

國立交通大學

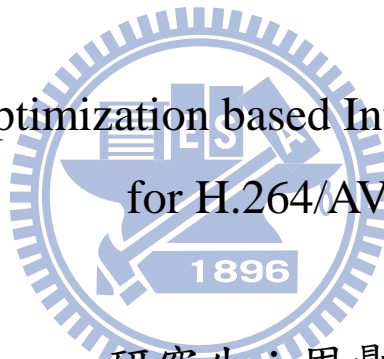
資訊科學與工程研究所

碩士論文

基於拉格朗日最佳化

針對 H.264 畫面內編碼的位元率控制演算法

Lagrangian-Optimization based Intra Frame Rate Control  
for H.264/AVC



研究生：周鼎力

指導教授：蔡文錦 博士

中華民國 98 年 6 月

基於拉格朗日最佳化針對 H.264 畫面內編碼的位元率控制演算法  
Lagrangian-Optimization based Intra Frame Rate Control for H.264/AVC

研 究 生：周鼎力

Student : Ting-Li Chou

指 導 教 授：蔡文錦

Advisor : Wen-Jiin Tsai

國 立 交 通 大 學  
資 訊 科 學 與 工 程 研 究 所  
碩 士 論 文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年六月

## 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 編碼與 GOP 編碼皆適用的位元率控制演算法。首先我們提出基於 Lagrangian 最佳化的 QP 決定方式，藉由位元率與 PSNR 預測模型，我們利用 Lagrangian 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，為了解決場景變換對位元率控制所產生的影響，在找出場景變換的畫面後，針對其使用以梯度複雜度為基礎的位元率模型來決定適合的 QP，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、拉格朗日最佳化、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 not only for GOP compression but also for Intra-only compression. First, we present a Lagrangian-optimization based QP determination scheme for I-frames. By the estimation models for rate and PSNR of I-frames, the best quantization parameters can be determined by Lagrangian optimization method. In order to deal with the specific intra frames caused by scene transitions, a gradient complexity based QP determination method is proposed. After detecting scene change frames, the proposed gradient complexity based rate-QS model is adopted 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, Lagrangian optimization, Prediction model

## 誌謝

本論文是我人生的一大步，願為人類科技的一小步。

在這兩年的研究所生涯裡，首先要感謝的是我的指導教授蔡文錦博士，其不辭辛勞的教導使我獲益匪淺，在課業研究上，給予我建議；在人生旅途上，指引我方向，在此謹向我的指導教授蔡文錦博士致上最高的敬意，感謝老師這兩年來對我的用心指導。

我要感謝實驗室裡的學長姐蕭家偉、陳信良、李威邦、黃重輔、吳秉承和陳佩詩，在我的研究過程中給予我指導和寶貴的意見。還要感謝我的同學黃子娟、陳建裕、林宜政、林詩凱和吳漢倫，謝謝你們兩年來不管是在課業上的討論、生活上的分享或戰場上的廝殺(笑)，都使我獲益良多。還有謝謝學弟王世明、許智為、潘益群和游顯榆，接下來要靠自己了，願你們明年都能順利畢業。

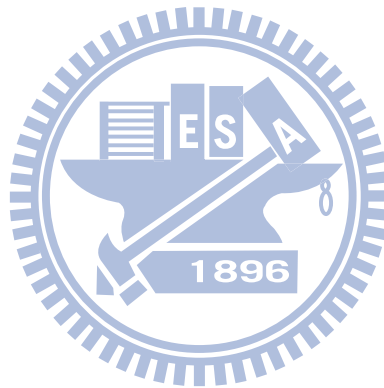
感謝我的父母和弟弟對我的支持，讓我能夠完成大學與研究所學業，步入下一個人生階段。最後感謝我的女朋友奕萱，謝謝你對我的陪伴與鼓勵，能彼此分享生活中的一切，是最棒的一件事。接下來，我將邁入人生的另一個里程，相信有你們的陪伴，未來的路會更寬廣、更美好。

謹以此論文獻給我的師長、父母和所有關心我的人

# CONTENTS

中文摘要.....	i
<b>ABSTRACT .....</b>	<b>ii</b>
誌謝.....	iii
<b>CONTENTS.....</b>	<b>iv</b>
<b>LIST OF FIGURES .....</b>	<b>vi</b>
<b>LIST OF TABLES.....</b>	<b>viii</b>
<b>Chapter 1 Introduction and Motivation .....</b>	<b>1</b>
1.1 Introduction to Rate Control .....	1
1.1.1 The Chicken Egg Dilemma for H.264 Rate Control .....	3
1.1.2 Main Criteria of Rate Control.....	4
1.2 Introduction to H.264 Intra-coded Frames .....	5
1.2.1 H.264 Intra Compression.....	6
1.2.2 H.264 Intra-only Profiles .....	6
1.3 Motivation .....	7
<b>Chapter 2 Related Works.....</b>	<b>10</b>
2.1 G012 Rate Control for H.264.....	10
2.1.1 Terminology.....	10
2.1.2 Overview to G012 Rate Control.....	12
2.2 Cauchy Density based Rate Control for H.264.....	15
2.3 Frame Complexity based Intra only Rate Control .....	17
2.4 Intra Frame Bit Allocation Algorithm .....	18
2.5 Adaptive Distortion based Intra Frame Rate Control.....	19
2.6 Summary .....	20
<b>Chapter 3 Proposed Rate Control Algorithm for Intra-only Compression .....</b>	<b>23</b>
3.1 Lagrangian-Optimization based QP Determination for Intra Frames .....	23
3.1.1 Taylor Series based Rate-QS Model.....	24
3.1.2 Gradient Complexity based PSNR-QP Model.....	27
3.1.3 Estimation of $\lambda$ .....	30
3.1.4 QP Determination Method for Intra Frames .....	31
3.2 Gradient Complexity based QP Determination for Scene Change Intra Frames .....	32
3.2.1 Gradient based Scene Change Detection.....	32

3.2.2	Gradient Complexity based Rate-QS Model .....	35
3.3	Description of the Proposed Rate Control Algorithm for Intra-only Compression.....	37
<b>Chapter 4</b>	<b>Proposed Rate Control Algorithm for GOP Compression .....</b>	<b>39</b>
4.1	Target Bit Allocation for Intra Frames .....	39
4.2	Description of Proposed Rate Control Algorithm for GOP Compression ...	41
<b>Chapter 5</b>	<b>Experiment Results .....</b>	<b>43</b>
5.1	Results of Intra-only Compression.....	43
5.2	Results of GOP Compression.....	47
<b>Chapter 6</b>	<b>Conclusion .....</b>	<b>51</b>
<b>REFERENCE</b>	<b>.....</b>	<b>52</b>



## LIST OF FIGURES

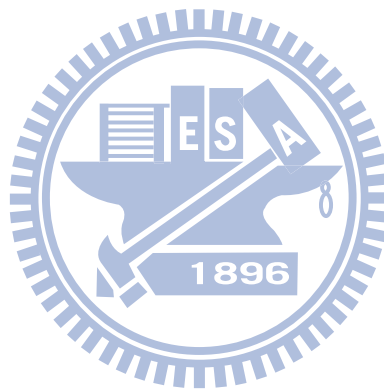
Fig. 1-1	Video transmission system .....	2
Fig. 1-2	(a) Variable bit rate vs. (b) Constant bit rate .....	3
Fig. 1-3	Basic rate control flow .....	3
Fig. 1-4	The chicken egg dilemma for H.264 rate control.....	4
Fig. 1-5	4x4 block intra prediction mode direction[9].....	6
Fig. 1-6	(a) Bits, (b) PSNR and (c) the percentage of intra coded MB for the qcif sequence <i>Akiyo-Foreman</i> which cascaded at 50 <sup>th</sup> frame, and the GOP size is 40.....	8
Fig. 2-1	The G012 rate control diagram .....	13
Fig. 2-2	Comparison of Laplacian model vs. Cauchy model[16].....	16
Fig. 2-3	Intra coded bits vs. gradient per pixel (a) <i>Foreman</i> , QP=36 (b) <i>Carphone</i> , QP=25 .....	17
Fig. 2-4	Curve fitting results of QS versus MSE for different sequences[19].....	20
Fig. 2-5	The relation between $SC_1$ and general I-frames .....	21
Fig. 2-6	The value of parameter $a$ from <i>akiyo</i> sequence.....	22
Fig. 3-1	The curves between $R_{norm}$ and QS of <i>Foreman</i> .....	24
Fig. 3-2	The curves between $R_{norm}$ and QS of <i>Akiyo</i> .....	25
Fig. 3-3	Prediction error comparison of proposed model and Jing's model. Analysis from (a) <i>Foreman</i> @QCIF-512kbps and (b) <i>Mobile</i> @QCIF-512kbps.....	27
Fig. 3-4	The relation curves between PSNR and QP of several sequences .....	28
Fig. 3-5	The relation between model parameter, $m$ and frame complexity, $G$ .....	28
Fig. 3-6	The prediction accuracy of <i>Foreman</i> at 512kbps (up), 1024kbps (down) ...	29
Fig. 3-7	Diagram of the proposed non $SC_1$ QP determination algorithm.....	32
Fig. 3-8	MDOG value of the QCIF sequence <i>Trevor-Stefan-Silent-Coastguard</i> .....	33
Fig. 3-9	FDs of eight QCIF cascaded sequences when the threshold is set to 35 .....	34
Fig. 3-10	The relation curve between $G$ and $G*a$ .....	35
Fig. 3-11	Prediction error comparison from a cascaded scene change sequence. ....	36
Fig. 3-12	Flow charts for Intra-only compression .....	37
Fig. 4-1	Relation curves between the best initial QP and bpp for <i>News</i> , <i>Foreman</i> , and <i>Mobile</i> [21].....	40
Fig. 4-2	Flow charts for GOP compression .....	41
Fig. 5-1	PSNR v.s. frames for (a) <i>Mobile</i> @512kbps (b) <i>Combo2</i> @512kbps.....	45



Fig. 5-2 Buffer fullness v.s. frames for *Foreman*@1024 kbps ..... 46

Fig. 5-3 Buffer fullness v.s. frames for *Combo2*@768kbps ..... 47

Fig. 5-4 (a) PSNR v.s. frames (b) Buffer fullness v.s. frames for *Mobile*@96kbps .. 49



# LIST OF TABLES

Table 1-1	Comparison between Intra-only and GOP compression[10] .....	7
Table 3-1	Detection correctness of two advertisements .....	35
Table 5-1	Performance comparisons for Intra-only scheme.....	44
Table 5-2	Result comparisons of normal sequences for GOP compression scheme....	48
Table 5-3	Result comparisons of scene change sequences for GOP compression scheme.....	48



# Chapter 1 Introduction and Motivation

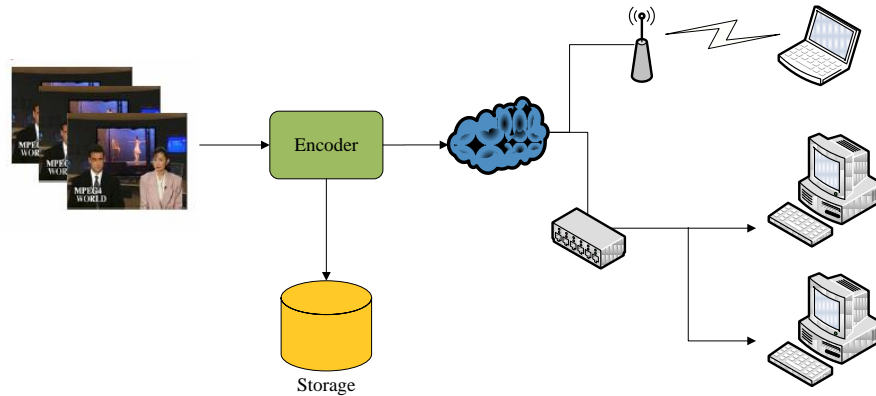
For the coming of digital multimedia communication, the demand for the storage and transmission of visual information has stimulated the development of video coding standards, including MPEG-1[1], MPEG-2[2], MPEG-4[3], H.261[4], H.263[5], and H.264/AVC[6].

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

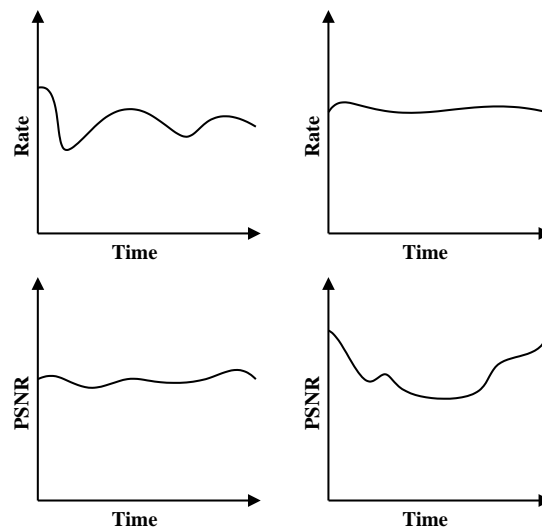
A rate control algorithm which meets a constrained channel rate by controlling the number of generated bits is necessary to encoder. Either the coded video is transmitted over the Internet or stored in a storage device, there is a bandwidth constraint to limit the bit rate of videos. Although the transmission 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

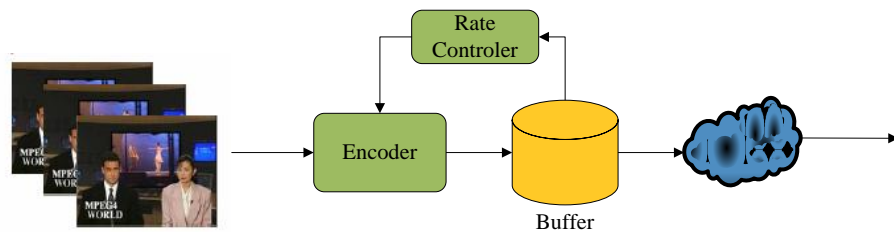


the video content varying.

