THE UNIVERSITY OF
WARWICK

Research Report
(CS-RR-423)

IMPROVED SCHEMES FOR MULTI-MODE CODING IN THE H.264/AVC STANDARD AND SIMPLIFIED LAGRANGIAN EVALUATION FOR VIDEO CODING

Andy C. Yu

Department of Computer Science
University of Warwick, Coventry CV4 7AL, UK

Supervisor: Dr. Graham Martin
February 2006

IMPROVED SCHEMES FOR MULTI-MODE CODING IN THE H.264/AVC STANDARD AND SIMPLIFIED LAGRANGIAN EVALUATION FOR VIDEO CODING

by

Andy C. Yu

ABSTRACT

This report consists of two parts. In the first part, an efficient algorithm for inter-frame coding in the H.264/AVC standard is extended to provide more significant speedup in computational performance for sequences containing high spatial correlation and motion. The proposed scheme features a more sophisticated search process and robust predictions to achieve better PSNR-rate performance for a large range of compression levels. Extensive simulation results demonstrate speedups of between 41% and 68%, with no noticeable deterioration in picture quality or compression ratio, even for the coding of complex video sequences. Next, the report proceeds to examine the properties of Lagrangian evaluation in the case of high Qp. Supported by extensive statistics and experiments, a cost evaluation is then proposed especially for real-time video coding operating at very low bit rates. Simulation results show that the proposed cost evaluation is able to achieve reductions in terms of coding time and bits spent.

Keywords: H.264/AVC, multi-mode selection, Lagrangian rate-distortion optimisation.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

Chapter 1
Overview of texture coding in H.264/AVC

1.1 Introduction

Multi-mode coding has proven to be efficient and is widely employed in many video coding standards. A cost evaluation based on residue energy for mode selection was recommended by Test Model 5 (TM5) [1]. However, this strategy does not promise an optimal coding performance due to the inconsideration of rate constraint. Later, Wiegand et al. [2] addressed the aforementioned problem and further combined rate-distortion optimisation with the H.263 standard. To this end, Lagrangian cost evaluation becomes a mainstream method for cost measurement for macroblock-based mode selection. Furthermore, it is adapted in the recent H.264/Advance Video Coding (AVC) standard to select optimal modes for both intra- and inter-macroblocks. According to the definition, Lagrangian optimisation is a process to search the best combination from a set of available coding options (modes) $\mathbf{m} = (m_1, \ldots, m_k)$ for a group of blocks (GOB) $\mathbf{B} = \{B_1, \ldots, B_N\}$. The aim is to minimise the entire distortion of the GOB at a given constraint rate, R':

$$\min_{\mathbf{m}} D(\mathbf{B}, \mathbf{m}) \quad \text{subject to } R(\mathbf{B}, \mathbf{m}) \le R' \tag{1-1}$$

Where $D(\mathbf{B}, \mathbf{m})$ and $R(\mathbf{B}, \mathbf{m})$ reflect the total distortion and rate costs of a GOB, respectively. (1-1) can be solved with an unconstrained framework [3, 10] as

$$\min_{\mathbf{m}} \sum_{i=0}^{N} \{D(\mathbf{B}, \mathbf{m}) + \lambda . R(\mathbf{B}, \mathbf{m})\} \tag{1-2}$$

where λ is a Lagrangian parameter. The sum of $D(\mathbf{B}, \mathbf{m})$ and weight $R(\mathbf{B}, \mathbf{m})$ is known as the Lagrangian cost.

However, (1-2) is computationally expensive for any practical video implementation. Thus, a simplified form of optimisation was proposed by Shoham et al. [4] to search independently the best mode for each block.

$$\sum_{i=0}^{N} \min_{\mathbf{m}} \{L(B_i, \mathbf{m})\} \tag{1-3}$$

where $L(B_i,\mathbf{m})$ is the Lagrangian cost for the i^{th} block

$$L(B_i, \mathbf{m}) = D(B_i, \mathbf{m}) + \lambda . R(B_i, \mathbf{m}) \tag{1-4}$$

and the Lagrangian parameter, λ, is approximated to 0.85 [5]. In the H.264/AVC standard [6], a non-linear relationship between λ and quantisation factor, Qp, is recommended as

$$\lambda = 0.85 \times 2^{(\mathrm{Qp}-12)/3} \tag{1-5}$$

In addition, $D(\mathbf{B}, \mathbf{m})$, the distortion measurement, which quantifies the difference between $\mathbf{B}_t$ and $\mathbf{B}'_{t-\tau}$, the reconstructed block taken from the $(t\text{-}\tau)^{\text{th}}$ frame, is separately defined in intra and inter cases as:

$$D(\mathbf{B}_t, \mathbf{m}_{\text{intra}}) = \sum_x \sum_y \left| \mathbf{B}_t(x,y) - \mathbf{B}_t'(x,y,\mathbf{m}_{\text{intra}}) \right| \tag{1-6}$$

$$D(\mathbf{B}_t, \mathbf{m}_{\text{inter}}) = \sum_x \sum_y \left| \mathbf{B}_t(x,y) - \mathbf{B}'_{t-\tau}(x+m_x, y+m_y, \mathbf{m}_{\text{inter}}) \right| \tag{1-7}$$

Where (m_x, m_y) is the motion vector; $R(\mathbf{B}_i, \mathbf{m})$ in (1-4) further reflects the number of bits associated with choosing the mode and *Qp* including the bits for the macroblock header, the motion vector(s) and all the DCT residue blocks. $\mathbf{m}$ indicates a mode chosen from the set of potential prediction modes, the respective possibilities of which are:

$$\mathbf{m}_{\text{intra}} \in \{I4MB, I16MB\}, \tag{1-8}$$

$$\mathbf{m}_{\text{inter}} \in \{SKIP, I4MB, I16MB, INTER\}. \tag{1-9}$$

Intra-frame mode has two modes, *I4MB* and *I16MB*. *I4MB* consists of 9 members which extrapolate elements of a 4×4 block with the neighbouring encoded pixels in 8

