國立交通大學

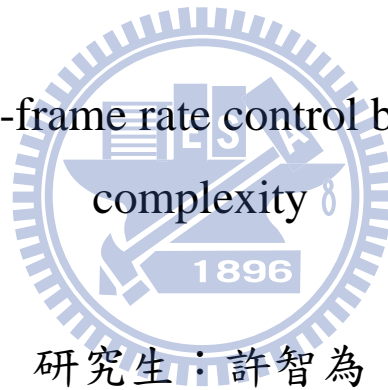## 多媒體工程研究所

碩士論文

基於偵幅複雜度畫面內編碼在 H.264 的位元率控制演算法

A H.264 intra-frame rate control based on frame complexity

研究生：許智為

指導教授：蔡文錦 博士

中華民國 99 年 7 月

基於偵幅複雜度畫面內編碼在 H.264 的位元率控制演算法

A H.264 intra-frame rate control based on frame complexity

研 究 生：許智為　　　　　Student：Chih-Wei Hsu

指導教授：蔡文錦　　　　　Advisor：Wen-Jiin Tsai

國 立 交 通 大 學
多 媒 體 工 程 研 究 所
碩 士 論 文

A Thesis
Submitted to Institute of Multimedia Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Computer Science

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

中文摘要

　　為了讓視訊編碼後的位元率能維持在頻寬的限制之內，並且達到良好與穩定的畫面品質，位元率控制是相當重要的。然而目前大多數的研究都集中在畫面間編碼(inter frames)而不是容易造成緩衝區溢位的畫面內編碼(intra frames)。此外 Intra-only 這種編碼方式也已經納入 H.264 新的 profile 當中，它較傳統 GOP 編碼更適合應用在特別要求畫面品質的產品上。

　　在此論文中，我們提出一個於 Intra-only 編碼的位元率控制演算法。首先我們提出基於 R-D-Q model 最佳化的 QP 決定方式，藉由位元率與 PSNR 預測模型，我們利用 R-D-Q model 最佳化來找出可以平衡畫面品質與編碼效能的 QP 最佳解。另外，在 Intra-only 編碼上為了解決場景變換對位元率控制所產生的影響，我們所利用的 R-D-Q model 可以直接偵測場景變換並做適當 QP 選擇，避免緩衝區溢位。實驗結果顯示，提出的方法可以達到更好更穩定的畫面品質，且緩衝內含量也都維持在較低的水平。

關鍵字： 位元率控制、H.264、畫面內編碼、R-D-Q model、預測模型

ABSTRACT

Rate control serves as an important technique to constrain the bit rate of video transmission over a limited bandwidth and to control the bit allocations within a video sequence to maximize its overall visual quality. However, most of rate control researches focus on inter coding frames instead of intra coding frames which are more possible to cause buffer overflow problem. Besides, H.264 Intra-only compression scheme has been standardized as H.264 profiles which are more proper for professional applications than traditional GOP compression scheme.

In this thesis, we propose an improved rate control scheme which is appropriate for Intra-only compression. First, we present a R-D-Q model based QP determination scheme for I-frames and P-frames . By the estimation models for rate and PSNR of I-frames, the best quantization parameters can be determined by R-D-Q model method. In order to deal with the specific intra frames caused by scene transitions, our R-D-Q model can directory detect scene change without extra detection to determine appropriate QPs for avoiding buffer overflow and saving bit budget. Simulation results show, that compared to other reference algorithms, our approach achieves better and stable quality with low buffer fullness.

Index Terms: Rate control, H.264, intra frames, R-D-Q model, Prediction model

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1   Introduction and Motivation

For the coming of digital multimedia communication, the demand for the storage and transmission of visual information has stimulated the development of video coding standards, including MPEG-1[1], MPEG-2[2], MPEG-4[3], H.261[4], H.263[5], and H.264/AVC[6].

H.264 is an up-to-date coding standard approved by ITU-T as MPEG 4 - Part 10 Advanced Video Coding (AVC). It includes the latest advances of video coding techniques. H.264 is designed in two layers: a video coding layer (VCL), and a network adaptation layer (NAL). Although H.264/AVC basically follows the framework of prior video coding standards such as MPEG-2, H.263, and MPEG-4, it contains new features that enable it to achieve a significant improvement in compression efficiency.

1.1   Introduction to Rate Control

A rate control algorithm which meets a constrained channel rate by controlling the number of generated bits is necessary to encoder. Either the coded video is transmitted over the Internet or stored in a storage device; there is a bandwidth constraint to limit the bit rate of videos. Although the transmission bandwidth is growing larger over the years, more exquisite videos with high resolutions, such as HD and Full HD, are becoming popular. These high definition videos consume much more bit rate than the traditional definition videos. Encoding video without rate control will suffer from several serious problems. For example, when the coded video transmits through a weak wireless access point (AP), network congestion and packet loss will occur if the bit rate of the video is higher than the bandwidth of the AP. In another example, suppose the generated bits are not constrained carefully, the fact that out of storage capacity will

happen. Fig 1-1 shows the two mentioned examples. Hence, rate control is a key issue of the modern video coding researches.
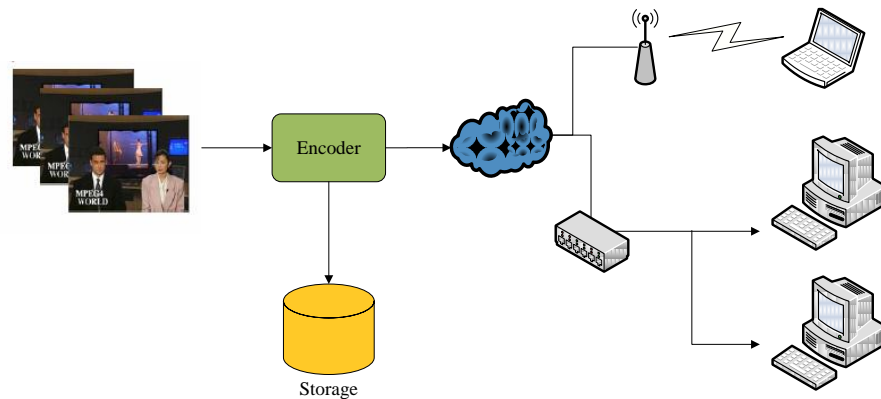
Fig 1-1   Video transmission system

The generated bits and video quality of an encoder highly rely on several coding parameters, especially the quantization parameter (QP). In particular, choosing a large QP reduces the resulting bit rate and meanwhile the visual quality of the encoded video is reduced. For illustration, Fig 1-2(a) shows that if the QP is constant, the resulting video is at a stable quality with a variable bit rate (VBR). However, a predetermined constant bit rate (CBR) is desired in most applications, such as CD, DVD, or video broadcast. Fig 1-2(b) shows the quality of a coded video with CBR floats because of the
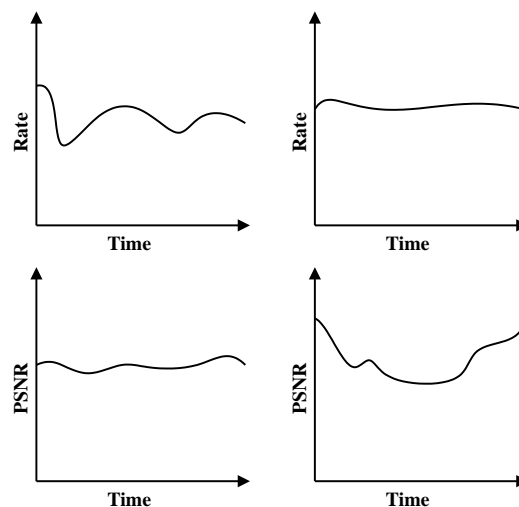
video content varying.

(a)   (b)

The task of controlling output bit rate by selecting an appropriate quantization parameter for each coding unit is performed by the rate control module. The goal of rate control is to keep the generated bit rate within the constrained bandwidth while achieving maximum video quality uniformly. A simple approach of rate control is shown in Fig 1-3. Basically, the encoder buffer smoothes out the bit rate so that the averaged output bit rate matches the channel bit rate.
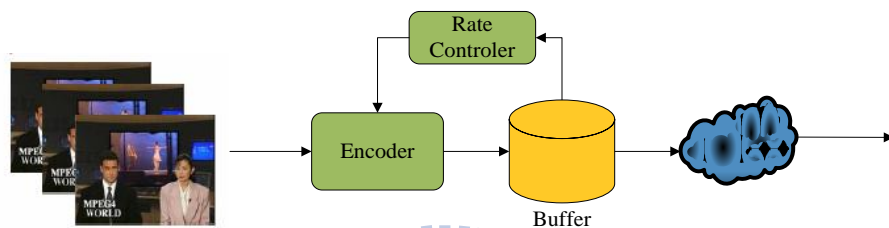
Fig 1-3   Basic rate control flow

The loss of synchronization with buffer in-coming rate and out-going rate usually causes buffer overflow or underflow. When the encoder generates more bits than the amount of bits the buffer can hold, a buffer overflow happens. The encoder then either re-encodes the current frame with coarser QP or simply drops it (frame skip) to avoid the overflow. A buffer underflow is the situation while there is no bit available in the encoder buffer. It wastes the available channel bandwidth. By monitoring the status of buffer, the rate controller can adjust the quantization parameters, which affects the output bit rate, to prevent the buffer from overflow and underflow.

1.1.1   The Chicken Egg Dilemma for H.264 Rate Control

One important property of H.264 is the implementation of rate distortion optimization (RDO)[7] for both motion estimation and mode decision. With RDO, the Lagrangian method is utilized to optimize the trade-off between distortion and bit rate consumed. For example, the Lagrangian cost function of motion estimation[7] is

$$J_{Motion}\left(MB_i, MV_i \mid QP, \lambda\right) = D\left(MB_i, MV_i \mid QP\right) + \lambda \cdot R\left(MB_i, MV_i \mid QP\right) \quad \textbf{(1-1)}$$

where MB_i and MV_i stand for the ith macro block (MB) and the motion vector (MV) of ith MB in the current frame, respectively; λ denotes the Lagrangian multiplier which depends on

$$\lambda = \sqrt{0.85 \times 2^{(QP-12)/3}} \quad \textbf{(1-2)}$$

According to (1-1) and (1-2), the cost calculation for each MV of the current MB takes QP as an important input parameter.
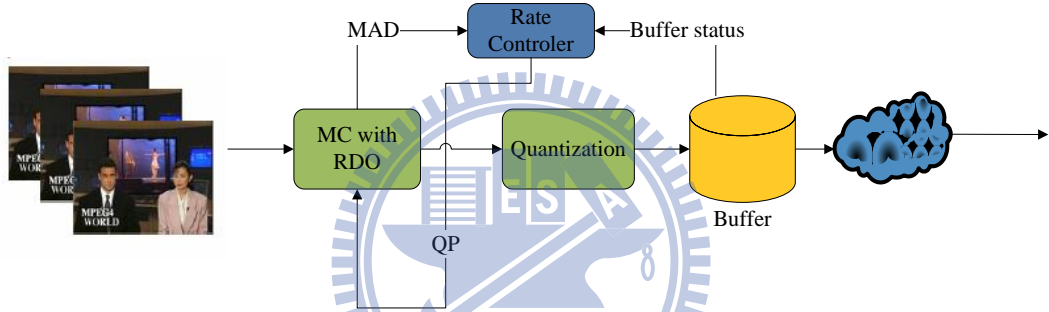
Fig 1-4   The chicken egg dilemma for H.264 rate control

Therefore, in H.264, QP affects both rate distortion optimization and residual quantization. In this way, the statistical information of the residual frame, such as mean absolute difference (MAD), varies with the QP adjustment, and the QP decision is also influenced by the statistical information. As shown in Fig 1-4, the rate control unit requires the MAD value from RDO to determine the QP value, but the RDO procedure also needs QP as an input parameter. This is the chicken egg dilemma for H.264 rate control.

1.1.2   Main Criteria of Rate Control

Rate control algorithms concentrate on keeping the encoded video quality as consistent and excellent as possible for each frame and constraining the bit rate within limited bandwidth. For grading rate control algorithms, there are four main criteria of

4

rate control:

A. *Mismatch between the target bit rate and the output bit rate.*

Because the main purpose of rate control is to constrain the output bit rate within the target bit rate, the mismatch between both should be minimized.

B. *Average PSNR of whole sequence.*

The generated video quality should be at the highest possible level for a better watching experience.

C. *Standard deviation of PSNR between frames.*

This criterion implies the quality variation of the video produced by the rate control algorithm. A good rate control should keep the deviation low, i.e., keep the quality variation small.

D. *Maximum buffer fullness.*

Lower maximum buffer occupancy implies that a small buffer is sufficient for preventing from buffer overflow. Further, a small buffer only takes few buffer delay while transmission. A good rate control algorithm should minimize the maximum buffer fullness.
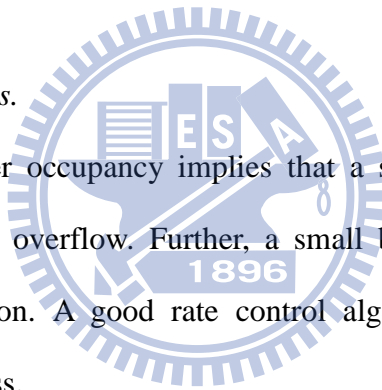
1.2 Introduction to H.264 Intra-coded Frames

H.264 exploits both temporal and spatial redundancy to increase its coding gain. It supports intra prediction mode to exploit the spatial domain correlation which helps reduce the residual energy of intra frames.

Recently, H.264 intra-only coding scheme for professional applications has been standardized as H.264 profiles[8]. These intra-only profiles take the advantages of H.264 intra coded frames and make H.264 as another great selection for intra compressed video.
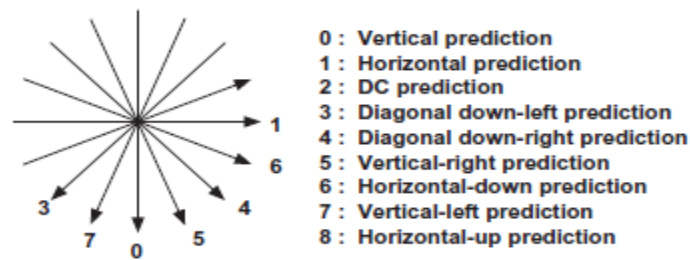
Fig 1-5   4x4 block intra prediction mode direction[9]

1.2.1   H.264 Intra Compression

H.264 utilizes the intra prediction to reduce the spatial redundancy within frames. Fig 1-5 shows the prediction options of 4x4 block intra prediction. Each pixel in the current 4x4 block is predicted from the neighboring reconstructed pixels, where nine prediction modes can be selected by the encoder, and the residue between the current block and the predicted block will be quantized for entropy coding. The key to the success of intra coding on improving the performance is that the entropy of the residual block is much less than the original block. Hence, the coding gain after intra prediction will be significantly superior.

1.2.2   H.264 Intra-only Profiles

In the seventh edition specification of H.264, there are three new profiles, e.g., *High 10 Intra*, *High 4:2:2 Intra*, and *High 4:4:4 Intra*, which are designed for professional applications. For the reason that the intra-only profile does not exploit the temporal correlation, there is no temporal dependency between consecutive frames. It is more convenient for editing and parallel processing, even less error propagation. Table 1-1 summaries the differences between intra-only scheme and the standard GOP compression. Because of the features of intra-only compression, it is greatly appropriate for the high-end applications.

6
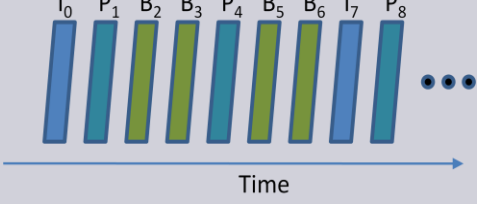
	Intra-only Compression		GOP Compression	
Compression scheme				
Bit rate saving	**Smaller**	Use spatial correlation only	**Greater**	Use spatial and temporal correlations
Process delay	**Smaller**	1 frame	**Greater**	Multiple frames
Edit easiness	**Easier**	frame by frame	**More difficult**	GOP
Error propagation	**Smaller**	Max. 1 frame	**Greater**	Multiple frames
Parallel processing	**Easier**	Frame independent	**More difficult**	GOP independent

Table 1-1 Comparison between Intra-only and GOP compression[10]

1.3 Motivation

Rate control aims at providing highest possible video quality while satisfying the limited bandwidth. Although various rate control algorithms have been proposed for H.264 (see Chapter 2), most of them focus on inter coding instead of intra coding, because the output number of bits of an intra coding frame is much higher than that of an inter frame. It is also more possible that the intra coded frame causes buffer overflow when the generated bits exceed the amount of bits that buffer can hold.

In the H.264 original rate control algorithm[11], the QP for each I-frame is decided by the average QP of all coded P-frames in the previous GOP. This approach does not take the buffer status and the frame complexity into consideration, and usually allocate too much bits for the I-frame, which degrades the video quality of the following

P-frames due to insufficient bits. In addition, because the intra coded DCT coefficients are not Laplacian distributed, the quadratic model which is used to predict the relation between bit rate and quantization parameter is not appropriate for intra frames.

We observed that the QP determination by pre-frame information's rate control algorithm would lead to the unpredictable result, for example the buffer overflow, PSNR drop, and the PSNR standard derivation larger, such as in more activity sequence or in the sequence which happened scene change often. Fig 1-6 shows the information predict by pre-frame. It really tells us the prediction by pre-frame in scene change shows the prediction error lager, which can lead to over of under estimation QP makes the generated bits larger or smaller than target bits.

Since most existing rate control algorithms for H.264 cannot handle the intra frames and scene change frames well, we need to find out a new scheme to determine the QPs for both kinds of frames. Instead of using the average QP of P-frames in the previous GOP, in this thesis, we propose an improved rate control algorithm that takes frame complexity into consideration to decide proper QPs for the intra frames.

