &= -23.62 m_{avg}^{0.25} - 0.0351\delta + 46.44 \qquad (18) \end{aligned}$$

For each video sequence, the Q-F-R model can be built based on (5) and the four model parameters can be obtained according to (11)–(14) or (15)–(18).

III. VIDEO ENCODING WITH THE PROPOSED MECHANISM OF FRAME RATE OPTIMIZATION

The ideal implementation of the proposed mechanism of frame rate optimization is a two-pass process. The first pass gathers the video characteristics of a video sequence and obtains the optimal frame rate according to our algorithm. The second pass encodes the same video sequence with the optimal frame rate. This implementation yields best performance but the complexity overhead is high.

To save the memory and the computational complexity, we suggest a Lower-Complexity QM-Based (LCQMB) mechanism of frame rate optimization. The LCQMB mechanism gathers the video characteristics during video encoding for each processing time unit, defined as the group of frames (GoF), where the GoF is set to be one second in this work. The optimal frame rate for the next GoF of video is predicted from the optimal frame rate for the present GoF. Due to temporal correlation, this design generally works fine for the video within a scene. For the video with scene change, the frame rate estimated from the previous GoF might not be optimal for the GoF with scene change and the frame rate estimated from the GoF with scene change may not be optimal for the next GoF. Consequently, the coding errors of

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

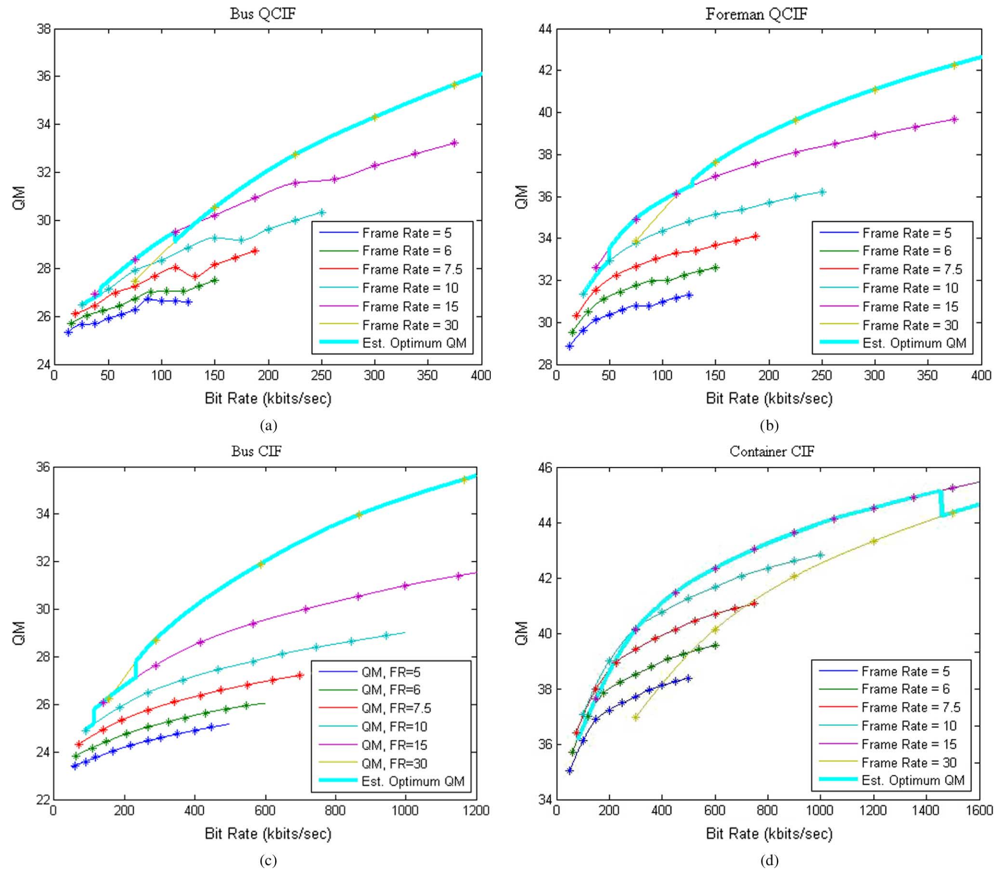


Fig. 6. Objective quality metrics for fixed frame rates under different bit rate constraints for the training sequences: (a) Bus QCIF; (b) Foreman QCIF; (c) Bus CIF; and (d) Container CIF.

