## AN IMPROVED SATD-BASED INTRA MODE DECISION ALGORITHM FOR H.264/AVC

*Yu-Ming Lee, Jyun-De Wu and Yinyi Lin* Department of Communication Engineering National Central University, Taiwan 32054, R.O.C.

# ABSTRACT

In the previous work [3] we presented a SATD-based intra mode decision algorithm for H.264/AVC to reduce the computational complexity, based upon SATD coefficients. In this paper, we extend the previous work and propose an improved SATD-based intra mode decision, which uses low-frequency components to provide local characteristics for intra mode selection. The experimental results reveal that our proposed algorithm can achieve significant reduction of computation time compared to the reference program, while still maintaining good coding performance. The proposed algorithm is also compared with two distinct algorithms and the results indicate that the proposed algorithm brings out higher reduction in computation that these two algorithms.

*Index Terms*— H.264/AVC, intra prediction, mode decision, discrete cosine transform, AClow

# **1. INTRODUCTION**

The latest H.264/AVC video coding standard achieves significantly better performance in both PSNR and visual quality at the same bit rate compared with prior video coding standards. One important technique is the uses of Lagrangian rate-distortion optimization (RDO) for inter mode decision as well as intra mode decision. The RDO technique is used to check all possible inter modes and intra modes to find the best coding result, but the computational load is far beyond any existing video coding algorithm.

Many efforts have been made to investigate the fast intra mode decision algorithm to reduce the computation complexity of the H.264/AVC encoder [1]-[7]. In general, there are two ways to reduce computation cost. One is to speed up the mode decision using SAD/SATD properties to reduce the candidate directional modes of intra prediction for RDO [1]-[3]. In [2], a rate estimation method using the  $\rho$ -domain model was suggested for rate-distortion cost function, while in [3] the standard deviation information of SATD coefficients was employed to estimate the real rate. The other one is to employ the edge detection algorithm to select a small part of the intra prediction for RDO calculation [4]-[7]. In [4] and [5], the Sobel operator was suggested in edge detection algorithm for intra prediction, based on edge direction histogram and image structure tensor (IST) respectively. The algorithm using image structure analysis achieves better RD performance than that using the edge histogram. In [6] and [7], dominant edge direction (DED) or strength algorithm was proposed to select a subset of intra prediction for RDO calculation, based on AC components of discrete cosine transform (DCT) coefficients and edge histogram descriptor.

In the previous work [3], we have presented a SATDbased intra mode decision for H.264/AVC to alleviate the encoder complexity while maintaining picture quality. The SATD-based intra mode decision in the previous work makes use of the SATD value and its variance information to filter out some directional modes for prediction. In this paper, we investigate a improved algorithm to reduce the computation complexity of the H.264/AVC encoder, in which we use the low-frequency components (denoted as  $AC_{low}$ ) to determine intra 4x4 prediction (denoted as I4MB) or intra 16x16 prediction (denoted as I16MB) for a macroblock (MB).

# 2. INTRA MODE PREDICTION AND DECISION IN H.264/AVC

H.264/AVC intra coding provides two intra mode predictions in luma components: I4MB and I16MB. For I4MB, each 4x4 luma block selects one of nine prediction modes, including dc prediction mode and eight directional prediction modes. The nine prediction modes are illustrated in Fig. 1(a). For I16MB, prediction modes as depicted in Fig. 1(b), each 16x16 luma block selects one of following modes.

For chroma components, each 8x8 block is predicted from its neighboring chroma samples. There are four prediction modes in chroma 8x8 mode prediction (denoted as C8MB), which are similar to those of I16MB except with a difference prediction mode order: Mode 0 (DC), Mode 1 (horizontal), Mode 2 (vertical) and Mode 3 (plane).