The remainder of this thesis is organized as follows: Chapter 2 introduces the related researches about rate control issue. Chapter 3 presents the proposed rate control scheme for Intra-only compression. Chapter 4 provides the simulation results compared to other rate control schemes. Finally, Chapter 5 concludes this thesis.
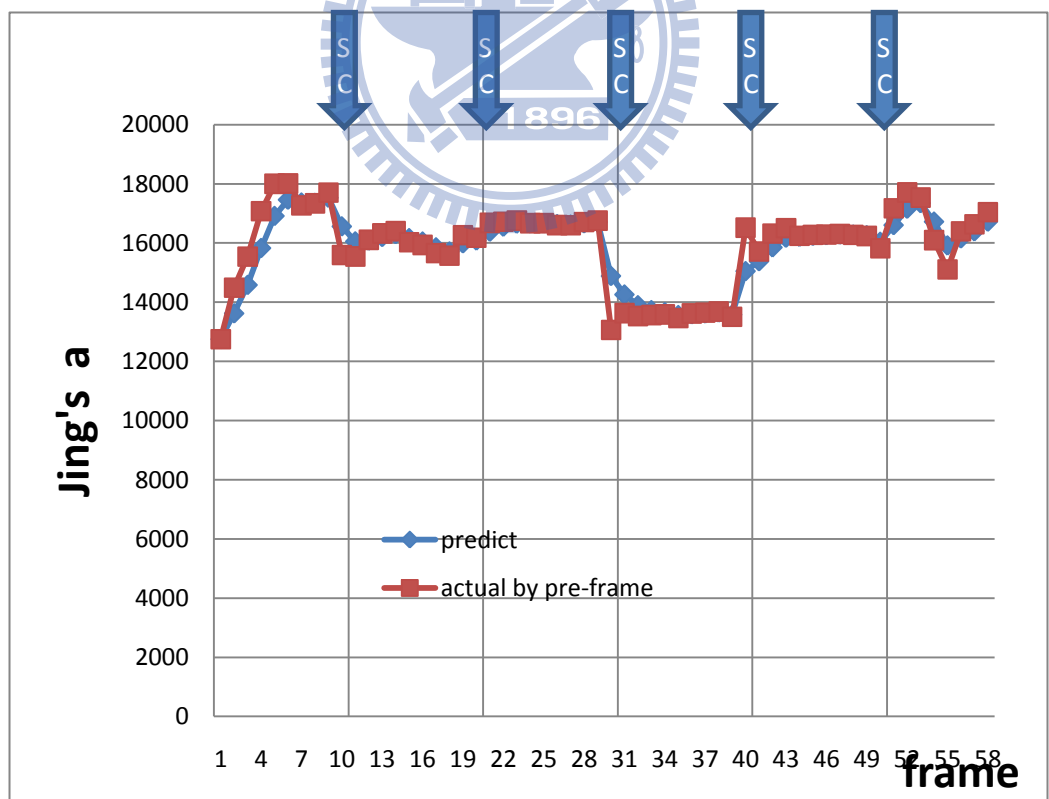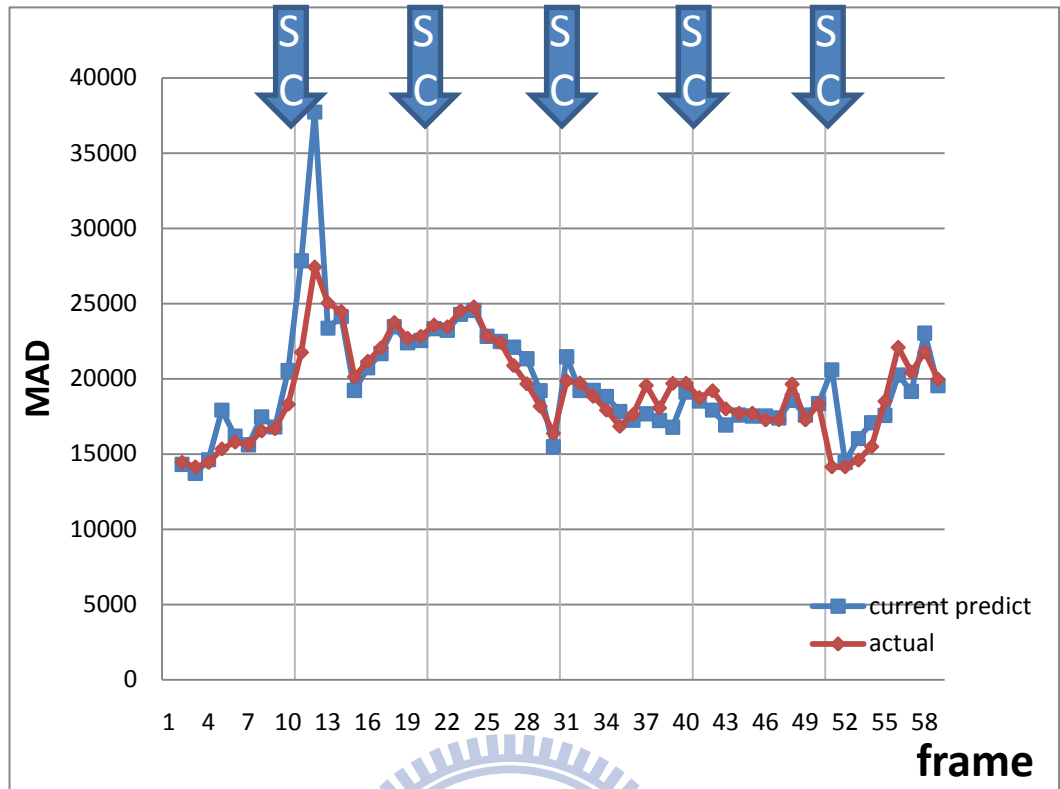
Fig 1-6 (a) JM's MAD predict and actual curve (b) Jing's a predict and actual curve @ test sequence scene change(SC) pre 10 frame

Chapter 2   Related Works

Rate control techniques have been studied intensively for many standards. The challenge of rate control in video encoding is to determine an appropriate quantization parameters to achieve the best video quality within the given application constraints. In this chapter, we will introduce the most famous rate control algorithm which is adopted in the official reference coding software of H.264[12] and other improved schemes for H.264 intra rate control.

2.1   G012 Rate Control for H.264

Li *et al.* proposed an one pass rate control algorithm, JVT-G012[11], which used the rate-quantization (R-Q) quadratic model in the standard MPEG4 rate control, and introduced the linear mean absolute difference (MAD) prediction model to solve the dilemma that we have mentioned in the previous chapter. Due to its efficiency, this scheme was adopted by JVT in the latest H.264 reference software.

2.1.1   Terminology

Before we introduce this algorithm, there are three terminologies we have to mention first.

A.   *Definition of A Basic Unit*

Suppose that a frame is composed of N_{mbpic} macroblocks (MBs). A basic unit is defined as a group of continuous MBs which is composed of N_{mbunit} macroblocks where N_{mbunit} is a fraction of N_{mbpic}. Denote the total number of basic units in a frame by N_{unit}, which is given by

$$N_{unit} = \frac{N_{picunit}}{N_{mbunit}} \tag{2-1}$$

A basic unit can be selected as a frame or some consecutive MBs. Note that, a smaller basic unit is needed in some low-delay applications which require stricter buffer regulations, less buffer delay, and better spatially perceptual quality. However, it is costly at low bit rate since there is additional overhead if the quantization parameter is varying frequently within a frame. On the other hand, by using a bigger basic unit, a higher PSNR can be achieved but the bit fluctuation is also larger.

B. Linear Model for MAD Prediction

MAD is the mean absolute difference between the reference frame and the current frame which describes the residue information and is given by

$$MAD(x, y) = \frac{1}{HW} \cdot \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \left| C(x+i, y+j) - R(x+i, y+j) \right| \tag{2-2}$$

where C and R stand for the original and referenced pixel, respectively.

In order to solve the chicken egg dilemma in H.264 rate control, the linear model is used to predict the MADs of the basic units in the current frame by using the MADs of the co-located basic units in the previous frame. The linear prediction model is then given by

$$MAD_{pb} = a_1 \times MAD_{cb} + a_2 \tag{2-3}$$

where a_1 and a_2 are two coefficients of the prediction model; MAD_{pb} and MAD_{cb} stand for the predicted MAD of the current basic unit and the real MAD of the co-located basic unit, respectively. The initial values of a_1 and a_2 are set to 1 and 0, respectively. They are updated after each basic unit has been encoded.

C. The MPEG4 quadratic rate distortion model

In order to illustrate the quadratic rate distortion model, we summarize the results in [13][14]. Assume that the source statistics satisfy a Laplacain distribution

$$P(x) = \frac{\alpha}{2} e^{-\alpha|x|} \quad \text{where} \ -\infty < x < \infty$$

(2-4)

and the distortion measure is defined by, $D(x, \bar{x}) = |x - \bar{x}|$, where x is the original sample and $\bar{x}$ is the reconstruction of x. Then, a closed solution for R-D function was derived as

$$R(D) = \ln\left(\frac{1}{\alpha D}\right) \quad \text{where} \ D_{\min} = 0, D_{\max} = \frac{1}{\alpha}, 0 < D < \frac{1}{\alpha}$$

(2-5)

Based on the R-D function, a quadratic rate-control model was proposed in [13] as

$$R = \frac{X_1}{QP} + \frac{X_2}{QP^2}$$

(2-6)

where R is the target number of bits used for encoding the current frame, and X_1 and X_2 are model parameters which are updated by linear regression method from previous coded information.

Lee *et al.*[14] improved the model with content scalability and achieved more accurate bit allocation within limited target bits. The improved model has been adopted as a part of the MPEG4 standard, and known as MPEG4 Q2 algorithm. The quadratic rate distortion model is defined by

$$R = \frac{MAD \cdot X_1}{QP} + \frac{MAD \cdot X_2}{QP^2} - H$$

(2-7)

where H is the number of bits used for the header, the motion vectors, and other non-texture information. Here, MAD is used to measure the coding complexity for

accomplishing the scalability of this model.

2.1.2 Overview to G012 Rate Control

As shown in Fig 2-1, G012 partitioned the rate control problem into three layers: 1) GOP layer; 2) frame layer, and 3) basic unit layer. There are two sub-problems, bit allocation and QP determination, for each layer.

GOP layer
　　Bit allocation for the GOP
　　Calculate the intra QP for the GOP

Frame layer
　　Bit allocation for the frame
　　Calculate the QP for the frame

Basic unit layer
　　Bit allocation for the basic unit
　　Calculate the QP for the basic unit
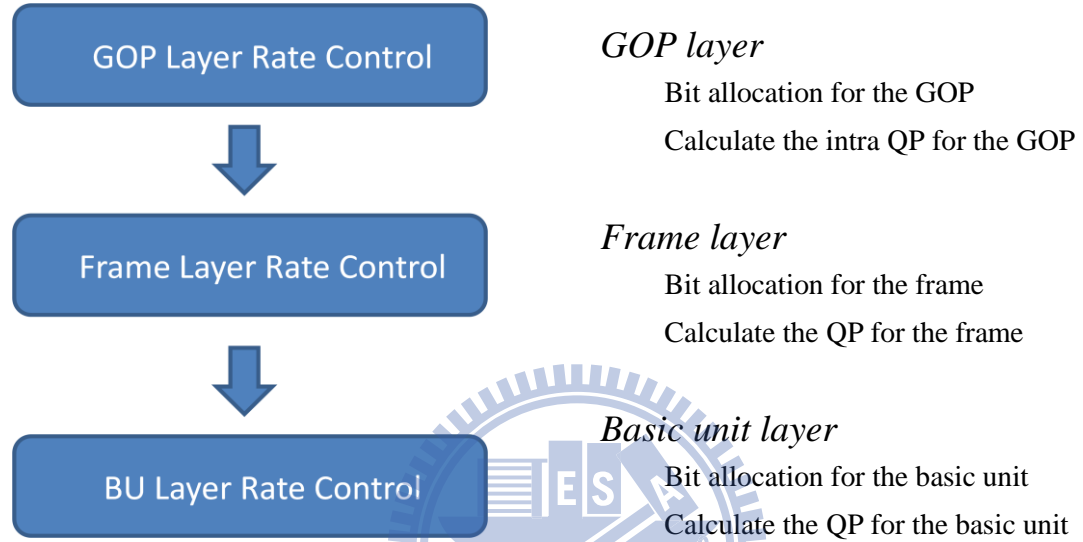
Fig 2-1　The G012 rate control diagram

In GOP layer rate control, it calculates the total bits R_r for all non-coded frames within the current GOP, and selects the QP for the starting I-frame. In the beginning of each GOP, the total number of bits is computed as follows

$$R_r = \frac{u}{F_r} \cdot N_{GOP} - B_c$$

(2-8)

where u is the channel bit rate; F_r indicates the frame rate; N_{GOP} denotes the number of frames in a GOP, and B_c is the occupancy of the buffer after coding the previous frame. In the case of constant bit rate, R_r is updated frame by frame as

$$R_r = R_r - b$$

(2-9)

where b is the number of bits generated from the previous coded frame.

The starting QP of the first GOP, $QP_{I,first}$ depends on the channel bit rate and the value of bit per pixel (bpp). On the other side, the starting QP of other GOPs, $QP_{I,other}$ is determined as the average QP of the P-frames of the previous GOP. Summarily, the starting QP is selected as follows

$$\begin{cases} QP_{I,first} = \begin{cases} 40 & bpp \leq l_1 \\ 30 & l_1 < bpp \leq l_2 \\ 20 & l_2 < bpp \leq l_3 \\ 10 & l_3 < bpp \leq l_4 \end{cases} & ,where \quad bpp = \dfrac{u}{F_r \times N_{pixel}} \\ QP_{I,other} = \dfrac{SumQP}{N_p} \end{cases}$$

$$\quad (2\text{-}10)$$

where N_{pixel} is the number of pixels within a frame; N_p indicates the number of P-frames of a GOP, and $SumQP$ stands for the summation of QPs of all P-frames of the previous GOP. l_i, $1 \leq i \leq 4$, are the predefined thresholds.

The approach of frame layer involves distributing the GOP budget among the frames and determines the QP of each frame to achieve the allocated budget. The target number of bits of ith P-frame in the current GOP is determined as

$$R_i = \beta \cdot \hat{R}_i + (1 - \beta) \cdot \tilde{R}_i$$

$$\quad (2\text{-}11)$$

where β is a weighted constant; $\hat{R}_i$ and $\tilde{R}_i$ are defined as

$$\hat{R}_i = \frac{R_r}{N_{remain}} \quad (2\text{-}12)$$

$$\tilde{R}_i = \frac{u}{F_r} + \gamma \cdot (Tbl_i - V_i) \quad (2\text{-}13)$$

where N_{remain} is the number of non-coded frames in the current GOP; γ is a constant, and Tbl_i and V_i are the target buffer level and the virtual buffer fullness of the ith frame, respectively.

After accomplishing the bit allocation, the linear MAD prediction model (2-3) and the quadratic rate distortion model (2-7) are utilized to determine the QP of the current frame, and then RDO procedure is performed for mode decision. At the last, the parameters of the quadratic model, and those of the MAD prediction model are updated based on the coding results.

If frames are not selected as basic units, basic unit layer rate control should be performed after frame layer bit allocation. In basic unit layer, it is almost the same as that in frame layer. It predicts MADs of all basic units in the current frame by equation (2-3) and calculates the target number of bits of them by

$$\tilde{b}_i = R_{c,remain} \cdot \frac{MAD_{i,pred}^2}{\sum\limits_{j=i}^{N_{unit}} MAD_{j,pred}^2} \qquad \textbf{(2-14)}$$

where $R_{c,remain}$ is the remaining target number of bits of current frame; $MAD_{i,pred}$ stands for the predicted MAD of ith basic unit in the current frame. Then, the quadratic model (2-7) is proposed to determine the QP of the current basic unit.

2.2   Source Model for Transform Video Coder and Its Application-Part I: Fundamental Theory

A source model describing the relationship between bits, distortion, and quantization step sizes of a large class of block-transform video coders is proposed in [15]. This model is initially derived from the rate-distortion theory and then modified to match the practical coders and real image data. The realistic constrains such as quantizer

dead-zone and threshold coefficient selections are included in Hang's formulation. The most attractive feature of this model is its simplicity in its final form. Its enables us to predict the bits needed to encode a picture at a given distortion or to predict the quantization step size at a given bit rate.

The well known rate distortion function of a discrete stationary Gaussian process $\{x(n)\}$ under the mean square distortion criterion is given as:

$$R_\theta = \frac{1}{4\pi} \int_{A1} \log_2 \frac{\Phi(\omega)}{\theta} d\omega$$

and

$$D(R_\theta) = \frac{1}{2\pi} \int_{A1} \theta d\omega + \int_{B1} \Phi(\omega) d\omega \qquad (2\text{-}15)$$

where $\theta \geq 0$

$$\begin{cases} \text{Re } gion \ A_1 : \{\omega \in (-\pi, \pi] \ and \Phi(\omega) \geq \theta\} \\ \text{Re } gion \ B_1 : (-\pi, \pi] - A_1 \end{cases}$$

where $\Phi(\omega)$ is the power spectrum density function of $\{x(n)\}$. That is R_θ is the minimum bit necessary to achieve an average distortion D by an ideal coder of possibly unbounded complexity and time delay. D(R) is the minimum average distortion that can possibly be achieved at bit rate R. In reality, we cannot use infinite length transforms to decompose a signal sequence in to non-overlapped block and perform block transformation on each data block separately. Then the simple and popular spectrum estimation method is the periodogram that computes the spectrum based on the weighted average of the Fourier transforms of non-overlapped data blocks . This data compression if the blocks transform components is the discrete approximation of the ideal continuous power spectrum. Assuming a uniform sampling grid in the frequency domain, the equation (2-15) can be approximated by the following discrete formula:

$$R_\theta = \frac{1}{2L} \sum_{\omega_i \in A_1} \log_2 \frac{\Phi(\omega_i)}{\theta}$$

and

$$D(R_\theta) = \frac{1}{L} \sum_{\omega_i \in A_1} \theta + \sum_{\omega_i \in B_1} \Phi(\omega_I)$$

(2-16)

Hang's shows that a case of interest is that at low distortion when $A_1 = (-\pi, \pi]$ (or B_1 is empty) then $D(R_\theta) = \theta$. Then the equation (2-16) can be rewritten as:

$$R(D) = \frac{1}{\alpha} \ln \frac{E}{D}$$

where   **(2-17)**

$$E = \left[\prod_{i=0}^{L-1} \Phi(\omega_i) \right]^{\frac{1}{L}}$$

where L is the number of samples in a data block. The bits and distortion of a signal are decided by a single parameter E, which is the product of all the components' variances. It represents the complexity of the signal. In theory, two signals of the same ordinary variance require different numbers bits of in coding if their E are significantly different.