(a) (b)

**Fig. 1-2 (a) Variable bit rate vs. (b) Constant bit rate**

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

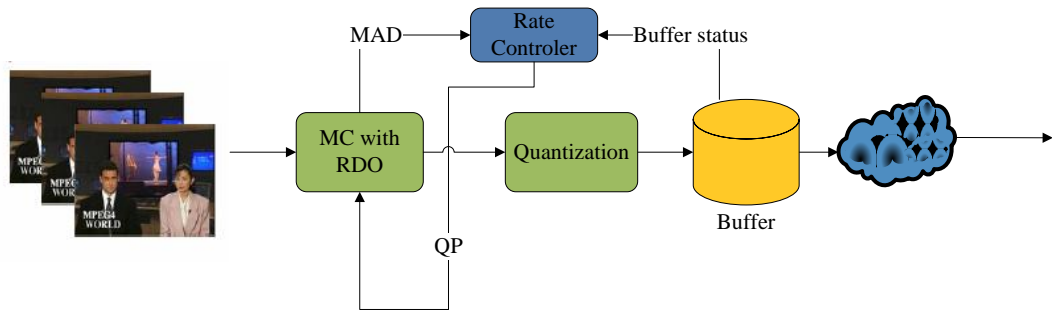
One important property of H.264 is the implementation of rate distortion optimization (RDO)[7] for both motion estimation and mode decision. With RDO, the Lagrangian method is utilized to optimize the trade-off between distortion and bit rate consumed. For example, the Lagrangian cost function of motion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

$$J_{Motion}(MB_i, MV_i | QP, \lambda) = D(MB_i, MV_i | QP) + \lambda \cdot R(MB_i, MV_i | QP) \quad (1.1)$$

where  $MB_i$  and  $MV_i$  stand for the  $i^{\text{th}}$  macro block (MB) and the motion vector (MV) of  $i^{\text{th}}$  MB in the current frame, respectively;  $\lambda$  denotes the Lagrangian multiplier which depends on

$$\lambda = \sqrt{0.85 \times 2^{(QP-12)/3}} \quad (1.2)$$

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

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.

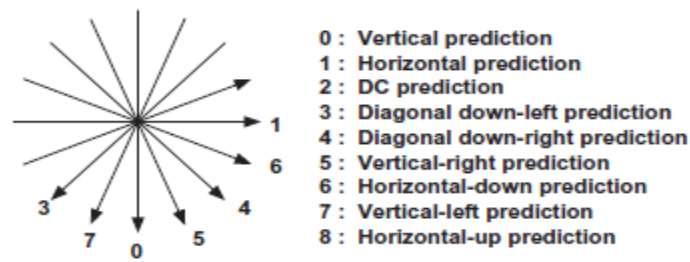
- D. *Maximum buffer fullness.*

A 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

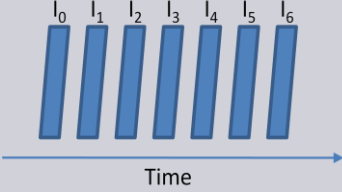
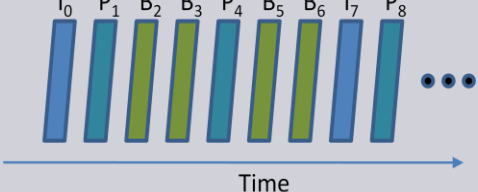
H.264 utilizes the intra prediction to reduce the spatial redundancy within frames. Fig. 1-5 shows the prediction options of 4x4 block intra prediction. Each pixel in the current 4x4 block is predicted from the neighboring reconstructed pixels, where nine prediction modes can be selected by the encoder, 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 profiles 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.



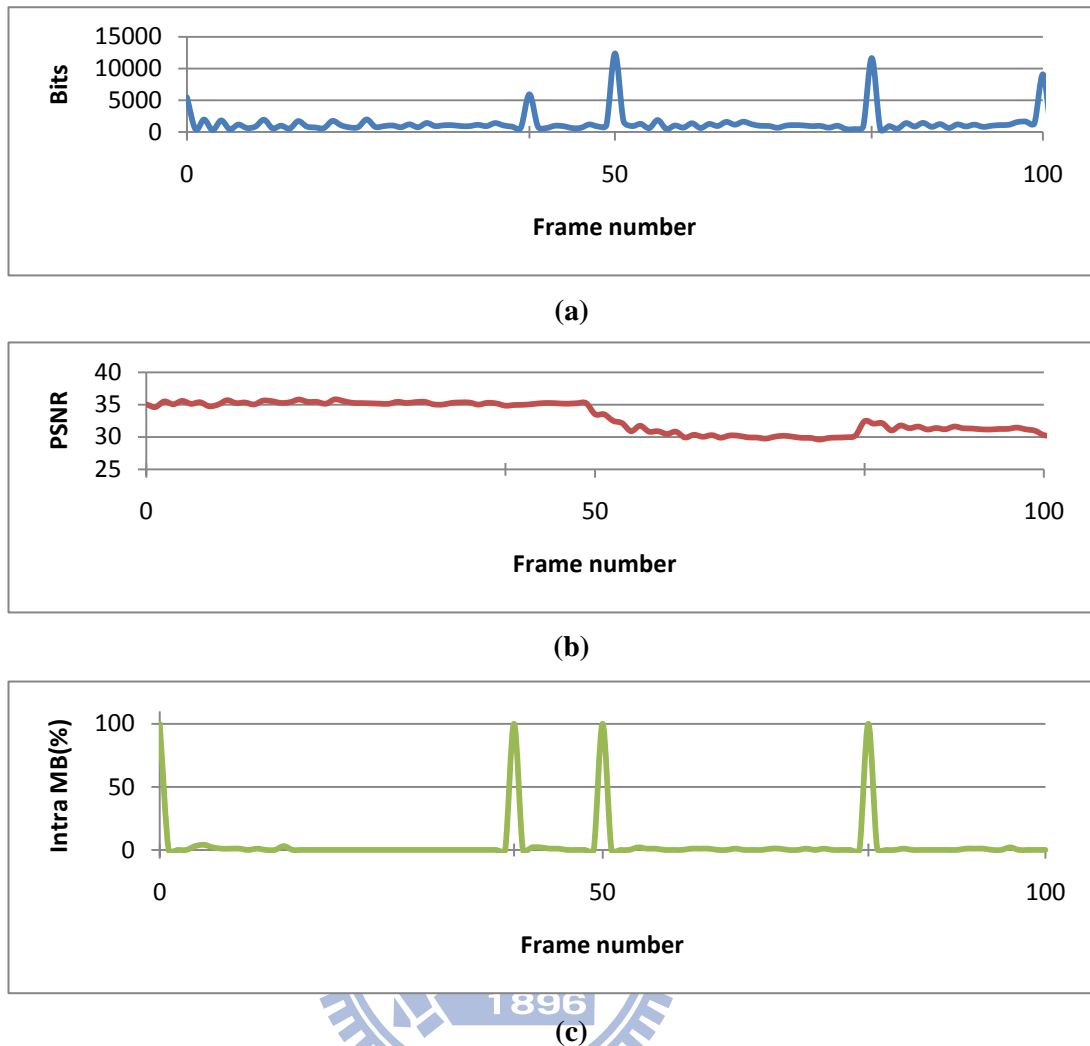
**Table 1-1 Comparison between Intra-only and GOP compression[10]**

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



### 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, even 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.



**Fig. 1-6 (a) Bits, (b) PSNR and (c) the percentage of intra coded MB for the qcif sequence *Akiyo-Foreman* which cascaded at 50<sup>th</sup> frame, and the GOP size is 40**

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 also observed that the abrupt scene change usually results in buffer overflow due to the fact that most of MBs in the scene change frame are intra coded. It often

produces more bits than the target bits and degrades the visual quality of the following frames. Fig. 1-6 illustrates the fact mentioned above. In Fig. 1-6(c), the percentage of intra coded MBs at the scene change frame 50<sup>th</sup> is 100% which means all the MBs are encoded with intra mode. Fig. 1-6(a) and (b) shows the output bits of the scene change frame are much more than those of other frames and the PSNRs of the following frames are degraded until the start of the next GOP.

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 take frame complexity into consideration to decide proper QPs for the both types of 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 and Chapter 4 for GOP compression. Chapter 5 provides the simulation results compared to other rate control schemes. Finally, Chapter 6 concludes this thesis.

## 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

$$N_{unit} = \frac{N_{picunit}}{N_{mbunit}} \quad (2.1)$$

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

$$MAD(x, y) = \frac{1}{HW} \cdot \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} |C(x+i, y+j) - R(x+i, y+j)| \quad (2.2)$$

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

$$MAD_{pb} = a_1 \times MAD_{cb} + a_2 \quad (2.3)$$

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 Laplacian distribution

$$P(x) = \frac{\alpha}{2} e^{-\alpha|x|} \quad \text{where } -\infty < x < \infty \quad (2.4)$$

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

$$R(D) = \ln\left(\frac{1}{\alpha D}\right) \quad \text{where } D_{\min} = 0, D_{\max} = \frac{1}{\alpha}, 0 < D < \frac{1}{\alpha} \quad (2.5)$$

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

$$R = \frac{X_1}{QP} + \frac{X_2}{QP^2} \quad (2.6)$$

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

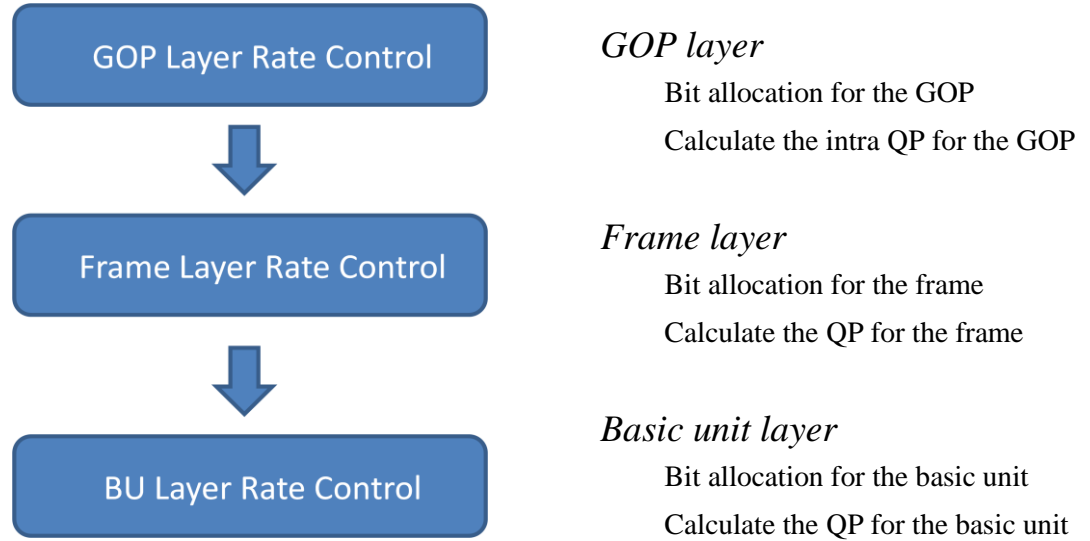
$$R = \frac{MAD \cdot X_1}{QP} + \frac{MAD \cdot X_2}{QP^2} - H \quad (2.7)$$

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.



**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

$$R_r = \frac{u}{F_r} \cdot N_{GOP} - B_c \quad (2.8)$$

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

$$R_r = R_r - b \quad (2.9)$$

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

$$\left\{ \begin{array}{l} QP_{I,first} = \begin{cases} 40 & bpp \leq l_1 \\ 30 & l_1 < bpp \leq l_2 \\ 20 & l_2 < bpp \leq l_3 \\ 10 & l_3 < bpp \leq l_4 \end{cases} \\ QP_{I,other} = \frac{SumQP}{N_p} \end{array} \right. , \text{where } bpp = \frac{u}{F_r \times N_{pixel}} \quad (2.10)$$

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  $i^{\text{th}}$  P-frame in the current GOP is determined as

$$R_i = \beta \cdot \hat{R}_i + (1 - \beta) \cdot \tilde{R}_i \quad (2.11)$$

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

$$\hat{R}_i = \frac{R_r}{N_{remain}} \quad (2.12)$$

$$\tilde{R}_i = \frac{u}{F_r} + \gamma \cdot (Tbl_i - V_i) \quad (2.13)$$

where  $N_{remain}$  is the number of non-coded frames in the current GOP;  $\gamma$  is a constant, and  $Tbl_i$  and  $V_i$  are the target buffer level and the virtual buffer fullness of the  $i^{\text{th}}$  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, 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.

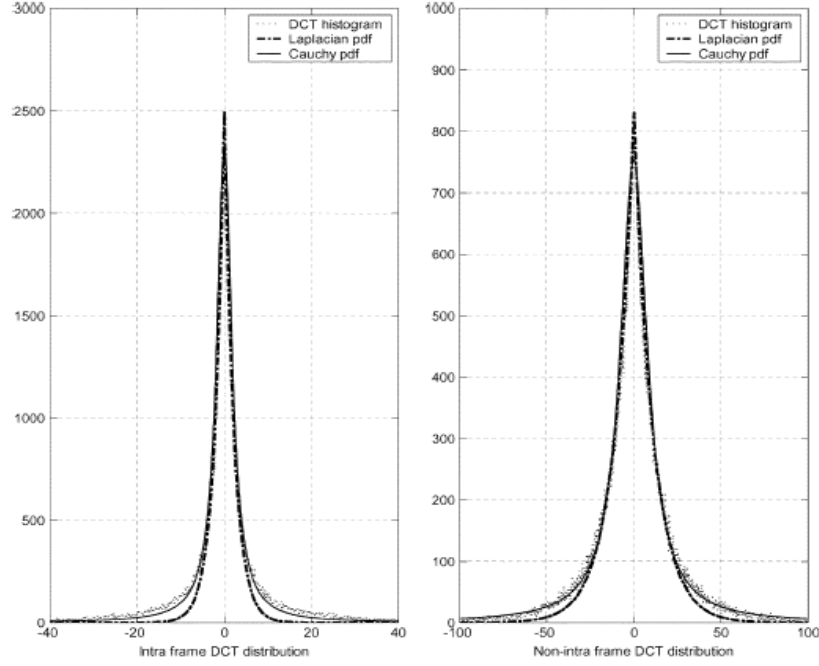
If frames are not selected as basic units, basic unit layer rate control should be performed after frame layer bit allocation. In basic unit layer, it is almost the same as that in frame layer. It predicts MADs of all basic units in the current frame by equation (2.3) and calculates the target number of bits of them by

$$\tilde{b}_i = R_{c,remain} \cdot \frac{MAD_{i,pred}^2}{\sum_{j=i}^{N_{unit}} MAD_{j,pred}^2} \quad (2.14)$$

where  $R_{c,remain}$  is the remaining target number of bits of current frame;  $MAD_{i,pred}$  stands for the predicted MAD of  $i^{\text{th}}$  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 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 [15], the coefficients are conjectured to have Laplacian distribution. In [16], 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[16]**

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

$$R = a \cdot QS^b \quad (2.15)$$

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

$$QS = \sqrt[b]{\frac{R}{a}} \quad (2.16)$$

Finally, the QP used for RDO can be calculated by

$$QP = 6 \cdot \log_2(QS) + 4 \quad (2.17)$$

where  $\bullet$  denotes the rounding operation.

