MODEL-BASED EARLY TERMINATION SCHEME FOR H.264/AVC INTER PREDICTION

Lien-Fei Chen, Shin-Ping Yang, and Yeong-Kang Lai

Department of Electrical Engineering, National Chung Hsing University
No. 250, Kuo Kuang Road, Taichung, Taiwan

ABSTRACT

Inter prediction is the most power-consumed component in H.264/AVC encoder. For the power-aware design, it is necessary to reduce the power consumption by using fast inter prediction techniques. In this paper, a system-level power-aware algorithm based on the early termination scheme of H.264/AVC inter prediction is proposed. The proposed early termination scheme for H.264 motion estimation is based on the statistical modeling of the motion compensated residual data. We develop a power-aware adaptive mechanism with multiple thresholds derived from the statistical model to early terminate the motion estimation (ME). According to the experimental results, our early termination scheme for H.264/AVC inter prediction not only preserves fine RD performance but also eliminates the unnecessary operation in both IME stage and FME stage to realize the power-aware H.264 encoder system.

***Index Terms*—**Early Termination, Skip Mode, Motion Estimation, VLSI, H.264/AVC

1. INTRODUCTION

The emerging video coding standard H.264/AVC is developed to achieve higher compression efficiency of moving pictures for various applications. Many advanced features are adopted in H.264/AVC, such as space-domain intra prediction, multiple reference frames, variable block-size motion estimation (VBSME), deblocking filter, integer transform, …, etc. Hence, the compression performance in H.264 is improved with a huge increase of the computational complexity. Therefore, the low power techniques are the significant bottleneck in reducing computations for H.264 encoder.

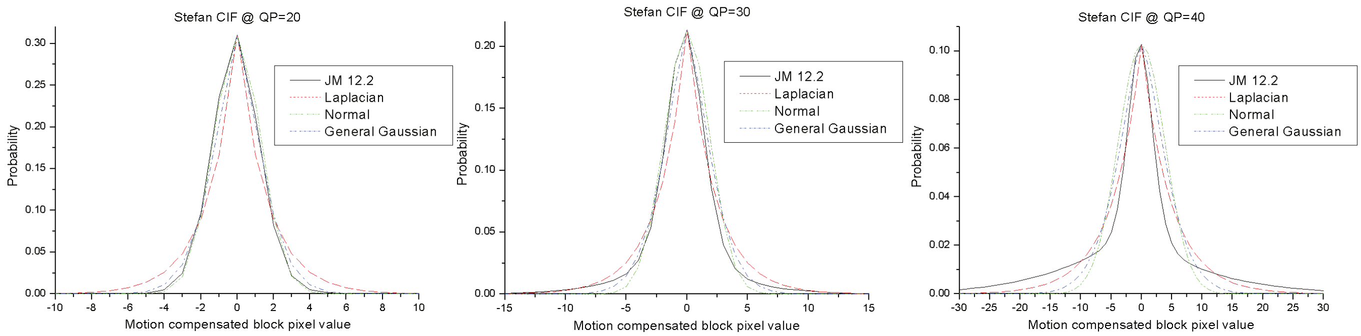
According to the power profile of H.264 baseline profile encoder chip, Chen *et al.* [1] shows the motion estimation architecture, including IME and FME, consume more than 70% power of the whole encoder. Therefore, the low power fast inter prediction techniques are the significant design challenges in reducing computation for H.264 encoder. Many fast IME and FME algorithms are proposed to save the computational power of the IME engine and FME architecture respectively. However, both of these fast algorithms have their own limit for the low-power design. The system-level fast inter prediction techniques are necessary to go a step future to reduce the computation of both IME and FME, such as the fast mode decision algorithms. Some hardware-oriented fast mode decision algorithms have been proposed to reduce the computation of the FME unit [2]-[4]. According to their decision criterions, a set of the probable modes are selected after IME stage to save 43%–73% computing cycles of FME engine.

TABLE I
THE PROBABILITY OF THE SKIP MODE MB IN H.264 BASELINE

	QP=20	QP=25	QP=30	QP=35	QP=40	QP=45
akiyo	74,27	80.46	86.79	91.45	94.92	97.42
foreman	7.04	17.63	35.99	52.45	64.29	73.41
stefan	8.62	14.30	23.08	36.00	50.44	62.16
mobile	0.93	2.42	7.90	21.28	42.97	57.98
hall	0.13	18.57	61.00	84.04	91.59	95.17
coastguard	1.92	8.41	24.10	44.71	64.65	79.23
container	27.38	53.45	77.12	87.04	91.94	95.46
news	61.26	70.13	78.13	83.76	88.77	92.52
average	**22.69**	**33.17**	**49.26**	**62.59**	**73.70**	**81.67**