The reference software of the H.264/AVC encoder proposes RDO technique to select the best intra mode, by calculating a Lagrangian cost function, which includes both distortion and rate.

$$J(c, p, MODE | QP, \lambda_{MODE}) = SSD(c, p, MODE | QP) + \lambda_{MODE} \cdot R(c, p, MODE | QP)$$
(1)

This work was supported in part by the MediaTek Fellowship and in part by the National Science Council of R.O.C. under Grant Number NSC 96-2221-E-008-013-MY2

where *QP* is the quantization parameter with  $0 \le QP \le 51$ , the Lagrange multiplier  $\lambda_{MODE}$  is a *QP* dependent variable with  $\lambda_{MODE} = 0.85 \cdot 2^{(QP-12)/3}$ , and *SSD* is the sum of the squared differences between the current block *c* and the reconstructed block *p*. R(c, p, MODE | QP) represents the truly encoded bits associated with the chosen mode and *QP*.





H.264/AVC uses the RDO technique to achieve the best coding performance. This indicates that the H.264/AVC encoder must encode all intra modes and select the one that gives the minimum RD cost. Since the selection of the prediction modes in chroma components is independents to that of luma components, for a MB it has to perform 592  $((9 \times 16 + 4) \times 4)$  separate RDO calculations the **RDO-optimal** before obtaining mode. The computational load of H.264/AVC encoder is immoderately high. In reference software JM12.2, a very efficient intra mode decision for C8MB has been proposed in which only one out of four prediction modes is selected based on SATD values. As a result, only 148 separate RDO calculations in luma components are performed and about 70~75% of computation time can be avoided, with negligible performance degradation.

# 3. AN IMPROVED SATD-BASED INTRA MODE DECISION USING DCT AND SATD COEFFICIENTS

**3.1. I4MB/I16MB mode selection using DCT coefficients** The H.264/AVC encoder executes both I4MB and I16MB for luma components to find the best mode. In most video sequences a lot of homogeneous areas exist, and the MB in these areas ends up with being resolved as I16MB after the computationally expensive RDO. In the video sequence belonging to non-homogeneous or high detail areas, I4MB usually brings out the best coding efficiency. The homogeneous characteristics of a video sequence can be described by its low-frequency DCT components, in which flat areas commonly possess small low-frequency components of the DCT block while high detail areas possess large low-frequency components. In this section, we employ the low frequency components of DCT coefficients of a MB to classify a MB as a homogeneous or nonhomogeneous MB, and MB is then predicted using only 116MB or I4MB.

The 16x16 DCT coefficients of a MB is given by

$$X(u,v) = \frac{1}{8}E(u)E(v)\sum_{l=0}^{15}\sum_{k=0}^{15}x(l,k)\cos\frac{(2l+1)u\pi}{32}\cos\frac{(2k+1)v\pi}{32}$$
 (2)

where  $E(u) = E(v) = 1/\sqrt{2}$  for u = v = 0 and E(u) = E(v) = 1 for other *u* and *v*. The low-frequency components X(u,0) and X(0,v) can be expressed as

$$X(u,0) = \frac{1}{8\sqrt{2}} \sum_{l=0k=0}^{15} \sum_{k=0}^{15} x(l,k) \cos\frac{(2l+1)u\pi}{32}$$
(3)

$$X(0,\nu) = \frac{1}{8\sqrt{2}} \sum_{k=0}^{15} \sum_{l=0}^{15} x(l,k) \cos\frac{(2k+1)\nu\pi}{32}$$
(4)

The sum of absolute low-frequency components |X(u,0)| and |X(0,v)| in some degree can reflect the flatness of a MB, which is given by

$$AC_{low} = \sum_{u=1}^{15} |X(u,0)| + \sum_{\nu=1}^{15} |X(0,\nu)|$$
(5)

The cumulative distribution function (CDF) of  $AC_{low}$  for MBs finally determined as I4MB and I16MB is depicted in Fig. 2, conducted on the *flower* sequence. As demonstrated, the MBs encoded with I16MB have small  $AC_{low}$ , which belong to flat or homogeneous areas. On the other hand, most of MBs that have large  $AC_{low}$  are finally encoded with I4MB.