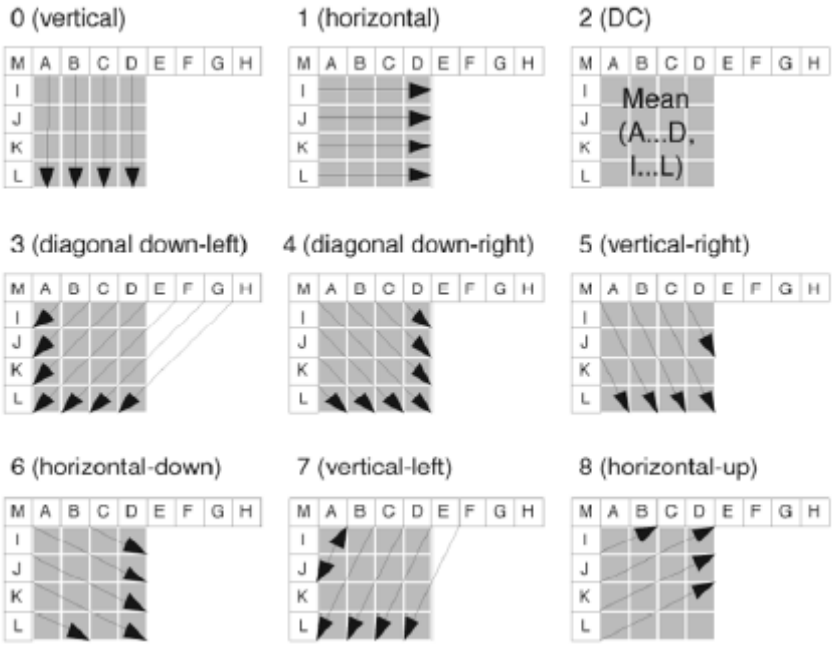

Fig. 1-1 *I4MB* consists 9 approaches to extrapolate elements of a 4×4 block with the neighbouring encoded pixels

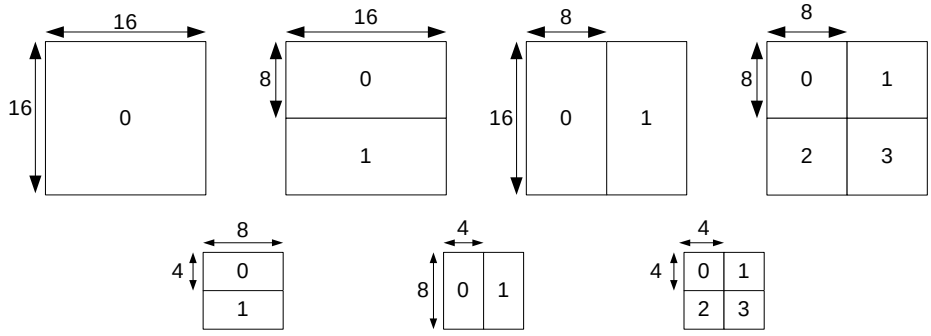

Fig. 1-2 *INTER* modes with 7 different block sizes ranging from 4×4 to 16×16.

directions as depicted in Fig. 1-1. For instance, *VERT*, the vertical member, extrapolates a 4×4 block vertically with 4 neighbouring pixels, *A*, *B*, *C*, *D*, whereas the horizontal member, *HORT*, utilizes the horizontal adjacent pixels to *I*, *J*, *K*, *L* to do the prediction. The other modes operate the same way according to their corresponding orientations, except for *DC*, the directionless member, which extrapolates all pixels with $(A+B+C+D+I+J+K+L)/8$. *I16MB* resembles *I4MB* but is less time- consuming, comprising 4 members to

predict a 16×16 macroblock as a whole. As for the inter-frame mode, it contains *SKIP* (direct copy), *I4MB*, *I16MB*, and *INTER*, the most time-consuming mode, which consists of 7 different block sizes as shown in Fig. 1-2.

In intra-frame coding, the final mode decision is selected by the member (either from *I4MB* or *I16MB*) that minimizes the Lagrangian cost in (1-1). In inter-frame coding, motion estimations with 7 different block-size patterns, as well as the other members in three members (*I4MB*, *I16MB*, and *SKIP*), are calculated. The final decision is determined by the mode that produces the least Lagrangian cost among the available modes. Currently, the H.264/AVC standard employs a brute force algorithm to search through all the possible candidates and its corresponding members to find an optimum motion vector [6]. Since the exhaustive search method is employed in all the modes to acquire a final mode decision, the computational burden of the search process is far more significant than any existing video coding algorithm.

1.3 Contributions and organisation of this report

The contributions of this report are to develop fast mode selection algorithms to reduce the computational time for multi-mode coding. This report extends the previously proposed algorithm to provide more significant speedup in computational performance for sequences containing high spatial correlation and motion. The details of the search process are included in the next chapter. In chapter 3, we examine the properties of Lagrangian evaluation in the case of high Qp factors to derive a cost evaluation for the real-time video coding operating at very low bit rates. Finally, some conclusions are drawn.

Chapter 2
Improved schemes for inter-frame coding in H.264/AVC standard

2.1 Introduction

In previous work, a fast algorithm, MFinterms, [7] has been proposed to alleviate inter-frame encoder complexity, while still maintaining picture quality and coding efficiency at certain compression levels. Success of the MFinterms algorithm is achieved by discriminating two kinds of encoded macroblocks: (a) macroblocks encoded with SKIP mode; (b) macroblocks encoded by inter-modes with the larger decomposed partition sizes (greater than 8x8 pixels). The MFInterms algorithm assigns different categories of inter-modes to each macroblock with a complexity measurement defined as follows:

$$\text{Complexity} = \begin{cases} \text{high}, & \ln(E_{AC})/15.25 > Th_c \\ \text{low}, & \ln(E_{AC})/15.25 < Th_c \end{cases} \qquad (2\text{-}1)$$

where Th_c is a spatial complexity threshold and E_{AC} is the total energy of the high-frequency component (AC component) in the current macroblock. If the result in (2-1) is low, the algorithm checks those inter modes with partition sizes of larger than 8x8. Otherwise, the macroblock is decomposed into four 8-by-8 blocks and a measurement of block-based motion consistency is made to determine whether inter-modes with a smaller partition size are required.

The MFInterms algorithm was shown to achieve a significant reduction in computation time [7]. However, the performance gains were dependent on the spatial content and extent of motion in the test sequences, for a limited range of compression levels. In this chapter, we propose improved schemes for the MFinterms algorithm

which achieve a better PSNR-rate performance and a reduced computational requirement for any coding condition.

The remainder of this chapter is organized as follows. In Section 2.2, the improved schemes for the MFinterms algorithm are presented. Extensive simulation results in terms of PSNR-rate performance and algorithm efficiency are presented in Section 2.3.

2.2. The proposed schemes for the MFinterms algorithm

The proposed scheme contains a more sophisticated search process and robust predictions in selecting an optimal inter-mode for each macroblock. The scheme is described as comprising three levels. Each level targets a different category of inter-modes according to the complexity of the search processes. The following subsections introduce these three levels.