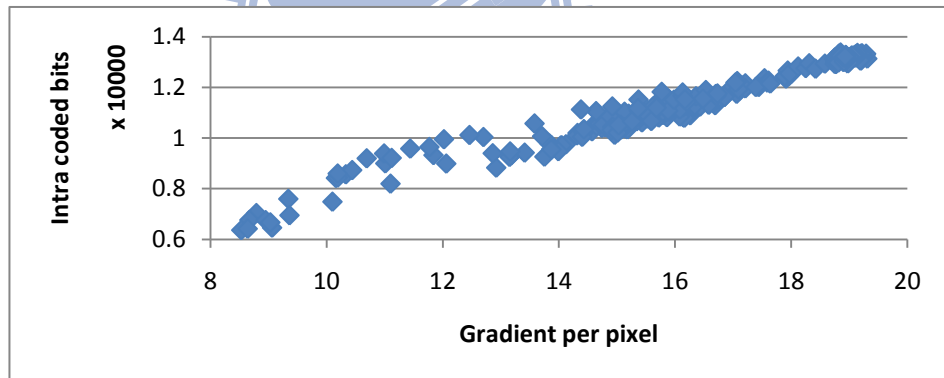
## 2.3 Frame Complexity based Intra only Rate Control

Based on Kamaci *et al.*'s rate estimation model, Jing *et al.*[17] proposed an improved model which is applied on intra frames and has sufficient adaptability to the varying of intra frame complexity.

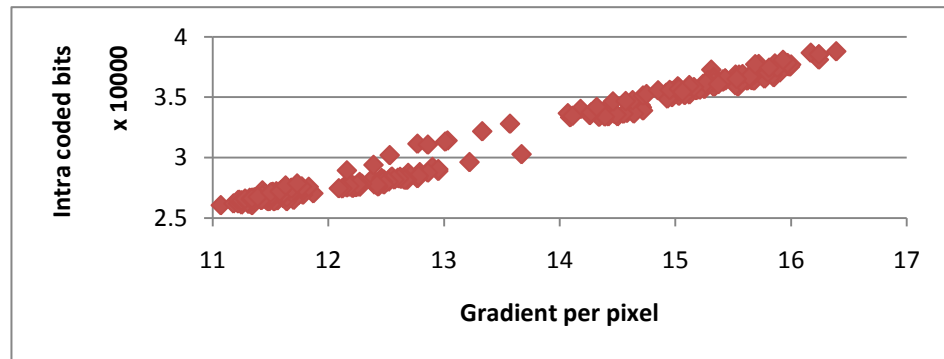
In their proposed algorithm, they defined the complexity measure of intra frames as the average gradient per pixel of the frame. The calculation of gradient complexity is defined by

$$G = \frac{1}{M \cdot N} \left( \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |I_{i,j} - I_{i+1,j}| + |I_{i,j} - I_{i,j+1}| \right) \quad (2.18)$$

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  $(i, j)$ .



(a)



(b)

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

$$R = G \times a \cdot QS^b \quad (2.19)$$

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

$$a_k = \begin{cases} \frac{R_0}{G_0 \cdot QS_0^b} & k = 0 \\ \alpha \cdot a_{k-1} + (1 - \alpha) \cdot \frac{R_{k-1}}{G_{k-1} \cdot QS_{k-1}^b} & \text{otherwise} \end{cases} \quad (2.20)$$

After frame layer bit allocation, QS can be calculated by (2.19), and QP can be derived from (2.17).

## 2.4 Intra Frame Bit Allocation Algorithm

Sun *et al.*[18] exploited prediction and feedback control to achieve accurate rate control while maximizing the picture quality and smoothing buffer fullness. Their algorithm estimates the bit budget for the I-frame of  $i^{\text{th}}$  GOP based on its global coding complexity with the following equation

$$R_i = R_r \cdot \frac{W_{Intra,i}}{W_{Intra,i} + W_{Inter} \cdot N_p} \quad (2.21)$$

where  $N_p$  is the number of P-frames in GOP;  $W_{Inter}$  is the weighting of inter coded frames which is set to 1.  $W_{Intra}$  stands for the weighting of intra coded frames which is calculated as follows

$$W_{Intra,i} = \frac{Bit_{avg(Intra),i-1}}{Bit_{avg(Inter),i-1}} \cdot e^{\left(\frac{PSNR_{avg(Inter),i-1} - PSNR_{avg(Intra),i-1}}{\delta}\right)} \quad (2.22)$$

where  $Bit_{avg(Intra),i-1}$  and  $PSNR_{avg(Intra),i-1}$  are the average number of bits and PSNR of the I-frame in the previous GOP, respectively,  $Bit_{avg(Inter),i-1}$  and  $PSNR_{avg(Inter),i-1}$  denote those of P-frames, and  $\delta$  is a model parameter which is set to 8 in their experiments. In equation (2.21), the target number of bits of the I-frame in the current GOP is determined by the intra weighting value which relies on the coding results of the previous GOP.

They also proposed a novel buffer controller based on the proportional integral derivative (PID) technique used in automatic control systems, and used (2.7) to determine QP.

## 2.5 Adaptive Distortion based Intra Frame Rate Control

Yan *et al.*[19] presented an adaptive distortion based intra rate estimation (ADIE) algorithm for H.264/AVC rate control. In this algorithm, a new rate control model is established according to the distortion which is predicted by taking image complexity, buffer status and scene change into considerations. From the quadratic rate model (2.5), they supposed the output bit rate is related to the output MSE, and is given by

$$R = \ln\left(\frac{1}{\lambda \cdot MSE}\right) \quad (2.23)$$

where  $MSE$  measures the coding distortion, i.e., the mean squared error between original frame and re-constructed frame;  $\lambda$  is the coefficient for Laplacian distribution. Further, they observed that MSEs of intra coded frames are linear correlated with the QPs used for encoding, which is shown in Fig. 2-4.

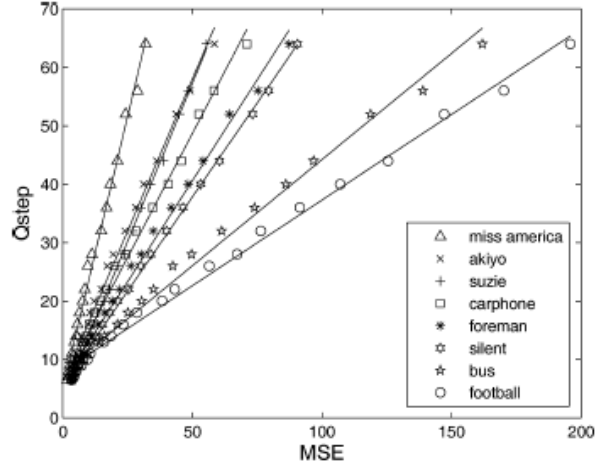


Fig. 2-4 Curve fitting results of QS versus MSE for different sequences[19]

Based on the above observation, they proposed that the relation between QS and MSE can be approximately modeled as

$$QS = \alpha \times MSE + \beta \quad (2.24)$$

where the value of  $\alpha$  can be obtained after coding the first I-frame, and  $\beta$  is a constant. Finally, their proposed MSE prediction model which based on gradient complexity and buffer status is

$$MSE_{pred,i} = MSE_{i-1} \times \frac{G_i}{MG} \times \frac{1}{1 - BR_{i-1}} + \theta \quad (2.25)$$

where  $MG$  is the mean gradient value of previous I-frames in this sequence, and  $\theta$  is a model constant.  $BR_i$  is the current buffer fullness ratio derived by  $BF_i / BufferSize$  where  $BF_i$  is the buffer fullness after encoding the  $i^{th}$  GOP. After the MSE prediction using (2.25), the model (2.24) is employed to determine the appropriate QS value.

## 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 as observed in Fig. 1-6, 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. 2-5.

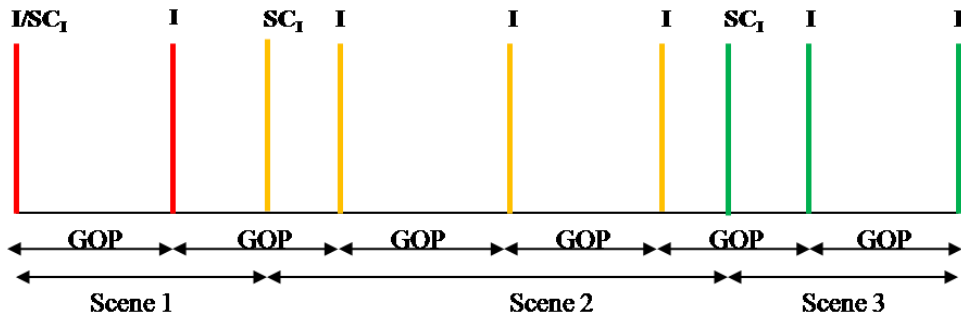
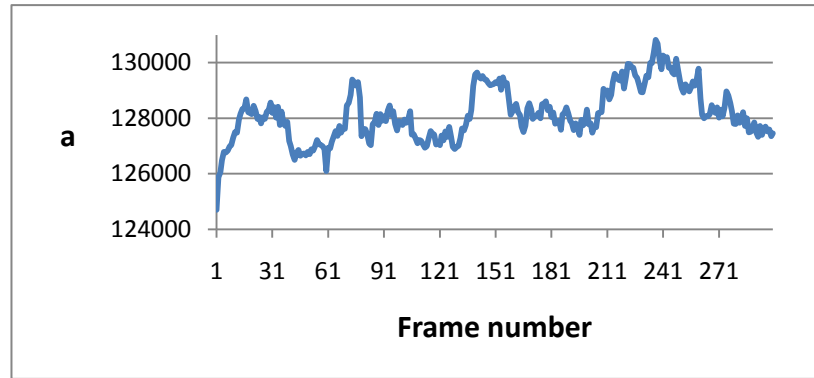


Fig. 2-5 The relation between  $SC_I$  and general I-frames

Similar to general I-frames, these  $SC_I$  can cause serious buffer overflow problem if no appropriate QP is determined for them. Although rate control algorithms have been widely studied [11][16], most of them didn't deal with the scene change intra frames. Yan *et al.*[19] had their mechanisms to detect scene change. However, they calculated the QP by using equation (2.10) which is not appropriate.

B. *Poor QP Determination for General Intra Frames*

In G012, the QP of each I-frame is decided by the average QP of all coded P-frames in the previous GOP. This simple approach which does not take frame complexity and buffer fullness into considerations may suffer from buffer overflow.



**Fig. 2-6** The value of parameter  $a$  from *akiyo* sequence

*C. No Accurate Rate Quantization Model for Intra Frames*

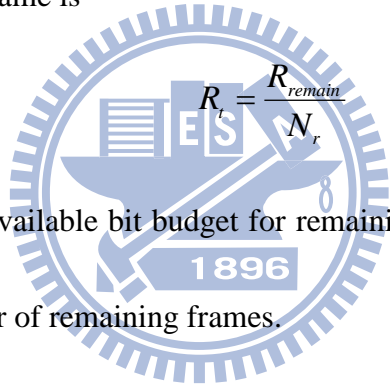
The quadratic model (2.7) is designed for inter coded frames whose source statistics are assumed satisfying Laplacain distribution. However, this assumption is inappropriate to intra coded frames. Jing *et al.*[17] proposed a novel rate quantization model for intra frames, but the parameter,  $a$  used in their model 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 as illustrated in Fig. 2-6 where the value of  $a$  in the figure varies frequently.



## Chapter 3 Proposed Rate Control Algorithm for Intra-only Compression

In this chapter, we present the proposed rate control algorithm for Intra-only compression. As mentioned in section 2.6, there are two kinds of intra frames should be dealt with. We first describe a Lagrangian-optimization based rate control scheme for intra frames, and then a gradient complexity based scheme is proposed for scene change intra frames ( $SC_I$  frames).

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

$$R_t = \frac{R_{remain}}{N_r} \quad (3.1)$$


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

### 3.1 Lagrangian-Optimization based QP Determination for Intra Frames

The proposed Lagrangian-optimization based rate control scheme is for QP determination of intra frames. First, we define the Lagrangian cost function as

$$J(QP) = PSNR(QP) - \lambda \cdot |R(QS) - R_t| \quad (3.2)$$

where  $\lambda$  is the Lagrangian multiplier;  $QS$  can be derived from the substituted QP;  $PSNR(\bullet)$  and  $R(\bullet)$  denote the proposed gradient complexity based PSNR-QP model and Taylor series based rate-QS model, respectively.

It is obvious that the higher the cost value  $J(QP)$  is, the better tradeoff between quality and rate can be obtained. In this section,  $PSNR(\bullet)$ ,  $R(\bullet)$ , and  $\lambda$ , are introduced first. Then, a QP determination algorithm based on equation (3.2) for general intra frames is proposed.

### 3.1.1 Taylor Series based Rate-QS Model

In section 2.6-C, we have mentioned the drawback of Jing's intra rate quantization model. In order to solve this problem, we modified the equation (2.19) by defining the normalized bit rate of  $i^{\text{th}}$  frame,  $R_{norm,i}$  as follows

$$R_{norm,i} = \frac{R_i}{G_i} = aQS^b \quad (3.3)$$

We gather statistics of  $R_{norm}$  from different frames with all QS in intra coding mode. Fig. 3-1 and Fig. 3-2 show the curves of  $R_{norm}(x)$  for the first five frames in foreman and akiyo sequences, respectively. It indicates that the  $R_{norm}(x)$  curves in neighboring frames or frames in the same scene are closely identical. In other words, the  $R_{norm}(x)$  curve of the current intra frame can be predicted from that of the previous intra frame if it is available.