video with scene change might be erroneous and will even propagate to the proceeding frames. To solve this problem, we can use scene change detection such as detecting the proportion of Intra mode and set the first frame of the GoF after scene change to I frame, then the coding error propagation will be terminated. Without error propagation and with the optimal frame rate estimated from the next GoF using our method, the following GoF can have excellent quality.

The optimal frame rate of the first GoF depends on the bit rate constraint R_S. From our experimental results using various kinds of video sequences as shown in Figs. 6 and 7, the optimal frame rate for QCIF is 15 fps for bit rate between 50 and 175 kbps and 30 fps for bit rate higher than 175 kbps. For bit rate lower than 50 kbps, the optimal frame rate is 10 or lower. With the statistic results described above, this method set the optimal frame rate of the first GoF accordingly.

The proposed method of frame rate optimization operates as follows:

Step 1) Gather video characteristics. The motion factors m and m_{avg} are obtained after motion estimation with 16×16 block size. The horizontal and vertical edge strength of each frame, $G_h(x,y,t)$ and $G_v(x,y,t)$, are obtained when the frame is read to memory for encoding. The texture factor δ is obtained by (8). MCD is obtained by (7).

Step 2) Build the Q-F-R model. The four model parameters of QCIF sequences are obtained according to (11)–(14) and the parameters of CIF sequences are obtained according to (15)–(18). For CIF sequences, the value of R_S is normalized as one quarter of total bits per second for modeling. Based on the four parameters and the given rate constraint, the model is built according to (5).

Step 3) Predict the optimal frame rate for the next GoF. Place the possible frame rates into the model (5). The frame rate that makes the QM value greatest is the optimal frame rate for this GoF. The optimal frame rate for the next GoF is predicted by it.

Step 4) Encode video with the predicted frame rate. The video frames for the next GoF are down-sampled for encoding based on the predicted frame rate. The bit rate of each frame is controlled to R_S/fr, where the value of R_S is the actual one instead of the normalized one for actual CIF sequences coding.

TABLE III
VIDEO SEQUENCES FOR THE SIMULATION

Sequences in the training set	Bus, Foreman, Flower Garden, Highway QCIF Bus, Container CIF
New test sequences	Car phone, Ice, Suzie, Stefan QCIF Football, Hall CIF

IV. EXPERIMENTAL RESULTS

This section compares the optimal frame rate obtained by the proposed mechanism and the actual optimal frame rate obtained by extensive experiments in which we try each possible frame rate to encode video sequences for each given bit rate. The frame