2.2.1. The first level of the proposed algorithm

The first level of the proposed algorithm distinguishes the macroblocks encoded with SKIP (denoted as skipped macroblocks). Fig. 2-1 illustrates the detailed flowchart for this level. The skipped macroblocks can be found in P-frames where the pixel information is almost identical to that in the corresponding same position in the previous frame. Thus, they can be detected by means of the temporal similarity

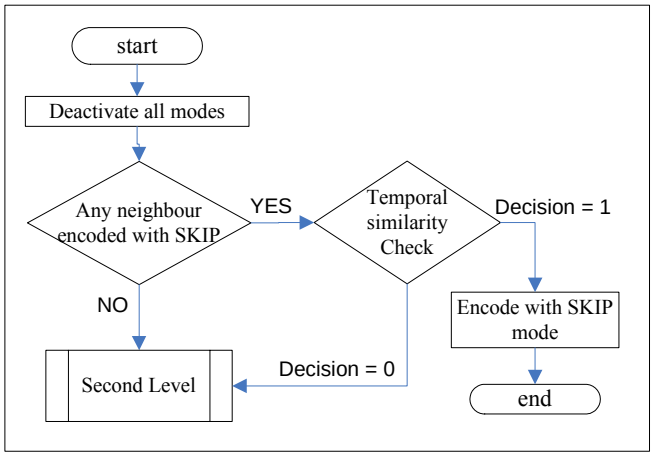

Fig. 2-1 Block diagram of the 1st level of the improved scheme.

between the two macroblocks. As skipped macroblocks tend to occur in clusters, such as in a patch of static background, we suggest that the current macroblock undergoes temporal similarity detection if at least one of two possible valid skipped macroblocks is found above or to the left of the current macroblock. The decision for the temporal similarity detection is defined as

$$\text{Decision} = \begin{cases} 0 & \text{if } \text{SAD}_{\text{current}}(t;t-1) > \text{x}_{\text{SAD}} \\ 1 & \text{otherwise} \end{cases} \qquad (2\text{-}2)$$

where $\text{SAD}_{\text{current}}(t;t-1)$ is the sum absolute difference (SAD) of the macroblocks in frames t and $t-1$; x_{SAD} is the SAD of the available skipped neighbour. An average of the SAD values is computed if both skipped neighbours are valid. A non-zero outcome in (2-2) indicates that the current macroblock is encoded with SKIP only. Otherwise, further examinations in the second level are required. The algorithm for this level is summarised as follows:

1^o Deactivate all inter-modes.

2^o Determine whether one of the encoded neighbours located above or to the left is a skipped macroblock. Execute 3^o if the situation is true. Otherwise, proceed to next level.

3^o Perform the temporal similarity test described in (2-2). If a non-zero decision is obtained, encode the current macroblock with SKIP mode and record its SAD value. Otherwise, further examinations in the next level are required.

2.2.2. The second level of the proposed scheme

Fig. 2-2 illustrates the flowchart of the second level. This level targets those macroblocks encoded with inter modes with a large partition size (8×16, 16×8, and 16×16 pixels). The general tendency that inter modes with large partition sizes are more suitable for the encoding of homogeneous content has been verified by a number of authors [7, 8]. The reasons given are that: (a) homogeneous macroblocks tend to contain fewer moving features requiring multiple motion descriptors, and (b) owing to the homogeneous content, the distortion costs arising from incorrect predictions are often insignificant.

The spatial complexity measurement in (2-1) is employed to determine if the current macroblock requires further examination by inter-modes with smaller partition

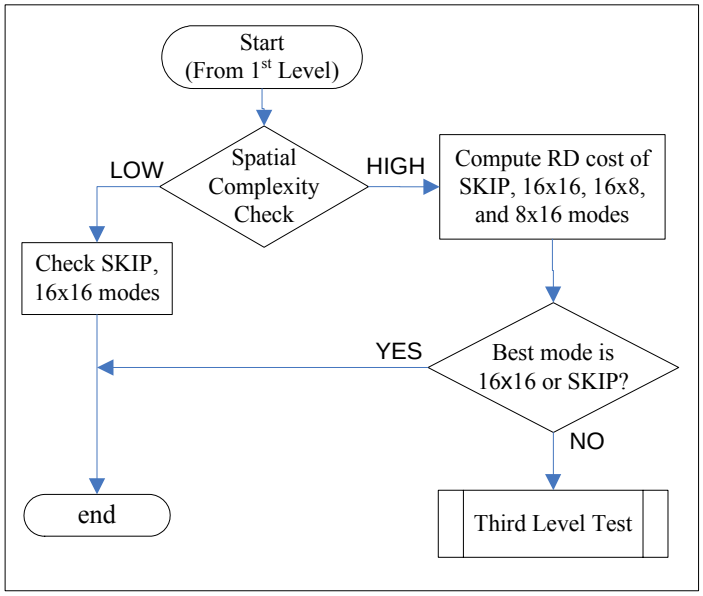

Fig. 2-2 Block diagram of the 2nd level of the improved scheme.

size. However, (2-1) excludes the case of macroblocks with spatial content in between the two-tiered classifications. Furthermore, it is observed that the mode decision for the aforementioned macroblocks varies according to what level of compression is applied. Since the mode decision for a macroblock is determined by the lowest RD cost, we suggest computing the RD cost of a few inter-modes before the entire search process is performed. If the best mode for a high-detailed macroblock in this level is in favour of using inter-modes with partition size of 16x8 and 8x16, a more thorough search in the next level is required. Otherwise, the mode decision for the current macroblock is made. The detailed algorithm at this level is described as follows:

1º Determine the complexity of the macroblock using (2-1). If the content of the current macroblock is homogeneous, select the best mode for the current macroblock either from SKIP or the inter-mode with partition size of 16×16. Otherwise, continue to 1º.

2º Activate the inter-modes with partition size of 8×16, 16×8, and 16×16 and SKIP. Check the current macroblock with the activated inter-modes.

3º Compute the RD cost of the four modes and obtain a mode decision for this level. If the optimal mode is not SKIP or 16×16, proceed to the third level. Otherwise, the mode decision for the current macroblock is decided.