To lessen the computational complexity associated with intra mode decision, a binary hypothesis testing is employed to encode a MB, based on  $AC_{low}$  value. If the  $AC_{low}$  value is less than a threshold  $\eta$ , it is encoded with I16MB. Otherwise, it is encoded with I4MB. Assume  $P(H_0)$  and  $P(H_1)$  as the priori probabilities of a MB resolved as I16MB and I4MB respectively, and  $f(x|H_0)$  and  $f(x|H_1)$  as the associated conditional probability density function of  $AC_{low}$  when  $H_0$  and  $H_1$  are true. Define the cost function C as

$$C = \begin{bmatrix} C_{00} & C_{01} \\ C_{10} & C_{11} \end{bmatrix}$$
(6)

where  $C_{ij}$  represents the cost of choosing  $H_i$  when  $H_j$  is true. The optimal mode decision rule is to pick a strategy to minimize the average cost  $\overline{C}$ , given by

$$\overline{C} = P(H_0) \{ C_{00} (1 - P(E \mid H_0) + C_{10} P(E \mid H_0)) \} + P(H_1) \{ C_{11} (1 - P(E \mid H_1) + C_{01} P(E \mid H_1)) \}$$
(7)

where  $P(E | H_0)$  is a type I error, representing a MB is encoded with I4MB when the true mode is I16MB; and  $P(E | H_1)$  is a type II error. They are given by

$$P(E \mid H_0) = \int_{\eta}^{\infty} f(x \mid H_0) dx$$
(8)

$$P(E \mid H_1) = \int_{-\infty}^{\eta} f(x \mid H_1) dx$$

where  $\eta$  is a threshold. Substituting (8) into (7) and using  $\frac{\partial \overline{C}}{\partial \eta} = 0$ , we can obtain the optimal threshold  $\eta_{opt}$ . To simplify the problem, we assume  $C_{00} = C_{11} = 0$  and  $C_{01} = C_{10} = 1$ , and both conditional probability density functions are Gaussian random variables with variances  $\sigma_{16}$  and  $\sigma_4$ , and means  $\mu_{16}$  and  $\mu_4$ , respectively. The optimal threshold  $\eta_{opt}$  is then given as

$$\eta_{opt} = \frac{(\mu_4 \sigma_{16}^2 - \mu_{16} \sigma_4^2) - 2\sigma_4 \sigma_{16} \sqrt{(\mu_4 - \mu_{16})^2 + 2(\sigma_{16}^2 - \sigma_4^2) \ln(\frac{P(H_1)\sigma_{16}}{P(H_0)\sigma_4})}}{\sigma_{16}^2 - \sigma_4^2} \quad (9)$$

The parameters in (10) vary depending on QP; particularly on priori probabilities  $P(H_0)$  and  $P(H_1)$ , I4MB is prevailing in a high bit-rate video (low QP) while I16MB becomes dominant in a low bit-rate video (high QP). An exhaustive experiment was conducted on many CIF and QCIF video sequences to evaluate the optimal threshold  $\eta_{opt}$  for various QPs.  $\eta_{opt}$  is experimentally taken as

$$\eta_{opt} = \mu_{16} + \frac{1}{6}(\mu_4 - \mu_{16})\exp[(QP - 40)/12]$$
(10)

which is a function of means and QP.

An experiment was conducted with JM12.2 reference software to evaluate the performance, in which the fast intra mode decision for C8MB is used. Each video sequence is coded with all I frames. In experiment we employ  $AC_{low}$ and  $\eta_{opt}$  to encode a MB, in which  $\eta_{opt}$  is calculated using means  $\mu_{16}$  and  $\mu_4$ , dynamically computed using moving average. Fig. 3 displays the hit or correct rate using our proposed algorithm. As shown, a very large number of MBs are encoded with correct modes. The results of PSNR loss, bit-rate increment and time saving are displayed in TABLE I (fifth column,  $AC_{low}$ ) for QP = 28. As shown, a reduction of average 24% of total encoding time can be accomplished, while still preserving good video quality.