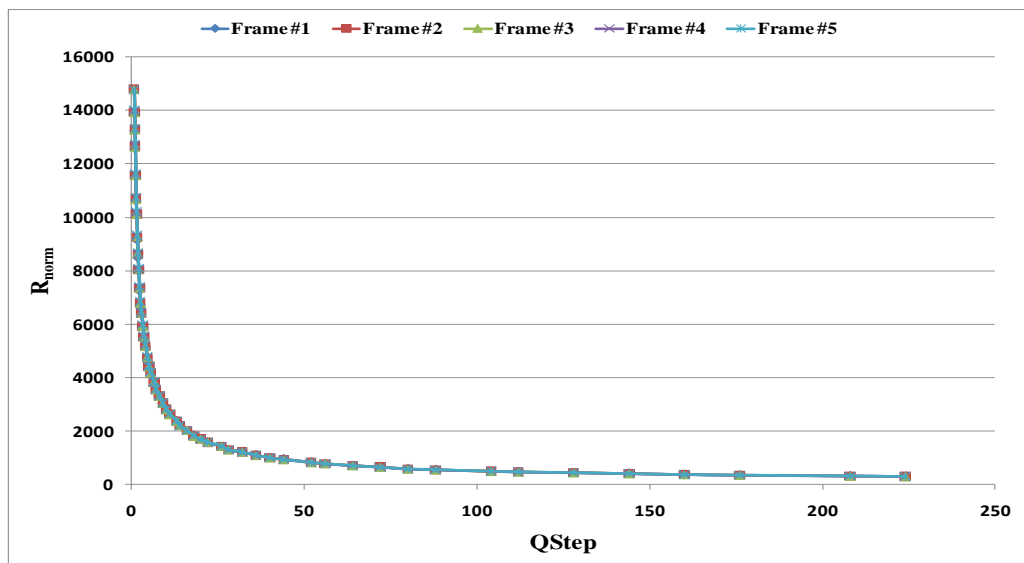
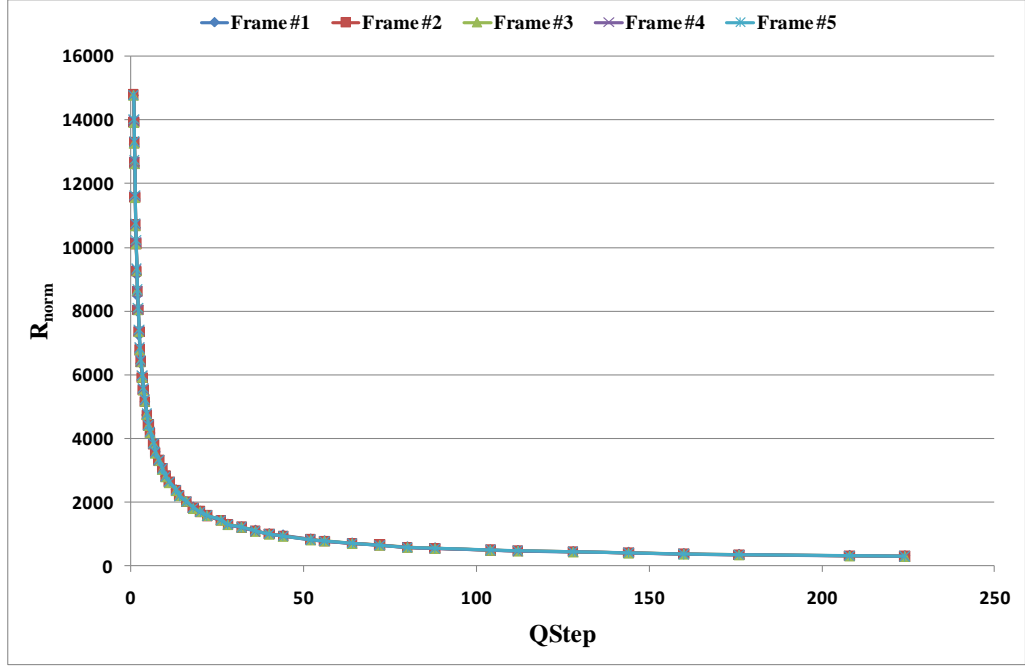


Fig. 3-1 The curves between  $R_{norm}$  and QS of *Foreman*



**Fig. 3-2 The curves between  $R_{norm}$  and QS of Akiyo**

In fact, there is one single point available in the curve of the previous intra frame because the QS used and the number of bits encoded for the previous intra frame can be obtained after its encoding procedure. Taylor series theory[22] indicates that any infinitely differentiable function,  $f(x)$ , can be represented as an infinite sum of terms calculated from all the values of derivation at a single point. Based on Taylor series theory, the formula of  $R_{norm,i}$  can be represented as<sup>1</sup>

$$\begin{aligned}
 R_{norm,i}(x) &= aQS^b \\
 &\cong R_{norm,i}(QS_i) + \frac{R'_{norm,i}(QS_i)}{1!}(x - QS_i) + \frac{R''_{norm,i}(QS_i)}{2!}(x - QS_i)^2
 \end{aligned} \tag{3.4}$$

where  $R'_{norm}(QS_i)$  and  $R''_{norm}(QS_i)$  can be derived from

---

<sup>1</sup> For simplification, we only expand the series to the second order.

$$\begin{cases} R'_{norm}(QS) = \frac{d(R_{norm}(QS))}{d(QS)} = b \times a \cdot QS^{b-1} = b \left( \frac{R_{norm}(QS)}{QS} \right) \\ R''_{norm}(QS) = \frac{d^2(R_{norm}(QS))}{d(QS)^2} = b \cdot (b-1) \left( \frac{R_{norm}(QS)}{QS^2} \right) \end{cases} \quad (3.5)$$

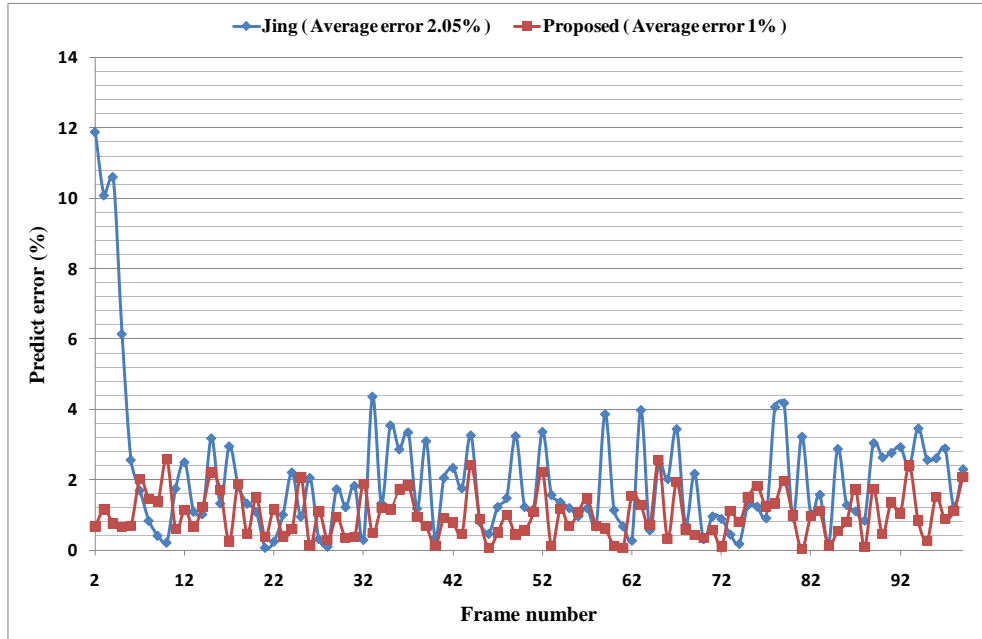
According to equations (3.3), (3.4), and (3.5) as well as the property that successive frames has identical  $R_{norm}$  curves, the proposed Taylor series based rate-QS model is

$$\begin{aligned} R_i(x) &= G_i \times R_{norm,i}(x) = G_i \times R_{norm,i-1}(x) \\ &= G_i \times \left( R_{norm,i-1} + \frac{R'_{norm,i-1}(QS_{i-1})}{1!} (x - QS_{i-1}) + \frac{R''_{norm,i-1}(QS_{i-1})}{2!} (x - QS_{i-1})^2 \right) \\ &= G_i \times \left[ R_{norm,i-1} + \frac{b \left( \frac{R_{norm,i-1}}{QS_{i-1}} \right)}{1!} (x - QS_{i-1}) + \frac{b \cdot (b-1) \left( \frac{R_{norm,i-1}}{QS_{i-1}^2} \right)}{2!} (x - QS_{i-1})^2 \right] \end{aligned} \quad (3.6)$$

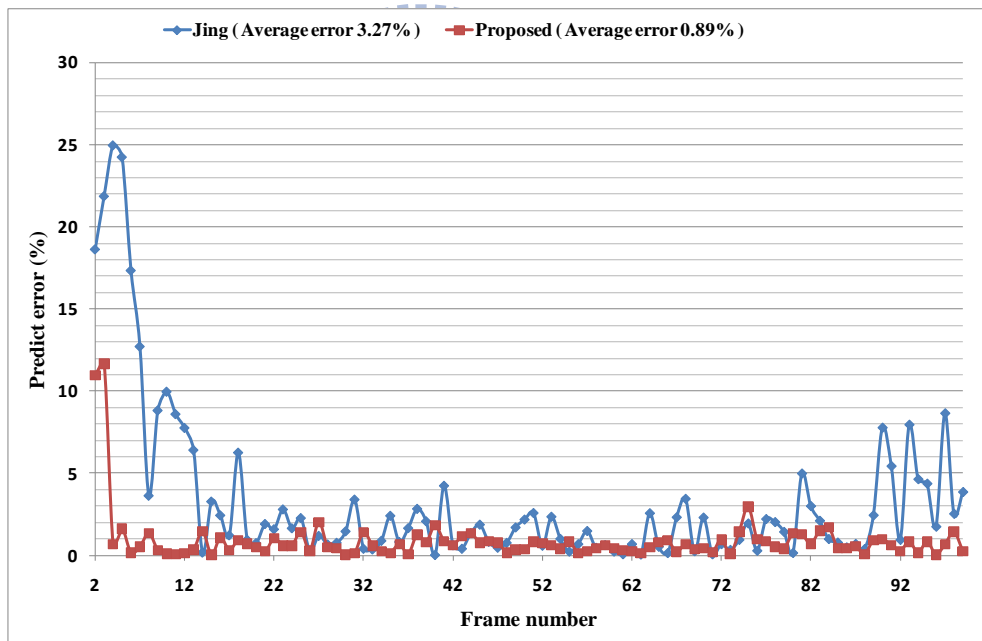
where  $b$  is a constant set to -0.76 in this thesis;  $R_{norm,i-1}$  and  $QS_{i-1}$  are the normalized number of bits encoded and the QS used in of the previous intra coded frame, respectively. Note that,  $R_{norm,0}$  and  $QS_0$  can be obtained after the coding procedure of the first intra frame.

Two experiments for the comparison between the proposed model and Jing's model (2.19) were conducted and the results were shown in Fig. 3-3. Note that, prediction error is calculated by equation (3.7). It is observed that the proposed model can achieve more accurate prediction, especially for the beginning frames. Compared with Jing's rate quantization model, the proposed Taylor series based model is more reliable due to its independence to the unstable model parameter,  $a$ . The proposed model can reduce the bit rate prediction error by up to 73%.

$$Error_{predict} = \frac{|Bits_{actual,i} - Bits_{predict,i}|}{Bits_{actual,i}} \times 100 \quad (3.7)$$



(a)



(b)

**Fig. 3-3 Prediction error comparison of proposed model and Jing’s model. Analysis from (a) *Foreman@QCIF-512kbps* and (b) *Mobile@QCIF-512kbps***

### 3.1.2 Gradient Complexity based PSNR-QP Model

In order to predict PSNR of the current intra coded frame, we encode the first frame of several sequences with a large range of QPs in intra coding mode and plot the results in Fig. 3-4. It indicates that there is a closely linear relation between QP and PSNR for intra coded frames. According to the observation, we propose to predict

PSNR by the following model

$$PSNR = m \cdot QP + k \quad (3.8)$$

where  $m$  and  $k$  are parameters relying on the content of sequence.

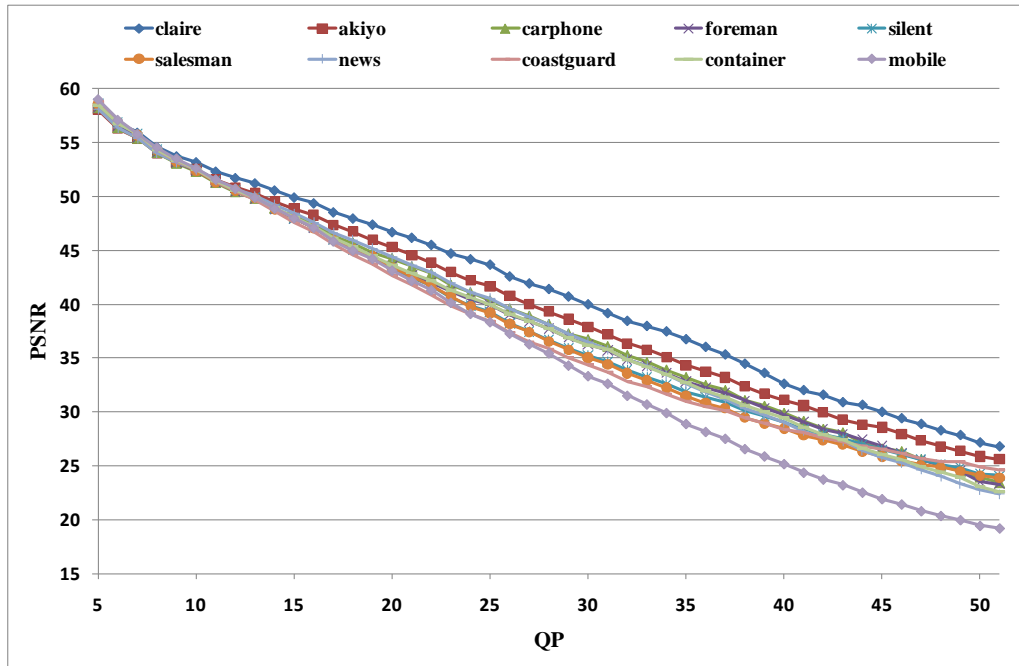


Fig. 3-4 The relation curves between PSNR and QP of several sequences

Fig. 3-4 also shows the slope,  $m$ , of each curve is different from others. The tendency of the slopes is related to frame complexity: the larger frame complexity, the more titled slope. After analyzing data from over 3000 intra coded frames, we realize the relation between the slope and the gradient based frame complexity,  $G$ , is also linear. Fig. 3-5 shows this relation. Based on the observation, we modelize the relation of  $m$  and gradient based frame complexity,  $G$ , with a linear training line.

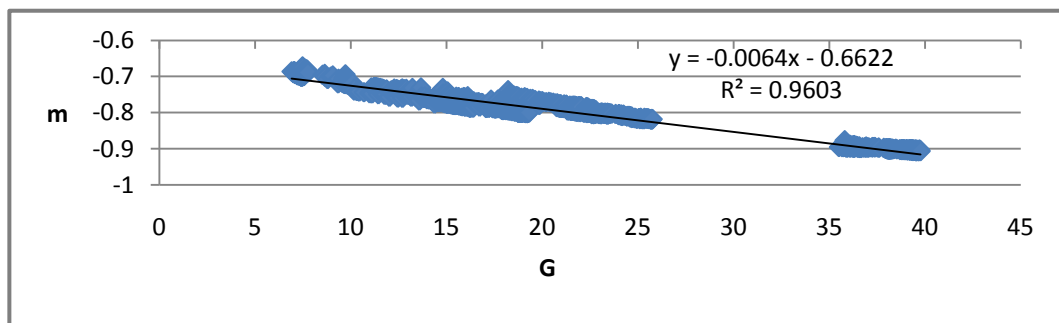


Fig. 3-5 The relation between model parameter,  $m$  and frame complexity,  $G$

During our model parameter updating procedure, the value of  $m$  and  $k$  are updated frame by frame with

$$\begin{cases} m_i = \frac{(\alpha \cdot G_i + \beta) + m_{i-1}}{2} \\ k_i = PSNR_{i-1} - m_{i-1} \cdot QP_{i-1} \end{cases} \quad (3.9)$$

where  $PSNR_{i-1}$  and  $QP_{i-1}$  are from the previous intra coded frame;  $\alpha = -0.0064$  and  $\beta = -0.6622$  are used for QCIF sequences.

