Fast Mode Decision for Intra-only H.264/AVC Coding

Jun Xin, Anthony Vetro

TR2006-034   May 2006

Abstract

In this paper, we propose an efficient mode decision algorithm for intra-only H.264/AVC video coding. We exploit the correlation between optimal coding mode decisions of temporally adjacent pictures to reduce the computational efficiency of the encoding. Compared to the RD optimized mode decision algorithm in JM10.1, the proposed algorithm can reduce the computational complexity up to 33% with negligible loss of compression efficiency (less-than 0.2dB).

PCS 2006

Fast Mode Decision for Intra-only H.264/AVC Coding

Jun Xin, Anthony Vetro

Mitsubishi Electric Research Laboratories
201 Broadway, Cambridge, MA 02139 USA

Abstract. In this paper, we propose an efficient mode decision algorithm for intra-only H.264/AVC video coding. We exploit the correlation between optimal coding mode decisions of temporally adjacent pictures to reduce the computational efficiency of the encoding. Compared to the RD optimized mode decision algorithm in JM10.1, the proposed algorithm can reduce the computational complexity up to 33% with negligible loss of compression efficiency (<0.2dB).

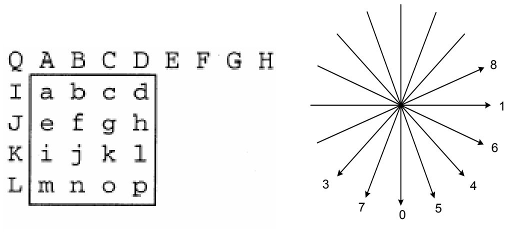
Index Terms—H.264/AVC, fast mode decision, video coding, intra video coding

1. INTRODUCTION

Intra-only video coding is a widely used coding method in professional and surveillance video applications partly due to its ease of editing. Prior experiments and demonstrations have shown that intra H.264/AVC coding [1] has an excellent performance, even compared to state of the art still image coding schemes such as JPEG 2000 [2]. To acknowledge such needs, JVT is currently working on an intra-only 4:4:4 profile [3].

In H.264/AVC intra prediction, previously decoded pixels in blocks to the left and/or above the current block are used to form spatial prediction signals. For coding the luma signal, one 16x16 macroblock may be predicted as a whole using Intra_16x16 modes, or the macroblock can be predicted as individual 4x4 blocks using nine Intra_4x4 modes. In the profiles that support Fidelity Range Extension (FRExt) tools, a macroblock may also be predicted as individual 8x8 blocks using nine Intra_8x8 modes. Intra prediction for the chroma signal uses similar techniques as those for luma Intra_16x16 prediction. Given all available prediction modes, an AVC encoder selects one for encoding. The mode decision may be performed with or without rate-distortion optimization (RDO). Better coding efficiency will be attained when RDO is enabled at the cost of higher complexity.

This paper aims at reducing the computational complexity of the intra mode decision process, while maintaining coding efficiency. We note that there are several existing works. These works aim to either reduce the number of candidate modes based on pre-analysis [4][5][6][7] or use transform-domain techniques [8]. In this paper, we propose to use the correlation between mode decisions of temporally adjacent pictures to speed up the mode decision process, which we believe has the potential for significant gains.



(a) (b)

Fig. 1. Intra_4x4 prediction illustration: (a) Prediction of samples a-p using neighboring samples A-Q. (b) Prediction mode directions except DC_Pred.

The rest of this paper is organized as follows. We first provide some background on the conventional rate-distortion optimized mode decision process in Section 2. The proposed mode decision algorithm is described in Section 3. Experimental results are presented in Section 4. Finally, concluding remarks are provided in Section 5.

2. H.264/AVC INTRA MODE DECISION

This section reviews conventional rate distortion optimized mode decision in H.264/AVC JM reference software. In particular, we focus on optimal decisions for the Intra_4x4 modes. In Fig. 1(a), the current 4x4 block pixels a-p are predicted from neighboring pixels A-Q that have already been decoded. The directional prediction options are illustrated in Fig. 1(b), where the directions of eight prediction modes except DC_Pred are plotted.