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

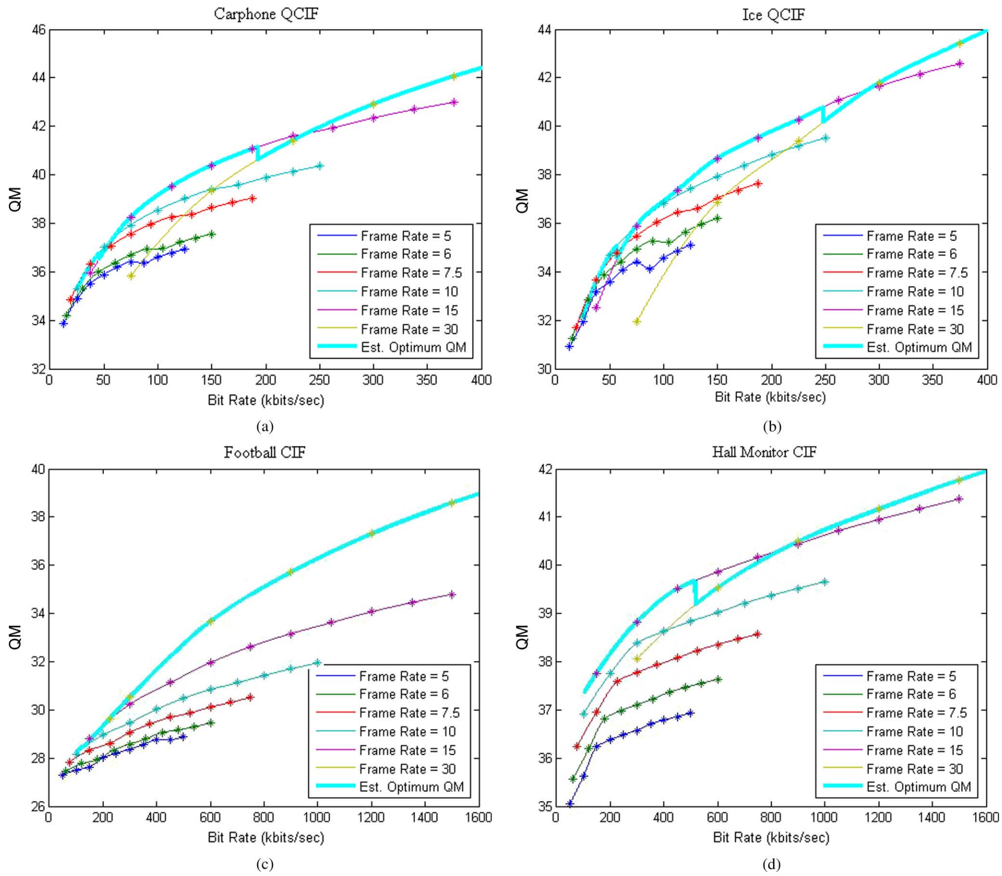


Fig. 7. The objective quality metrics for fixed frame rates under different bit rate constraints for the new test sequences: (a) Car phone QCIF; (b) Ice QCIF; (c) Football CIF; and (d) hall monitor CIF.

TABLE IV
THE VALUES OF QM AND THE OPTIMAL FRAME RATE WITH THE PROPOSED LCQMB MECHANISM WHEN THE BIT RATE PER SECOND IS SET TO 100 KBPS

Sequence (QCIF)	Method	QM (frame 1~30)	QM (frame 31~60)	QM (frame 61~90)	QM (frame 91~120)	QM (frame 121~150)
Bus	LCQMB	27.75 (fr=15)	27.41 (fr=15)	28.44 (fr=15)	28.91 (fr=15)	30.49 (fr=15)
	2pass	27.75 (fr=15)	27.41 (fr=15)	28.44 (fr=15)	28.91 (fr=15)	30.49 (fr=15)
Foreman	LCQMB	35.56 (fr=15)	35.42 (fr=15)	34.83 (fr=15)	36.76 (fr=15)	34.93 (fr=15)
	2pass	35.56 (fr=15)	35.42 (fr=15)	34.83 (fr=15)	36.76 (fr=15)	34.93 (fr=15)
Suzie	LCQMB	39.81 (fr=15)	36.88 (fr=15)	36.06 (fr=15)	39.14 (fr=15)	40.37 (fr=15)
	2pass	39.81 (fr=15)	36.88 (fr=15)	36.06 (fr=15)	39.14 (fr=15)	40.37 (fr=15)
Ice	LCQMB	32.73 (fr=15)	35.73 (fr=15)	36.48 (fr=15)	36.06 (fr=15)	35.85 (fr=15)
	2pass	32.73 (fr=15)	35.73 (fr=15)	36.48 (fr=15)	36.06 (fr=15)	35.85 (fr=15)
Stefan	LCQMB	26.85 (fr=15)	27.10 (fr=15)	26.11 (fr=15)	25.64 (fr=15)	26.28 (fr=15)
	2pass	26.85 (fr=15)	27.10 (fr=15)	26.11 (fr=15)	25.64 (fr=15)	26.28 (fr=15)
Carphone	LCQMB	37.36 (fr=15)	39.25 (fr=15)	37.43 (fr=15)	39.44 (fr=15)	39.63 (fr=15)
	2pass	37.36 (fr=15)	39.25 (fr=15)	37.43 (fr=15)	39.44 (fr=15)	39.63 (fr=15)
Flower	LCQMB	30.44 (fr=15)	29.72 (fr=15)	29.22 (fr=15)	28.43 (fr=15)	29.00 (fr=15)
	2pass	30.44 (fr=15)	29.72 (fr=15)	29.22 (fr=15)	28.43 (fr=15)	29.00 (fr=15)
highway	LCQMB	40.16 (fr=15)	40.14 (fr=15)	39.52 (fr=15)	38.69 (fr=15)	39.72 (fr=15)
	2pass	40.16 (fr=15)	40.14 (fr=15)	39.52 (fr=15)	38.69 (fr=15)	39.72 (fr=15)