### 3.2. Summary of proposed algorithm

Success of the SATD-based intra mode decision [3] (denoted as SATD) is achieved by discarding the directional prediction modes of intra prediction according to SATD and its variance information. To improve the computation efficiency, we employ low-frequency components of DCT

coefficients to describe the homogeneous characteristics of a MB, and they are used to select either I4MB or I16MB for mode prediction. The block diagram of our proposed improved SATD-based intra mode decision algorithm is illustrated in Fig. 4.



Fig. 2 Cumulative distribution function (CDF) of  $AC_{low}$  for I4MB and I16MB







Fig. 4 Improved SATD-based intra mode decision algorithm

## 4. EXPERIMENTAL RESULT

Several fast and effective algorithms for intra mode decision in the H.264/AVC encoder have been proposed recently. Two individual approaches, the IST algorithm [5] and the DED algorithm [6] are used for comparison with the proposed algorithm. These mentioned algorithms are implemented into the JM12.2 encoder to evaluate the performance, which the fast intra mode decision for C8MB is used. Each video sequence is coded with all I frames. The experimental results for QP = 28 are summarized in TABLE I. As displayed, the degradation in PSNR and bit-rate increment in both proposed algorithm and IST algorithm are trifling, compared to the original algorithm; while the DED algorithm runs into average 3.9% of bit-rate increment. The proposed algorithm, to whatever extent, accomplishes a reduction of 52% of encoding time, 22% from I4MB/I16MB mode decision and 30% from I4MB/I16MB prediction mode selection. The computation saving in proposed algorithm is much higher than those of IST algorithm (with an average 22% reduction only) and DED algorithm (with an average 39% reduction). The rate-distortion curve as well as encoding time for flower.CIF video sequence is also shown in Fig. 5 for comparison for various *QPs*. As can be seen, the proposed algorithm brings out a momentous reduction in computation.

## **5. CONCLUSION**

In this paper, we extended the previous work and presented an improved SATD-based intra mode decision for H.264/AVC to reduce the computational load. The proposed algorithm is accomplished in two stages. In the first stage, we employed  $AC_{low}$  to select an I4MB or I16MB mode for prediction; while in second stage we used SATD coefficients to eliminate unlikely directional prediction modes. The results in all intra frame experiments reveal that average 52% of computation time can be saved while maintaining good coding performance. The coding performance is also compared with IST and DED algorithm, and the results show that the computation saving in the proposed algorithm is much higher than both IST and DED algorithms.



(a) Rate-distortion curve



(b) Encoding time Fig. 5 Performance comparison for flower.CIF (III...)

#### 6. REFERENCES

[1] C. Kim, H. H. Shih, and C. C. J. Kuo, "Feature-based intra-prediction mode decision for H.264," in *Proc. IEEE ICIP*, pp. 769-772, 2004.

[2] H. Kim and Y. Altunbasak, "Low-complexity macroblock mode selection for H.264/AVC encoder," in *Proc. IEEE ICIP*, pp. 765-768, 2004.

[3] Y. T. Sun, and Y. Lin, "SATD-based intra mode decision for H.264/AVC video coding," in *Proc. IEEE ICME*, pp. 61-64, June 2008.

[4] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast mode decision algorithm for intraprediction in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 813-822, July 2005.

[5] C. Huang, S. Zhung, and S. H. Lai, "Efficient intra mode selection using image structure tensor for H.264/AVC," in *Proc. IEEE ICIP*, pp. V-289-V-292, Sept. 2007.

