Fast Direct Mode Decision Algorithm based on Optimal Mode Pattern Searching

Seok-Pil Lee[1], Neungsoo Park[2] and Yungho Choi[3*]

[1] *Korea Electronics Technology Institue, Seoul, Korea*
[2] *Dept. Of Computer Science & Engineering, Konkuk University Seoul, Korea*
[3] *Dept. Of Electrical Engineering, Konkuk University Seoul, Korea*
[1]*ispio@keti.re.kr,* [2] *neungsoo@konkuk.ac.kr,* [3] *yunghoch@konkuk.ac.kr*

Abstract

In H.264 video compression standard, B frame bi-directional predictions and its DIRECT mode coding are included to enhance video compression efficiency. However, such tools require a complex mode decision process and result in a long computation time. Therefore, fast algorithms are crucial in implementing H.264. In this paper, to save B frame encoding time, a fast DIRECT mode prediction algorithm is proposed, which determines DIRECT mode at an early stage by optimal mode pattern matching. Simulations show that the proposed algorithm can determine DIRECT mode without a complex mode decision process for 42% more macroblocks and, on mis-determination for DIRECT mode, cut down its rate-distortion (RD) cost increment by half, compared with Jin's algorithm. This enables a fast B frame encoding with less quality degradation.

Keywords: *B frame encoding, DIRECT mode, H.264, video compression, mode prediction.*

1. Introduction

These days, many multimedia services such as Internet TV, video conference, etc. have been widely provided [1]. In these services, image compression techniques play a critical role by reducing data size and thus, allowing to transmit and to store image data with less resources[2, 3]. Among such techniques, H.264 [4] is one of the recent image compression standards and has been used for many digital image devices such as CCTV, camcorders, smart phones, etc. For the H.264 standard, many new compression tools are introduced in order to significantly enhance compression efficiency. Such tools include B frame encoding [5] and its DIRECT mode [6]. B frame encoding enables to reduce image data redundancy by supporting bi-directional predictions. However, although this scheme reduces image data redundancy, it may suffer from larger side information caused by encoding multiple motion vectors and multiple reference picture indexes. To resolve this problem, DIRECT mode for B frame encoding is introduced in H.264. DIRECT mode does not encode motion vectors by allowing decoders to predict those vectors utilizing previously coded information temporally or spatially.

Despite the high coding efficiency of B frame encoding and its DIRECT mode, these tools are considered to be one of the most time-consuming tasks since this requires complicated motion estimations and bi-directional predictions. Thus, if DIRECT mode can be determined at an early stage, a considerable coding time can be saved by skipping all the rest of complex mode decision processes.

[*] Yungho Choi is the corresponding author. TEL: +82-2-2049-6009

To enable an early decision on DIRECT mode, some research works have been proposed [7, 8, 9]. One of such work is Jin's algorithm [9]. This scheme exploits the strong relation between optimal modes and rate-distortion costs of the currently encoding macroblock (MB) and ones of its co-located macroblock in previously encoded frames. This enables to determine DIRECT MODE coding for the current macroblock without a complex mode decision process, increasing B frame encoding speed significantly. However, Jin's algorithm leaves out the cases where co-located macroblock's best mode is not either DIRECT_16x16 or INTER_16x16 mode although, even in these cases, macroblock's optimal mode can be DIRECT mode. Therefore, this algorithm has a limitation in speeding up B frame encoding.

To resolve the limitation of the Jin algorithm, this paper proposes a fast DIRECT mode decision scheme that extends the Jin algorithm and can determine DIRECT MODE even for macroblocks that Jin's algorithm leaves out. Because, when macroblocks can satisfy the conditions of Jin's algorithm, the proposed algorithm just follows Jin's scheme, not hampering the performance of Jin's algorithm. Indeed, it always increases the number of DIRECT MODE macroblocks that the algorithm can determine at an early stage, saving B frame encoding time.

The rest of the paper is organized as follows. Section 2 introduces a fast direct mode decision algorithm based on the mode pattern matching. Section 3 shows the performance of the proposed algorithm and compares it with the one of Jin's algorithm. The conclusion is given in the last section.

2. Fast Direct Mode Decision Algorithms

The section introduces a novel fast DIRECT mode decision algorithm which enhances the performance of Jin's algorithm [9] by taking care of macroblocks (MB) which the Jin's algorithm need to run all the complex mode decision process for.

The following subsection describes the Jin algorithm and Subsection 2.2 proposes a new fast DIRECT mode decision algorithm which extends the Jin algorithm by enabling encoders to predict DIRECT mode even in the case where the Jin algorithm cannot.