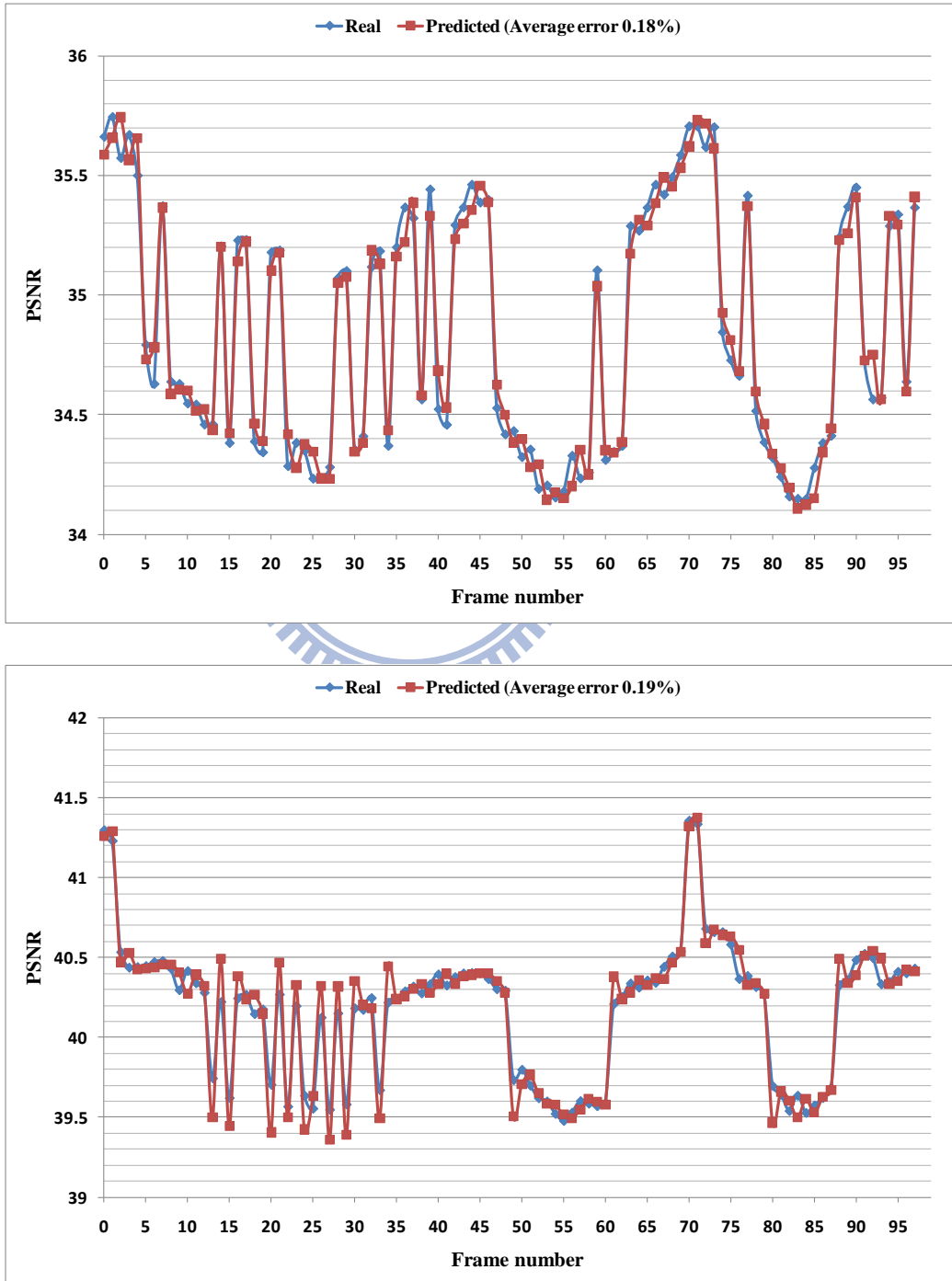


Fig. 3-6 The prediction accuracy of *Foreman* at 512kbps (up), 1024kbps (down)

To demonstrate the accuracy of the proposed gradient complexity based PSNR-QP model, Fig. 3-6 shows the experimental results, where real PSNR value and predicted PSNR value of each frame in the Foreman sequence are presented. It clearly illustrates the proposed model is reliable because the predicted PSNR curve closely fits the real one.

### 3.1.3 Estimation of $\lambda$

In Lagrangian cost function (3.2), the Lagrangian multiplier,  $\lambda$ , plays an important role to balance the weight of visual quality and departure from the target bit rate. If the value of  $\lambda$  is too small, the second term of the cost function has no influence against to the first term. If it is too large, the result is severely affected by the target bit rate departure, so the cost function cannot determine the best QP for intra coded frames.

In order to derive a fair Lagrangian multiplier, we substitute the proposed PSNR-QP model (3.8) and Jing's Rate-QS model<sup>2</sup> (2.19) into the cost function which can be written as

$$\begin{aligned}
 J(QP) &= PSNR(QP) - \lambda \cdot |R(QS) - R_t| \\
 &= (m \cdot QP + k) - \lambda \cdot |G \cdot aQS^b - R_t| \\
 &= (m \cdot QP + k) - \lambda \cdot \left| G \cdot a \cdot 2^{\frac{b(QP-4)}{6}} - R_t \right|
 \end{aligned} \tag{3.10}$$

According to Lagrangian optimization method, the optimized solution happens while  $\partial J / \partial QP = 0$ . It indicates that  $\lambda$  under optimized condition can be derived with<sup>3</sup>

---

<sup>2</sup> Due to the complexity of the proposed rate-QS model, we adopt Jing's model to derive  $\lambda$  for simplicity.

<sup>3</sup> Note that,  $\frac{1}{aG} = QS^b / R$  where  $R$  is the output bits of the previous intra frame.



$$\frac{\partial J}{\partial QP} = \frac{\partial PSNR(QP)}{\partial QP} - \lambda \cdot \frac{\partial |R(QP) - R_t|}{\partial QP} = 0$$

$$\Rightarrow \lambda = \begin{cases} \frac{\frac{\partial PSNR(QP)}{\partial QP} \div \frac{\partial R(QP)}{\partial QP}}{\frac{\partial R(QP)}{\partial QP}} = \frac{m}{\frac{1}{6} \cdot \log 2 \cdot abG \cdot 2^{\frac{b(QP-4)}{6}}} & , \text{if } R(QP) > R_t \\ = \frac{19.96 \cdot m}{abG \cdot 2^{\frac{b(QP-4)}{6}}} = \frac{19.96 \cdot m \cdot QS^b}{Rb \cdot 2^{\frac{b(QP-4)}{6}}} & \\ \frac{\frac{\partial PSNR(QP)}{\partial QP} \div -\left(\frac{\partial R(QP)}{\partial QP}\right)}{\frac{\partial R(QP)}{\partial QP}} = \frac{-19.96 \cdot m \cdot QS^b}{Rb \cdot 2^{\frac{b(QP-4)}{6}}} & , \text{otherwise} \end{cases}$$

Finally, the value of  $\lambda$  can be calculated with

$$\lambda = \frac{19.96 \cdot m \cdot QS^b}{Rb \cdot 2^{\frac{b(QP-4)}{6}}} \quad (3.11)$$

while the estimated number of bits is larger than the number of target bits. On the other hand, while the estimated number of bits is smaller than the target number of bits,  $\lambda$  is

$$\lambda = \frac{-19.96 \cdot m \cdot QS^b}{Rb \cdot 2^{\frac{b(QP-4)}{6}}} \quad (3.12)$$

### 3.1.4 QP Determination Method for Intra Frames

After the introduction of the above three main components in the cost function, we propose a novel QP determination algorithm for intra frames. To obtain the best QP, we can substitute all possible QPs into Lagrangian cost function (3.2) and calculate the cost value of each QP. The optimized QP is the one with the largest cost value. In order to take PSNR deviation constraint into consideration, we propose that only QPs within the range of  $[QP_{i-1} - k, QP_{i-1} + k]$  are used to determine the best one, where  $QP_{i-1}$  is the one used for the previous intra frame encoding. Fig. 3-7 illustrates the proposed concept. Note that,  $k$  is set to 4 in this thesis.

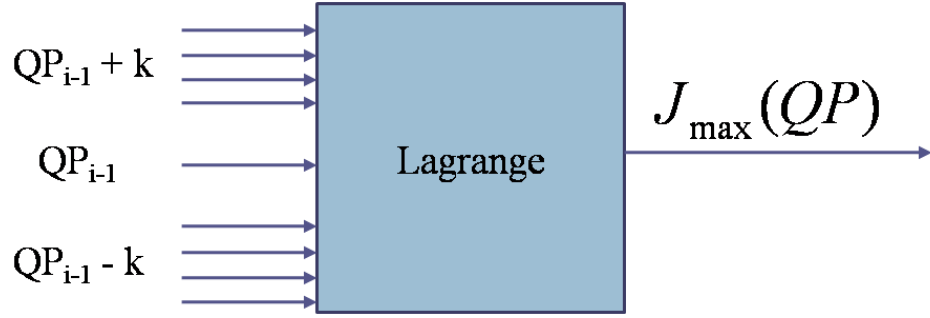


Fig. 3-7 Diagram of the proposed non SC<sub>I</sub> QP determination algorithm

## 3.2 Gradient Complexity based QP Determination for Scene Change Intra Frames

The gradient complexity based rate control scheme for scene change intra frames (SC<sub>I</sub> frames) is proposed in this section. First, we present a gradient based scene change detection algorithm, and then the QP determination method is described.

### 3.2.1 Gradient based Scene Change Detection

In order to prevent the buffer overflow problem caused by SC<sub>I</sub> frames, a scene change detection algorithm is essential. If a remarkable difference between consecutive frames can be detected by a appropriate metric which describes the frame characteristic perfectly, a scene transition can be declared whenever that metric exceeds a given threshold.

Various such metrics have been studied over years. In [5], Kim *et al.* have classified these frame complexity measures into four categories and suggested that the gradient based method is more reliable. According to Kim's research, we propose a gradient based scene change detection algorithm.

First, the pixel gradient at the location of  $(i, j)$  in the  $n^{\text{th}}$  frame is defined as

$$g_n(i, j) = |I(i, j) - I(i, j-1)| + |I(i, j) - I(i-1, j)| \quad (3.13)$$

where  $I(i, j)$  denotes the luminance value of the pixel at the location of  $(i, j)$ . And

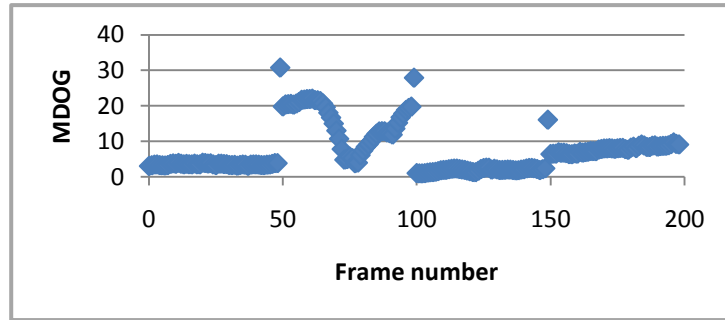
the frame complexity of  $n^{\text{th}}$  frame is measured as

$$G_n = \frac{1}{H \cdot W} \left( \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} g(i, j) \right) \quad (3.14)$$

where  $W$  and  $H$  are the horizontal and vertical dimensions of the frame, respectively. Then, the average gradient difference of the co-located pixel between consecutive frames, named mean difference of gradient ( $MDOG$ ), is given by

$$MDOG_n = \frac{1}{WH} \sum_{i=1}^H \sum_{j=1}^W |g_n(i, j) - g_{n-1}(i, j)| \quad (3.15)$$

Intuitively, the value of  $MDOG$  should be distinguishable while the scene change happens. Fig. 3-8 shows the  $MDOG$  values of a cascaded test sequence which is composed of Trevor, Stefan, Silent, and Coastguard sequences. There are three scene change frames at  $50^{\text{th}}$ ,  $100^{\text{th}}$ , and  $150^{\text{th}}$ , respectively. Although  $MDOG$  values at three scene change frames are relatively higher than their respective neighboring frames, a high-motion sequence would get an over estimated  $MDOG$  of non scene change frames due to its fast action. The second cut of the cascaded sequence, Stefan, is a classical high-motion example.

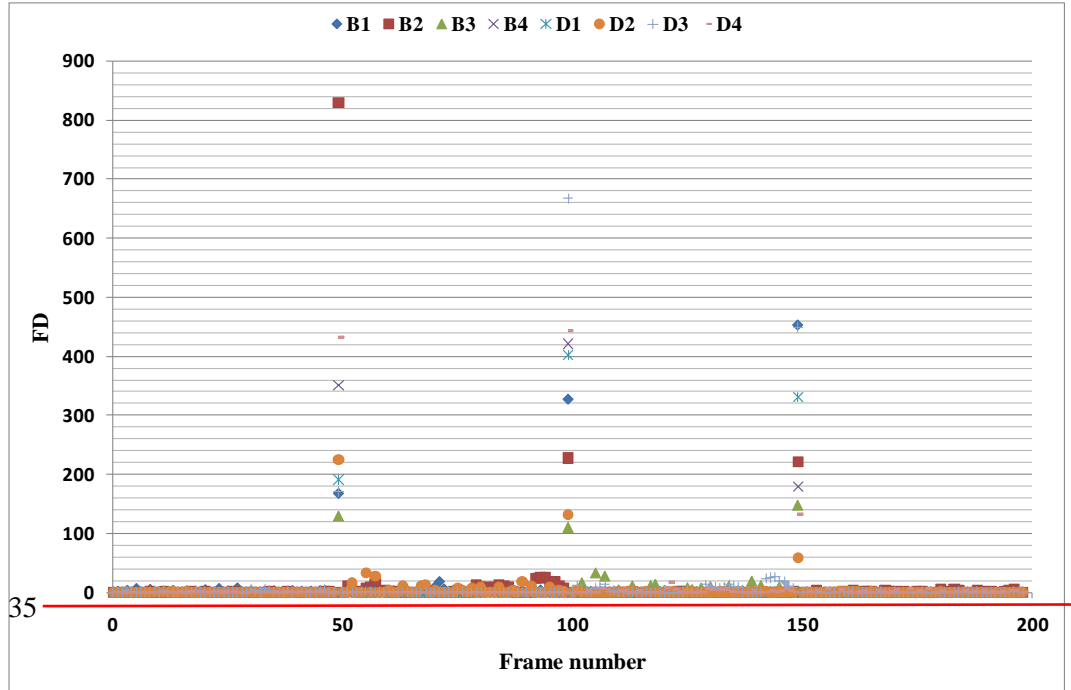


**Fig. 3-8**  $MDOG$  value of the QCIF sequence *Trevor-Stefan-Silent-Coastguard*

After carefully observing  $MDOG$  values of many test sequences, we found out that  $MDOG$  of the current frame is similar with that of the previous frame, even in a high-motion sequence. Based on the observation, we propose an improved  $MDOG$  metric, named frame distance ( $FD$ )

$$FD_n = |MDOG_n - MDOG_{n-1}| \times MDOG_n \quad (3.16)$$

Note that, the second term,  $MDOG_n$ , is a scalar used to dynamically enhance the effect of the first difference term while current frame is unlike the previous one. Fig. 3-9 shows  $FD$  values of eight cascaded test sequences and it indicates a threshold of 35 is a good choice to decide whether a scene transition occurs. Note that, the starting I-frame which is the first frame of the sequence is considered a scene change frame.