2.3   Cauchy Density based Rate Control for H.264

Knowledge of the probability distribution of discrete cosine transform (DCT) coefficient is important in the design and optimization of rate control algorithms. In the early studies [16], the coefficients are conjectured to have Laplacian distribution. In [17], Kamaci *et al.* proposed a better solution using a Cauchy probability density function (pdf) for DCT coefficients estimation. As shown in Fig 2-2, Cauchy model actually outperforms traditional Laplacian model in both intra and inter coded frames.
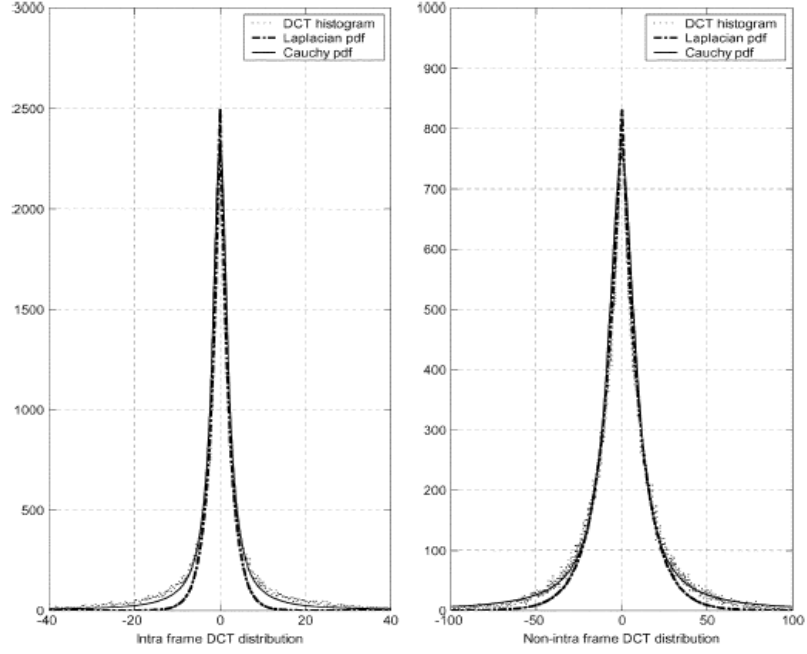
Fig 2-2   Comparison of Laplacian model vs. Cauchy model[17]

Kamaci *et al.* further presented the Cauchy density based rate estimation models by approximating the entropy function of quantization. The rate model was applied in frame layer to determine the QP of each frame based on the given target number of bits of current frame R.

Their Cauchy based rate estimation models is

$$R = a \cdot QS^b \qquad \textbf{(2-18)}$$

where QS is the quantization step; a and b are model parameters which depend on the content of the coding sequence and different types of coding mode, i.e., I-, P-, and B-frames. Then, the QS is determined as following

$$QS = \sqrt[b]{\frac{R}{a}} \qquad \textbf{(2-19)}$$

Finally, the QP used for RDO can be calculated by

$$QP = \left[6 \cdot \log_2(QS) + 4 \right] \qquad \textbf{(2-20)}$$

where [ ] denotes the rounding operation.

2.4   Frame Complexity based Intra only Rate Control

Based on Kamaci *et al.*'s rate estimation model, Jing *et al.*[18] proposed an improved model which is applied on intra frames and has sufficient adaptability to the varying of intra frame complexity.

In their proposed algorithm, they defined the complexity measure of intra frames as the average gradient per pixel of the frame. The calculation of gradient complexity is defined by

$$G = \frac{1}{M \cdot N} \left(\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left| I_{i,j} - I_{i+1,j} \right| + \left| I_{i,j} - I_{i,j+1} \right| \right) \quad \textbf{(2-21)}$$

where M and N are the horizontal and vertical dimensions of the frame, respectively; $I_{i,j}$ denotes the luminance value of the pixel at the location of (i, j).
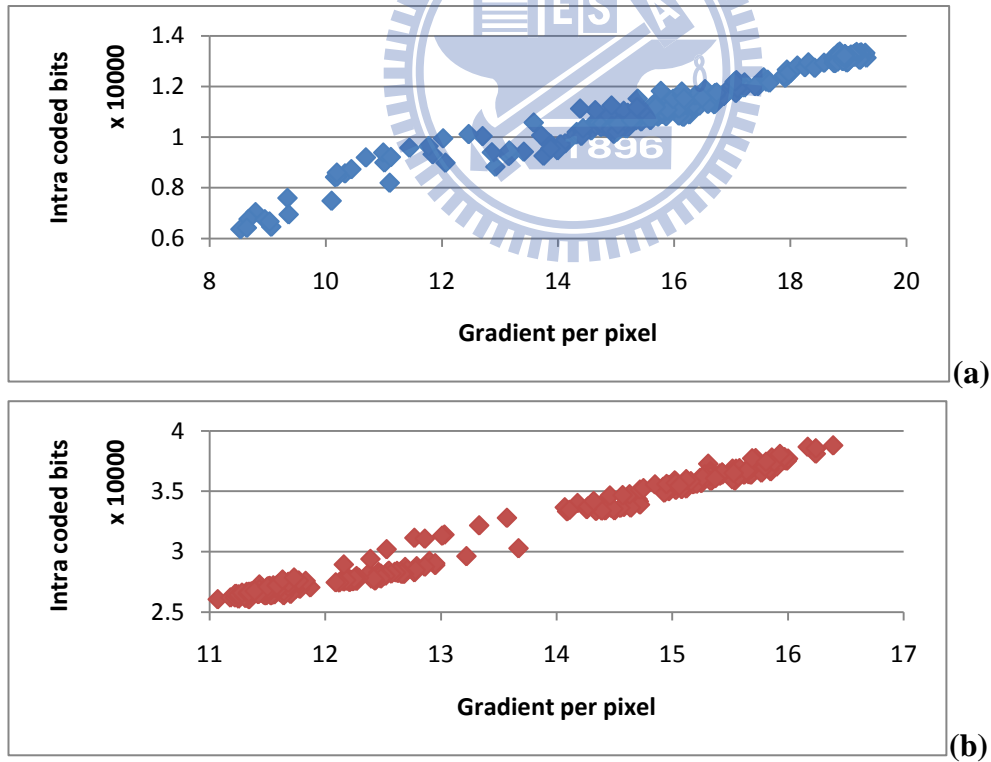
Fig 2-3   Intra coded bits vs. gradient per pixel (a) *Foreman*, **QP=36 (b)** *Carphone*, **QP=25**

They observed that the number of coding bits of an intra coded frame is highly correlated with its gradient value, as shown in Fig 2-3. From the linear correlation

between these two factors, they assumed that for a fixed QP, the output number of bits of one intra frame is proportional to the value of its average gradient per pixel. Based on the assumption, they revised Cauchy rate estimation model as follows

$$R = G \times a \cdot QS^b \qquad \textbf{(2-22)}$$

where b is a constant which is set to -0.8 and a is updated frame by frame as

$$a_k = \begin{cases} \dfrac{R_0}{G_0 \cdot QS_0^b} & k = 0 \\[4mm] \alpha \cdot a_{k-1} + (1-\alpha) \cdot \dfrac{R_{k-1}}{G_{k-1} \cdot QS_{k-1}^b} & otherwise \end{cases} \qquad \textbf{(2-23)}$$

After frame layer bit allocation, QS can be calculated by (2-23), and QP can be derived from (2-22).

2.5   Effective Intra-only Rate Control for H.264/AVC

In Tian's proposed algorithm   [19], they would like to increase the performance of intra-only rate control, so they introduces the geometry gradient information as a new complexity measure to accurately represent the complexity of an intra-frame and to develop a linear rate-complexity model. In their paper, a RD model also proposed.

The Fig 2-4 shows the R-Q relation as follows:
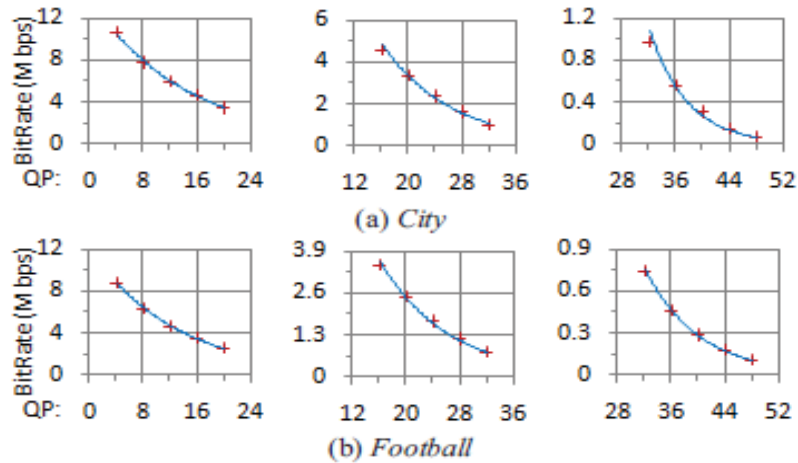
Fig 2-4 R-Q relation

The R-Q model with an exponential form as follows:

$$R = \alpha \cdot e^{-\beta Q}$$ **(2-24)**

The R-C model is also proposed, for intra-frames the motion estimation between consecutive frames is not required, and the output bits only relate to the content complexity of the intra-frame itself. In Tian's paper, they proposed a new complexity measure called geometry gradient, to estimate the content complexity for a picture. The formula as follows:

$$GeoGrad = \sum_{i=0}^{N-2} \sum_{j=0}^{M-2} \frac{\sqrt{\left|P_{i,j} - P_{i,j+1}\right| \cdot \left|P_{i,j} - P_{i+1,j}\right|}}{(N-1)(M-1)}$$ **(2-25)**

where P is luminance value of a pixel, N and M denote the columns and rows of a frame. The Fig 2-5 shows the relationship between R and GeoGrad..
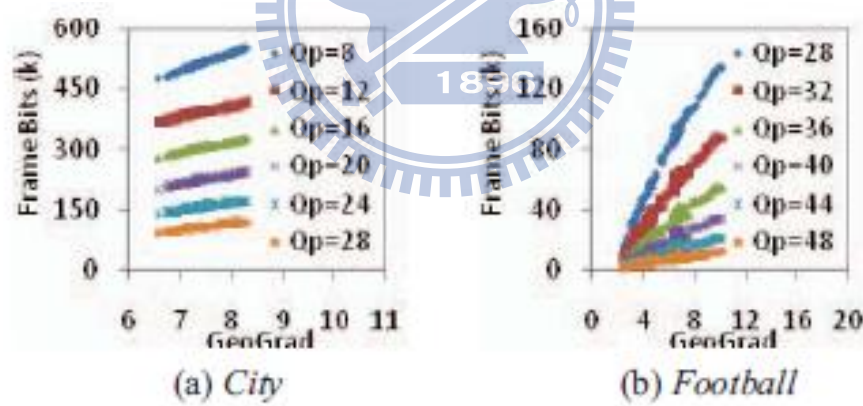
Fig 2-5 relationship between R and GeoGrad

The R-C model formula as follows:

$$R = a \cdot G + b$$ **(2-26)**

The QP calculated by:

$$Q_t = \left\lfloor Q_{t-1} + \frac{\ln \hat{R}_t(G_t, Q_{t-1}) - \ln A_t}{\beta} + 0.5 \right\rfloor$$

where

$$\hat{R}_t(G_t, Q_{t-1}) = R_{t-1} + a \cdot (G_t - G_{t-1})$$

and

$$A_t = \hat{R}_t(G_t, Q_{t-1}) \cdot e^{-\beta \cdot (Q - Q_{t-1})}$$

(2-27)

where the β and a are update by pre-five frames by linear regression as following formula:

$$\beta = -\frac{w \cdot \sum_{i=0}^{w-1} Q_{t-i} \cdot \ln R_{t-i} - (\sum_{i=0}^{w-1} Q_{t-i})(\sum_{i=0}^{w-1} \ln R_{t-i})}{w \cdot \sum_{i=0}^{w-1} Q_{t-i}^2 - (\sum_{i=0}^{w-1} Q_{t-i})^2}$$

and

$$a = \frac{w \cdot \sum_{i=0}^{w-1} G_{t-i} \cdot \tilde{R}_{t-i} - (\sum_{i=0}^{w-1} G_{t-i})(\sum_{i=0}^{w-1} \ln \tilde{R}_{t-i})}{w \cdot \sum_{i=0}^{w-1} G_{t-i}^2 - (\sum_{i=0}^{w-1} G_{t-i})^2}$$

(2-28)

2.6   Summary

In the above sections, we have introduced several researches for H.264 rate control and intra coded frame rate control. However, they still have some problems which can be organized as follows:

A.   Without Dealing with Scene Change Intra Frames

Due to that all MBs within a scene transition frame will be intra coded, we regard such a frame as a special kind of intra frame, called scene change intra frame (SC_I). The locations of SC_I frames and general intra frames in a video sequence can be illustrated by Fig 2-6.
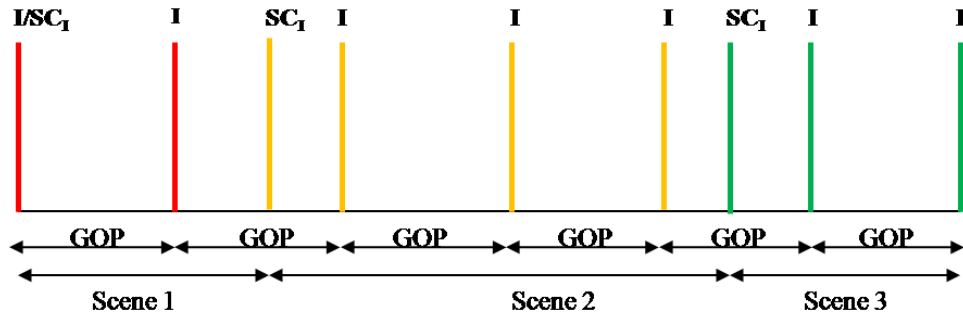
Fig 2-6 The relation between SC_I and general I-frames

Similar to general I-frames, these SC_I can cause serious buffer overflow problem if no appropriate QP is determined for them. Actually the prediction result in their model parameter even won't get a reasonable result, like Fig 1-6 shows the prediction by pre-frame. Although rate control algorithms have been widely studied [11][17], most of them didn't deal with the scene change intra frames.

B. No Accurate Rate Quantization Model for Intra Frames

The quadratic model (2-7) is designed for inter coded frames whose source statistics are assumed satisfying Laplacain distribution. However, this assumption is inappropriate to intra coded frames. Jing *et al.*[18] and Tian *et al.*[19] are proposed a novel rate quantization model for intra frames, but their parameter cannot be estimated precisely. In order to determine this parameter, they employed an update procedure which assumes that its value is stationary frame by frame. However, this assumption is not always true where the parameter value in the figure varies frequently.

Chapter 3 Proposed Rate Control Algorithm for Intra-only Compression

This section presents the proposed rate control algorithm for intra-only compression. We first describe a rate-distortion-QS model for bit-rate prediction, and then a QP determination algorithm is proposed.

3.1   R-D-Q Model

The proposed R-D-Q model is based on Hang's source model [15], formula (3-1), where E is the entropy variance of a signal, D is signal distortion, and α is a model parameter.

$$R(D) = \frac{1}{\alpha} \log \frac{E}{D} \qquad \textbf{(3-1)}$$

When applying to intra-only coding, we replace E with the E' in equation (3-2), which represents the MSE between original pixel and residual value after intra prediction; replace D with the D' in equation (3-3), which represents the MSE between original frame and reconstructed frame; and redefine R(D) as the bit-rate needed to encode this frame, subject to the distortion D':

$$E'' = \sum_{j=0}^{j=H} \sum_{i=0}^{i=W} (f(i, j) - f(\tilde{i}, j))^2 \qquad \textbf{(3-2)}$$

$$D' = \sum_{j=0}^{j=H} \sum_{i=0}^{i=W} \frac{(f(i, j) - f\hat{(i, j)})^2}{H * W} \qquad \textbf{(3-3)}$$

From equations (3-2) and (3-3), it can be seen that both E' and D' depends on the reconstructed frame which is unknown before encoding, and therefore we need a way to estimate E' and D' accurately. And the $f(i, j)$ means original frame pixel, the $\tilde{f}(i, j)$ means the pixel value after intra prediction, and the $\hat{f}(i, j)$ means the reconstruct

frame. The estimates of E' and D' are presented in section 3.1.1 and section 3.1.2, respectively. The model parameter α subject to the re-defined E, D, and R is presented in section 3.1.3.

3.1.1 Estimation of E'

To estimate E', experiments were conducted for various intra-coded sequences and finally we found that E' value is a function of *frame complexity* (G), *frame resolution* (S), and quantization step size (QS). That is, E'= f(G, S, QS). The frame complexity, G, is measured using pixel gradient. Let $I(i, j)$ denotes the luminance value of the pixel at the location of (i, j), the pixel gradient at the location of (i, j) in the nth frame is defined as:

$$g_n(i,j) = \left| I(i,j) - I(i,j-1) \right| + \left| I(i,j) - I(i-1,j) \right| \qquad \textbf{(3-4)}$$

And the frame complexity of nth frame, say G_n, is defined by average gradient per pixel of that frame, which is measured as:

$$G_n = \frac{1}{H \cdot W} \left(\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} g(i,j) \right) \qquad \textbf{(3-5)}$$

where W and H are the horizontal and vertical dimensions of the frame, respectively. The experimental results are shown in Fig 3-1 where E'' as a function QS, ranging from (0.6875 to 208) are plotted; and E'' denotes the 10-based logarithm values of the E' divided by frame complexity G and frame resolution H×W. From Fig 3-1 which shows that E'' is linearly correlated to QS, our E' estimation model is obtained as follows

$$\log_{10}\left(\frac{E'}{H \times W} \times \frac{1}{G_n} \right) = aQS + b \qquad \textbf{(3-6)}$$

where a=-0.0006 and b=1.55*log10(G). From equation (3-6), it is observed that the E' value can be derived from G, H and W for a given QS. It is worth mentioning that since all the G, H and W can be calculated before the frame is encoded, they can be obtained from current coding frame. Namely, the E' of the current frame can be accurately estimated by following formula, without any prediction from the previous

frame.