2.1 Adaptive Fast Direct Mode Decision Algorithm using Mode and Lagrangian Cost Prediction for B Frame in H.264/AVC

This algorithm [9] was proposed by Jin to exploit the strong relations of optimal modes and rate-distortion costs (RDcost) between the macroblock (MB) to be currently encoded and its co-located macroblock. This enables Jin's algorithm to determine DIRECT MODE coding for macroblocks without a complex mode decision process, increasing B frame encoding speed significantly. In this algorithm, DIRECT mode for a macroblock can be determined by checking (1),(2):

$$\text{Mode: } (MD_{PCO.MB} \leq 1 \ \&\& \ MD_{NCO.MB} \leq 1) \ || \ (MD_{BCO.MB} = 0) \quad (1)$$

$$\text{RDcost: } J_{CURRENT\_MB} < \alpha \times J_{PCO\_MB} \ || \ J_{CURRENT\_MB} < \beta \times J_{NCO\_MB} \quad (2)$$

where $MD_{PCO.MB}$, $MD_{NCO.MB}$ and $MD_{BCO.MB}$ are best mode numbers for co-located macroblocks (MBs) in the previous P frame, the next P frame, and the previous B frame, respectively. $J_{CURRENT\_MB}$, J_{NCO_MB} and J_{PCO_MB} are RD costs for the current macroblock, the co-located macroblocks in the next P frame and in the previous P frame.

The key rationale of this work is shown in Figure 1. In this figure, many macroblocks' best modes are shown to be DIRECT modes when the best modes of the corresponding co-located blocks (i.e., PCO.MB, NCO.MB, and BCO.MB) are either DIRECT_16x16 mode or INTER16x16 mode. However, this figure also shows that many macroblocks (MBs) can have DIRECT mode as their optimal modes even when G2 condition, i.e., $MD_{PCO.MB} > 1 \parallel MD_{NCO.MB} > 1$, is satisfied. Therefore, this paper proposes a fast DIRECT mode prediction algorithm that can determine DIRECT mode for macroblocks which do not satisfy the condition given above: (1),(2). In the next subsection, the proposed algorithm will be introduced.

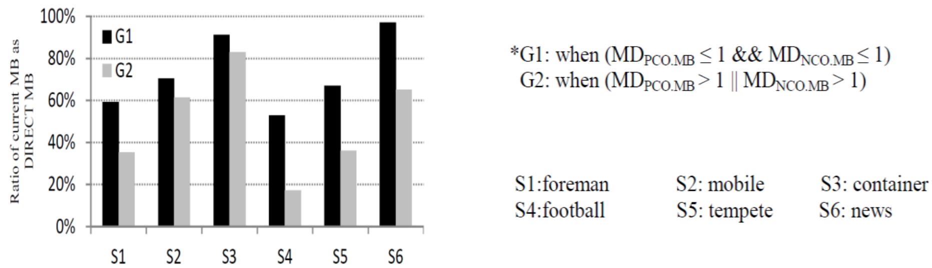

Figure 1. Analysis of mode selection relation [9].

2.2 Fast Direct Mode Decision Algorithm based on Optimal Mode Pattern Searching

The key motivation of this new algorithm lies in the fact that the optimal mode of considerable macroblocks can be DIRECT mode although they do not satisfy Jin's algorithm conditions: (1), (2). This is true specially when video images are complicated and move fast. In such cases, Jin's algorithm may not perform well because many macroblocks do not satisfy (1) and (2).Thus, it needs to run a long complex mode decision process for the macroblocks, increasing B frame encoding time. To resolve this problem, the proposed algorithm extends the Jin's algorithm so that it can determine DIRECT mode not only for macroblocks satisfying (1) and (2) but also for macroblocks not satisfying. The algorithm is summarized as follows.

First, in the case of macroblocks satisfying (1) and (2), the proposed algorithm does the same thing of the Jin's algorithm. The main contribution of this algorithm lies in the macroblocks not satisfying (1) and (2). To determine DIRECT mode coding for such macroblocks, the proposed algorithm locates macroblocks which can correspond to the currently encoding macroblock, in the previous P frame, the next P frame, and the previous B frame. If the backtracked macroblocks' optimal mode is DIRECT mode, the possibility that the current macroblock's optimal mode is DIRECT mode can be high, allowing the proposed algorithm to code the current macroblock in DIRECT mode. The backtracking process can be done by pattern matching as shown in Figure 2.