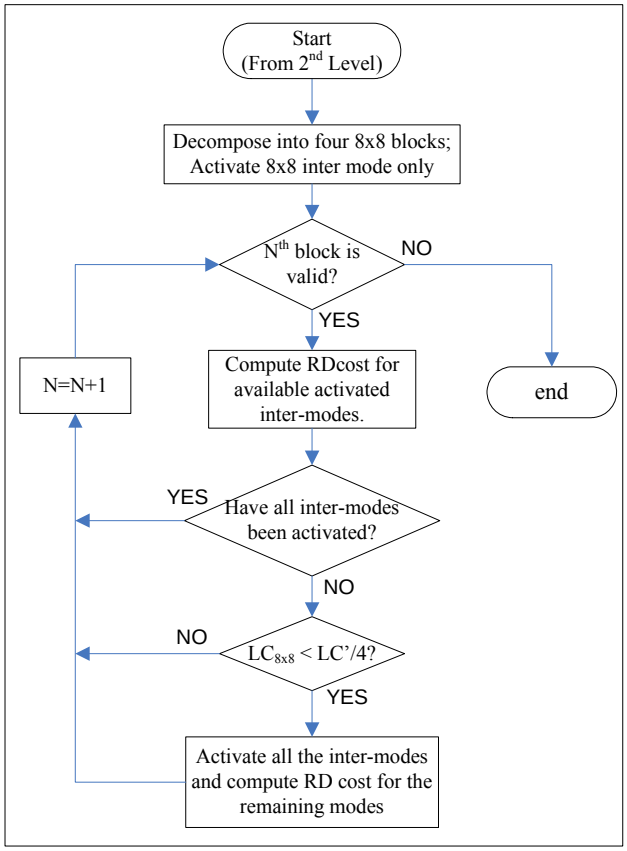

Fig. 2-3 Block diagram of the last level of the improved scheme.

2.2.3. The last level of the improved scheme

The third level computes searches within the P8x8 mode (including the inter-modes with smaller partition sizes of 8×8, 8×4, 4×8, and 4×4 pixels). Since each decomposed 8-by-8 block has to undergo search operations, a very large computational overhead is expected. Fig. 2-3 illustrates the proposed scheme for this level. The current macroblock is decomposed into four non-overlapping 8×8 blocks. Each block is then checked by the inter mode with partition size of 8×8. The other inter-modes with smaller partition sizes are activated if the following condition is true

$$LC_{8\times8}(N^{th}) < LC'/4 \qquad (2\text{-}3)$$

where $LC_{8x8}(N^{th})$ is the RD cost of the N^{th} current 8×8 block and LC' represents the RD cost of the best mode obtained from the last level. Note that (2-3) does not need to be revised if all inter-modes are activated. Thus, a computational saving is achieved for the first few blocks which do not satisfy the condition in (2-3). The detailed algorithm of this level is described as follows:

1^o Decompose the current macroblock into four non-overlapping 8×8 blocks.

2^o Activate on the inter mode with partition size of 8×8 only.

3^o Check and compute RD cost for the activated inter-modes for the current 8×8 block until all the blocks have been examined.

4^o Detect if the other available inter-modes have been activated. If the situation is true, proceed to 3^o. Otherwise, examine the condition (2-3) and implement 5^o.

5^o If the condition in (2-3) is satisfied, activate all the available inter-modes in this level. Otherwise, return to 3^o.

2.3. Simulation results

This section compares the results of the proposed algorithm incorporating the improved schemes with the previously reported MFInterms algorithm [7]. Results are presented as improvements over the standard H.264 benchmark, software version JM6.1e [9]. Except were stated, the selected sequences are of QCIF resolution (176×144 pixels). The other settings are as follows: in total 100 frames of each sequence were processed. The frame rate and GOP are 30 frames/sec and 10 frames, respectively. The precision and search range of the motion estimation is set to ¼ pixel and ±8 pixels, and finally, CABAC coding is utilised.

Fig. 2-5 and Fig. 2-6 exhibit two PSNR-rate relationship diagrams for the Flower sequence in CIF resolution and the Stefan sequence in QCIF resolution. Both sequences are classified as Class C sequences. It is clear that the PSNR-rate performance of the new algorithm is virtually identical to that provided by the JM6.1e software, at all bit rates. However the previously reported MFinterms algorithm shows a decreased PSNR, particularly at the higher bit rates. Consequently we conclude that the new algorithm provides better performance than the MFinterms algorithm at all bit rates.

Table 2-1 summarises the speed up over the JM6.1e benchmark for the two algorithms. Qp was fixed at 32. Generally, both algorithms provide different degrees of speed up depending on the class of the selected test sequences. Both algorithms reduce the computation time for Class A and Class B sequences by in excess of 57%. For Class C sequences (that generally are more difficult to encode), the new scheme

performs almost twice as well as the previously reported MFinterms algorithm, providing speedups of 41%-55% compared with 23%-30%.

Fig. 2-4 illustrates the speed up achieved by both algorithms for different Qp settings, for the Stefan sequence. Significantly, the proposed algorithm maintains the same speed up of around 42% for low Qp values (high bit rates) and then increases to over 60%; while a slow increase of between 17% and 34% is shown for the MFinterms algorithm. Fig. 2-4 illustrates that the proposed algorithm is superior to MFinterms at all bit rates.

Table 2-1
The speed up over the JM6.1e benchmark for the two algorithms at fixed Qp =32.

Class / Sequence		MFinterms [1]	Proposed scheme
A	Container Ship	72.94%	68.10%
B	Silent Voice	60.22%	57.61%
C	Car Phone	30.97%	51.60%
	City	22.85%	54.74%
	Crew	24.39%	51.76%
	Flower (CIF)	27.54%	48.40%
	Ice	29.91%	44.10%
	Stefan	28.72%	41.45%
	Tempete (CIF)	20.87%	44.37%

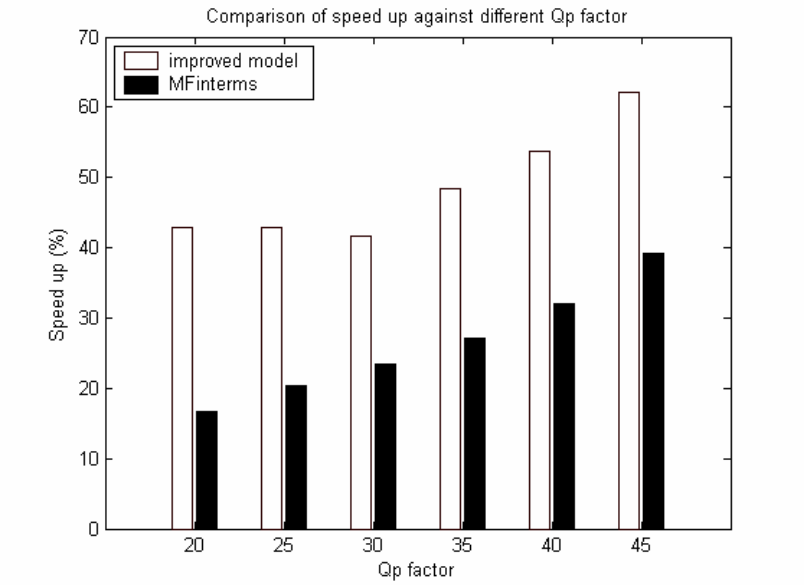

Fig. 2-4 Speedup for Stefan (QCIF) sequence for different Qp values.