$$E' = G_n \times W \times H \times 10^{a \times QS + b} \qquad \textbf{(3-7)}$$
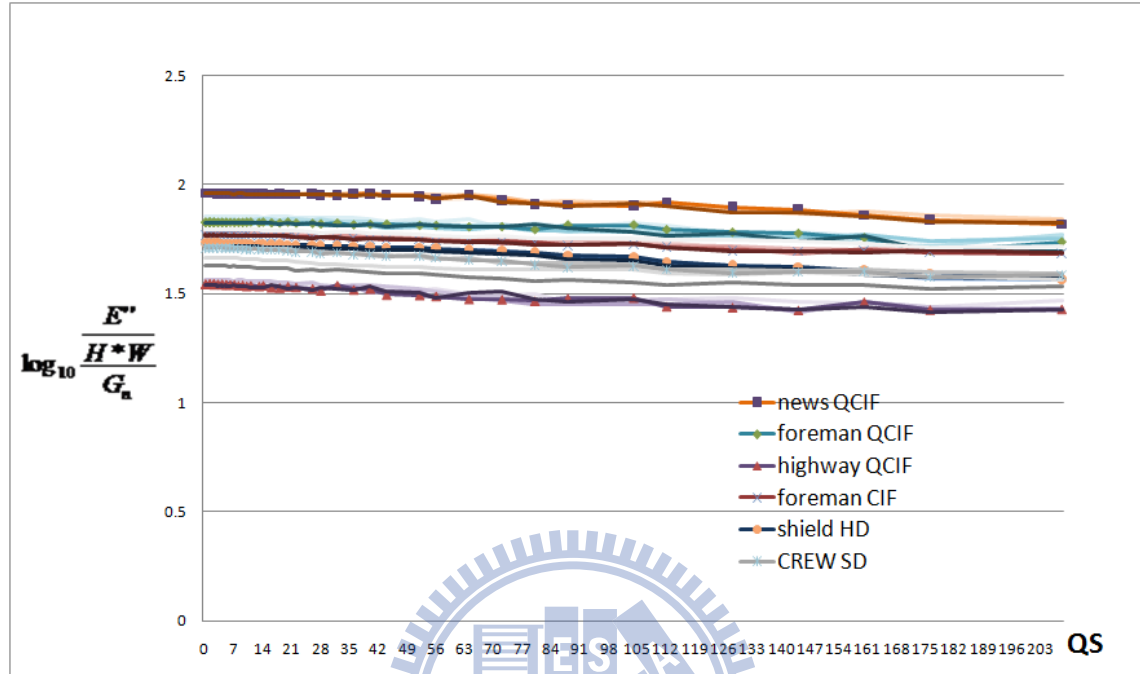
Fig 3-1: the relation of E and QS @news qcif, foreman qcif, highwayqcif, foreman CIF, crew SD, shield HD from frame#2 to frame#6

And the Table 3-1 the correlation of test sequence shows the correlation of test sequence

Test sequence	correlation	Test sequence	Correlation
Foreman QCIF	0.887	Foreman CIF	0.975
News QCIF	0.938	Crew SD(4CIF)	0.955
Highway QCIF	0.918	Shield(HD 720P)	0.971

Table 3-1 the correlation of test sequence

To show the correctness of the proposed E'' estimation model, experiments were conducted for four frames of foreman sequence QCIF and the results were shown in Fig 3-2, where E'' as a function of QS is presented. In Fig 3-2 the average prediction error equals to 0.0552, which is calculated by:

$$\frac{abs(acture\ E\ term - predict\ E\ term)}{acture\ E\ term} \qquad (3\text{-}8)$$

26
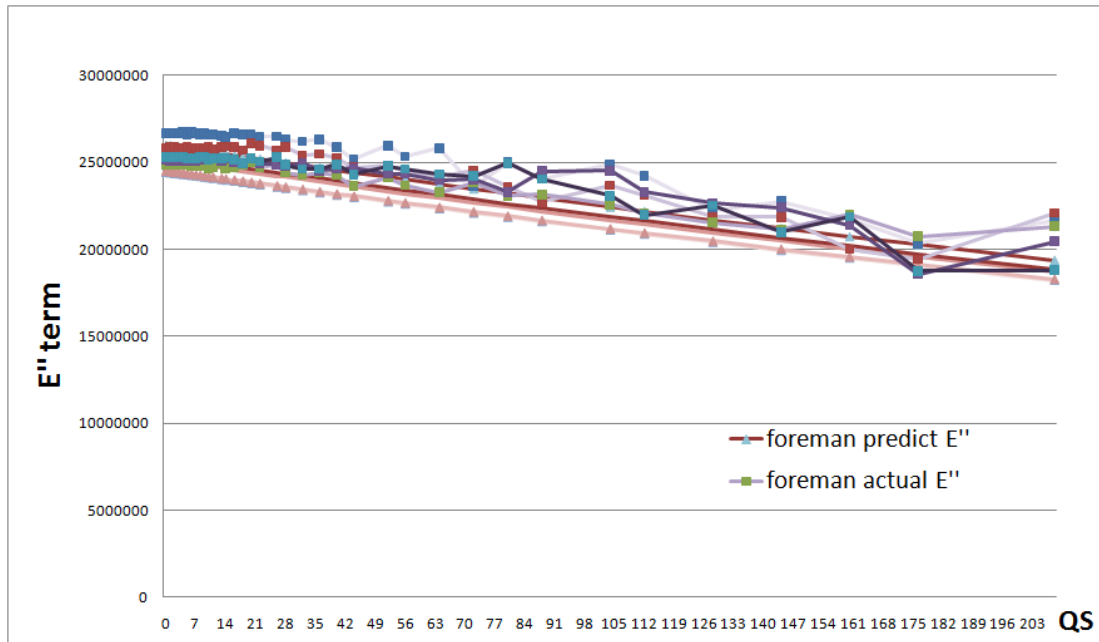
Fig 3-2 : E term prediction error from frame#2 to frame#6

From Fig 3-2, it is observed that E'' predict correctly without any information from previous frame.

3.1.2 Estimation of D

The distortion (D) means the different between original frames and reconstructed frame as described in equation (3-3). To estimate D before encoding, we employ PSNR formula as follow:

$$PSNR = 10 \times \log \frac{255^2}{MSE} \qquad \textbf{(3-9)}$$

where the MSE is the same to the definition of our D. Namely, we can rewrite the equation (3-9) as :

$$D = \frac{255^2}{10^{\frac{PSNR}{10}}} \qquad \textbf{(3-10)}$$

Now we conclude that if we know the PSNR value then we could get the D'. So we have to estimate the PSNR value. To estimate PSNR, experiments were performed for various intra-coded sequences: foreman QCIF, mobile QCIF, NEWs QCIF, foreman

CIF, and shield HD. The results were shown in Fig 3-3 where the PSNR of frame#2 to frame#6 in each sequence are plotted as a function of QS.
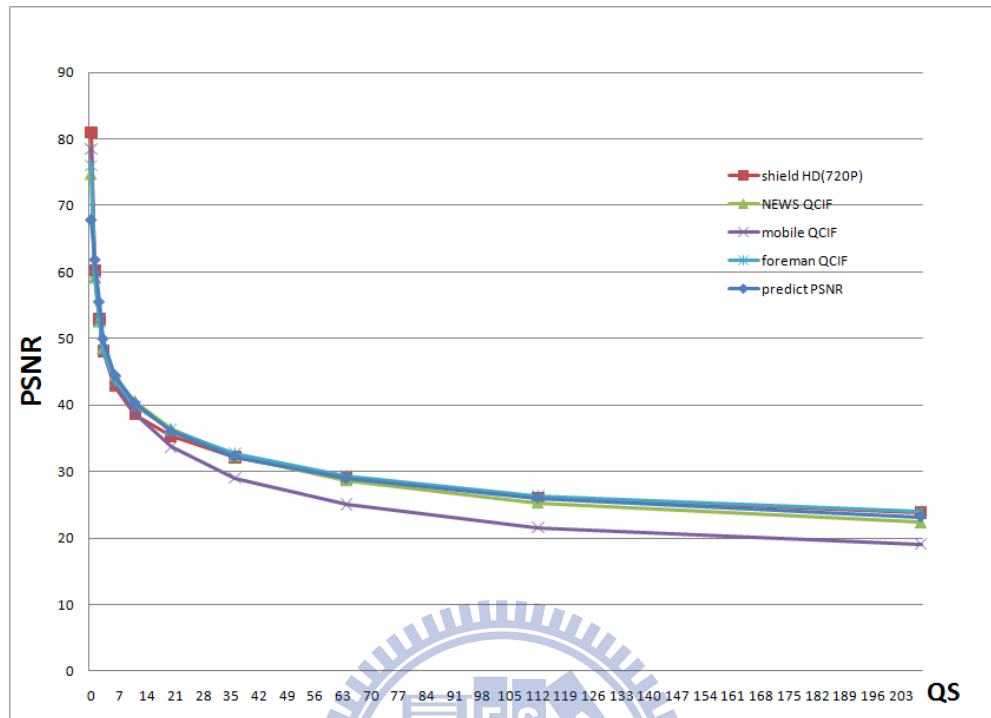
Fig 3-3 the relation of PSNR and QS in different sequences from frame#2 to frame#6

From $^{Fig\ 3-3}$, it can be seen that the curves of the sequences are highly correlated and can be model as:

$$PSNR = cQS^{d} \qquad (3\text{-}11)$$

The correlations between these sequences and equation (3-11) are listed in Table 3-2, where the parameter c is set to 63.168, parameter d set to -0.188.

Test sequence	correlation	Test sequence	Correlation
Foreman QCIF	0.812	Mobile QCIF	0.944
News QCIF	0.874	Shield HD	0.808

Table 3-2 the correlation of PSNR and QS

3.1.3 Estimation of α

In [15], the α parameter should be a constant that equals to 1.386, but the model

28

parameter was experimentally determined by Gaussian signals. In the real case, these kinds of signal source seldom appear, so we have to modify this parameter to fit in our model. In our model, the α no longer remains constant.

To determine α, we conducted experiments for some sequences, get the real value of encoded bit rate, PSNR, and E term to calculate the real α in different test sequences with different resolution. We found that in our model, the α term is related to QP, frame resolution, as well as G.

By normalizing the α with gradient and showing its logarithm (with base 10) as a function of QP, we found that the curves of different sequences are highly correlated as shown in Fig 3-4 and can be modeled as following formula:

$$\log 10(\frac{\alpha}{G_n}) = eQP + f \qquad \textbf{(3-12)}$$
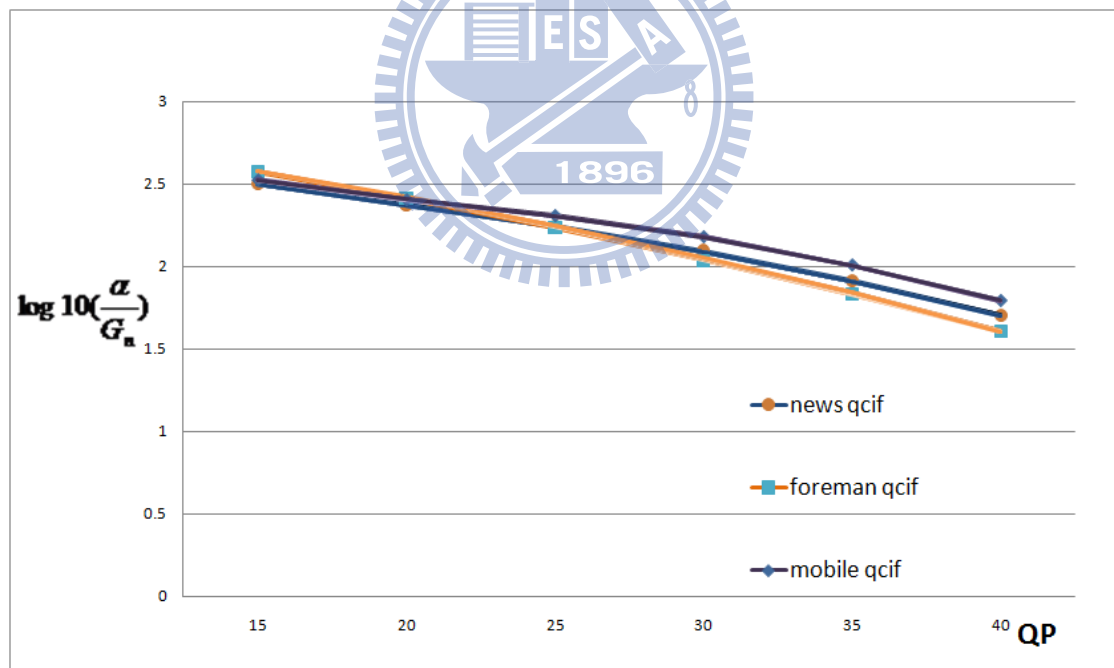
Fig 3-4 linear relation between QP and alpha term after processing from frame#2 to frame #6

where e=-0.0329 and f=3.08, and the correlation of those experiment sequences are shown in Table 3-3:

Test sequence	Correlation
Foreman QCIF	0.995
Mobile QCIF	0.978
News QCIF	0.99

Table 3-3 the correlation between QP and predict alpha term

Fig 3-5 shows the different resolution result:
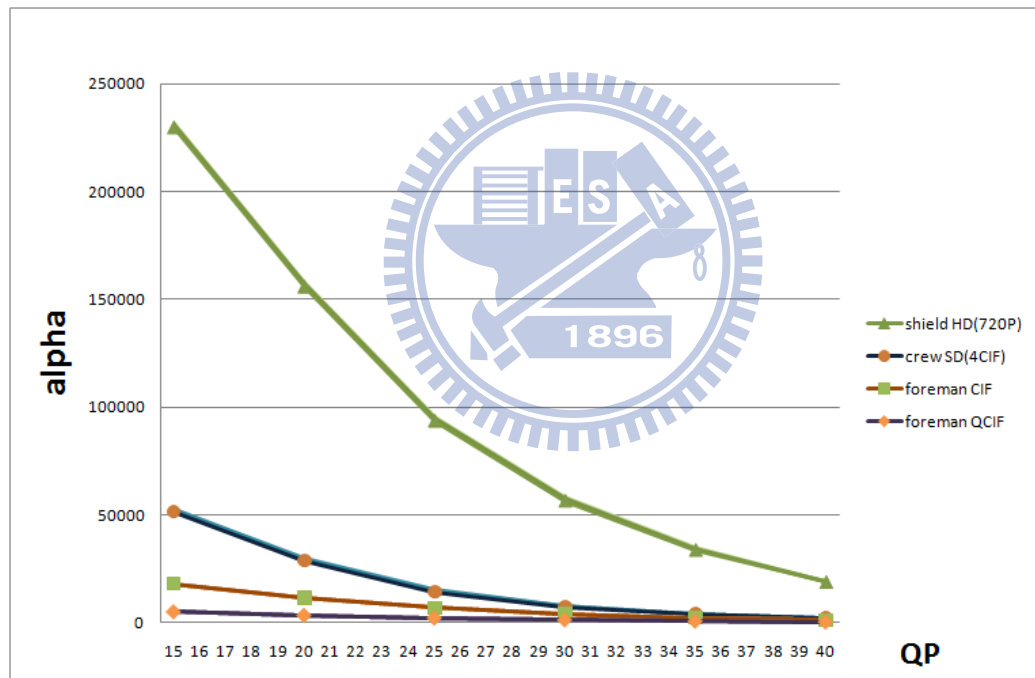
Fig 3-5 different resolution alpha term

So we can predict the α term by the following formula:

$$\alpha = G_n \times 10^{(eQP+f)} \qquad \textbf{(3-13)}$$

By testing with sequences with different resolutions, we found that e and f should be set different for different resolutions which list in Table 3-4.

resolution	QCIF	CIF	SD	HD(720P)
e, f multiply		1.24	1.465	1.66

Table 3-4 different resolution multiply

If there are higher resolutions, this parameter should be updated accordingly, but still can predict the current frame without any pre-frame information. Fig 3-6 shows the comparison of predicted α and experimental α.
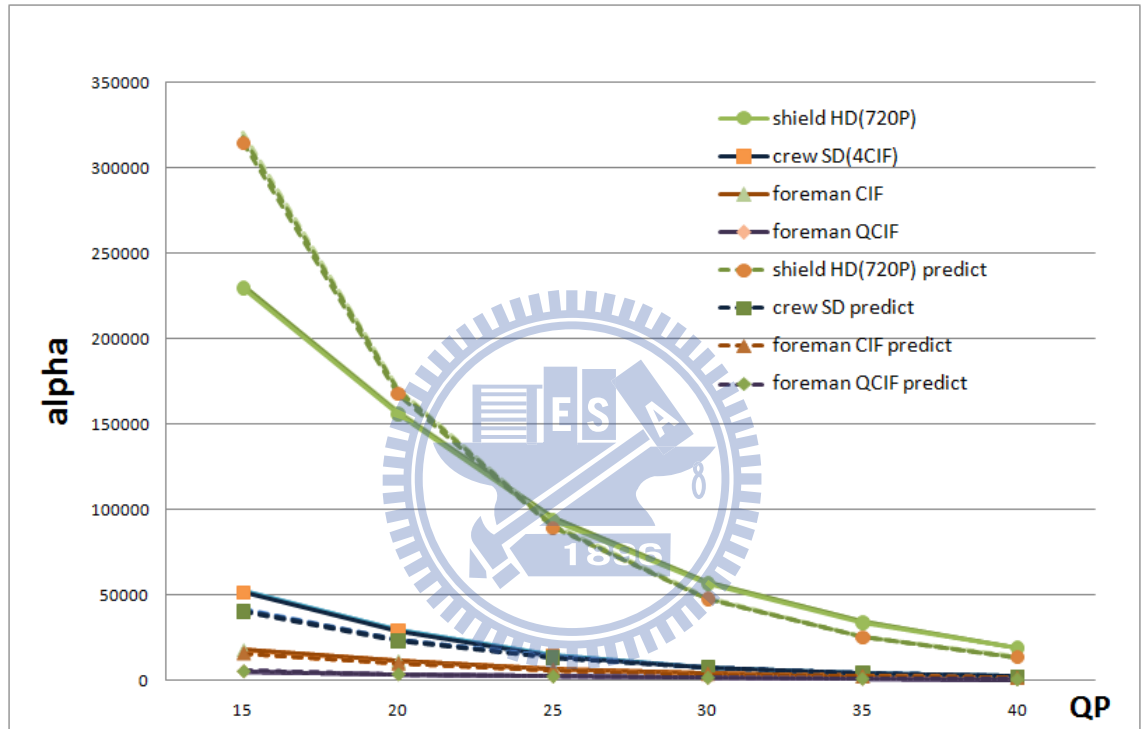
Fig 3-6 shows the predict alpha term and actual alpha term

3.1.4   QP Determination Method for Intra Frames

In this section, we propose a QP determination algorithm for intra frames using the R-D model of equation (3-1). By replacing the E'', D and α in equation (3-1) by equation (3-7), equation (3-10), and equation (3-13), our R-D model becomes:

$$R = \frac{1}{2} \times G_n \times 10^{(eQP+f)} \times \log \frac{G_n \times W \times H \times 10^{a \times QS+b}}{255^2} \qquad \textbf{(3-14)}$$
$$10^{\frac{cQS^d}{10}}$$

It can be seen that the bit rate for a given QP can be derived easily because all the parameters in equation (3-14) can be obtained from current coding frame. To obtain the best QP, we can substitute all possible QPs into R-Q model, equation (3-14) and calculate the prediction bit rate of each QP. The optimized QP is the one associated with the prediction bit rate closest to target bits. In order to take instantaneous scene change constrain into consideration, we propose that QPs within the range of QP from 1 to 51 are used to determine the best one. Fig 3-7 illustrates the proposed concept.
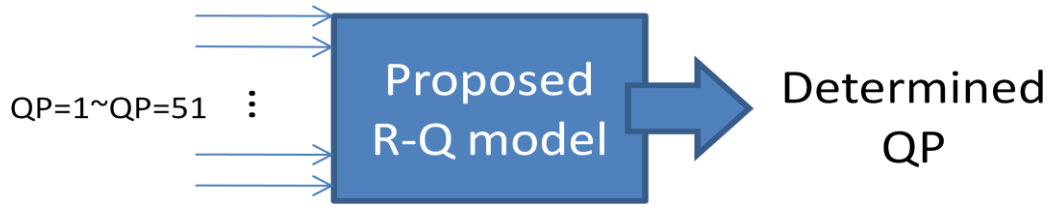
Fig 3-7 diagram of the proposed QP determination model

3.2 Description of the Proposed Rate Control Algorithm for Intra-only Compression

3.2.1 Target-bits Allocation

For Intra-only compression, since all frames are intra coded, there is no need to consider the difference between coding modes. A simple and efficient bit allocation for the current general I-frame is

$$R_t = \frac{R_{remain}}{N_r} \qquad \textbf{(3-15)}$$

where R_{remain} is the available bit budget for remaining frames within the current GOP, and N_r is the number of remaining frames.