Considering the rate distortion calculation in a video encoder with RDO on, the process of computing the Lagrange cost for one coding module (in this case for one 4x4 luma block) is shown in Fig. 2. The prediction residual is transformed, quantized and then entropy encoded to calculate the rate, $R(m)$, for a given mode m. Then, the inverse quantization, inverse transform are performed and then compensated with the prediction block to get the reconstructed signal. The distortion, denoted $SSD_{REC}(m)$, is computed as the sum of squared distance between the original block, s, and the reconstructed block for mode, m, denoted as $\tilde{s}(m)$:

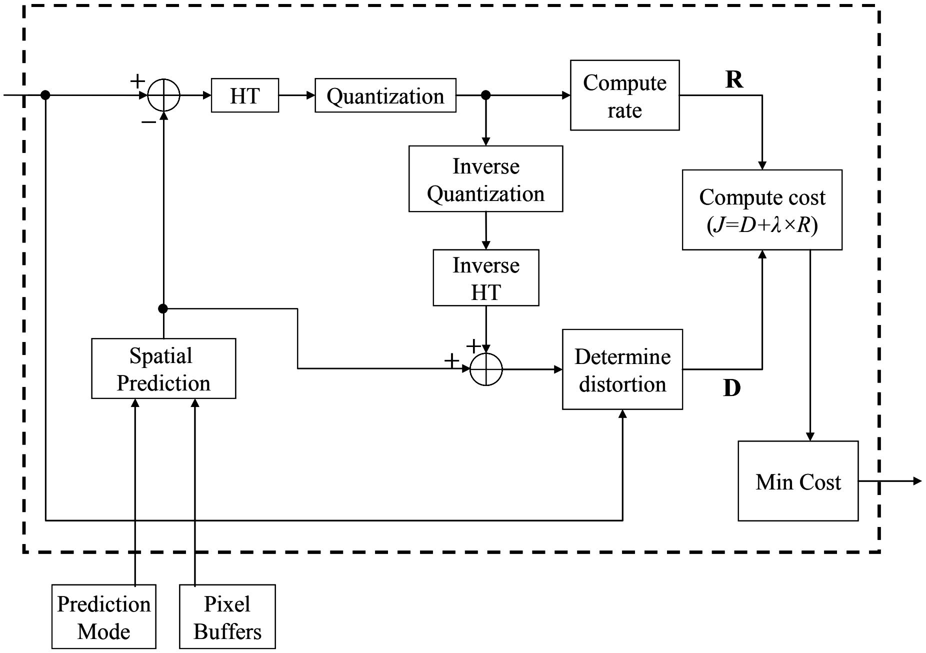


Fig 2. Conventional process of RD cost computation

$$SSD_{REC}(m) = \|s - \tilde{s}(m)\|_2^{\,2} \qquad (1)$$

where $\|\cdot\|_p$ is the matrix p-norm. The Lagrange cost is computed using the rate and distortion for each mode as follows:

$$Cost_{4x4} = SSD_{REC}(m) + \lambda_{MODE} R(m) \qquad (2)$$

where λ_{MODE} is a Lagrange multiplier, which may be calculated as a function of the quantization parameter. The optimal coding mode corresponds to the mode with the minimum cost.

Note that other distortion measures may also be used. However the *SSD* is often adopted for its superior performance.

3. FAST MODE DECISION USING TEMPORAL CORRELATION

It is well known that strong correlations exist between adjacent pictures. It is natural to believe that the optimal mode decision results of collocated macroblocks in two adjacent pictures are also strongly correlated. The basic idea of the proposed approach is to measure the difference between current macroblock and its collocated macroblock in the previous (already coded) picture. If they are close enough, the current macroblock will reuse the mode decision of its collocated macroblock and the entire mode decision process is skipped.

To measure the degree of correlation between two macroblocks for the purpose of reusing mode decision, we define a simple difference measure between two macroblocks, b_2 and b_1, as defined in (3).