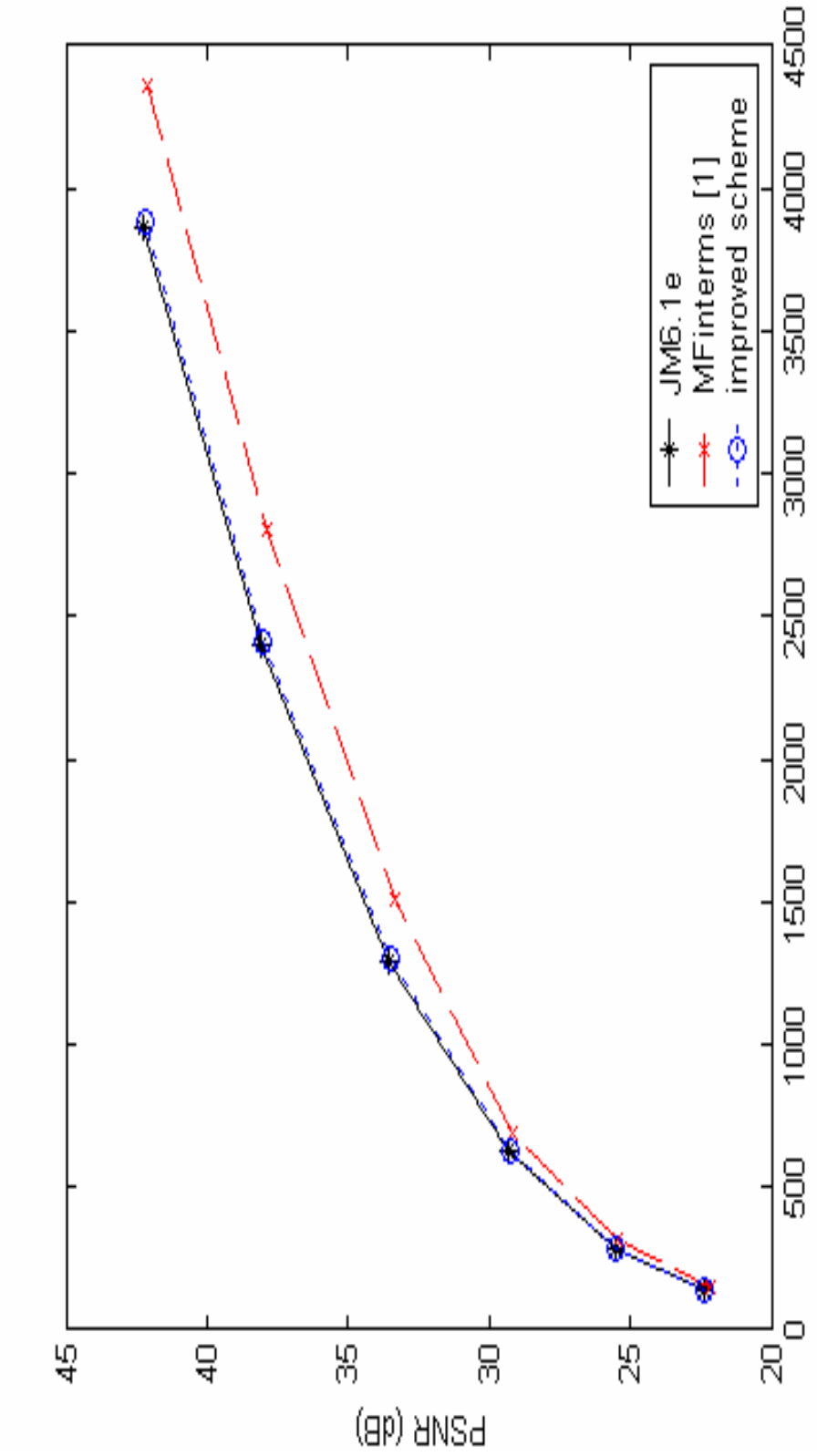

Fig 2-5 PSNR-rate relationship diagram for Flower sequence in CIF resolution.