3.2.2 QP refinement Algorithm

In this section, we proposed a QP refinement algorithm to prevent the bit rate prediction error from amplifying. In our model, we almost can find the suitable QP in

many cases, but for some extreme case accidently cause our prediction incorrectly, furthermore we don't need to update our model parameter. These kinds of case would be influence the following coding frames, we use actual encoded bits and predicted bits of previous frame to modify the QP determination result of current frame. We defined a measure formula by:

$$predict\_error = \frac{prediction\,Bits - actualBits}{actualBits}$$

The algorithm is shown as the following:

$$if\,(predict\_error) > 0.2$$
$$QP_R = QP_M + 2$$
$$else\,\,if\,(0.05 < predict\_error < 0.2)$$
$$QP_R = QP_M + 1$$
$$else\,\,if\,(-0.05 > predict\_error > -0.2)$$
$$QP_R = QP_M - 1$$
$$else\,\,if\,(predict\_error) < -0.2$$
$$QP_R = QP_M - 2$$
$$else$$
$$QP_R = QP_M$$

where QP_R is QP after refined, and QP_M is QP after our model.

3.2.3   Proposed Rate Control Algorithm

With bit allocation for intra frames, and QP determination algorithms for intra frames, the detailed block diagram of the proposed rate control algorithm for Intra-only
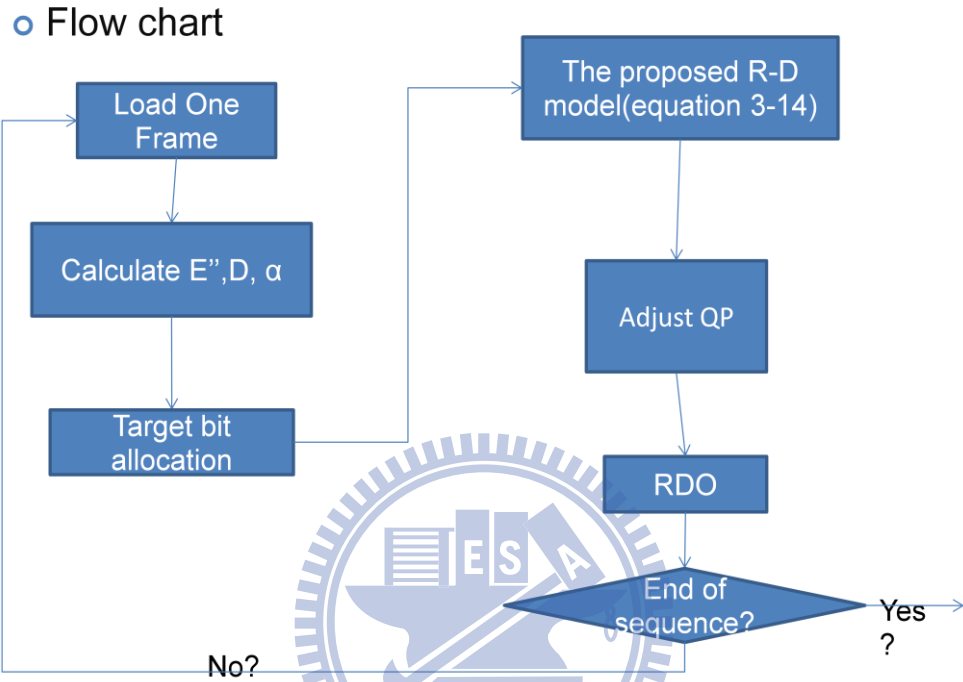
Fig 3-8 Flow charts for Intra-only compression

compression is shown in Fig 3-8. We summarize it with the following six steps:

Step 1.   Calculate the E'',D ,and α, according to equations (3-7), (3-10), and (3-13)

Step 2.   The intra frame bit allocation is calculated based on (3-15)

Step 3.   Calculate the predicted bit rate for the given QP using the proposed R-D-Q model equation (3-14)

Step 4.   Find QP meet the minimum difference predict bits and target bits.

Step 5.   QP adjusts.

Step 6.   Go to Step 1 until the end of the sequence.

3.2.4 Proposed Rate Control Algorithm Verification

To show the correctness of the proposed model, In this section, we set a test to verification our model. This test method is that choice a frame here is the #5, then let all the frame QP before frame #5 fix into X and the following frame #6 QP fix into frame #5 X+4. Then QP from X+4 to X+1. The X from 5 to 47 Collecting all the bits then shows the predict error

Fig 3-9 to Fig 3-12 shows the above experiment verification at QP from X+4 to X+1 at football qcif:
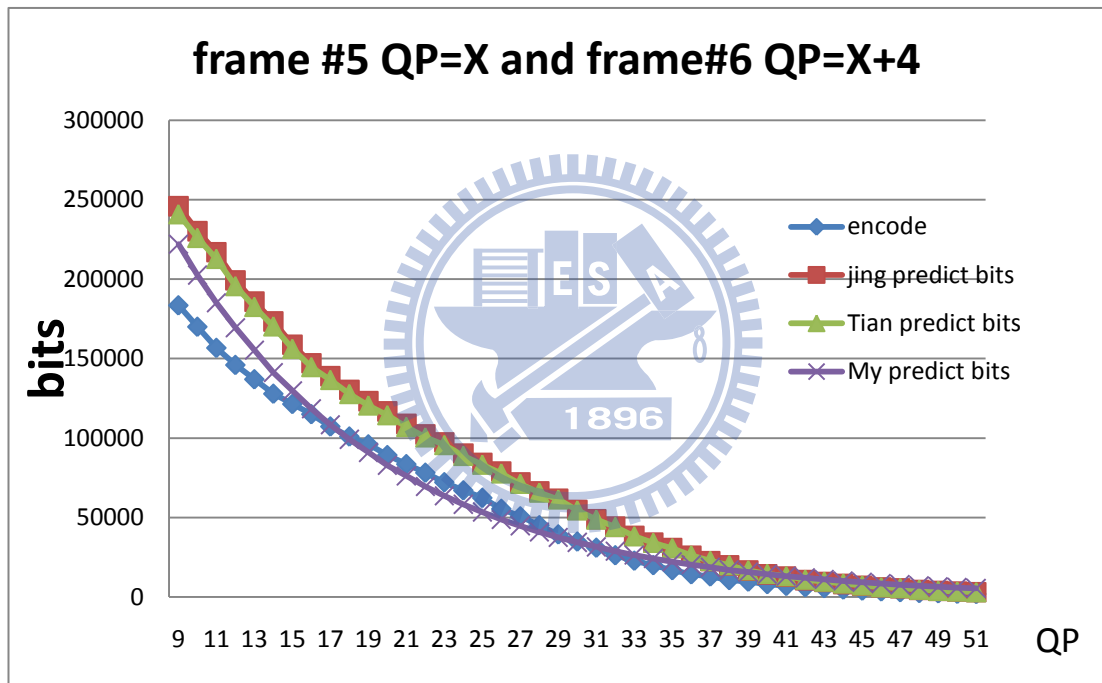
Fig 3-9   football QCIF model verification at QP+4
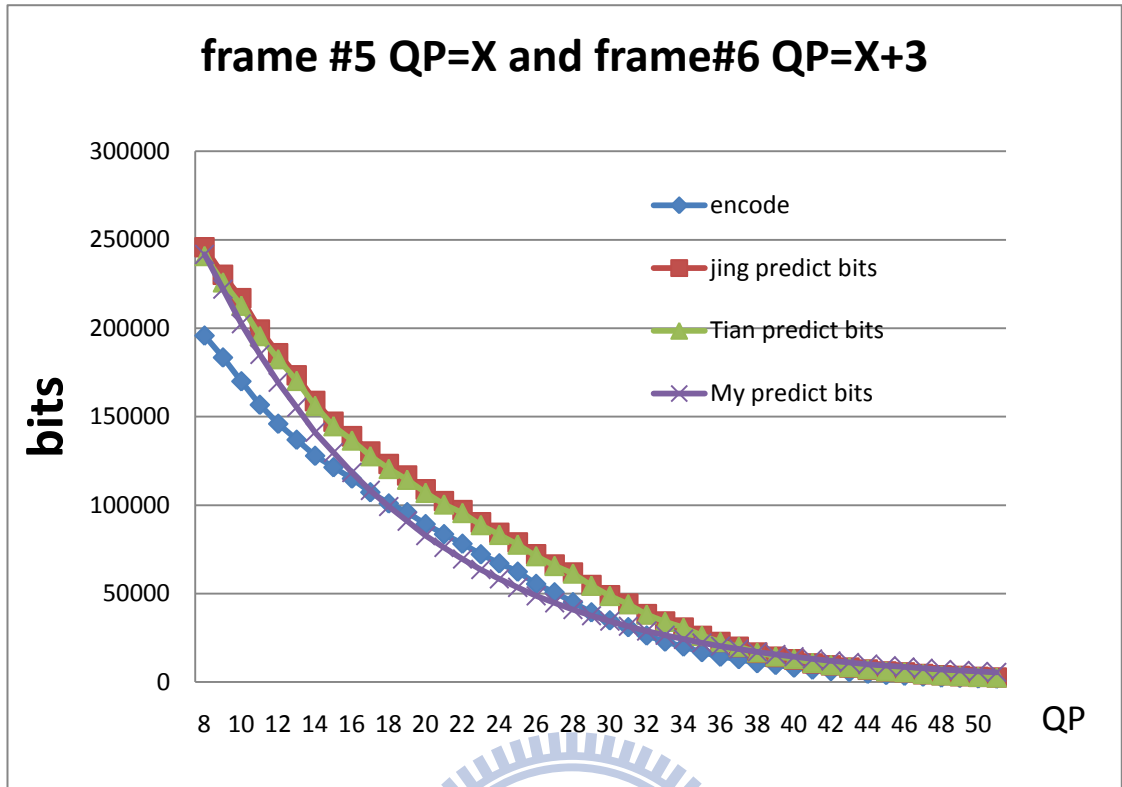
Fig 3-10 football QCIF model verification at QP+3
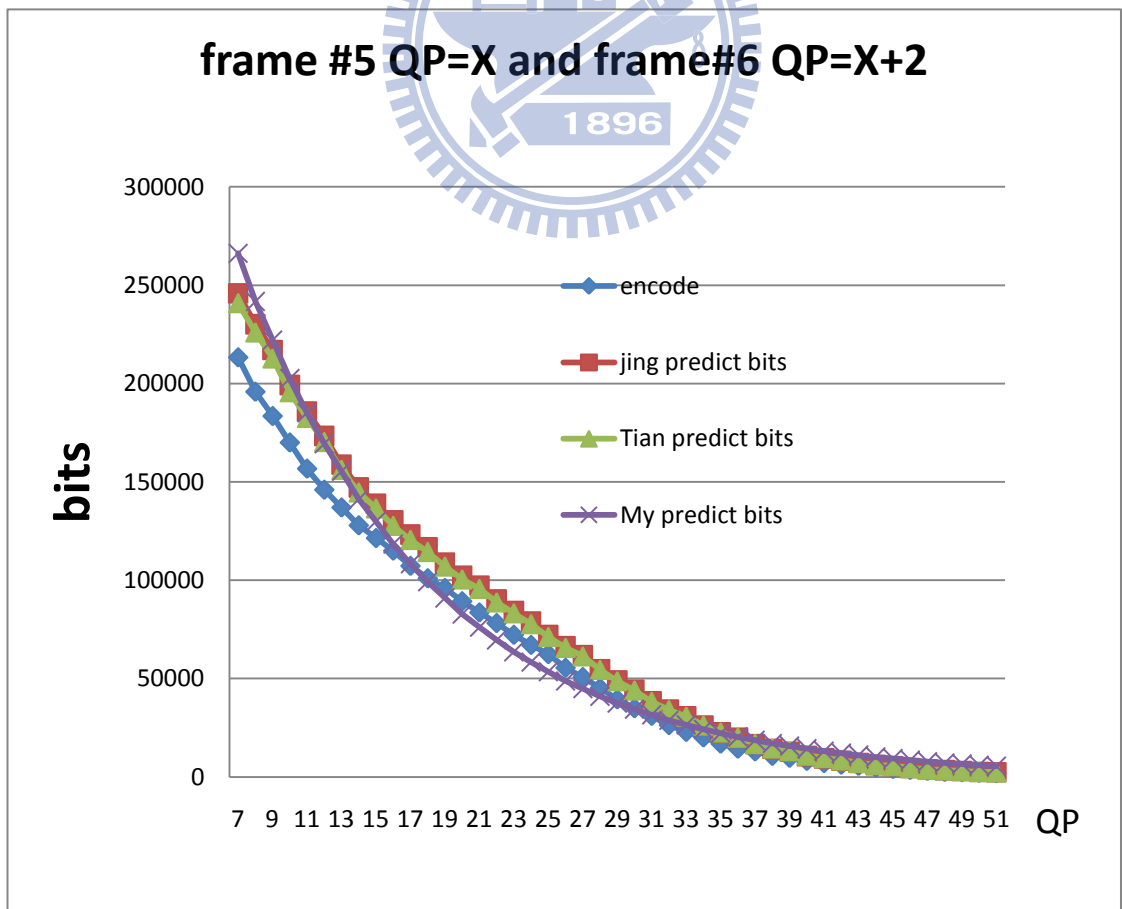
Fig 3-11 football QCIF model verification at QP+2
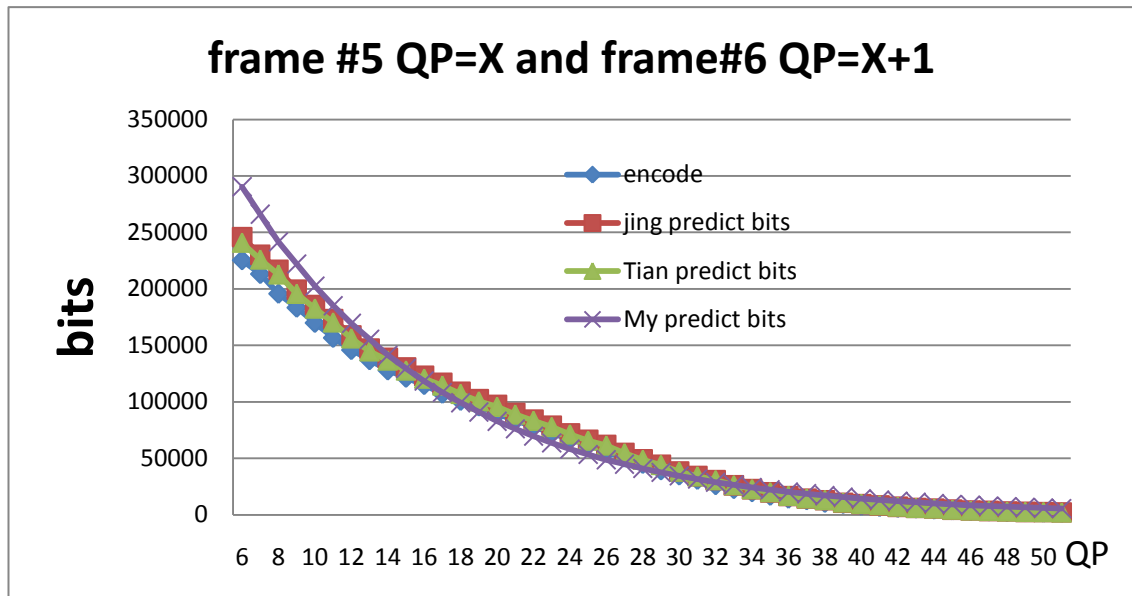
Fig 3-12 football QCIF model verification at QP1

The Fig 3-9 to Fig 3-12 shows us if the pre-frame QP differ larger, the predict bits by pre-frame information would be imprecise. And the Fig 3-13 to Fig 3-16 shows us if the pre-frame QP plus larger, the predict bits by pre-frame information would be imprecise, too. In our model, it is no need to consider the pre-frame information to decide current frame QP.