In this equation, p_2 and p_1 are the two pictures containing b_2 and b_1 respectively. b_y and b_x are the vertical and horizontal coordinates of b_2 and b_1. This difference measure includes all pixels that could be used for intra prediction for a macroblock. Specifically, it includes the contributions from not only the collocated macroblock, but also its spatial neighbors that may be used to form intra predictions. Fig. 3 illustrates the adjacent neighboring pixels that may be used to predict the current macroblock.

To prevent accumulation of difference from affecting the mode reuse, we use a data structure to store macroblocks and their adjacent spatial neighboring pixels necessary for intra decision:

<short RefMBStore[PicSizeInMBs][297]>

where PicSizeInMBs is the number of macroblocks in a picture. A macroblock with address mbAddr is associated with RefMBStore[mbAddr][]. For any macroblock, if it is decided to reuse the intra prediction modes from previous picture, its associated RefMBStore remains unchanged. Otherwise, the associated RefMbStore is updated by the current macroblock and neighbors. In other words, for any macroblock location, the corresponding RefMBStore stores the macroblock and neighbors (at that location)

$$D(b_2,b_1)=\sum_{j=b_y-1}^{b_y+15}\sum_{i=b_x-1}^{b_x+15}\left|p_2(j,i)-p_1(j,i)\right|+\sum_{i=b_x+16}^{b_x+23}\left|p_2(b_y-1,i)-p_1(b_y-1,i)\right| \qquad (3)$$

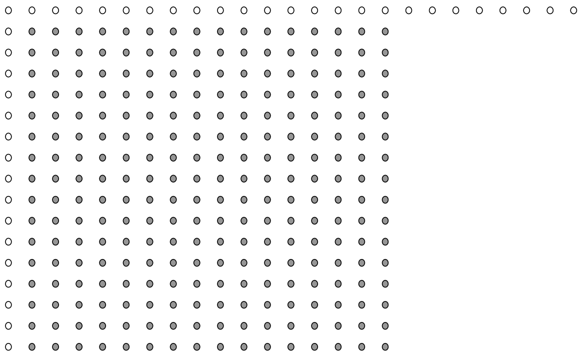


Fig. 3 - MBStore stores a macroblock (filled circles) and its adjacent spatial neighboring pixels necessary for forming intra prediction (open circles)

of the latest picture for which the RefMbStore was updated, i.e. for which the mode was re-decided. The complete algorithm is stated as follows.

```
For (CurrMbAddr = 0; CurrMbAddr<PicSizeInMbs;
CurrMbAddr++) {
    Compute the difference (D) between the
    current macroblock and its associated
    RefMBStore[CurrMbAddr][].
    if (D>=TH) {
        Perform full mode decision for the current
        macroblock;
        Update RefMBStore[CurrMbAddr][] with
        current macroblock and neighbors;
    }
    else {
        Reuse mode decision of the collocated
        macroblock in previous picture;
    }
}
```

The threshold TH can be used to control the quality-complexity tradeoff. A larger TH leads to lower quality, but faster mode decision and hence lower computational complexity.

4. SIMULATION RESULTS

Simulation settings are summarized in Table 1. Figs 4 and 5 plot the RD performance for BigShips and Crew sequences, respectively. The run time results are presented in Table 2.

As expected, different threshold values provide different tradeoffs of RD performance and computational complexities. For instance, in the case of BigShips, computational complexity is reduced by 33% with less than 0.2 dB loss, and 67% with a loss of 0.5dB. Such reductions in complexity could be quite significant for encoding of high resolution 4:4:4 pictures.

5. CONCLUDING REMARKS

In this paper, we proposed a fast mode decision algorithm for intra-only H.264/AVC video coding. The computational efficiency is achieved by exploiting temporal correlations between temporally adjacent pictures. Compared to the RD optimized mode decision algorithm in JM10.1, the proposed algorithm can reduce the computational complexity up to 33% with negligible loss of compression efficiency.