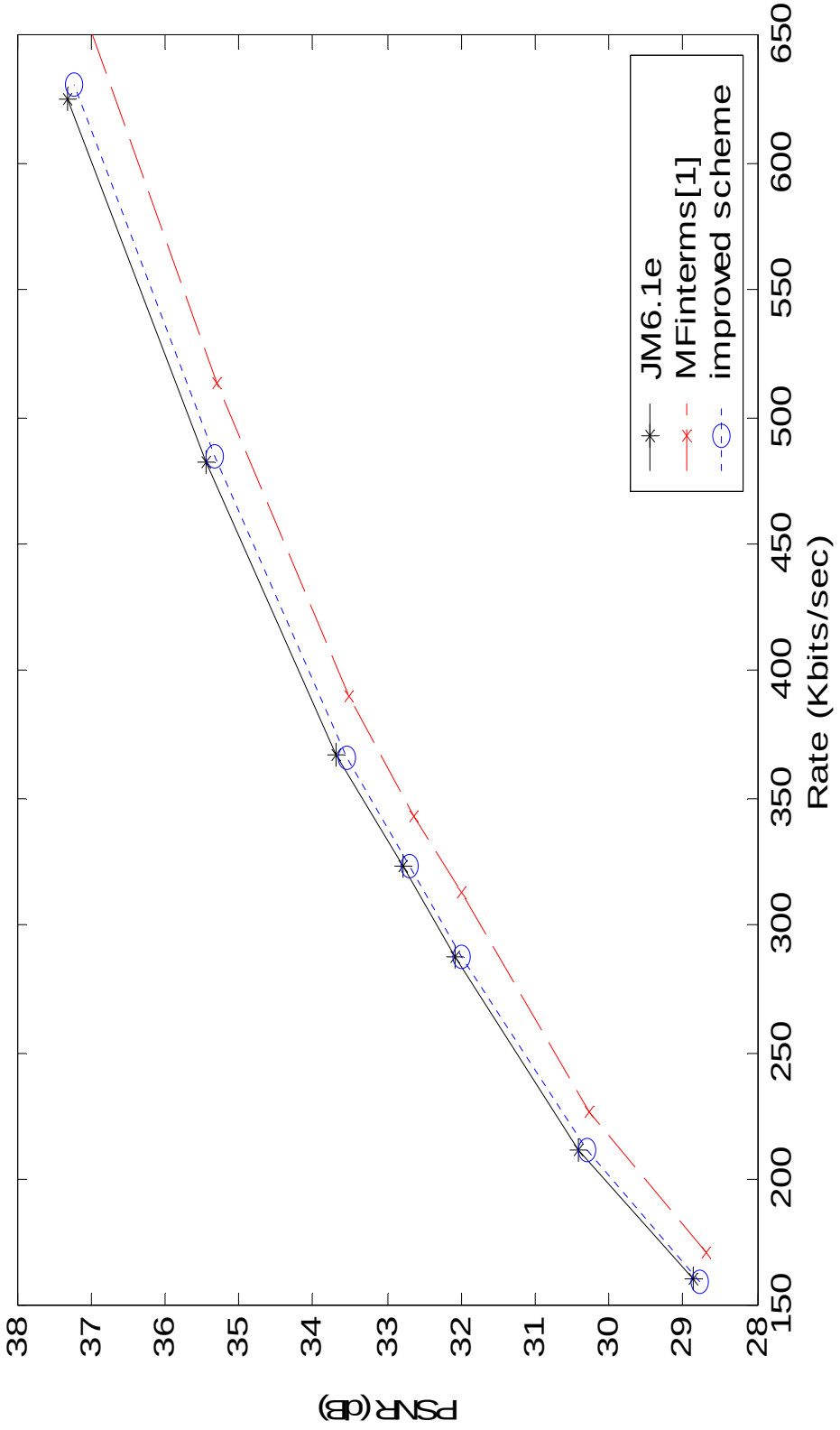

Fig 2-6 PSNR-rate relationship diagram for Stefan sequence in QCIF resolution.

Chapter 3
A simplified Lagrangian evaluation for video coding at very low bit rate

3.1 Introduction

One of popular implementation for very low bit rate video coding is to increase the value of the Qp factor. By performing this, a larger pixel-based distortion between the original and reconstructed frame is produced. In contrast, the rate is reduced due to the coarser quantisation process. However, since the Lagrangian parameter, λ, grows exponentially with the increase in Qp value (as shown in Fig. 3-1), the product of the Lagrangian parameter and rate increase accordingly. Thus, a large Lagrangian cost is expected when the value of Qp factor is high.

The implementation of the optimisation process in the H.264/AVC standard requires the availability of the reconstructed image for all *k* modes. It implies that the

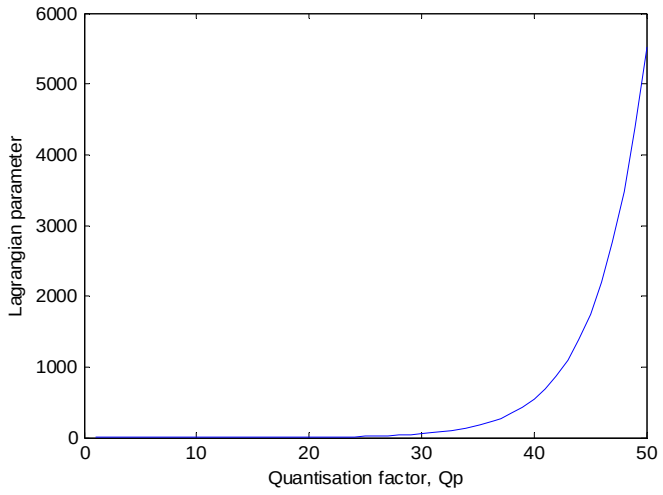

Fig.3-1 The non-linear relationship between the λ and Qp, recommended by the H.264/AVC standard

completion of the encoding-decoding cycle has to be repeated k times for each block. Thus, it is still computationally expensive for real-time video applications. Motivated by this, the aim of this chapter is to derive a simple cost evaluation for multi-mode video coding employing high Qp factors. The other arrangements of this chapter are as follows: the properties of the Lagrangian evaluation are described in next section. Then, a cost evaluation is proposed in Section 3.3 for multi-mode coding operating at very low bit rates. The simulation results are presented in Section 3.4.

3.2 The properties of Lagrangian evaluation in high Qp case

Generally, the distributions of distortion cost and weight rate cost on the Lagrangian space are distinct. Fig. 3-2 and Fig. 3-3 illustrate these relationships

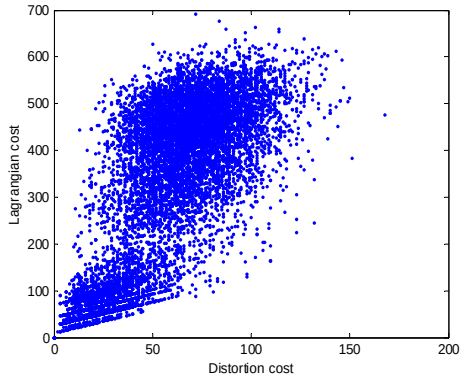

Fig.3-2 The distribution of distortion cost on the Lagrangian plane

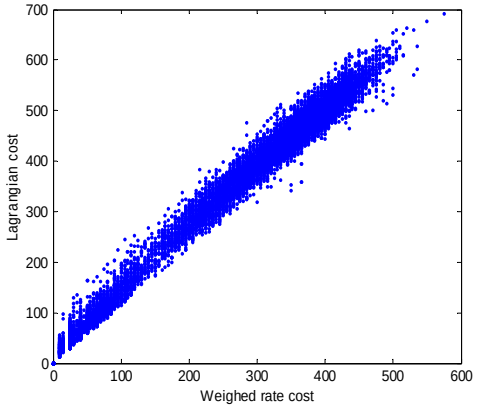

Fig.3-3 The distribution of weighed rate cost on the Lagrangian plane

respectively. Fig. 3-2 demonstrates that the distortion costs tend to be random regardless of the values of Lagrangian cost. On the other hand, the weight rate cost possesses a plausible linear relationship with relatively smaller variance. These two figures further reflect that the weight rate cost is a better indication to address the general tendency of Lagrangian cost.

Fig. 3-4 depicts an average composition of Lagrangian costs in terms of distortion component and weight rate component with respect to the various Qp values. Regardless of the selection of Qp values, weight rate components dominate distortion components by approximately 3:1. In other words, weight rate components have a more significant role in the construction of Lagrangian cost. It further implies that the mode which possesses the least weight rate is more likely to have minimum Lagrangian cost.