The other part is the #5, and then let all the frame QP before frame #5 fix into X and the following frame #6 QP fix into frame #5 X-4. Then QP from X-4 to X-1. The X from 1 to 47 Collecting all the bits then shows the predict error:
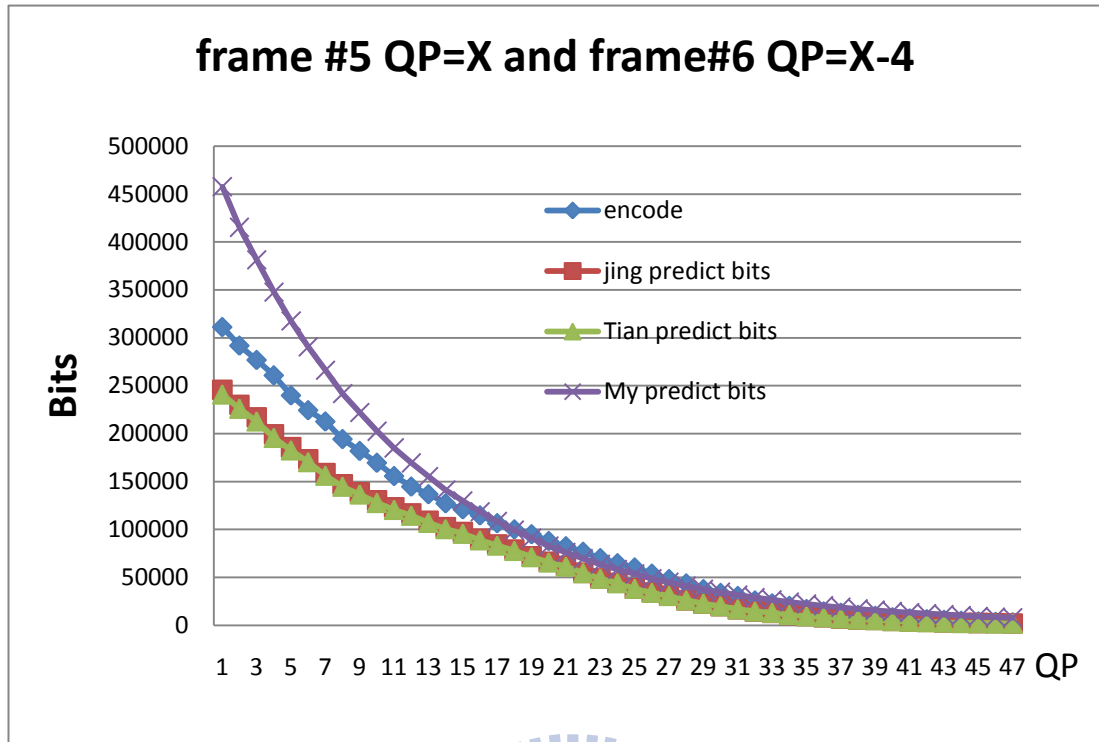
Fig 3-13 football QCIF model verification at QP-4
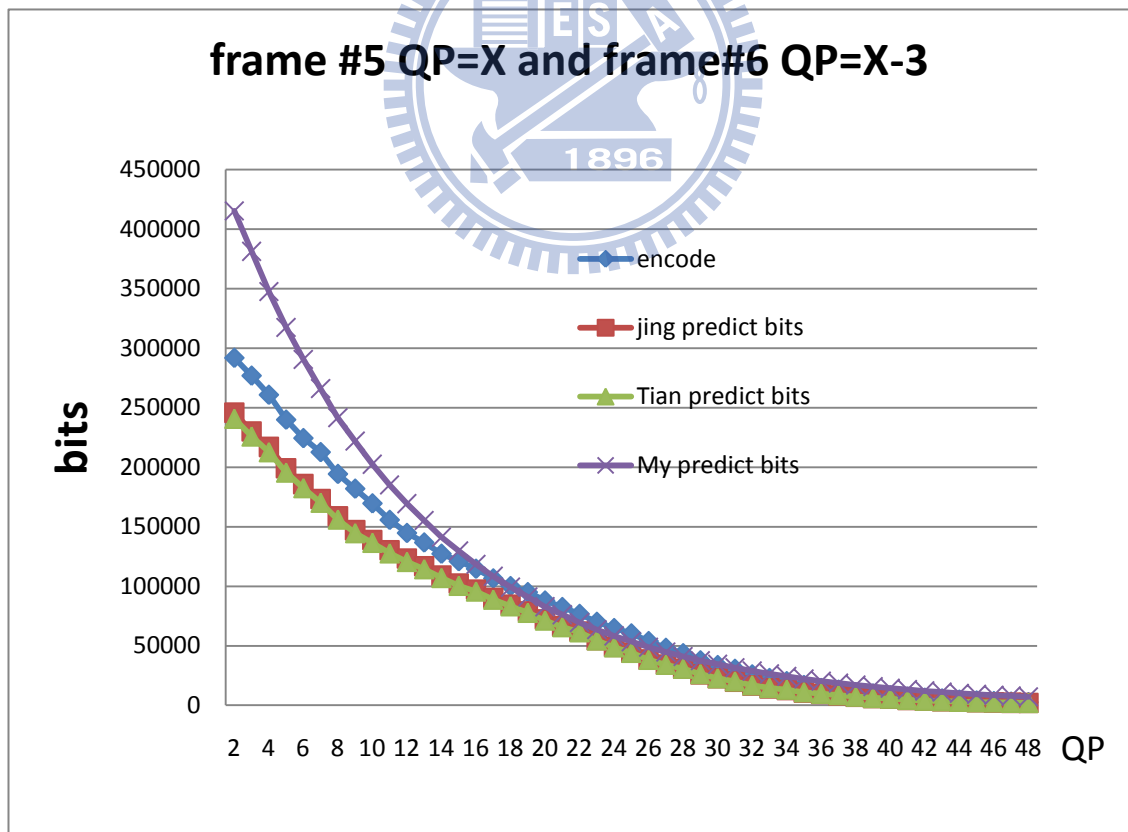
Fig 3-14 football QCIF model verification at QP - 3
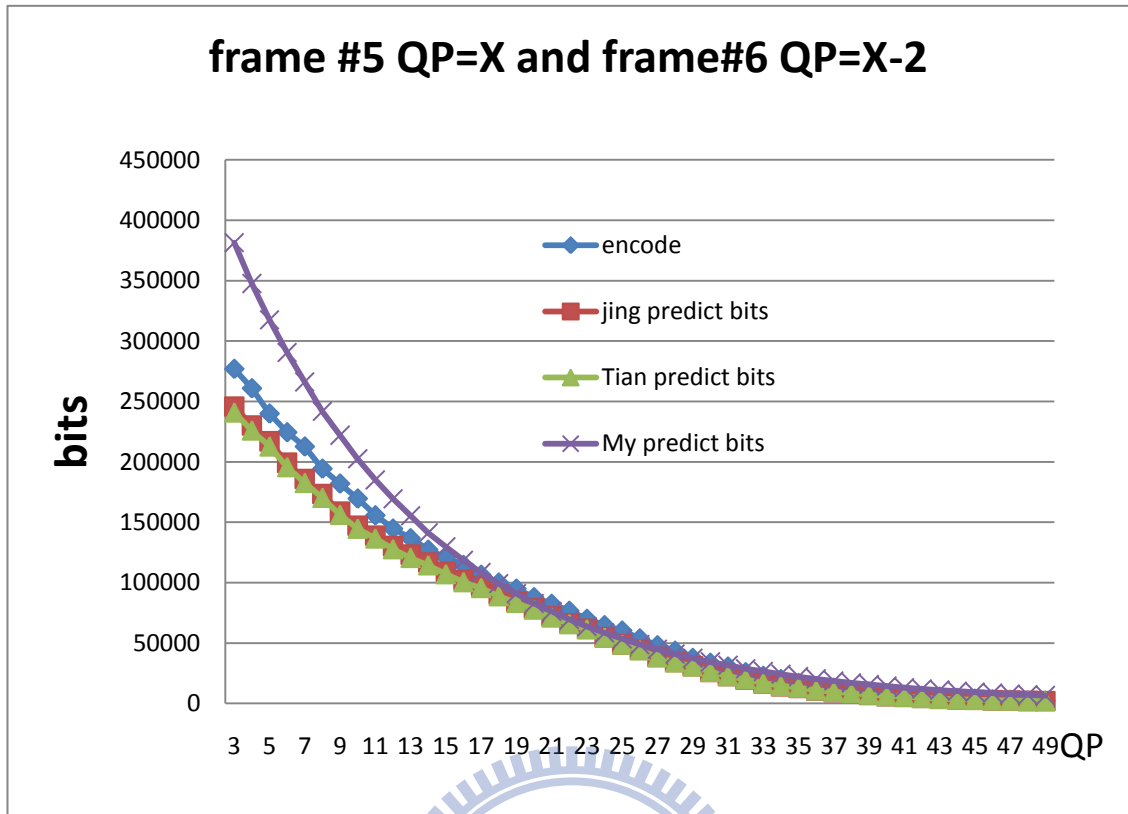
Fig 3-15 football QCIF model verification at QP-2
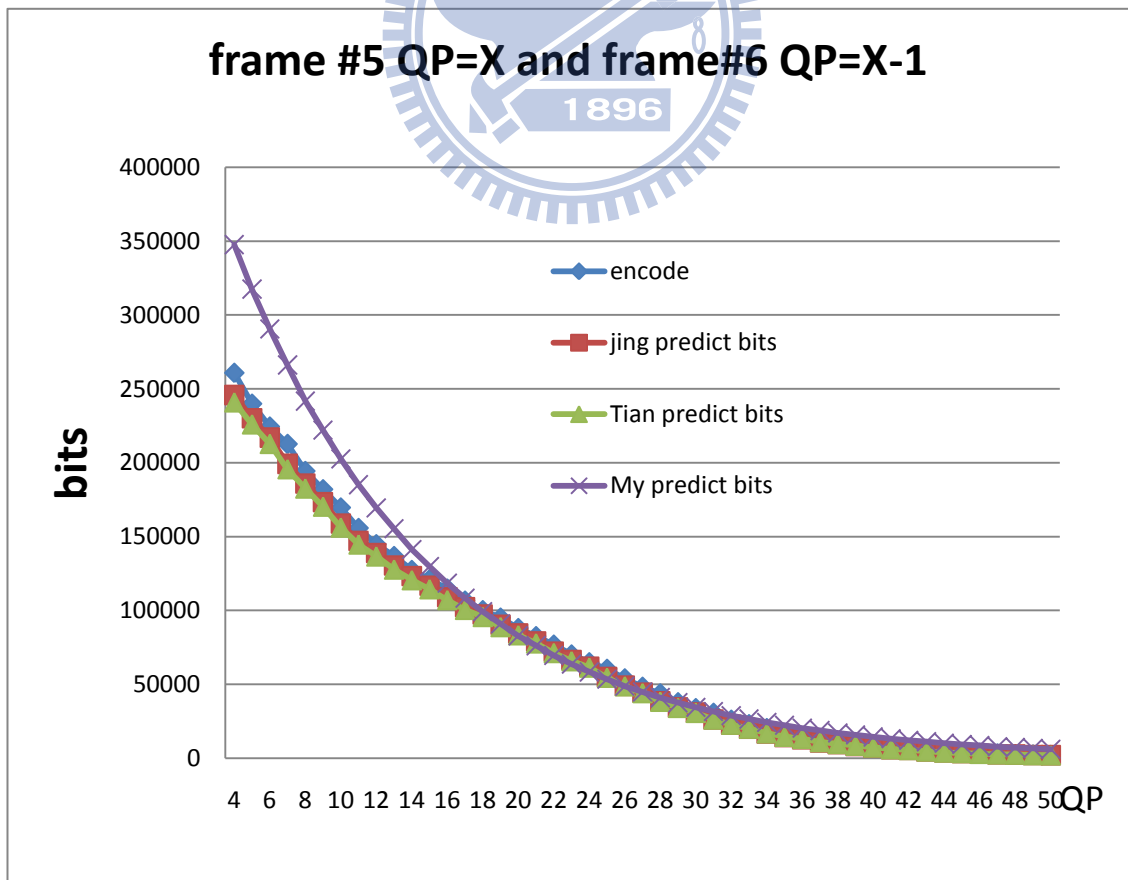
Fig 3-16 football QCIF model verification at QP-1

Chapter 4　　Experiment Results

The proposed rate control algorithm is integrated into the latest JVT reference software JM15.0[12]. The simulation test sequences was conducted with following :

Test sequence	Frame No.	Frame start No.	Test sequence	Frame No.	Frame start No.
Football(Q)	125	1	Combo_cif_2	200	1
Akiyo(Q)	200	1	Crew(S)	200	1
Highway(Q)	200	1	City(S)	200	1
Coastguard(Q)	200	1	Harbor(S)	200	1
Combo1(Q)	200	1	Combo_SD	200	1
Combo2(Q)	100	1	Mobcal(H)	200	1
Coastguard(C)	200	1	Parkrun(H)	200	1
Tennis(C)	150	1	Shield(H)	200	1
Stefan(C)	200	1	Vidyo(H)	200	1
Silent(C)	200	1	Combo_HD	200	1
Football(C)	125	1			
Combo_cif_1	150	1			

Table 4-1 test sequences

In addition, in order to test the proposed algorithm under scene change condition, two scene change sequences "Combo1" (Trevor-Stefan-Silent-Coastguard) and "Combo2" (Akiyo-Mobile), were created by cascading corresponding sequences, and the intervals of every two consecutive scene cuts are 50 frames long. Two scene change sequences "combo_cif_1" (football_news_bus_container_city) were created by cascading corresponding sequences, and the intervals of every two consecutive scene cuts are 30 frames long. And "combo_cif_2" (foreman_ mobile_ news_coastguard) are

intervals by 50 frames. One scene change sequence "combo_SD" (ice_crew_harhour_city) were created by cascading corresponding sequences, and the intervals of every two consecutive scene cuts are 50 frames long. One scene change sequence "combo_HD" (mobcal_parkrun_stockholm_shields) were created by cascading corresponding sequences, and the intervals of every two consecutive scene cuts are 50 frames long.

In Intra-only compression, we compare the proposed algorithm with Jing's method [18], Tian's method [19] and JM15.0 Intra-only rate control algorithm which is a modified version based on G012[11]. In our compression schemes, CAVLC and RDO are enabled, and the size of basic unit is set to 99(QCIF), 396(CIF), 1584(SD), 3600(720P). All parameters are selected equivalently for all algorithms.

4.1   Results of Intra-only Compression

In Intra-only compression, Table 4-2 summarizes the overall performance results including actual bit rate, average PSNR, and PSNR deviation at QCIF. The proposed algorithm is cable of increasing average PSNR by up to 1 dB (0.51 dB on average) and 0.91 dB (0.27 dB on average) and 0.92 dB (0.33 dB on average) compared to JM and Jing's algorithm and Tian's algorithm, respectively. In addition, PSNR deviation is reduced by up to 79% (37% on average), 76% (33% on average) and 86% (37% on average) in contrast with JM and Jing's algorithm, and Tian's algorithm, respectively. Although, the mismatches of real bit rate and target bit rate among three methods are close, the proposed algorithm slightly reduce the mismatch compared to other three schemes.

QCIF	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
512kbps	football	512.48	512.51	512.48	510.91	**511.46**	26.18	26.27	26.27	27.17	**27.18**
	akiyo	511.66	511.91	511.95	512.79	**512.17**	38.29	38.74	38.75	39.04	**39.03**
	highway	512.06	512.04	512.09	512.44	**512.40**	40.13	40.47	40.47	40.73	**40.73**
	caustguard	512.02	512.06	511.93	511.19	**511.51**	31.49	31.71	31.71	31.98	**31.98**
	Average	**512.06**	**512.13**	**512.11**	**511.83**	**511.89**	**34.02**	**34.30**	**34.30**	**34.73**	**34.73**
1024kbps	football	1023.63	1023.47	1024.03	1023.48	1024.44	30.91	30.92	30.93	31.07	31.08
	akiyo	1022.53	1023.91	1023.99	1025.64	1024.37	44.69	45.02	45.02	45.08	45.08
	highway	1024.43	1024.05	1023.95	1025.40	1025.17	43.67	43.78	43.78	43.79	43.78
	caustguard	1023.39	1024.36	1024.14	1023.27	1023.75	36.39	36.59	36.59	36.65	36.66
	Average	**1023.50**	**1023.95**	**1024.03**	**1024.45**	**1024.43**	**38.92**	**39.08**	**39.08**	**39.15**	**39.15**

QCIF	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
512kbps	Combo1	512.33	512.20	512.03	511.01	**511.51**	31.13	31.41	31.36	31.77	**31.78**
	Combo2	512.11	511.90	512.27	511.42	**511.41**	30.51	30.88	30.53	31.45	**31.45**
1024kbps	Combo1	1024.20	1023.96	1023.90	1023.03	1023.49	36.60	36.79	36.72	36.90	36.90
	Combo2	1023.49	1023.57	1023.68	1024.42	1024.91	36.37	36.68	36.36	36.82	36.83

(a)

QCIF	Test seq.	PSNR StDev

		JM	Jing	Tian	propose	propose + QP adjust
	football	4.95	4.91	4.91	1.16	**1.17**
	akiyo	2.80	2.08	2.07	0.54	**0.56**
512kbps	highway	1.68	1.63	1.62	0.46	**0.48**
	caustguard	2.97	2.88	2.88	0.88	**1.07**
	Average	**3.10**	**2.87**	**2.87**	**0.76**	**0.82**
	football	3.86	3.84	3.83	0.97	1.00
	akiyo	1.13	0.89	0.90	0.40	0.35
1024kbps	highway	1.06	1.01	1.01	0.60	0.57
	caustguard	2.19	2.10	2.12	0.77	0.81
	Average	**2.06**	**1.96**	**1.97**	**0.68**	**0.68**

QCIF	Test seq.	PSNR StDev				
		JM	Jing	Tian	propose	propose + QP adjust
512kbps	Combo1	5.68	4.32	4.52	3.30	**3.89**
	Combo2	**8.18**	8.11	8.97	7.80	**7.71**
1024kbps	Combo1	3.95	3.64	3.97	3.20	3.13
	Combo2	**8.67**	8.63	9.65	8.41	8.31

(b)

Table 4-2 QCIF Preference Result (a) Bit rate and Average PSNR (b) PSNR standard derivation

Table 4-3 summarizes the overall performance results including actual bit rate, average PSNR, and PSNR deviation at CIF. The proposed algorithm is cable of increasing average PSNR by up to 0.62 dB (0.42 dB on average) and 0.16 dB (0.09 dB on average) and 0.33 dB (0.12 dB on average) compared to JM and Jing's algorithm and Tian's algorithm, respectively. In addition, PSNR deviation is reduced by up to 69% (43% on average), 67% (47% on average) and 67% (45% on average) in contrast with

JM and Jing's algorithm, and Tian's algorithm, respectively.