rate which yields maximal QM value in (1) is the actual optimal frame rate. We first examine the ideal implementation with two pass. The video sequences of the simulation are listed in Table III. The number of frames used to gather video characteristics and determined the optimal frame rate is 150. In addition to the training sequences, the new test sequences are also used to examine the mechanism. The other options of the simulation are set as Table I shows.

Fig. 6 shows the video quality measured by QM for the training sequences encoded with various frame rates under various bit rate constraints. Fig. 7 shows the video quality measured by QM for the new test sequences. The value of bit rate constraint for CIF sequences shown in these two figures is the actual one instead of the normalized one. Under a certain bit rate constraint, the frame rate that leads to the highest QM value in (1) is the actual optimal frame rate. The bold curve is the QM value (1) of sequences encoded with the optimal frame rate estimated by our mechanism. Note that some curves in the figures exhibit reverse relationship between bit rate and the value of QM, that is, higher bit rate does not always result in

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE V
THE VALUES OF QM AND THE OPTIMAL FRAME RATE WITH THE PROPOSED LCQMB MECHANISM WHEN THE BIT RATE PER SECOND IS SET TO 150 KBPS

Sequence (QCIF)	Method	QM (frame 1~30)	QM (frame 31~60)	QM (frame 61~90)	QM (frame 91~120)	QM (frame 121~150)
Bus	LCQMB	28.80 (fr=15)	29.03 (fr=30)	30.23 (fr=30)	29.32 (fr=30)	30.51 (fr=30)
	2pass	29.29 (fr=30)	29.03 (fr=30)	30.23 (fr=30)	29.32 (fr=30)	30.51 (fr=30)
Foreman	LCQMB	36.74 (fr=15)	36.55 (fr=15)	36.04 (fr=30)	37.97 (fr=15)	36.03 (fr=15)
	2pass	36.74 (fr=15)	36.96 (fr=30)	36.01 (fr=15)	37.97 (fr=15)	36.03 (fr=15)
Suzie	LCQMB	41.02 (fr=15)	37.65 (fr=15)	38.49 (fr=30)	40.23 (fr=15)	41.37 (fr=15)
	2pass	41.02 (fr=15)	38.40 (fr=30)	37.11 (fr=15)	40.23 (fr=15)	41.37 (fr=15)
Ice	LCQMB	34.22 (fr=15)	37.09 (fr=15)	38.02 (fr=15)	37.48 (fr=15)	37.35 (fr=15)
	2pass	34.22 (fr=15)	37.09 (fr=15)	38.02 (fr=15)	37.48 (fr=15)	37.35 (fr=15)
Stefan	LCQMB	28.79 (fr=15)	27.65 (fr=30)	27.20 (fr=30)	25.52 (fr=30)	28.21 (fr=30)
	2pass	26.80 (fr=30)	27.65 (fr=30)	27.20 (fr=30)	25.52 (fr=30)	28.21 (fr=30)
Carphone	LCQMB	38.53 (fr=15)	39.12 (fr=30)	38.64 (fr=15)	39.31 (fr=30)	40.91 (fr=15)
	2pass	37.91 (fr=30)	40.56 (fr=15)	38.54 (fr=30)	40.72 (fr=15)	40.91 (fr=15)
Flower	LCQMB	30.44 (fr=15)	31.20 (fr=30)	30.08 (fr=30)	30.53 (fr=30)	31.03 (fr=30)
	2pass	30.96 (fr=30)	31.20 (fr=30)	30.08 (fr=30)	30.53 (fr=30)	31.03 (fr=30)
highway	LCQMB	40.61 (fr=15)	40.73 (fr=15)	40.07 (fr=15)	39.11 (fr=15)	40.34 (fr=15)
	2pass	40.61 (fr=15)	40.73 (fr=15)	40.07 (fr=15)	39.11 (fr=15)	40.34 (fr=15)