**Fig. 3-9** FDs of eight QCIF cascaded sequences when the threshold is set to 35

To demonstrate the correctness of the proposed scene change detection algorithm with  $FD$  threshold 35, experiments were conducted for two advertisement sequences with many scene transitions. The results are shown in Table 3-1 where  $N_{SC}$  is the number of scene change frames;  $N_C$  is the number of correct detection;  $N_m$  presents the number of miss detection, and  $N_f$  stands for the number of false alarms. It is obvious that most of scene change frames of both sequences are detected, and the number of false alarms is low.

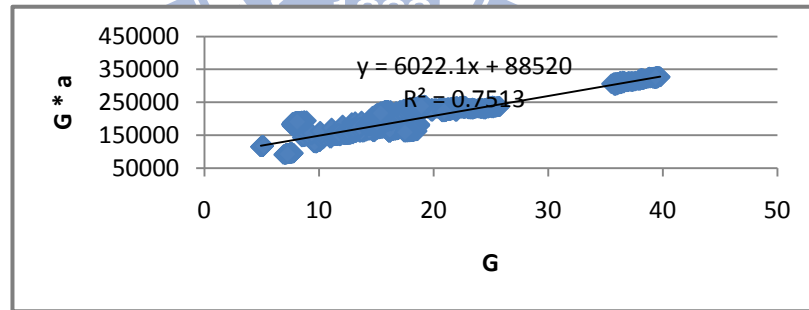
Sequences	Frames	$N_{SC}$	$N_C$	$N_m$	$N_f$
AD1	898	19	15	4	1
AD2	900	20	18	2	9

**Table 3-1 Detection correctness of two advertisements**

The proposed scene change algorithm is efficiency due to the low complexity gradient operation, and the value of pixel gradient is re-useable in the frame complexity measurement, shown in equation (3.14).

### 3.2.2 Gradient Complexity based Rate-QS Model

Because of the scene transition, the information from previous coded frames is not useful to predict the result of current  $SC_I$  frame. In order to solve the buffer overflow problem caused by  $SC_I$  frames, we propose a gradient complexity based rate-QS model for  $SC_I$  frames. The proposed model is based on Jing's rate-QS model (2.19), but it only takes information of current frame as predictor.



**Fig. 3-10 The relation curve between  $G$  and  $G*a$**

After analyzing data from over 3000 intra coded frames, we realize the relation between the gradient based frame complexity,  $G$  and the term of  $G \cdot a$  is closely linear. Fig. 3-10 shows this relation. Based on the collected data, we modelize the relation to a linear training line. Then, the original Jing's model can be written as

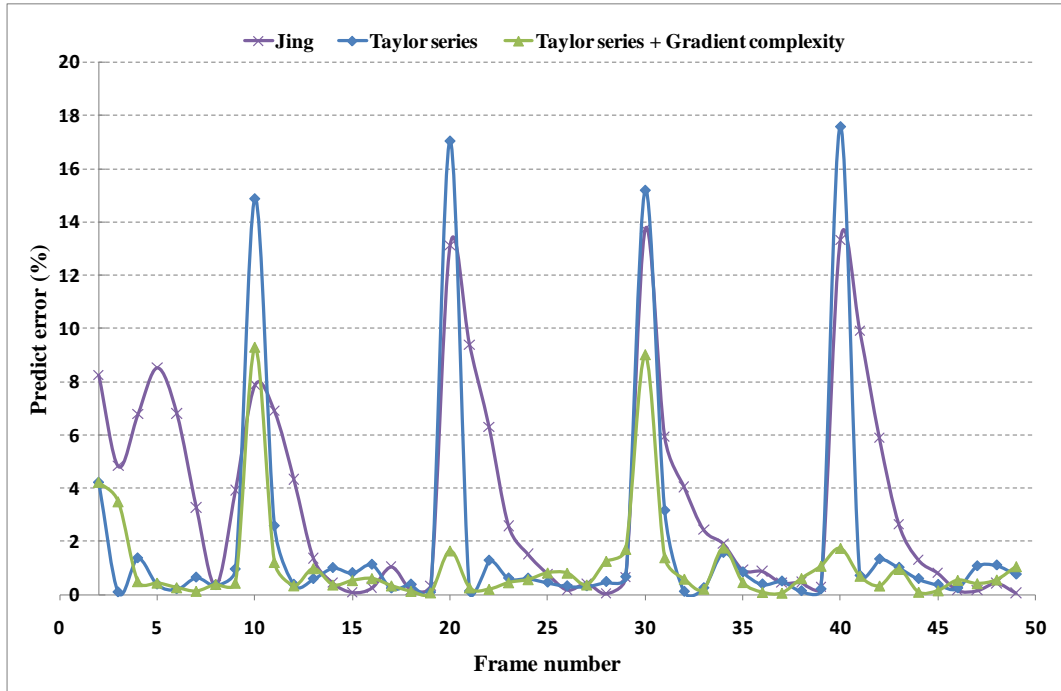
$$R = (\omega \cdot G_i + \mu) \times QS^b \quad (3.17)$$

where  $\psi = \{\omega, \mu, b\}$  is model parameter set. In QCIF sequences,  $\psi$  is set to

$\{6022.1, 88520, -0.76\}$ . According to the improved model (3.17), QS can be calculated with

$$QS_i = \sqrt[b]{\frac{R_t}{\omega \cdot G_i + \mu}} \quad (3.18)$$

, and QP can be derived from (2.17).

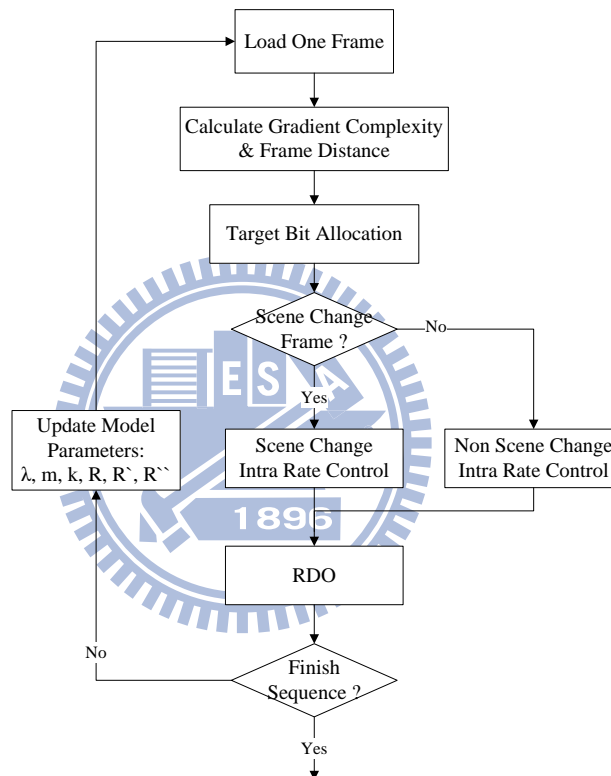


**Fig. 3-11 Prediction error comparison from a cascaded scene change sequence.**

To show the effects of taking account for scene changes in the rate control, Fig. 3-11 shows the prediction error for intra-only compression on a cascaded sequence with scene changes at frames 10, 20, 30, and 40. The rate control methods used for comparison include: Jing’s method, the proposed rate control method (the proposed method), the proposed method without scene change consideration (proposed w/o SC). The prediction error is calculated by equation (19). In Fig. 3-11, it is obvious to see that, compared with other two methods, the proposed method is more accurate on bit-rate prediction at the scene change frames.

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

With the scene change detection method, bit allocation for intra frames, and QP determination algorithms for both general intra frames and  $SC_I$  frames, the detailed block diagram of the proposed rate control algorithm for Intra-only compression is shown in Fig. 3-12. We summarize it with the following five steps.



**Fig. 3-12 Flow charts for Intra-only compression**

- Step 1. Calculate the gradient frame complexity,  $G$  and the frame distance,  $FD$  of the  $i^{\text{th}}$  frame using equation (3.14)–(3.16).
- Step 2. The intra frame bit allocation is calculated based on (3.1)
- Step 3. Detect whether the current frame is a scene change frame or not. If it is not a scene transition, initialize the Taylor series based rate-QS model and gradient based PSNR-QP model. Determine the optimized QP with the method mentioned

in section 3.1.4. On the other hand, calculate a appropriate QS using (3.18) and derive QP with (2.17) for scene change intra frames.

Step 4. Perform H.264 RDO for mode decision and the following coding procedures with the determined QP. After RDO procedure, update model parameters such as

$\lambda, m, k, R_{norm,i}, R'_{norm,i}$ , and  $R''_{norm,i}$  with the coding result of the  $i^{th}$  frame.

Step 5. Go to Step 1 until the end of the sequence.





# Chapter 4 Proposed Rate Control Algorithm for GOP Compression

For GOP compression, there are two kinds of frames, intra coded and inter coded frames. QP determination for intra coded frames is the same with that mentioned in the previous chapter. On the other hand, we adopt G012 algorithm for inter frame rate control. Because of the difference between both kinds of frames, we first propose a novel target bit allocation scheme for intra frames. Then, the overall description of the proposed rate control algorithm for GOP compression is presented.

## 4.1 Target Bit Allocation for Intra Frames

For GOP compression, the starting I-frame usually needs more bits and better quality for the following P-frames. The bit allocation for the first I-frame is calculated as


$$R_{t,0} = \kappa \cdot \frac{u}{F_r} \quad (4.1)$$

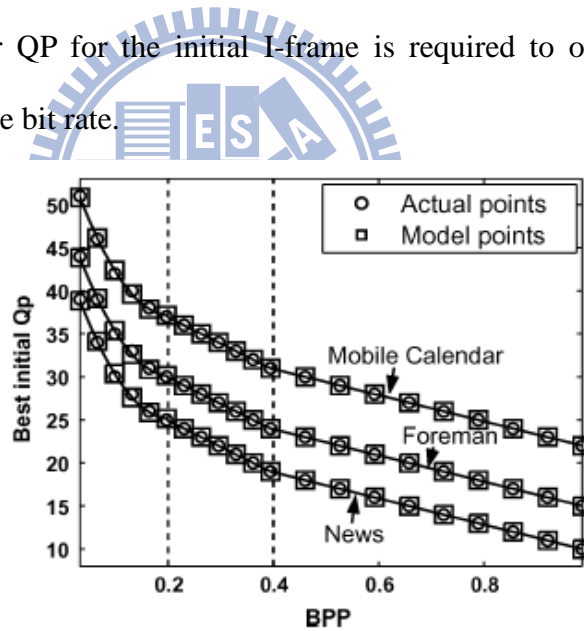
where  $R_{t,0}$  is the target bit of I-frame in the 0<sup>th</sup> GOP;  $u$  is the channel bit rate;  $F_r$  stands for the frame rate, and  $\kappa$  is a constant which is set to 8 experientially.

Since there are intra and inter coded frames in GOP compression, the relation between both is important for bit allocation. Yu's intra bit allocation formulas, (2.21) and (2.22) have several factors: average number of bits used in encoding previous frames, and average PSNR of previous frames. Yu tried to smooth the visual quality, but he did not take into account of frame complexity which is a significant factor for intra coded frames. Hence, we propose an improved intra bit allocation mechanisms for the I-frame in the  $i^{\text{th}}$  GOP:

$$R_{t,i} = R_{remain,i} \cdot \frac{W_{Intra,i}}{W_{Intra,i} + W_{Inter,i} \cdot N_p} \cdot \delta \quad (4.2)$$

$$\delta = \begin{cases} 1.8 & G \leq TH_1 \\ 1.6 & TH_1 < G \leq TH_2 \\ 1.4 & TH_2 < G \leq TH_3 \\ 1.2 & otherwise \end{cases}$$

where  $\varepsilon = \{TH_1, TH_2, TH_3\}$  is the threshold set;  $R_{t,i}$  is the target number of bits of I-frame in the  $i^{\text{th}}$  GOP.  $W_{Intra}$  and  $W_{Inter}$  are the weighting of intra frames and inter frames, respectively.  $\delta$  is an adaptive scalar depending on the complexity of current frame,  $G_n$ , and the threshold of  $\varepsilon$ . Wang *et al.*[21] proposed that the more complex sequences, the larger QP for the initial I-frame is required to obtain the best visual quality under the same bit rate.



**Fig. 4-1** Relation curves between the best initial QP and bpp for *News*, *Foreman*, and *Mobile*[21]

In Fig. 4-1, it is observed that large QPs are required for the initial I-frame of high complex sequences such as Mobile; while relatively small QPs are required for that of low complex sequences such as News and Foreman. Note that,  $BPP$  is calculated by equation (2.10) and actual points are determined by trying all possible QPs and recording the best one which results in the best R-D point. Based on this observation,

the scalar  $\delta$  is set adaptively, and the threshold set  $\varepsilon$  is set as below

$$\varepsilon = \{9.65, 15.59, 18.03\}$$

for QCIF sequences.

## 4.2 Description of Proposed Rate Control Algorithm for GOP Compression

Fig. 4-2 depicts the flow chart of the proposed rate control algorithm for GOP compression. It is similar with the original framework of G012, but QPs of both kinds of intra frames are determined by the proposed algorithm. We summarize it with the following five steps.