CIF	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
1024kbps	caustguard	1024.31	1024.42	1024.44	1022.44	1023.78	29.31	29.82	29.82	29.89	29.89
	tennis	1025.93	1023.99	1025.23	1023.06	1023.37	30.57	30.68	30.70	30.86	30.83
	stefan	1024.20	1024.13	1024.34	1024.57	1023.56	26.71	27.13	27.12	27.26	27.25
	silent	1024.03	1024.01	1024.40	1023.59	1023.48	32.02	32.37	32.39	32.42	32.42
	football	1024.25	1023.98	1024.63	1023.73	1023.99	27.38	27.61	27.62	27.77	27.75
	Average	**1024.54**	**1024.11**	**1024.61**	**1023.48**	**1023.64**	**29.20**	**29.52**	**29.53**	**29.64**	**29.63**
2048kbps	caustguard	2047.73	2047.60	2047.77	2046.37	2047.50	32.88	33.35	33.35	33.39	33.39
	tennis	2049.54	2047.93	2052.72	2045.39	2047.39	34.05	34.19	34.20	34.26	34.26
	stefan	2047.80	2048.61	2048.16	2050.02	2048.62	31.54	31.77	31.77	31.86	31.85
	silent	2045.65	2048.06	2048.19	2049.94	2048.35	35.31	35.57	35.58	35.63	35.62
	football	2040.75	2048.58	2048.83	2050.38	2047.72	30.70	30.89	30.90	30.98	30.98
	Average	**2046.29**	**2048.16**	**2049.13**	**2048.42**	**2047.92**	**32.90**	**33.15**	**33.16**	**33.22**	**33.22**
CIF	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
1024kbps	combo_cif_1	1024.45	1023.56	1024.20	1022.37	1024.22	29.90	30.28	30.27	30.46	30.44
	Combo_cif_3	1024.13	1023.69	1024.55	1022.19	1023.42	29.56	30.12	29.85	30.18	30.18
2048kbps	combo_cif_1	2048.49	2047.63	2048.86	2048.89	2046.90	33.88	34.25	34.15	34.36	34.36
	Combo_cif_3	2048.19	2047.76	2048.38	2045.40	2048.30	33.66	34.12	33.89	34.18	34.18

(a)

CIF	Test seq.	PSNR StDev				
		JM	Jing	Tian	propose	propose + QP adjust
1024kbps	caustguard	1.23	1.36	1.32	0.92	0.92
	tennis	1.89	1.87	1.86	1.56	1.64
	stefan	1.92	1.81	1.80	0.79	0.59
	silent	0.83	0.87	0.80	0.30	0.44
	football	1.85	1.82	1.82	1.04	1.10
	Average	**1.54**	**1.55**	**1.52**	**0.92**	**0.94**
2048kbps	caustguard	1.28	1.41	1.40	0.93	0.93
	tennis	2.25	2.12	2.12	1.67	1.80
	stefan	2.01	1.78	1.77	0.95	0.93
	silent	1.07	0.93	0.92	0.35	0.36
	football	2.00	1.91	1.90	0.91	0.89
	Average	**1.72**	**1.63**	**1.62**	**0.96**	**0.98**
CIF	Test seq.	PSNR StDev				
		JM	Jing	Tian	propose	propose + QP adjust
1024kbps	combo_cif_1	2.71	2.60	2.64	2.43	2.41
	Combo_cif_3	5.49	5.30	5.96	5.19	5.16
2048kbps	combo_cif_1	3.52	3.24	3.54	3.00	3.00
	Combo_cif_3	5.80	5.58	6.17	5.41	5.40

(b)

Table 4-3 CIF Preference Result (a) Bit rate and Average PSNR (b) PSNR standard derivation

Table 4-4 summarizes the overall performance results including actual bit rate, average PSNR, and PSNR deviation at SD. The proposed algorithm is cable of increasing average PSNR by up to 0.86 dB (0.54 dB on average) and 0.05 dB (0.01 dB on average) and 0.09 dB (0.02 dB on average) compared to JM and Jing's algorithm and Tian's algorithm, respectively. In addition, PSNR deviation is reduced by up to 45%

(11% on average), 37% (6% on average), and 40% (8% on average) in contrast with JM , Jing's algorithm, and Tian's algorithm respectively.

SD(4CIF)	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
4096 kbps	crew	4090.35	4095.51	4098.83	4104.29	4098.73	37.40	37.96	37.96	38.00	38.00
	city	4095.04	4094.70	4096.95	4100.59	4095.99	31.22	31.74	31.74	31.75	31.74
	ice	4085.42	4096.00	4095.22	4100.11	4096.94	42.61	42.71	42.71	42.76	42.75
	HARBOUR	4097.75	4096.24	4096.88	4098.64	4095.79	30.75	31.61	31.61	31.61	31.61
	Average	4092.14	4095.61	4096.97	4100.91	4096.86	35.50	36.01	36.01	36.03	36.03
8192 kbps	crew	8195.16	8192.92	8192.38	8209.11	8200.40	40.56	40.98	40.97	40.99	40.99
	HARBOUR	8190.84	8191.30	8192.89	8194.87	8190.41	35.13	35.95	35.95	35.97	35.96
	city	8190.44	8191.59	8190.57	8206.68	8194.26	35.08	35.64	35.63	35.66	35.65
	ice	8178.85	8193.76	8194.81	8197.60	8192.93	45.27	45.57	45.57	45.57	45.57
	Average	8188.82	8192.39	8192.66	8202.07	8194.50	39.01	39.54	39.53	39.55	39.54

SD(4CIF)	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
4096kbps	Combo	4094.50	4094.88	4095.79	4101.01	4094.56	35.46	35.99	35.95	36.05	36.04
8192kbps	Combo	8181.64	8192.83	8193.18	8204.21	8191.57	39.01	39.59	39.53	39.60	39.60

(a)

			PSNR StDev			
SD(4CIF)	Test seq.	JM	Jing	Tian	propose	propose + QP adjust
4096 kbps	crew	1.24	1.18	1.24	1.12	1.11
	city	0.73	0.75	0.73	0.68	0.72
	ice	1.18	0.96	0.92	0.62	0.65
	HARBOUR	0.30	0.36	0.43	0.39	0.36
	Average	**0.86**	**0.81**	**0.83**	**0.70**	**0.71**
8192 kbps	crew	1.06	1.04	1.04	1.01	0.98
	HARBOUR	0.56	0.51	0.54	0.32	0.32
	city	1.02	0.89	0.92	0.69	0.72
	ice	0.72	0.70	0.74	0.59	0.59
	Average	**0.84**	**0.79**	**0.81**	**0.65**	**0.65**
			PSNR StDev			
SD(4CIF)	Test seq.	JM	Jing	Tian	propose	propose + QP adjust
4096kbps	Combo	5.12	4.84	4.98	4.86	4.84
8192kbps	Combo	**4.58**	4.27	4.43	4.30	4.27

(b)

Table 4-4 SD Preference Result (a) Bit rate and Average PSNR (b) PSNR standard derivation

Table 4-5 summarizes the overall performance results including actual bit rate, average PSNR, and PSNR deviation at HD(720P). The proposed algorithm is cable of increasing average PSNR by up to 0.49 dB (0.33 dB on average) and 0.06 dB (0.01 dB on average) and 0.09 dB (0.02 dB on average) compared to JM and Jing's algorithm and

Tian's algorithm, respectively. In addition, PSNR deviation is reduced by up to 68% (15% on average), 66% (13% on average), and 65% (13% on average) in contrast with JM , Jing's algorithm, and Tian's algorithm respectively

HD(720P)	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
10240 kbps	mobcal	10238.77	10239.13	10239.39	10244.71	10238.50	28.73	29.14	29.14	29.15	29.15
	parkrun	10245.69	10241.13	10236.36	10249.44	10238.00	25.36	25.69	25.69	25.71	25.70
	stockholm	10241.61	10242.54	10240.34	10241.88	10235.62	33.11	33.40	33.41	33.41	33.41
	vidyo	10234.16	10242.05	10242.10	10228.70	10236.17	42.52	42.71	42.71	42.73	42.74
	Average	**10240.06**	**10241.21**	**10239.55**	**10241.18**	**10237.07**	**32.43**	**32.74**	**32.74**	**32.75**	**32.75**
20480 kbps	mobcal	20479.96	20478.44	20482.71	20486.74	20469.59	32.59	32.92	32.92	32.96	32.95
	parkrun	20498.12	20482.87	20482.17	20510.66	20490.38	28.95	29.27	29.27	29.33	29.32
	stockholm	20481.22	20482.10	20480.98	20481.21	20481.21	36.12	36.28	36.28	36.28	36.28
	vidyo	20481.31	20484.54	20477.44	20427.77	20466.99	45.75	45.95	45.95	45.95	45.95
	Average	**20485.15**	**20481.99**	**20480.83**	**20476.60**	**20477.04**	**35.85**	**36.11**	**36.11**	**36.13**	**36.13**
HD(720P)	Test seq.	Bit Rate					Avg. PSNR (db)				
		JM	Jing	Tian	propose	propose + QP adjust	JM	Jing	Tian	propose	propose + QP adjust
10240kbps	combo	10239.82	10236.41	10238.66	10256.63	10236.58	29.12	29.53	29.47	29.55	29.54
20480kbps	combo	20467.08	20480.43	20478.98	20501.15	20488.15	32.61	32.94	32.88	32.98	32.97

(a)

HD(720P)	Test seq.	PSNR StDev				
		JM	Jing	Tian	propose	propose + QP adjust
10240 kbps	mobcal	0.85	0.87	0.87	0.84	0.86
	parkrun	1.03	1.00	0.96	0.86	0.87
	stockholm	0.31	0.35	0.31	0.28	0.32
	vidyo	0.78	0.77	0.77	0.38	0.41
	Average	**0.74**	**0.75**	**0.73**	**0.59**	**0.62**
20480 kbps	mobcal	1.20	1.19	1.18	0.97	0.99
	parkrun	1.59	1.42	1.46	0.89	0.98
	stockholm	0.31	0.29	0.28	0.20	0.20
	vidyo	**0.54**	0.56	0.49	0.30	0.30
	Average	**0.91**	**0.86**	**0.85**	**0.59**	**0.62**

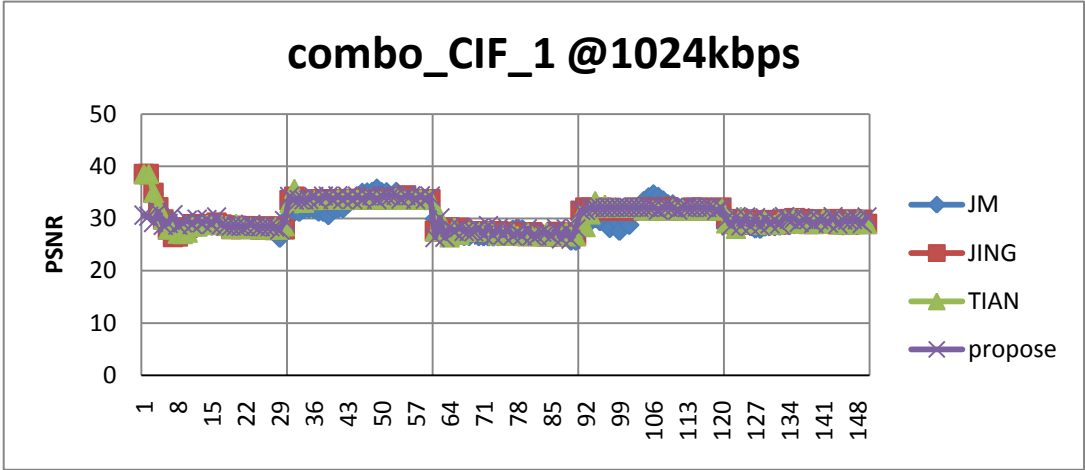| HD(720P)	Test seq.	PSNR StDev				
		JM	Jing	Tian	propose	propose + QP adjust
10240kbps	combo	3.29	3.24	3.26	3.22	3.24
20480kbps	combo	**3.25**	3.10	3.18	2.99	3.04

(b)

Table 4-5 HD Preference Result (a) Bit rate and Average PSNR (b) PSNR standard derivation
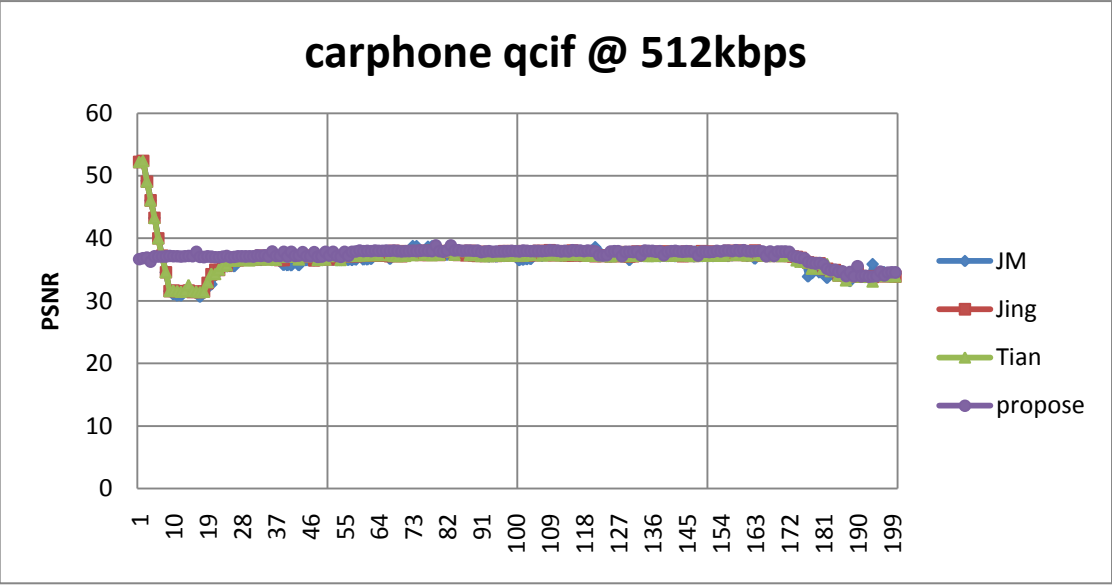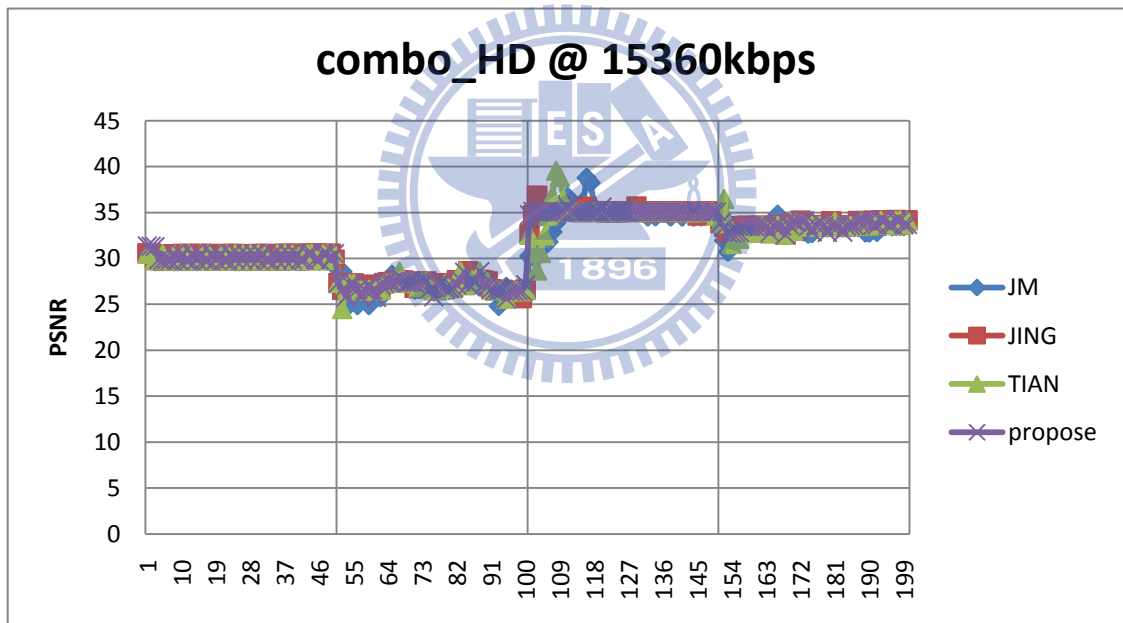
(a)
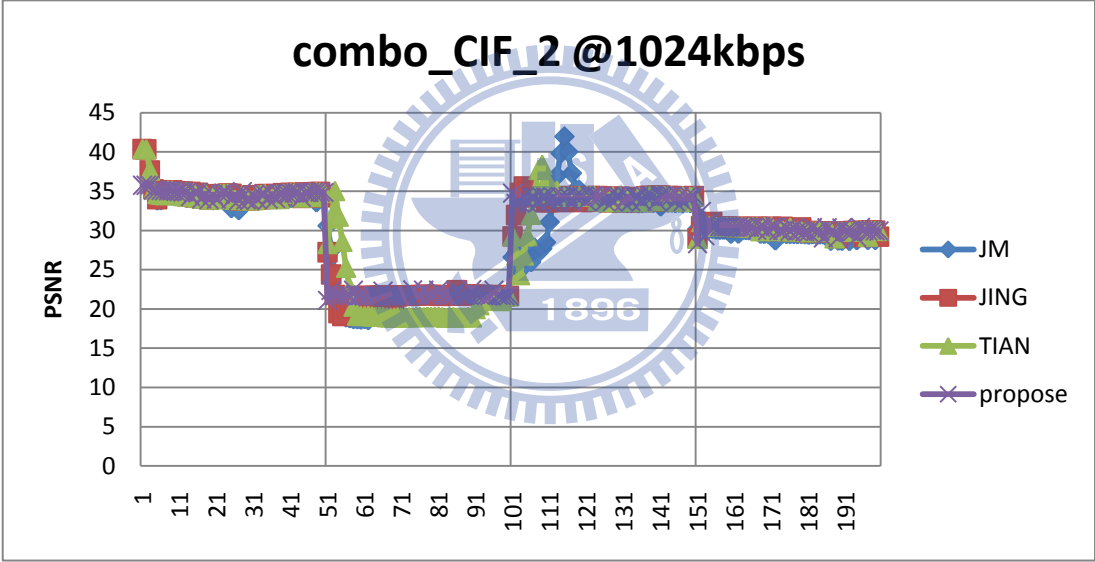

(b)

(c)
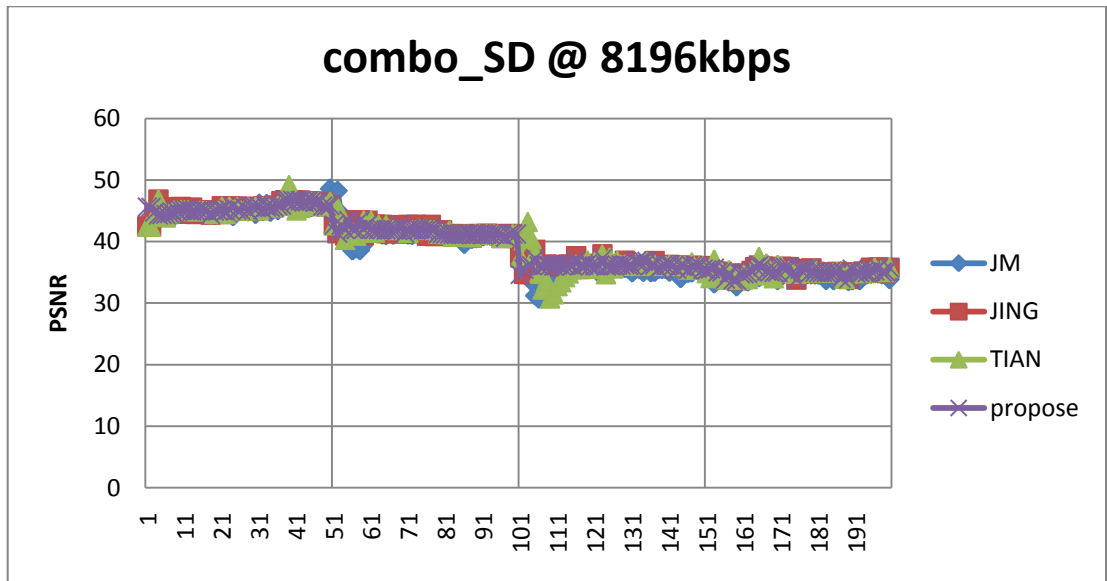
(d)