TABLE VI
THE VALUES OF QM AND THE OPTIMAL FRAME RATE WITH THE PROPOSED LCQMB MECHANISM WHEN THE BIT RATE PER SECOND IS SET TO 200 KBPS

Sequence (QCIF)	Method	QM (frame 1~30)	QM (frame 31~60)	QM (frame 61~90)	QM (frame 91~120)	QM (frame 121~150)
Bus	LCQMB	31.05 (fr=30)	30.42 (fr=30)	31.62 (fr=30)	30.86 (fr=30)	32.18 (fr=30)
	2pass	31.05 (fr=30)	30.42 (fr=30)	31.62 (fr=30)	30.86 (fr=30)	32.18 (fr=30)
Foreman	LCQMB	38.13 (fr=30)	38.33 (fr=30)	37.57 (fr=30)	39.07 (fr=30)	37.69 (fr=30)
	2pass	38.13 (fr=30)	38.33 (fr=30)	37.57 (fr=30)	39.07 (fr=30)	37.69 (fr=30)
Suzie	LCQMB	41.02 (fr=30)	39.98 (fr=30)	39.84 (fr=30)	40.90 (fr=30)	41.40 (fr=30)
	2pass	41.02 (fr=30)	39.98 (fr=30)	39.84 (fr=30)	40.90 (fr=30)	41.40 (fr=30)
Ice	LCQMB	38.50 (fr=30)	42.10 (fr=30)	39.03 (fr=15)	38.47 (fr=15)	38.47 (fr=15)
	2pass	38.50 (fr=30)	38.22 (fr=15)	39.03 (fr=15)	38.47 (fr=15)	38.47 (fr=15)
Stefan	LCQMB	28.60 (fr=30)	29.47 (fr=30)	28.60 (fr=30)	27.37 (fr=30)	30.00 (fr=30)
	2pass	28.60 (fr=30)	29.47 (fr=30)	28.60 (fr=30)	27.37 (fr=30)	30.00 (fr=30)
Carphone	LCQMB	39.15 (fr=30)	40.46 (fr=30)	39.93 (fr=30)	40.62 (fr=30)	40.79 (fr=30)
	2pass	39.15 (fr=30)	40.46 (fr=30)	39.93 (fr=30)	40.62 (fr=30)	40.79 (fr=30)
Flower	LCQMB	32.85 (fr=30)	32.37 (fr=30)	32.17 (fr=30)	32.06 (fr=30)	32.65 (fr=30)
	2pass	32.85 (fr=30)	32.37 (fr=30)	32.17 (fr=30)	32.06 (fr=30)	32.65 (fr=30)
highway	LCQMB	40.60 (fr=30)	40.69 (fr=30)	40.28 (fr=30)	40.10 (fr=30)	40.32 (fr=30)
	2pass	40.60 (fr=30)	40.69 (fr=30)	40.28 (fr=30)	40.10 (fr=30)	40.32 (fr=30)

higher QM. It is caused by the rate control mechanism. These simulations show that the optimal frame rate estimated by our mechanism is identical to the actual optimal frame rate under most bit rate constraints. In addition, the quality loss caused by the estimation error is limited within 0.8 dB.

Then we examine the accuracy of the LCQMB mechanism proposed in Section III. The results are listed from Tables IV–VI. As shown in Tables IV and VI, the optimal frame rate for each GoF of video estimated by the LCQMB mechanism is identical to the optimal frame rate estimated by the two-pass process when the bit rate per second R_S is set to 100 kbps and 200 kbps. When R_S is set to 150 kbps, the optimal frame rate for most GoF of video estimated by the LCQMB mechanism is also identical to the optimal frame rate estimated by the two-pass process as shown in Table V. These results reveal that the proposed LCQMB mechanism provide an accurate estimation of the optimal frame rate under various bit rate constraints.