Unfortunately, by applying the aforementioned property for block-based mode selection it does not always guarantee the minimum Lagrangian cost. The reasons are as follows: (a) the least weight rate cost may be found more than one in the available

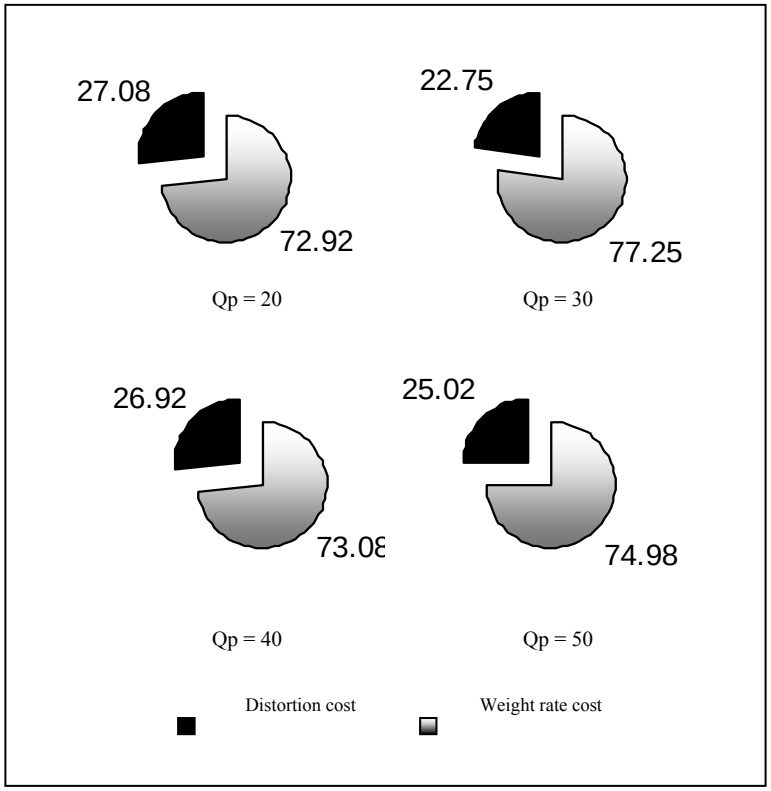

Fig.3-4 The average composition (%) of Lagrangian costs with respect to various Qp values.

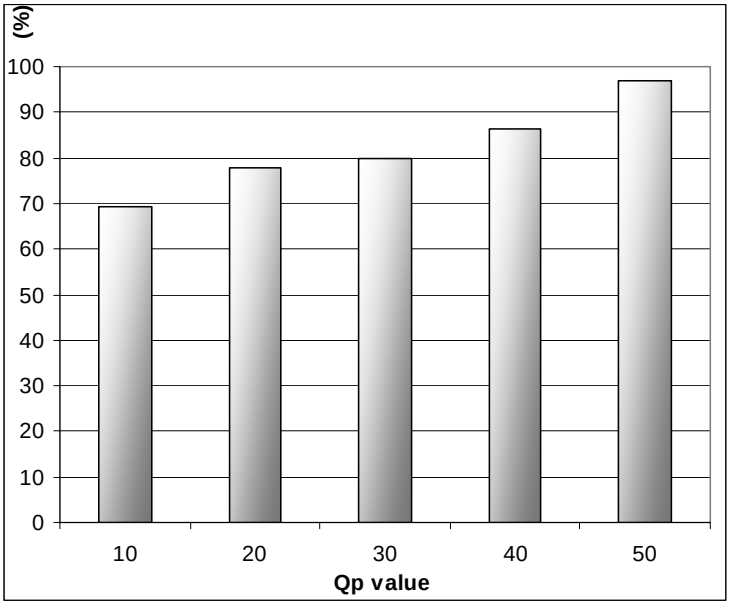

Fig.3-5 The matching percentage between the best mode selected by the least rate cost and those by the minimum Lagrangian cost. (Stefan sequence intra-coded at Qp = 45)

modes; (b) the distortion cost may greatly vary among the modes which generate quite similar weight rate cost. Fig. 3-5 illustrates the relationship between the best modes selected by the least weight rate cost and those by the minimum Lagrangian cost in intra-coding. One can see that a higher ratio is achieved for increased values of Qp factor. The explanation is the diminishment of high frequencies in the residue data by the coarse quantisation. It results in the similar distortion cost regardless of which mode is selected. In contrast, the rate cost may vary between the *MostProbableMode*, requiring fewer bits to be encoded, and the other modes. In conclusion, the mode selection for highly compressed video coding can be achieved by selecting the least rate cost.

3.3 Error-propagation control

The proposed cost evaluation tailored for highly compressed video coding makes use of the rate cost to predict the minimum Lagrangian cost. Since the least rate cost also promises the minimum bits for encoding a video block, it benefits video applications operating at very low bit rate. However, by applying the proposed evaluation, the mode selection process quite often encounters the choice among more

than one mode which possesses the least rate cost. Fig. 3-6 illustrates the relative occurrences of this situation for different sequences. Even though the proportion of the aforementioned problem seems not to be too significant, an error is unavoidable if a random selection is performed. Since the choice of the *MostProbableMode* depends on the previously encoded neighbours, one erroneous decision will propagate to the next few consecutive blocks and consequent picture degradation may be expected. Thus, error-propagation control is necessary to prevent such an occurrence.

Unfortunately, there is no better error-propagation control than computation of the Lagrangian cost, since distortion cost is almost unpredictable. In other words, the additional computation for distortion cost evaluation is required for fewer modes which possess the least rate cost. Nevertheless, it does not imply that the overall coding computation is increased significantly, since fewer candidates require a full Lagrangian evaluation. Furthermore, the evaluations of distortion cost and rate cost share many similar procedures, such as the integer discrete cosine transform (DCT) and the quantisation process.

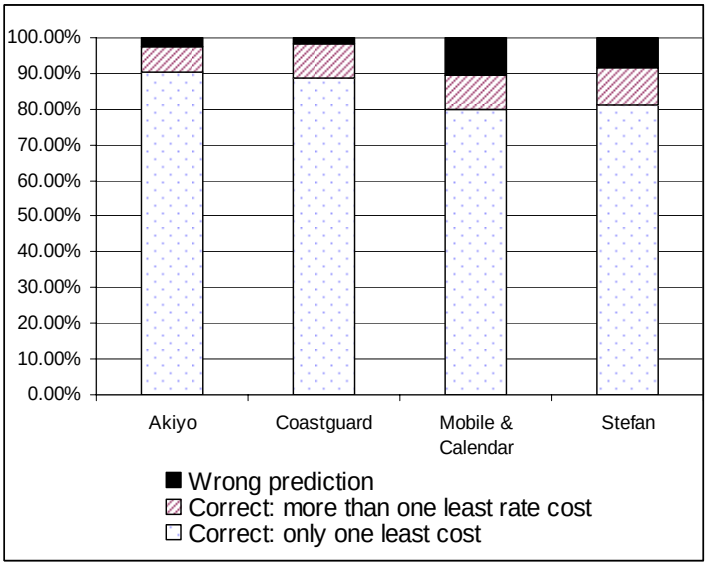

Fig.3-6 The percentages that more than one least rate cost is found with respect to the correct ratio by using rate cost evaluation (Qp =45).