(e)

**Fig 4-1 PSNR v.s. frames (a) carphone QCIF (b) combo_CIF_3 (c) combo_CIF_1 (d)
combo_SD (e) combo_HD**

For further evaluation, the curves of PSNR versus frames for five test cases are
shown in Fig 4-1. From the plot (a), it is observed that the proposed algorithm can
maintain a consistent video quality in contrast with other three algorithms which
consume too much bits for the first frame so that the quality of the succeeding frames is

decreased and unstable. Plot (b) shows that the proposed algorithm properly deal with the scene change frame (50th frame) in CIF, so the quality of the following frames is more stable and higher than JM and Jing's and Tian's algorithm. Plot (c) shows that the proposed algorithm properly deal with the scene change frame (30th frame) in CIF, so the quality of the following frames is more stable and higher than JM and Jing's and Tian's algorithm. Plot (d) shows that the proposed algorithm properly deal with the scene change frame (30th frame) in SD, so the quality of the following frames is more stable and higher than JM and Jing's and Tian's algorithm. Plot (e) shows that the proposed algorithm properly deal with the scene change frame (30th frame) in HD, so the quality of the following frames is more stable and higher than JM and Jing's and Tian's algorithm. Fig 4-2 shows the buffer occupancy versus frames for five test cases. The proposed algorithm shows superior performance by achieving consistent buffer fullness at a very low level. The reason is that, with our approach, the amount of generated bits of each frame are closely equivalent to the instantaneous channel bit rate. Hence, the buffer fullness is kept at a stable and low level which means the proposed scheme can achieve small buffer delay while real-time transmits and successfully avoid buffer overflow.
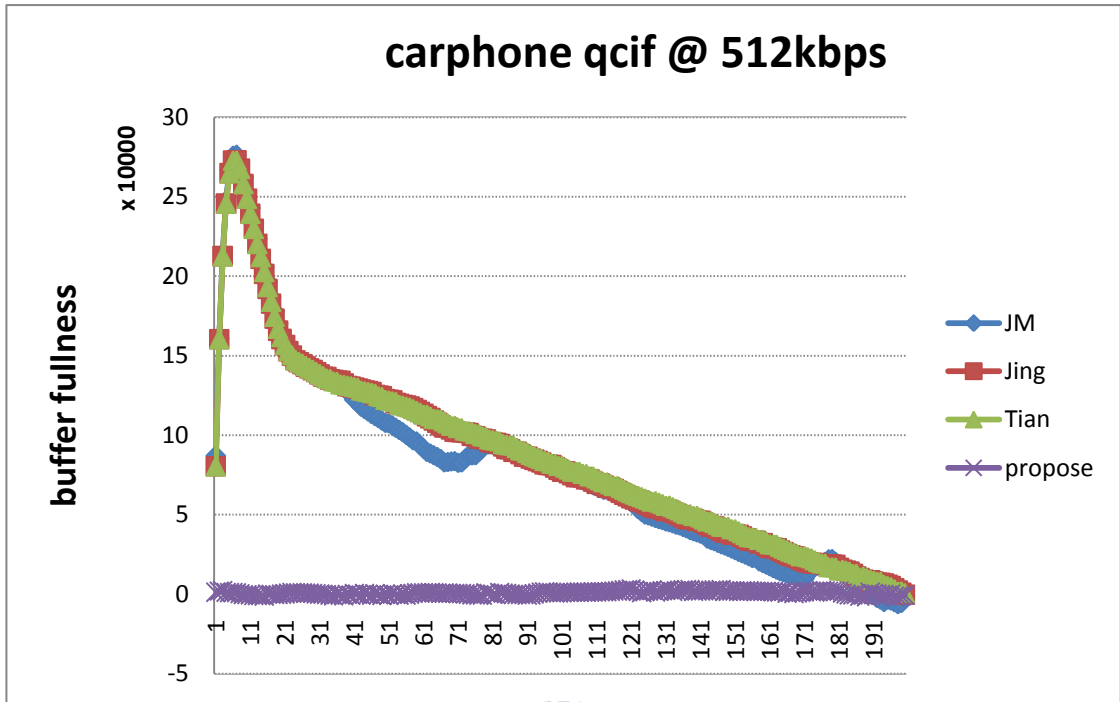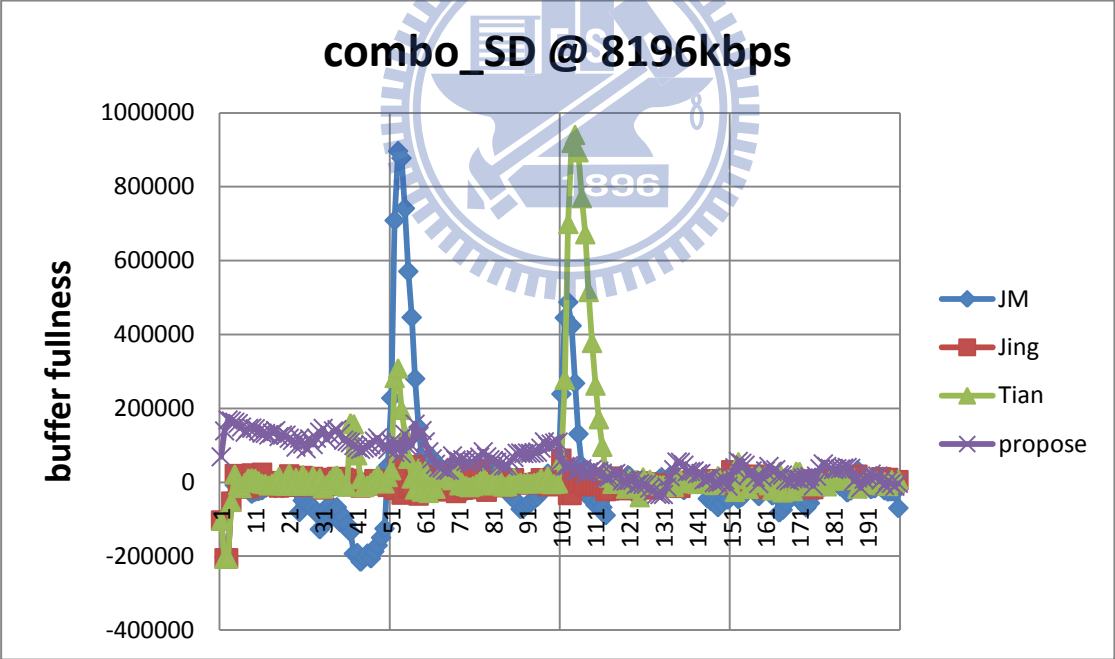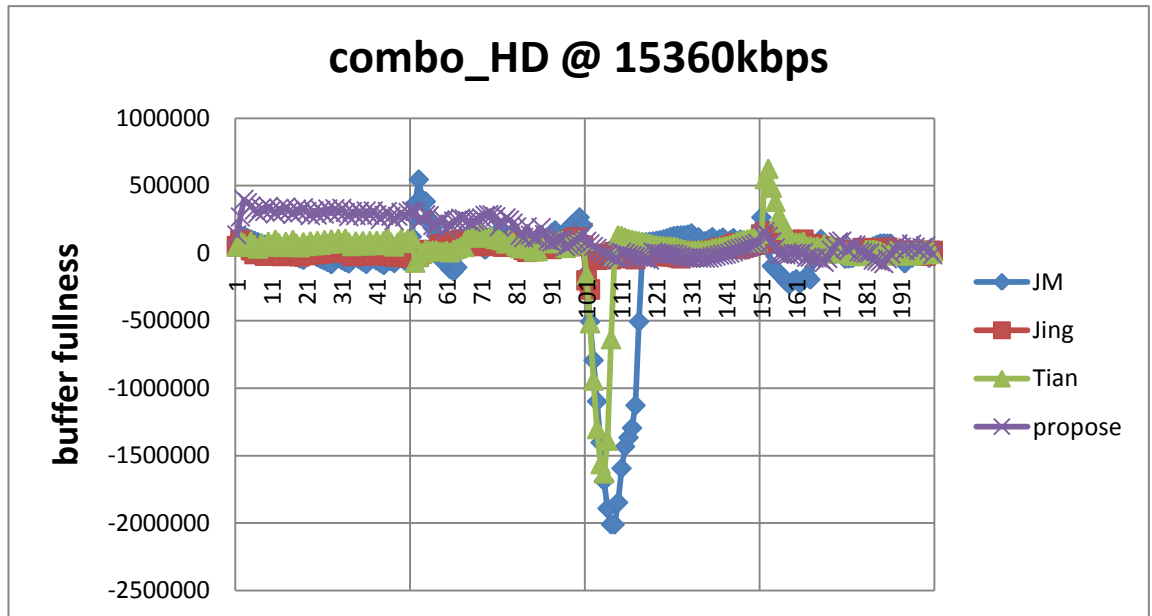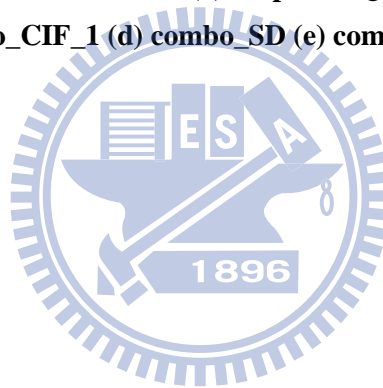
(a)
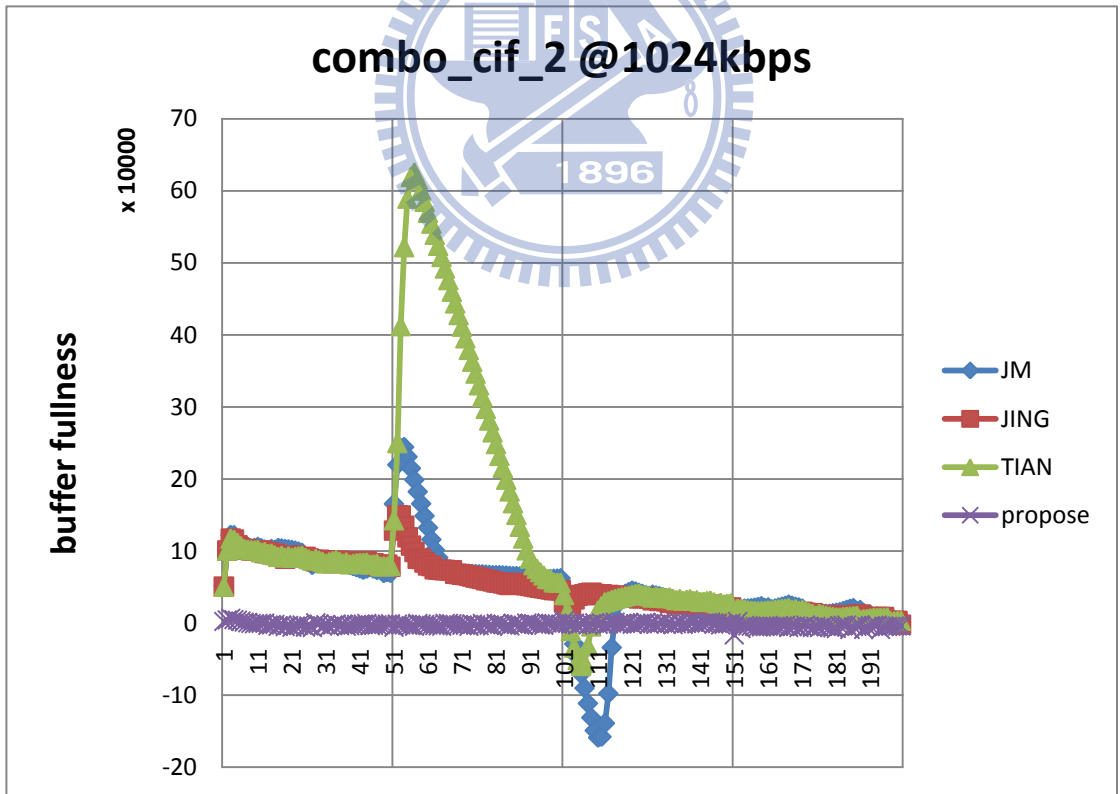
(b)
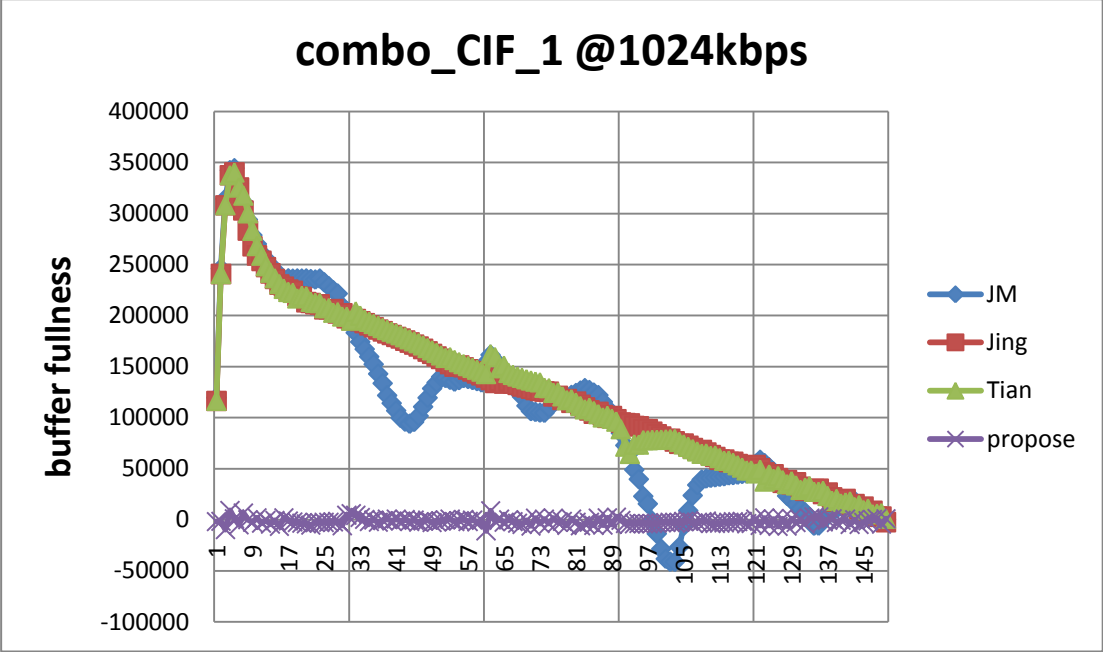
(c)

(d)

(e)

Fig 4-2 Buffer fullness v.s. frames (a) carphone QCIF (b) combo_CIF_3 (c) combo_CIF_1 (d) combo_SD (e) combo_HD

Chapter 5   Conclusion

We present an improved rate control algorithm for H.264 by controlling the QP of intra frames and SC_I frames. For intra frames and SC frame, we propose gradient based R-D-Q model. The cost value of each candidate QP is calculated to determine the optimized QP.

The simulation results show our approach is adequate for Intra-only compression. The proposed algorithm is cable of achieving an average of 0.51 dB in QCIF, 0.42 dB in CIF, 0.54 dB in SD, and 0.33 dB in HD PSNR gain compared to JM rate control algorithm for Intra-only compression, respectively. In contrast with Jing's and Tian's algorithm, our scheme has an average of 0.27 and 0.33 dB in QCIF, 0.09 and 0.12 dB in CIF, 0.01 and 0.02 dB in SD, 0.01 and 0.02 dB in HD PSNR gain for Intra-only compression, respectively. Our proposal also has better performance in buffer fullness and low PSNR standard derivation. Besides, the proposed algorithm is not only no need to deal with scene change frame but also suitable than other algorithms using pre-frame information in more activity sequences.

REFERENCE

[1] ISO/IEC 11172-2, "Information Technology – Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s – Part 2: Video," Edition 1, 1993. (MPEG-1 Video).

[2] ISO/IEC 13818-2, "Information technology – Generic coding of moving pictures and associated audio information: Video", Edition 2, 2000. (MPEG-2 Video).

[3] ISO/IEC 14496-2:2001, "Coding of Audio-Visual Objects - Part 2: Visual," 2nd Edition, 2001. (MPEG-4 Video).

[4] ITU-T Recommendation H.261, "CODEC for audio-visual services at p × 64 kbit/s," 1993.

[5] ITU-T Recommendation H.263, "Video coding for low bit rate communication," 1998.

[6] "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC," in Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, 2003.

[7] T. Wiegand, H. Schwarz, A. Joch, F. Lossentini, and G. J. Sullivan, "Rate constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst., Video Technol.*, vol. 13, pp.688–703, Jul. 2003.

[8] G. J. Sullivan, H. Yu, S. Sekiguchi, H. Sun, T. Wedi, S. Wittmann, Y. Lee, A. Segall, and T. Suzuki, "New standardized extension of MPEG-4AVC/H.264 for professional-quality video applications," *IEEE International Conference on Image Processing*, Sep. 2007.

[9] T. Tsukuba, I. Nagayoshi, T. Hanamura, and H. Tominaga, "H.264 fast intra-prediction mode decision based on frequency characteristic," *13th European Signal Processing Conference*, September 2005.

[10]   〝AVC-Intra (H.264 Intra) Compression,"

ftp://ftp.panasonic.com/pub/Panasonic/Drivers/PBTS/papers/WP_AVC-Intra.pdf

[11]   Z.G. Li, F. Pan, K.P. Lim, G.N. Feng, X. Lin and S. Rahardja, "Adaptive basic unit layer rate control for JVT," JVT-G012-r1, 7th Meeting, Pattaya II, Thailand, Mar. 2003.

[12]   JM 15.0, *H.264/AVC Reference Software*, http://iphome.hhi.de/suehring/tml/

[13]   T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate-distortion modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 246–250, Feb. 1997

[14]   H. J. Lee, T. Chiang, and Y.-Q. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 878–894, Sept. 2000

[15]   H.M Hang, and J.J.Chen, "Source Model for Transform Video Coder and Its Application—Part:1 Fundamental Theory," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, No. 2, April, 1997

[16]   A. N. Netravali and J. O. Limb, "Picture coding: A review," Proc. IEEE, vol. PROC-68, no. 3, pp. 7–12, Mar. 1960.

[17]   N. Kamaci, Y. Altunbasak, and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via a cauchy-density-based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 994–1006, Aug. 2005.

[18]   X. Jing, L.-P. Chau, and W.-C. Siu, "Frame complexity-based rate-quantization model for H.264/AVC intraframe rate control," *IEEE Signal Process. Lett.*, vol. 15, pp. 373–376, 2008.

[19]   L.Tian, Y Sun, I Ahmad and S. Sun, "Effective intra-only rate control for H.264/AVC," *IEEE International Conf. on Image Proc,* Nov. *2009.*