V. CONCLUSION AND FUTURE WORK

In this work, we propose a mechanism to determine the optimal frame rate in the sense of maximizing the objective video quality metric which can reflect the perceptual quality. Experimental results show that reducing frame rate may be more effective than increasing quantization step size when the given bit rate is not sufficiently high. Results also reveal that the frame rate provided by the proposed algorithm is close to the optimal frame rate obtained by the extensive experiments.

In this work, we set the default duration of determining frame rate to one second. Our future work is planned to study the optimal duration of changing frame rate. Once the optimal duration is figured out, the duration of the proposed algorithm can be set to the optimal duration. Another future work is planned to incorporate our research on spatial resolution [24] and address the frame rate and spatial resolution optimization of video coding under rate constraints.

REFERENCES

[1] G. Cermak, M. Pinson, and S. Wolf, "The relationship among video quality, screen resolution, and bit rate," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 258–262, Jun. 2011.

[2] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 165–182, Jun. 2011.

[3] I. P. Gunawan and M. Ghanbari, "Efficient reduced-reference video quality meter," *IEEE Trans. Broadcast.*, vol. 54, no. 3, pp. 669–679, Sep. 2008.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[4] D. Wang, F. Speranza, A. Vincent, T. Martin, and P. Blanchfield, “Towards optimal rate control: a study of the impact of spatial resolution, frame rate, and quantization on subjective Video quality and bit rate,” in *Proc. VCIP*, pp. 198–209.

[5] Y. Wang, S.-F. Chang, and A. Loui, “Subjective preference of spatio-temporal rate in video adaptation using multi-dimensional scalable coding,” in *Proc. ICME*, Jun. 2004, vol. 3, pp. 1719–1722.

[6] G. Yadavalli, M. Masry, and S. S. Hemami, “Frame rate preference in low bit rate video,” in *Proc. ICIP*, Nov. 2003, vol. 1, p. I–441–4.

[7] J. McCarthy, M. A. Sasse, and D. Miras, “Sharp or smooth?: Comparing the effects of quantization vs. frame rate for streamed video,” in *Proc. ACM CHI Human Factors Comput. Syst.*, Apr. 2004, pp. 535–542.

[8] J. Y. C. Chen and J. E. Thropp, “Review of low frame rate effects on human performance,” *IEEE Trans. Syst., Man, Cybern.*, vol. 37, pp. 1063–1076, Nov. 2007.

[9] Z. Lu, W. Lin, B. C. Seng, S. Kato, S. Yao, E. Ong, and X. K. Yang, “Measuring the negative impact of frame dropping on perceptual visual quality,” in *Proc. SPIE Human Vis. Electron. Imaging*, Jan. 2005, vol. 5666, pp. 554–562.

[10] K. C. Yang, C. C. Guest, K. El-Maleh, and P. K. Das, “Perceptual temporal quality metric for compressed video,” *IEEE Trans. Multimedia*, vol. 9, pp. 1528–1535, Nov. 2007.

[11] Q. Huynh-Thu and M. Ghanbari, “Temporal aspect of perceived quality of mobile video broadcasting,” *IEEE Trans. Broadcast.*, vol. 54, no. 3, pp. 641–651, Sep. 2008.

[12] E. Ong, X. Yang, W. Lin, Z. Lu, and S. Yao, “Perceptual quality metric for compressed videos,” in *Proc. ICASSP*, Mar. 2005, vol. 2, pp. 581–584.

[13] A. Bhat, I. Richardson, and S. Kannangara, “A new perceptual quality metric for compressed video,” in *ICASSP*, 2009, vol. 51, no. 3, pp. 933–936.

[14] Y. F. Ou, Z. Ma, T. Liu, and Y. Wang, “Perceptual quality assessment of video considering both frame rate and quantization artifacts,” *IEEE Trans. Circuits Syst. Video Technol.*, online available.

[15] Y.-F. Ou, T. Liu, Z. Zhao, Z. Ma, and Y. Wang, “Modeling the impact of frame rate on perceptual quality of video,” in *IEEE ICIP*, San Diego, CA, Oct. 2008, pp. 689–692.