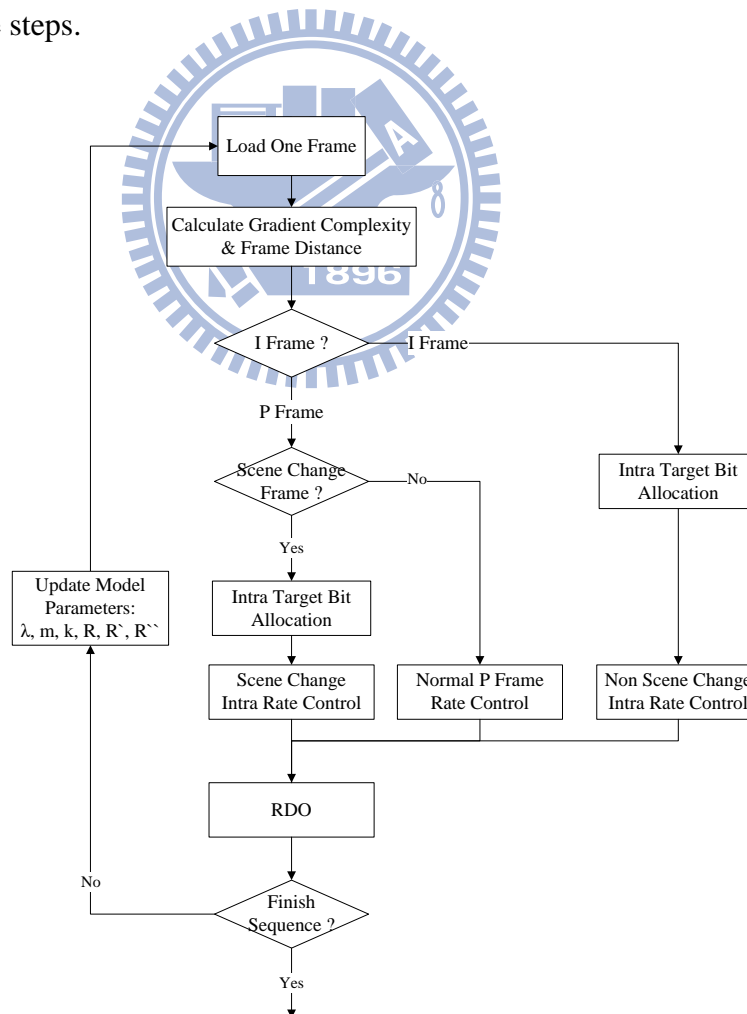


Fig. 4-2 Flow charts for GOP compression

- Step 1. This step is the same with the first step in Intra-only scheme.
- Step 2. If current frame is I-frame, allocate target bits using equation (4.2) and determine QP with the method in section 3.1. Then, perform H.264 RDO with the determined QP.
- Step 3. If current frame is P-frame, detect whether it is a scene change frame. If it is a  $SC_1$  frame, determine QP with the method in section 3.2. If not, the QP is calculated with the P-frame mode by G012 proposal. Then, H.264 RDO procedure is performed after the QP determination.
- Step 4. The updating stage is the same with the 4<sup>th</sup> step in Intra-only compression.
- Step 5. Go to Step 1 until the end of the sequence.



## Chapter 5 Experiment Results

The proposed rate control algorithm is integrated into the latest JVT reference software JM15.0[12]. The simulation was conducted with the first 200 frames of four standard QCIF test sequences, including “Carphone”, “Foreman”, “Mobile”, and “News”. 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.

In Intra-only compression, we compare the proposed algorithm with Jing’s method[17] and JM15.0 Intra-only rate control algorithm which is a modified version based on G012[11]. In GOP compression, each sequence is coded at 30 fps by GOP size 40 with structure IPPP. The JM G012 algorithm and Yan’s method[19] were selected as comparison references. MV resolution is 1/4 pixel precision with 32-pixel-length search window and the number of reference frame is set to 1. In both compression schemes, CAVLC and RDO are enabled, and the size of basic unit is set to 99. All parameters are selected equivalently for all algorithms.

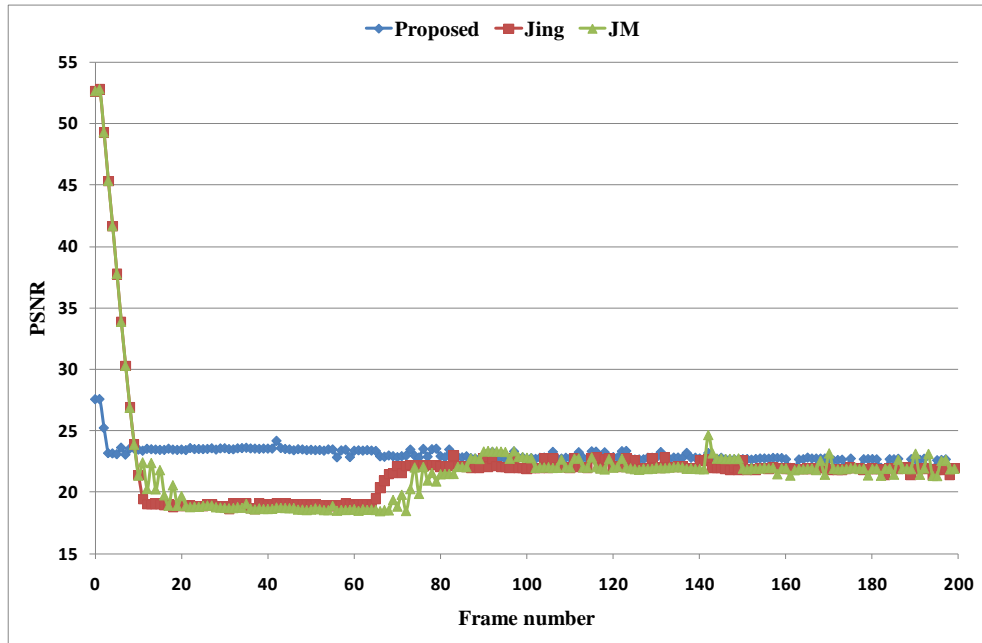
### 5.1 Results of Intra-only Compression

In Intra-only compression, Table 5-1 summarizes the overall performance results including actual bit rate, average PSNR, and PSNR deviation. Due to the accurate prediction of the proposed models, Lagrangian-optimization based QP determination for intra frames and gradient complexity based QP determination for  $SC_1$  frames, our approach can decide appropriate QPs for not only general intra frames but also scene change intra frames. The proposed algorithm is cable of increasing average PSNR by up to 0.99 dB (0.41 dB on average) and 0.97 dB (0.35 dB on average) compared to JM and Jing’s algorithm, respectively. In addition, PSNR deviation is reduced by up to 87%

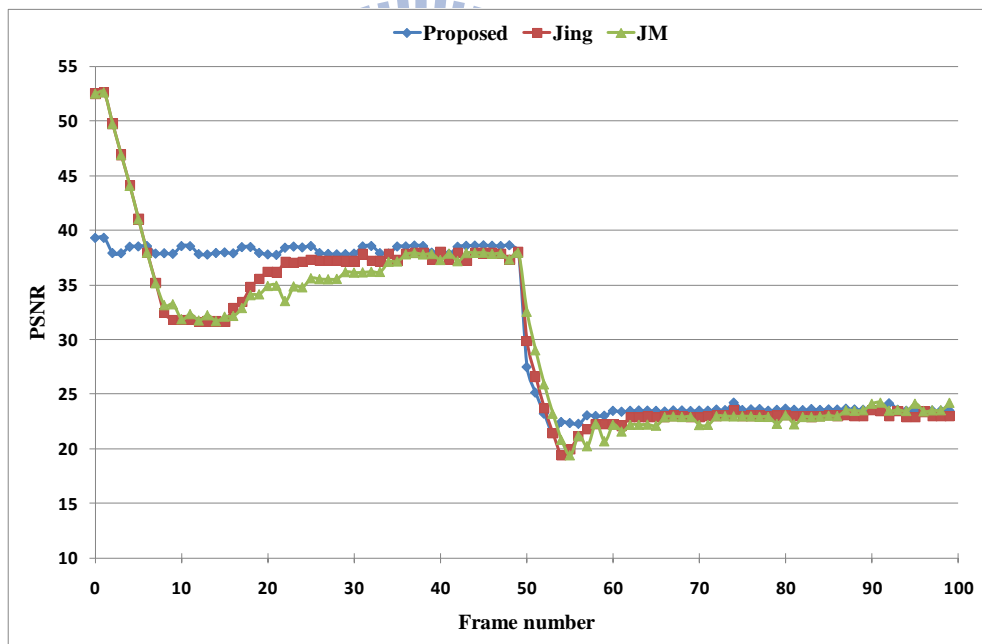
(51% on average) and 86% (49% on average) in contrast with JM and Jing's algorithm, respectively. Although, the mismatches of real bit rate and target bit rate among three methods are close, the proposed algorithm slightly reduce the mismatch compared to other two schemes.

**Table 5-1 Performance comparisons for Intra-only scheme**

Sequences	Target	Bit Rate			Avg. PSNR (db)			PSNR StDev		
		JM	Jing	Proposed	JM	Jing	Proposed	JM	Jing	Proposed
256 kbps	Foreman	255.80	256.07	256.16	29.67	29.75	30.13	2.470	2.276	0.888
	Mobile	275.15	275.15	256.16	19.05	19.05	19.48	3.367	3.367	0.555
	News	255.60	256.09	256.09	27.61	27.66	28.08	2.677	2.546	0.415
	Carphone	255.66	255.99	256.00	31.20	31.22	31.51	2.170	2.140	0.984
	Combo1	255.95	256.05	256.33	26.84	26.94	27.36	3.716	3.673	2.912
	Combo2	261.86	256.25	256.39	25.90	25.68	26.30	6.944	6.777	6.306
	<b>Average</b>	<b>260.00</b>	<b>259.27</b>	<b>256.19</b>	<b>26.71</b>	<b>26.72</b>	<b>27.14</b>	<b>3.557</b>	<b>3.463</b>	<b>2.010</b>
512 kbps	Foreman	511.40	512.11	512.14	34.66	34.71	35.13	2.931	2.849	0.924
	Mobile	511.12	512.18	512.14	22.03	22.05	23.02	4.846	4.784	0.651
	News	511.09	512.20	512.13	33.06	33.11	33.64	3.232	3.140	0.453
	Carphone	510.64	512.02	512.07	36.38	36.42	36.87	2.692	2.673	1.037
	Combo1	510.86	512.13	512.21	30.68	30.79	31.23	4.485	4.372	3.179
	Combo2	511.67	512.33	512.15	30.07	30.21	30.87	7.820	8.020	7.412
	<b>Average</b>	<b>511.13</b>	<b>512.16</b>	<b>512.14</b>	<b>31.15</b>	<b>31.22</b>	<b>31.79</b>	<b>4.334</b>	<b>4.306</b>	<b>2.276</b>
768 kbps	Foreman	766.40	768.08	767.95	38.03	38.07	38.24	2.256	2.205	0.929
	Mobile	767.57	768.21	767.89	25.16	25.21	25.64	4.291	4.204	0.786
	News	765.92	767.93	768.12	37.12	37.17	37.46	2.437	2.361	0.435
	Carphone	766.17	767.81	768.12	39.98	40.01	40.22	2.106	2.025	1.121
	Combo1	767.91	768.05	767.97	33.71	33.85	34.07	4.315	4.024	3.282
	Combo2	766.95	768.00	768.58	33.54	33.68	33.95	8.377	8.354	7.864
	<b>Average</b>	<b>766.82</b>	<b>768.01</b>	<b>768.11</b>	<b>34.59</b>	<b>34.67</b>	<b>34.93</b>	<b>3.964</b>	<b>3.862</b>	<b>2.403</b>
1024 kbps	Foreman	1021.48	1023.95	1024.26	40.47	40.51	40.58	1.904	1.771	0.942
	Mobile	1022.66	1023.90	1024.00	27.48	27.50	27.83	3.826	3.787	0.869
	News	1021.22	1024.17	1023.80	40.14	40.20	40.32	1.934	1.761	0.509
	Carphone	1021.13	1024.14	1023.95	42.63	42.67	42.76	1.728	1.620	1.117
	Combo1	1022.37	1024.22	1023.98	36.16	36.26	36.36	3.998	3.759	3.235
	Combo2	1022.36	1023.9	1023.90	36.09	36.28	36.43	8.636	8.533	8.160
	<b>Average</b>	<b>1021.87</b>	<b>1024.05</b>	<b>1023.98</b>	<b>37.16</b>	<b>37.24</b>	<b>37.38</b>	<b>3.671</b>	<b>3.539</b>	<b>2.472</b>



(a)

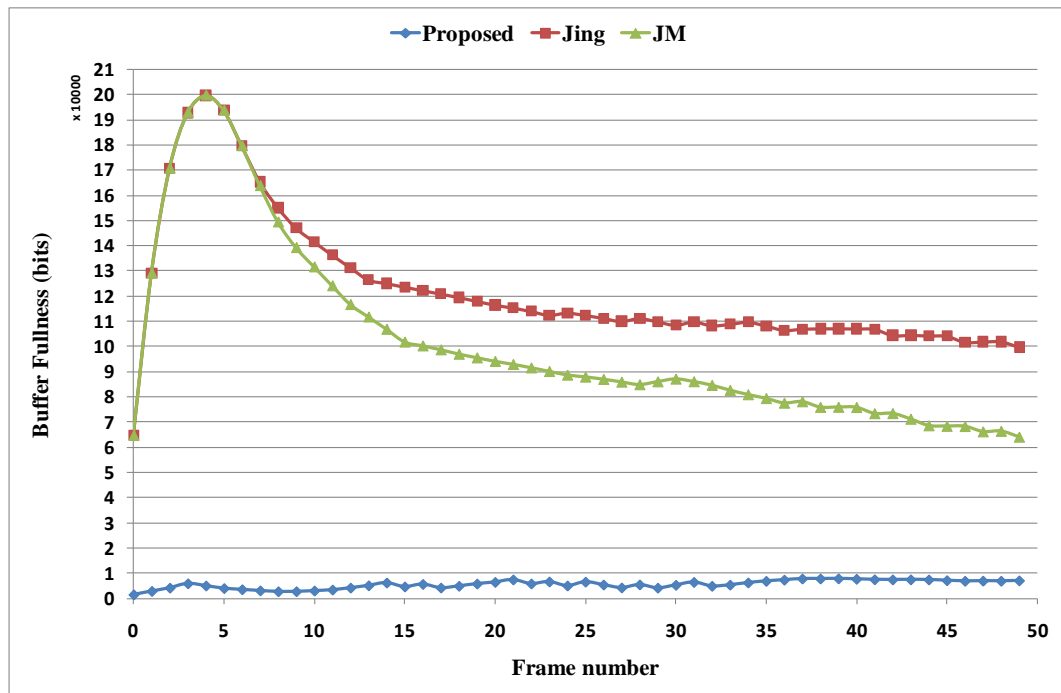


(b)

**Fig. 5-1 PSNR v.s. frames for (a) *Mobile@512kbps* (b) *Combo2@512kbps***

For further evaluation, the curves of PSNR versus frames for two test cases are shown in Fig. 5-1. From the plot (a), it is observed that the proposed algorithm can maintain a consistent video quality in contrast with other two 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 (50<sup>th</sup> frame), so the quality of the following frames is more stable and higher than JM and Jing’s algorithm. Fig. 5-2 and Fig. 5-3 show the buffer occupancy versus frames for two test cases. The proposed algorithm shows superior performance by achieving a 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.