REFERENCES

1. ITU-T Rec. H.264 | ISO/IEC 14496-10, “Advanced Video Coding,” 2003.
2. K. Bruce, L. Cardelli and B. Pierce, “Comparing object encodings,” in Abadi, M., Ito, T., eds., “Theoretical Aspects of Computer Software” Vol.0 1281 of Lecture Notes in Computer Science, Berlin Heidelberg New York, Springer-Verlag pp. 415–438, 1997.
3. H. Yu, and L. Liu, “Advanced 4:4:4 Profile for MPEG4-Part10/H.264,” JVT-P017 (rev1), July 2005.
4. F. Pan, et al., “Fast Mode Decision for Intra Prediction,” JVT-G013, 7th meeting, Pattaya, Thailand, March 2003.
5. B. Meng, O. Au, C.-W. Wong and H.-K. Lam, “Efficient Intra-Prediction Mode Selection for 4x4 Blocks in H.264,” Proc. ICME, Baltimore, USA, July 2003.
6. Y.-D. Zhang, F. Dai and S.-X. Lin, “Fast 4x4 Intra-prediction Mode Selection for H.264,” Proc. ICME, Taipei, Taiwan, June 2004.
7. F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, “A Directional Field Based Fast Intra Mode Decision Algorithm for H.264 Video Encoding,” Proc. ICME, Taipei, Taiwan, June 2004.
8. J. Xin, A. Vetro and H. Sun, “Efficient macroblock coding-mode decision for H.264/AVC video coding,” Proc. Picture Coding Symposium (PCS), December 2004.

Table 1. Simulation settings

	Item	**Description**
Code generation	Hardware	Pentium IV 3.0GHz, 2GB Memory
	Software	Windows XP, MS VC V6
	Intra AVC Encoder	{JM10.1, Proposed (based on JM10.1)}
Encoder configuration	Video sequences	120 frames of {BigShips, Crew}
	Chroma format	4:2:0
	Resolution	1280x720p, 60fps
	Frame rate (fps)	60
	QP	{26, 29, 32}
	Entropy coding	CABAC
	Profile	High
	TH (for proposed)	{128, 256, 512, 1024}

Table 2. Run time results

	Encoder	**Run time (normalized by JM10.1)**		
		QP=26	**QP=29**	**QP=32**
BigShips	JM10.1	100	100	100
	New-TH256	82.7	95.0	95.0
	New-TH512	67.5	82.2	77.1
	New-TH1024	32.8	64.0	38.0
Crew	JM10.1	100	100	100
	New-TH256	83.4	89.6	92.2
	New-TH512	58.4	61.5	60.5
	New-TH1024	33.4	37.3	38.7