[16] *Methodology for the Subjective Assessment of the Quality for Television Pictures*, ITU-R BT.500ITU-R Std., Rev. 11, Jun. 2002.

[17] R. Feghali, D. Wang, F. Speranza, and A. Vincent, “Video quality metric for bit rate control via joint adjustment of quantization and frame rate,” *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 441–446, Mar. 2007.

[18] S. H. Jin, C. S. Kim, D. J. Seo, and Y. M. Ro, “Quality measurement modeling on scalable video applications,” in *Proc. IEEE Workshop Multimedia Signal Process.*, Oct.. 2007, pp. 131–134.

[19] C. S. Kim, S. H. Jin, D. J. Seo, and Y. M. Ro, “Measuring video quality on full scalability of H.264/AVC scalable video coding,” *IEICE Trans. Commun.*, vol. E91-B, no. 5, pp. 1269–1278, 2008.

[20] D. Wang, L. Zhang, and A. Vincent, “Motion-compensated frame rate up-conversion—Part I: Fast multi-frame motion estimation,” *IEEE Trans. Broadcast.*, vol. 56, no. 2, pp. 441–446, Jun. 2010.

[21] D. Wang, A. Vincent, P. Blanchfield, and R. Klepko, “Motion-compensated frame rate up-conversion—Part II: New algorithms for frame interpolation,” *IEEE Trans. Broadcast.*, vol. 56, no. 2, pp. 142–149, Jun. 2010.

[22] C. T. Hsu, C. H. Yeh, C. Y. Chen, and M. J. Chen, “Arbitrary frame rate transcoding through temporal and spatial complexity,” *IEEE Trans. Broadcast.*, vol. 55, no. 4, pp. 767–775, Dec. 2009.

[23] A. M. Bruckstein, M. Elad, and R. Kimmel, “Down-scaling for better transform compression,” *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1132–1144, Sep. 2003.

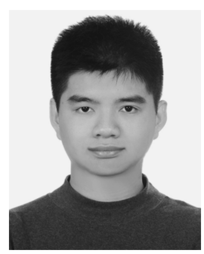
[24] R. J. Wang, M. C. Chien, and P. C. Chang, “Adaptive down-sampling video coding,” in *Proc. SPIE Conf. Electron. Imaging*, San Jose, CA, Jan. 2010, vol. 7542, pp. 1–8.

Ming-Chen Chien received the B.E., M.S., and Ph.D. degrees in electrical engineering from National Central University, Taiwan, in 1994, 1996, and 2011 respectively.

From 1996 to 2000, he was a software engineer of Accton Technology Corporation, where he developed network software for Ethernet switch hub. In 2000, he joined the faculty of the Electronic Engineering Department in Asia-Pacific Institute of Creativity, Taiwan. His research interests include video coding, image processing, and game development.

Ren-Jie Wang (S’10) received the B.S. degree from National Quemoy University and the M.S. degree from National Central University, Taiwan, in 2007 and 2009, respectively. He is currently studying toward Ph. D. degree in National Central University.

His research interests include video/image compression, scalable coding, video surveillance and image/video retrieval.



Chien-Hsun Chiu received the B.S. and M.S. degrees from National Central University, Taiwan, in 2008 and 2010.

He is currently on military service. His research interests include video/image compression, scalable coding.

Pao-Chi Chang (M’86) received the B.S. and M.S. degrees from National Chiao Tung University, Taiwan, in 1977 and 1979, respectively, and the Ph. D. degree from Stanford University, California, in 1986, all in electrical engineering.

From 1986 to 1993, he was a research staff member of the department of communications at IBM T. J. Watson Research Center, Hawthorne, New York. At Watson, his work centered on high speed switching systems, efficient network design algorithms, and multimedia conferencing. In 1993, he joined the faculty of National Central University (NCU), Taiwan, where he is presently a Professor in the Department of Communication Engineering and the Department of Electrical Engineering, as well as the Advisor of the Video-Audio Processing Laboratory (VAPLab). He was a visiting professor at Stanford University in 2000 and 2004, respectively. His main research interests include speech/audio coding, video/image compression, scalable coding, error resilient coding, digital watermarking and data hiding, and multimedia delivery over packet and wireless networks.