**Fig. 5-2** Buffer fullness v.s. frames for *Foreman*@1024 kbps



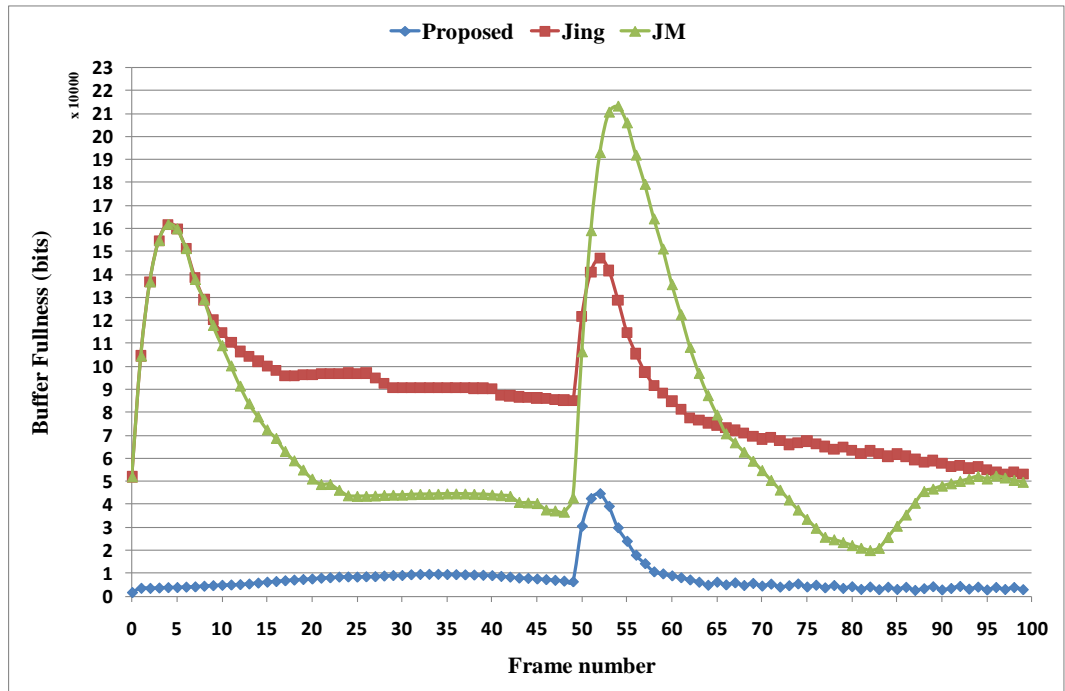


Fig. 5-3 Buffer fullness v.s. frames for *Combo2@768kbps*

## 5.2 Results of GOP Compression

Table 5-2 summarizes the overall performance results of four standard test sequences. The proposed algorithm selects appropriate QPs for both kinds of intra frames to save bit budget for the following P-frames. The proposed algorithm achieves average PSNR gain up to 0.99 dB (0.18 dB on average) and 2.47 dB (0.31 dB on average) compared to JM and Yan's algorithm, respectively. In addition, PSNR deviation is reduced by up to 50% (12% on average) and 59% (8% on average) in contrast with JM and Yan's algorithm, respectively.

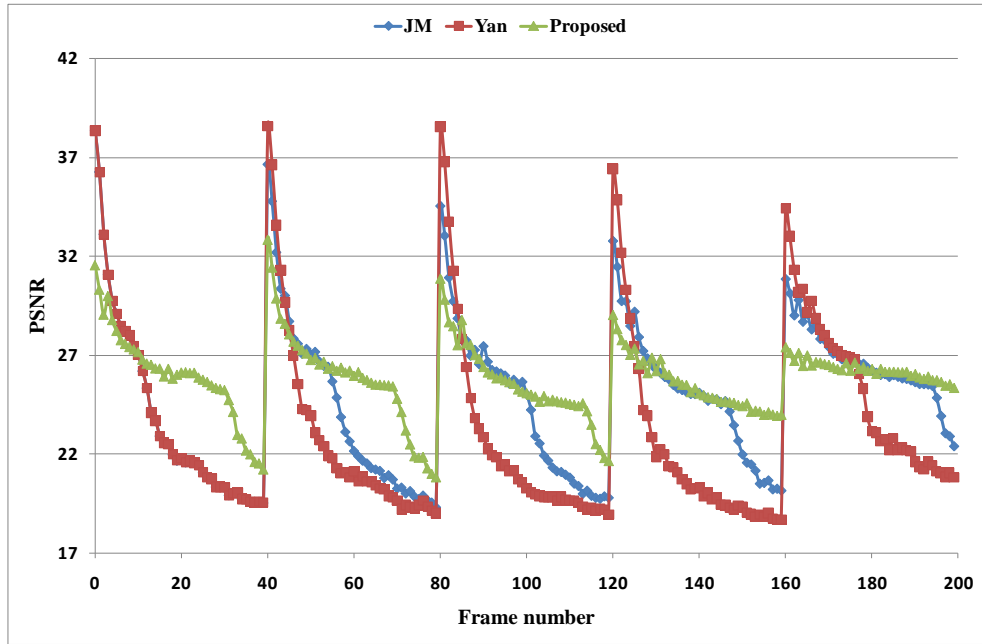
For scene change sequences, Table 5-3 shows the performance results of three methods. For scene transition, there is a trade-off between achieving accurate bit rate control and keeping stable visual quality. In Table 5-3, although the average PSNR deviation of the proposed algorithm is slightly higher than that of Yan's method, the proposed scheme have better performance in average PSNR and bit rate mismatch.

Target Rate	Sequences	Bit Rate			Avg. PSNR (db)			PSNR StDev		
		JM	Yan	Proposed	JM	Yan	Proposed	JM	Yan	Proposed
32 kbps	Carphone	33.33	33.74	33.46	30.96	31.05	31.04	3.164	3.467	3.138
	Foreman	33.17	33.61	33.70	27.91	27.95	27.90	2.687	2.924	2.493
	Mobile	34.73	33.64	32.93	20.89	20.67	21.10	2.284	2.216	1.592
	News	32.35	32.56	32.47	30.97	31.00	30.81	1.837	1.939	1.634
	<b>Average</b>	<b>33.4</b>	<b>33.39</b>	<b>33.14</b>	<b>27.68</b>	<b>27.67</b>	<b>27.71</b>	<b>2.493</b>	<b>2.637</b>	<b>2.214</b>
48 kbps	Carphone	49.11	49.69	49.00	32.93	33.01	32.98	2.813	2.916	2.936
	Foreman	49.02	48.83	49.01	30.28	30.3	30.30	2.478	2.028	2.61
	Mobile	48.86	49.58	48.38	22.04	22.91	22.91	2.767	2.964	1.459
	News	48.42	48.26	48.29	33.16	33.17	33.17	1.439	0.701	1.223
	<b>Average</b>	<b>48.85</b>	<b>49.09</b>	<b>48.67</b>	<b>29.6</b>	<b>29.51</b>	<b>29.84</b>	<b>2.374</b>	<b>2.152</b>	<b>2.057</b>
64 kbps	Carphone	64.47	64.92	64.65	34.23	34.22	34.33	2.592	2.565	2.555
	Foreman	64.7	64.43	64.48	31.87	31.67	31.89	2.142	1.64	2.204
	Mobile	64.33	64.49	64.33	23.53	23.91	24.04	2.732	2.672	1.555
	News	64.45	64.25	64.12	34.68	34.35	35.09	1.010	1.043	1.210
	<b>Average</b>	<b>64.49</b>	<b>64.52</b>	<b>64.4</b>	<b>31.08</b>	<b>31.04</b>	<b>31.34</b>	<b>2.119</b>	<b>1.98</b>	<b>1.881</b>
96 kbps	Carphone	96.16	97.1	96.66	36.16	36.12	36.14	2.654	2.981	2.669
	Foreman	96.02	96.11	96.10	34.09	34.02	34.10	1.844	2.25	1.876
	Mobile	96.48	97.58	96.17	24.81	23.33	25.80	3.851	4.689	1.926
	News	96.02	96.06	96.93	37.9	38.03	37.74	1.468	1.502	1.277
	<b>Average</b>	<b>96.17</b>	<b>96.71</b>	<b>96.47</b>	<b>33.24</b>	<b>32.88</b>	<b>33.45</b>	<b>2.454</b>	<b>2.856</b>	<b>1.937</b>

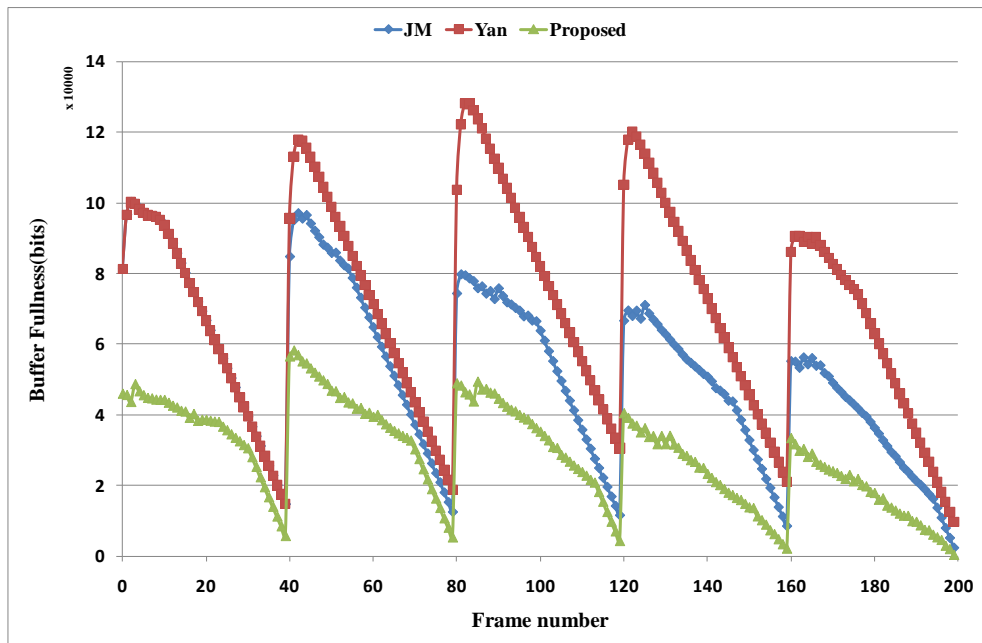
**Table 5-2 Result comparisons of normal sequences for GOP compression scheme**

Target Rate	Sequences	Bit Rate			Avg. PSNR (db)			PSNR StDev		
		JM	Yan	Proposed	JM	Yan	Proposed	JM	Yan	Proposed
150 kbps	Combo1	150.61	150.63	150.45	32.74	32.32	32.64	4.359	4.923	4.426
	Combo2	185.31	164.64	157.34	35.29	35.49	35.76	10.739	10.116	10.763
	<b>Average</b>	<b>167.96</b>	<b>157.64</b>	<b>153.9</b>	<b>34.02</b>	<b>33.91</b>	<b>34.2</b>	<b>7.549</b>	<b>7.52</b>	<b>7.595</b>
200 kbps	Combo1	200.77	200.66	200.71	34.09	34.09	34.15	4.761	4.761	4.813
	Combo2	206.80	218.82	203.99	35.79	36.79	37.13	11.224	10.081	10.918
	<b>Average</b>	<b>203.79</b>	<b>209.74</b>	<b>202.35</b>	<b>34.94</b>	<b>35.44</b>	<b>35.64</b>	<b>7.993</b>	<b>7.421</b>	<b>7.866</b>
250 kbps	Combo1	249.97	250.54	250.72	35.25	35.03	35.26	5.103	5.565	5.175
	Combo2	288.95	271.14	265.60	37.40	38.13	38.09	11.921	10.571	11.491
	<b>Average</b>	<b>269.46</b>	<b>260.84</b>	<b>258.16</b>	<b>36.33</b>	<b>36.58</b>	<b>36.68</b>	<b>8.512</b>	<b>8.068</b>	<b>8.333</b>

**Table 5-3 Result comparisons of scene change sequences for GOP compression scheme**



(a)



(b)

**Fig. 5-4 (a) PSNR v.s. frames (b) Buffer fullness v.s. frames for *Mobile@96kbps***

Fig. 5-4 shows the curves of PSNR and buffer fullness versus frames for mobile at 96 kbps. The plot (a) illustrates the visual quality within one GOP of the proposed algorithm is progressively stable with time. Because the proposed intra frames bit allocation method takes the qualities of I-frame and P-frames in the previous GOP into consideration, the PSNR deviation between I-frame and P-frames is gradually reduced

in the next GOP. The plot (b) shows the buffer occupancy of the proposed algorithm is more stable and lower than other algorithms.



## Chapter 6 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, we propose Taylor expansion based rate-QS model and gradient complexity based PSNR-QP model. The cost value of each candidate QP is calculated to determine the optimized QP. For  $SC_I$  frames, a gradient complexity based rate-QS model is proposed to determine appropriate QPs.

The simulation results show our approach is adequate not only for Intra-only compression but also for GOP compression. The proposed algorithm is cable of achieving an average of 0.41 dB and 0.18 dB PSNR gain compared to JM rate control algorithm for Intra-only compression and GOP compression, respectively. In contrast with Jing's and Yan's algorithm, our scheme has an average of 0.35 dB and 0.31 dB PSNR gain for Intra-only compression and GOP compression, respectively. Our proposal also has better performance in buffer fullness and bit rate mismatch control. Besides, the proposed algorithm is flexible to be integrated into other rate control algorithms which focus on inter frame issues.

## 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  $p \times 64$  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](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] A. N. Netravali and J. O. Limb, “Picture coding: A review,” *Proc. IEEE*, vol. PROC-68, no. 3, pp. 7–12, Mar. 1960.
- [16] 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.
- [17] 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.
- [18] Y. Sun and I. Ahmad, “A robust and adaptive rate control algorithm for object-based video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, pp. 1167-1182, Oct. 2004.
- [19] B. Yan and M. Wang, “Adaptive distortion-based intra-rate estimation for H.264/AVC rate control,” *IEEE Signal Process. Lett.*, vol. 16, pp. 145–148, 2009.
- [20] W. J. Kim, J. W. Yi, and S. D. Kim, “A bit allocation method based on picture activity for still image coding,” *IEEE Trans. Image Process.*, vol. 8, no. 7, pp.

974–977, Jul. 1999.

- [21] H. Wang, and S. Kwong, “Rate-Distortion Optimization of Rate Control for H.264 With Adaptive Initial Quantization Parameter Determination,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, pp. 140-144, 2008
- [22] Taylor series, [http://en.wikipedia.org/wiki/Taylor\\_series](http://en.wikipedia.org/wiki/Taylor_series)