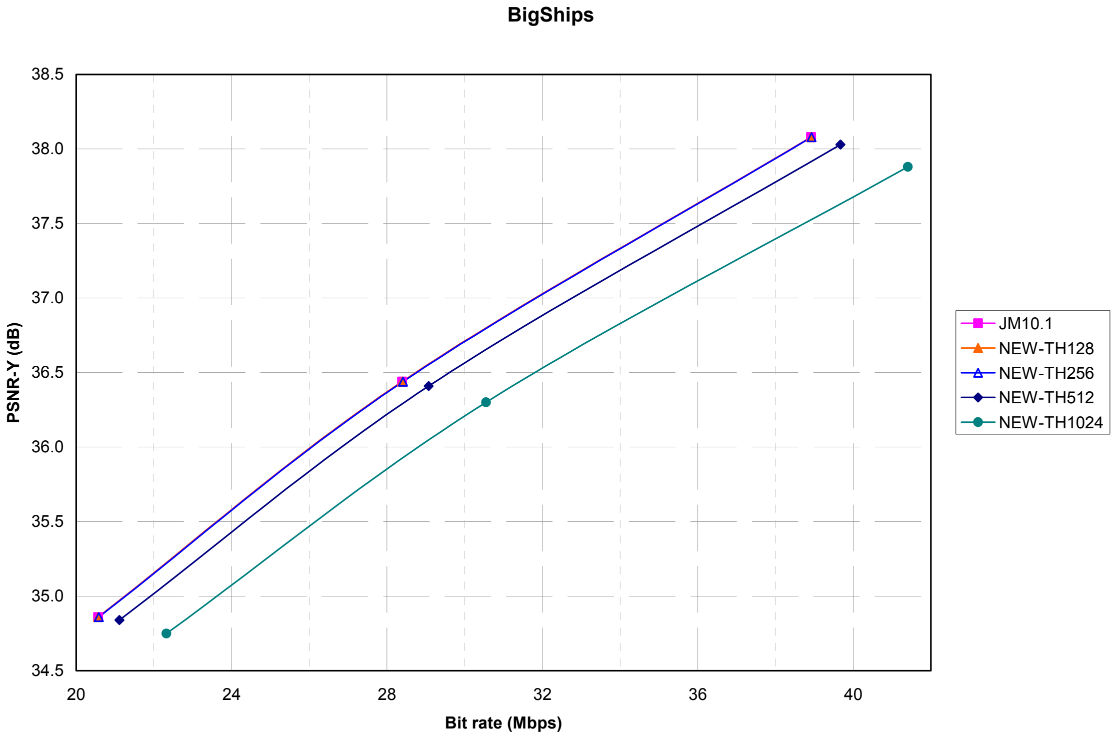


Fig. 4. Rate-distortion performance comparison of proposed algorithms versus JM10.1 for BigShips

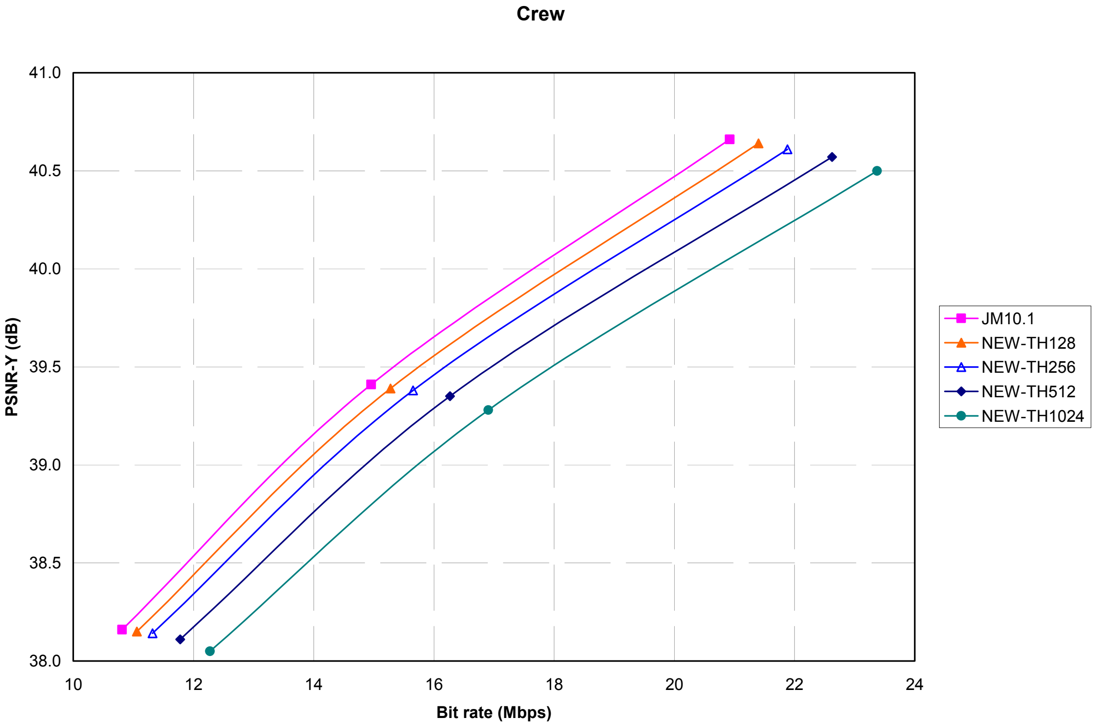


Fig. 5. Rate-distortion performance comparison of proposed algorithms versus JM10.1 for Crew