For this pattern matching, optimal modes of macroblocks surrounding the current macroblock are used as the pattern to be compared in previously encoded frames. In this figure, optimal modes of macroblock A, B,C,D in F frame are compared in F-1, F-2, F+1 frames, where F specifies the displaying order frame number and thus, F, F-1, F+1, and F-2 are the current B frame, the previous P frame, the next P frame, and the previous B frame, respectively. The searching scope for the pattern is limited to *R*. This is because, in most cases, objects in a frame cannot move too far in subsequent frames. Through pattern searching in the searching area, the macroblock corresponding to the current macroblock can be located in previous frames. If the located macro's optimal mode is DIRECT mode, the possibility that the optimal mode of the current macroblock is DIRECT mode can be high since the located macroblock may correspond to the current macroblock as shown in the figure. However, this cannot be always true. To increase the accuracy of this DIRECT mode decision algorithm, this scheme uses the following condition to filter out macroblocks having too high rate-distortion cost and thus, having a low possibility to be encoded in DIRECT mode.

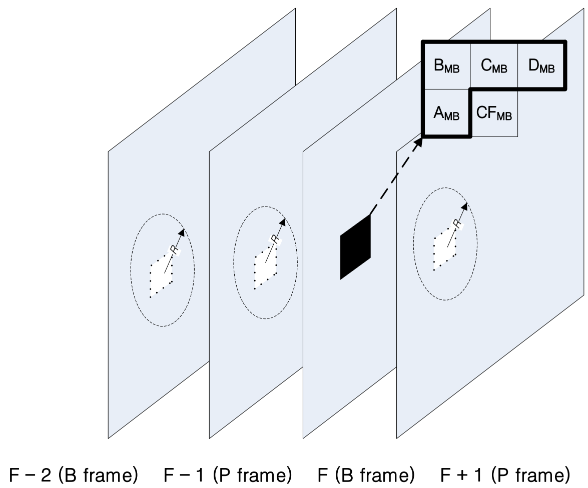

Figure 2. Mode Pattern Matching for Direct Mode Decision

When at least one macroblock in F-1, F-2, F+1 frames matches optimal mode pattern defined above, if the average RD cost of the matched macroblocks is less than the RD cost of the current macroblock, the proposed algorithm encodes the macroblock in DIRECT mode without a complex mode decision process, saving encoding time. Otherwise, all the mode decision calculations are required. In the next section, simulation results and their analysis for this algorithm are given.

3. Simulation Results

Table 1. Comparison of Jin's and the Proposed Algorithm (QP=20).

Sequences	Size	Total # DM	DM % in MBs	Jin's algorithm			The proposed algorithm		
				# DM	Hit %	RDC Error	#DM	Hit %	RDC Error
News	QCIF	3890	79%	4088	87%	356	4195	90%	114
Foreman	QCIF	1772	34%	1980	47%	440	2263	67%	224
Hall_monitor	QCIF	2456	51%	3108	55%	153	3463	62%	102
Container	QCIF	4284	83%	4516	85%	53	4815	88%	39
Carphone	QCIF	4075	82%	4341	84%	46	4629	86%	32
Bus	CIF	7084	36%	8138	57%	551	8952	73%	153
Tempete	CIF	7083	36%	8134	57%	550	8952	73%	153
Paris	CIF	8346	42%	12257	52%	375	12976	64%	198
Stefan	CIF	8064	47%	8878	63%	369	9742	79%	169
Waterfall	CIF	9800	51%	11287	55%	309	13182	72%	138

To measure the performance of the proposed algorithm, the proposed fast direct mode decision algorithm is incorporated into JM 14.2 [10]. Under the following conditions: high complexity mode (RDOptimization=1), GOP=IBPBP, encoding frames including I,B,P=199, MV search range=32, resolution={qcif, cif}, the accuracy of DIRECT mode decision and the average rate-distortion cost (RDcost) error of wrong mode decision of Jin's algorithm and the proposed algorithm are compared. The results are shown in Table 1. In the table, Total#DM is the total number of macroblocks whose optimal modes are DIRECT mode (DM) in all the B frames. DM% in MBs is the percentage of direct mode macroblocks in all the frames. #DM is the number of macroblocks chosen as direct mode macroblocks. Hit% is the ratio of the number of macroblocks correctly determined to #DM. RDC Error is the difference between the RD cost of the true optimal mode of a macroblock and the one of its direct mode on determination.

Simulation results show that the proposed algorithm improves the accuracy of direct mode decisions by up to 42%, compared to the accuracy of Jin's algorithm. Especially, when direct mode macroblock ratio to the total number of macroblocks is low, the proposed algorithm outperforms Jin's algorithm. This is because, in this case, many macroblocks cannot satisfy the condition of Jin's algorithm. Thus, Jin's algorithm needs to calculate complex mode decisions, suffering a long encoding time. Furthermore, the proposed does not hamper the performance of Jin's algorithm because, in the case satisfying (1) and (2), it does follow Jin's algorithm while it can additionally determine direct mode macroblocks even when not satisfying (1) and (2).