[6] B. La, M. Eom, and Y. Choe, "Fast mode decision for intra prediction in H.264/AVC encoder," in *Proc. IEEE ICIP*, pp. V-321-V-324, Sept. 2007.

[7] J. F. Wang, J. C. Wang, J. T. Chen, A. C. Tsai, and A. Paul," A novel fast algorithm for intra mode decision in H.264/AVC encoders," in *Proc. IEEE ISCAS*, pp. 3498-3501, July 2006.

TABLE I Performance comparison for QP=28

(a) PSNR loss

QP=28		$PSNR(dB, \triangle(dB))$						
Sequence		Orig.	IST [5]	DED [6]	SATD [3]	AC low	Proposed	
CIF	Tempete	36.635	-0.105	-0.204	-0.107	-0.001	-0.105	
	Flower	37.481	-0.139	-0.263	-0.133	0.001	-0.133	
	Football	35.332	-0.097	-0.157	-0.095	0.007	-0.084	
	Bus	37.537	-0.077	-0.137	-0.073	0.005	-0.073	
	Stefan	37.354	-0.097	-0.182	-0.108	-0.001	-0.111	
QCIF	Foreman	37.062	-0.025	-0.077	-0.050	-0.007	-0.050	
	Mobiel	35.654	-0.091	-0.222	-0.127	-0.003	-0.127	
	Grandma	37.557	-0.064	-0.141	-0.059	-0.004	-0.060	
	Mthr_dotr	37.805	-0.058	-0.119	-0.058	0.000	-0.065	
	Salesman	36.692	-0.059	-0.124	-0.070	-0.008	-0.075	
Avg.		-0.081	-0.163	-0.088	-0.001	-0.088		

(c) Bit fuite Comparison								
QP=28		Bit rate ( bps,△(%) )						
Sequence		Orig.	IST [5]	DED [6]	SATD [3]	AC low	Proposed	
CIF	Tempete	4883208	0.47	3.06	-0.10	0.02	-0.02	
	Flower	5882784	0.48	2.18	0.02	0.15	0.17	
	Football	4594963	0.25	2.83	-0.31	0.26	0.00	
	Bus	3567139	1.00	5.34	0.05	0.22	0.31	
	Stefan	4646645	0.72	4.29	-0.07	0.10	0.00	
QCIF	Foreman	791009	0.36	5.45	0.05	0.17	0.22	
	Mobiel	2099482	0.43	2.38	0.00	0.01	0.08	
	Grandma	657941	0.65	3.93	-0.11	0.17	0.02	
	Mthr_dotr	653018	0.67	5.45	0.08	0.15	0.18	
	Salesman	888768	0.55	5.03	-0.16	0.04	-0.05	
Avg.		0.56	3.99	-0.06	0.13	0.09		

(b) Bit rate comparison

		•
(C	) Timing	comparison

QP=28		Time ( ms,△(%) )						
Sequence		Orig.	IST [5]	DED [6]	SATD [3]	AC low	Proposed	
CIF	Tempete	18094	-24.59	-39.58	-31.71	-21.50	-52.29	
	Flower	18066	-24.86	-40.06	-28.24	-30.05	-52.25	
	Football	15529	-24.94	-38.87	-29.07	-23.32	-50.18	
	Bus	16484	-24.07	-39.13	-33.21	-23.85	-54.86	
	Stefan	17262	-24.53	-38.63	-31.84	-25.73	-55.10	
QCIF	Foreman	3998	-19.46	-39.02	-34.42	-20.91	-45.00	
	Mobiel	5203	-19.18	-39.42	-29.98	-20.73	-51.37	
	Grandma	3719	-19.01	-37.86	-27.80	-29.56	-52.49	
	Mthr_dotr	3711	-20.51	-39.85	-25.09	-26.39	-52.49	
	Salesman	4078	-18.78	-38.30	-30.58	-20.72	-52.89	
Avg.			-21.99	-39.07	-30.19	-24.28	-51.89	