3.4. Simulation results

This section compares the performances of the proposed cost evaluation and the full Lagrangian evaluation. Both cost evaluations are integrated in the H.264 reference software recommended by JVT [9]. In order to examine the efficiency of the proposed cost evaluation in an extreme case, the sequences are encoded with intra-modes [6] which result in larger residue energy. Furthermore, two sequences, Stefan and Mobile & Calendar, are selected due to their complicated composition in texture structure. Both sequences are of QCIF resolution (176x144 pixels) and encoded at 4.29Hz. Fig. 3-7 and Fig. 3-8 illustrate the PSNR-rate curves generated by the proposed cost evaluation (dotted line) and the full Lagrangian evaluation (solid line), respectively. Qp values are employed in the range 42 to 50, to generate low bit-rate video sequences. Due to the intra-coding, the rates are slightly higher than the proposed rates for video coding at very low bit rates.

In Fig 3-7 and Fig. 3-8, insignificant PSNR degradation is noticed due to the wrong prediction caused by the proposed evaluation cost. Nevertheless, the proposed cost evaluation is able to encode the sequences with fewer bits at the fixed Qp shown in the figures. That is because the proposed cost evaluation selects the best mode according to the least rate cost. This satisfies the requirement of low bit-rate video applications. As to speed up, the simulations show that the proposed cost evaluation can save computational time of between 60% and 70% when compared to full Lagrangian evaluation. The intra modes that require a full Lagrangian examination are reduced significantly.

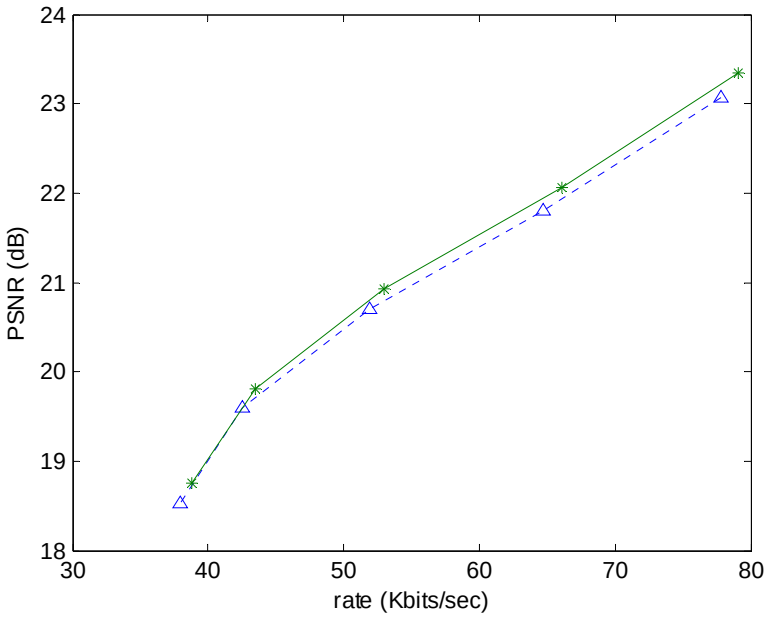

Fig.3-7 The PSNR-rate graph of Mobile and Calendar sequence in QCIF resolution at 4.29Hz.

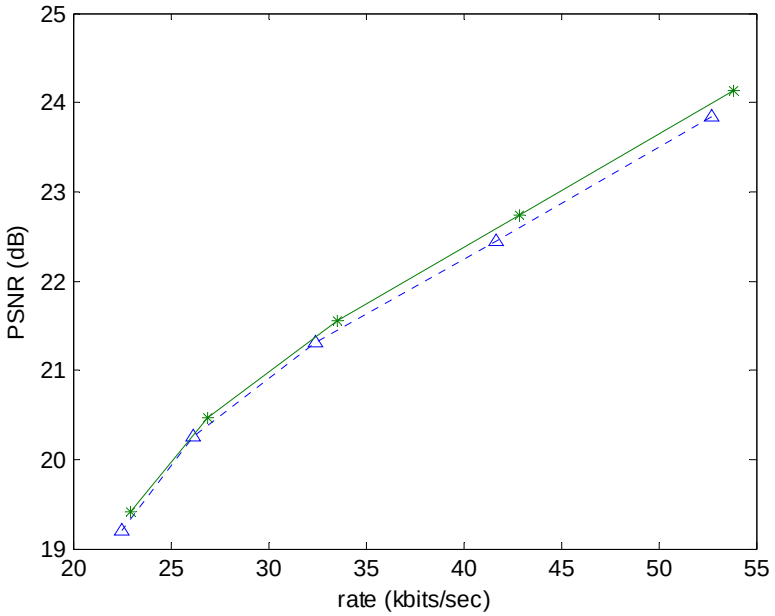

Fig.3-8 The PSNR-rate graph of Stefan sequence in QCIF resolution at 4.29Hz.

Conclusions

The details of the proposed improved schemes for inter-frame coding in the H.264/AVC standard have been recorded in Chapter 2. The new algorithm significantly improves the PSNR-rate performance of the previously reported MFinterms algorithm at all bit rates, while almost halving the computational requirement for Class C sequences. The simulation results further show that the proposed algorithm achieves the same coding performance in terms of picture quality and compression ratio as that of the H.264/AVC standard, yet reduces the computational requirement by up to 68%.

In Chapter 3, we have introduced a simple but efficient cost evaluation to challenge the full Lagrangian cost evaluation in high Qp case, required for low bit rate coding. The proposed cost evaluation is derived from the extensive statistics and experiments. The simulation results show that insignificant picture degradation is found in the sequences employing the proposed cost evaluation. However, the reductions in the encoding time and the number of encoded bits brought by the proposed cost evaluation are more practical for real-time video application operating at very low bit rates.

References

[1] ISO/IEC AVC-491:1993, Coded representation of picture and audio information – MPEG-2 Test Model 5, Apr 1993.

[2] T. Wiegand, M. Lightstone, D. Mukherjee, T. Campbell, and S. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE trans. on Circuit and System for Video Technology*, vol. 11, issue 3, Mar 2001.

[3] M. Lightstone, D. Miller, and S. Mitra, "Entropy-constrained product code vector quantization with application to image coding," *Proceedings of ICIP 94*, vol. I, pp. 623-627, Nov 1994.

[4] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 36, pp. 1445-1453, Sept 1988.

[5] T. Wiegand and B. Girod, Multi-frame motion-compensated prediction for video transmission, first edition, Kluwer Academic Publishers, 2001.

[6] ISO/IEC 14496-10:2003. Information technology – coding of audio-visual objects – Part 10: advance video coding, Dec. 2003

[7] A. Yu and G. Martin, "Advanced block size selection algorithm for inter-frame coding in H.264/AVC," *Proc. ICIP 2004*, Singapore, pp. 95-98, Oct. 2004.

[8] I. Richardson, "H.264/MPEG-4 video compression: video coding for next-generation multimedia," Wiley, London, 2003.

[9] JVT reference software, http://bs.hhi.de/~suehring/tml.

[10] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, pp. 399 – 417, 1963