Furthermore, since the proposed scheme locates similar macroblocks in previously encoded frames through pattern matching, the RD cost of the macroblock to be currently encoded can be estimated more accurately as shown in Table 1.

4. Conclusions

To enhance the performance, the work proposes a fast direct mode decision algorithm that extends Jin's algorithm. Thus, the proposed algorithm can make direct mode decisions not only for macroblocks satisfying (1) and (2) but also for macroblocks not satisfying the conditions. This allows to save B frame encoding time by encoding more macroblocks in DIRECT mode without executing long mode decision processes.

References

[1] M. Kim and I. Jee, "H.264 Encoding Technique of Multi-view Image expressed by Layered Depth Image", the Journal of IWIT **(2010)**, Vol. 10 No. 1, pp.81-90.

[2] B. Bae and Y. Kim, "A Study on Controlling of Target Transfer Rate in Terrestrial DMB Systems", the Journal of IWIT **(2010)**, Vol. 10 No. 3, pp.19-26.

[3] Y. Im and E. Kang, "MPEG-2 Video Watermarking in Quantized DCT Domain", the Journal of IWIT **(2011)**, Vol. 11 No. 1, pp.81-86.

[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Trans. Circuits Syst. Video Technol. **(2003)**, vol. 13, no. 7, pp. 560–576.

[5] M. Flierl and B. Girod, "Generalized B pictures and the draft H.264/AVC video-compression standard", IEEE Trans. Circuits Syst. Video Technol. **(2003)**, vol. 13, no. 7, pp. 587–597.

[6] A. M. Tourapis, F. Wu and S. Li, "Direct Mode Coding for Bipredictive Slices in the H.264 Standard", IEEE Trans. Circuits Syst. Video Technol. **(2005)**, vol. 15, no. 1, pp. 119-126.

[7] J. Lee, I. Choi, W. Choi and B. Jeon, "Fast Mode Decision for B slice", Proceedings of JVT-K021 **(2003)**, Antalya, Turkey.

[8] I. Choi, J. Lee and B. Jeon, "Fast Coding Mode Selection With Rate-Distortion Optimization for MPEG-4 Part-10 AVC/H.264", IEEE Trans. Circuits Syst. Video Technol. **(2006)**, vol. 16, no. 12, pp. 1557-1561.

[9] X. Jin, J. Sun, J. Zhou, Y. Huang, J. Su and T. Ikenaga, "Adaptive Fast Direct Mode Decision Algorithm using Mode and Lagrangian Cost prediction for B Frame in H.264/AVC", Proceedings of 2011 IEEE International Conference on Multimedia and Expo. **(2011)** July, pp. 1-6.

[10] H.264 Reference Software JM14.2 [Online] http://iphome.hhi.de/suehring/tml/download/old-jm.

Authors

Seok-Pil Lee

He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Yonsei University, Seoul, Korea, in 1990, 1992, and 1997, respectively. From 1997 to 2002, he has been working as a senior research engineer in Digital TV Research Center of DAEWOO Electronics Co., Ltd., Seoul, Korea. Since 2002, he has been working at Korea Electronics Technology Institute (KETI), where he is presently the Director of Digital Media Research Center. His research interests are in the areas of multimedia communication and multichannel audio processing, personalized broadcasting service. He is a chairman of TVA-Korea and a member of committee of MPEG-Korea.

Neungsoo Park

He received his Ph.D. in Electrical Engineering from the University of Southern California in 2002. He also received his BS and MS in Electrical Engineering from Yonsei University, Seoul, Korea, in 1991 and 1993 respectively. Currently, he is an associate professor in the Department of Computer Science and Engineering, Konkuk University, Seoul, Korea. Prior to joining Konkuk University, he was a senior engineer in Samsung Electronics Co. Ltd., Korea. His research interests include parallel computing, computer architecture, embedded system, high performance computing for signal processing, VLSI computations, and multimedia systems, etc.

Yungho Choi

He received his BS and PhD degrees from Yonsei University and University of Southern California, in 1991 and 2001, respectively. He was a senior Engineer in Intel, Shrewsbury, USA. Currently, he is an associate professor in the Department of Electrical Engineering, Konkuk University, Seoul, Korea. His research interests include multiprocessor systems, high performance codec architectures, interconnection networks as well as image recognition.