The early ME terminated mechanism is another significant interest to terminate the operation of the inter prediction including IME and FME. For the nature video sequence, there exist a lot of homogeneous or stationary regions which belong to the same video objects or exhibit similar motion even if still. Most of these homogenous blocks and stationary regions are coded as the *SKIP* mode, and their motion vectors often occur around (0, 0) or MVP. Table I shows the *SKIP* mode statistics with some popular test video sequences. According to this table, the probability of the *SKIP*-MB is appreciable especially for high *QP* applications. In addition, the early termination scheme is the context-adaptive algorithm. More than half of the MBs are usually coded as the *SKIP*-MB for the slow motion pictures or surveillance applications such as akiyo, hall, container, news. Therefore, many early ME-terminated algorithms are proposed based on the fast *SKIP* prediction algorithms [1][5]-[7]. However, these fast *SKIP* detection methods either depends on the constant threshold values [5] or are determined by a great quantity of experimental data [1][6][7].

In this paper, we will propose an early termination scheme to detect the homogeneous MB and the stationary MB efficiently. In view of the theoretical model, a mathematic method to determine the threshold value is proposed to early terminate motion estimation operation. By way of modeling the motion compensated residual data, a hybrid



(a) (b) (c)

Fig. 1. Distribution of the pixel values in the motion compensated blocks in H.264 by using stefan sequence with (a) QP = 20, (b) QP = 30, (c) QP = 40

model based on generalized Gaussian distribution (GGD) is utilized to develop the adaptive threshold values for our early termination mechanism.

2. Modeling of the Motion Compensated Residual Data in H.264/AVC

A. Statistics of the Motion Compensated Residual Pixels

Many statistical models of the motion-compensated residual data are analyzed and proposed. Pao *et. al.* [8] analyzes the distribution of the pixel values after motion-compensated prediction can be modeled by the zero-mean Laplacian distribution. In [9], Wang *et. al.* consider that recent advanced motion estimation algorithms have shown that the motion compensated video frames are very close to random noise. Thus, the motion compensated video frames are best to be modeled by Gaussian distribution.

In order to obtain the accurate mathematical model, we simulate the JM12.2 reference software, and gather the values of the motion compensated residual pixels. Fig. 1 shows the plot of the histogram of the motion compensated residual values. The video sequence used here is from the STEFAN sequence. General Gaussian means that the index number of the exponent is 1.5 for the GGD model. From Fig. 1, the distribution is approximated by Gaussian distribution at low QP (high bit-rate). But it is most commonly approximated by Laplacian distribution at high QP (low bit-rate). Other video sequences, are also simulated by JM12.2, have the same property. It is reasonable that the motion compensated residual data are very close to random noise because of the requirement of the high video quality at high bit-rate video. When the QP step is increases, the distortion issue is not the critical point for low bit-rate applications. The residual data can not be regarded as random noise and its distribution will approximate the GGD even if the Laplacian distribution at medium and low bit-rate applications.

According to the analysis, the motion compensated residual data can be approximated by the zero-mean generalized Gaussian distribution (GGD) probability density function:

$$p(x;\sigma,\alpha) = a \cdot \exp\left(-|x/b|^{\alpha}\right), x \in \mathrm{R} \tag{1}$$

where

$$a = \frac{1}{2\Gamma(1+1/\alpha)b};\ b = \sigma\sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} \tag{2}$$

In (1), the index number of the GGD model α is the shape parameter. When $\alpha = 1$, it is a Laplacian distribution, and $\alpha = 2$ corresponds to a normal distribution. According to Fig. 1, we find that residual data can be modeled by GGD model with the shape parameter from 1 to 2. In order to simplify the complexity of the modeling, the values 1, 1.5, and 2, are chosen for shape parameter to model the residual blocks for different QP parameters.

B. Detection of All-Zero Transform Quantized Block

Because the homogenous blocks and stationary regions are usually coded like the SKIP mode with the motion vectors around (0, 0) or MVP, and all transform quantized coefficients within the block are also usually quantized as zero. Therefore, we are confident that the MB is homogenous or stationary if its motion vector is around the (0, 0) or MVP and the quantized coefficients of all 4×4 blocks within the MB are zeros. However, we can only obtain the SAD values at the (0, 0) and MVP at IME stage. The all-zero block detection method is necessary to find the relation between the SAD value and the threshold value of the all-zero transform quantized coefficients.

According to the definition of the H.264 quantization, the transform quantized coefficient will be zero if the following inequality holds:

$$|\mathrm{TR}[i][j]| < \frac{2^{qp\_bits} - qp\_const}{quanti\_coef[qp\_rem][i][j]} \equiv \mathrm{TH}_{\mathrm{TR}}[\mathrm{QP}][i][j] \tag{3}$$

Because the threshold value TR in (3) is not available at IME stage. We should find the relation between the SAD value and TR. Thus, we can deduce the threshold value for the SAD according to this correlation.

The derivation of threshold value is very similar to [10]. Here, we only focus on the results of the threshold value based on the GGD model with three different shape parameters. For the Laplacian distribution, the all-zero situation holds if SAD value is smaller than predefined threshold as follows:

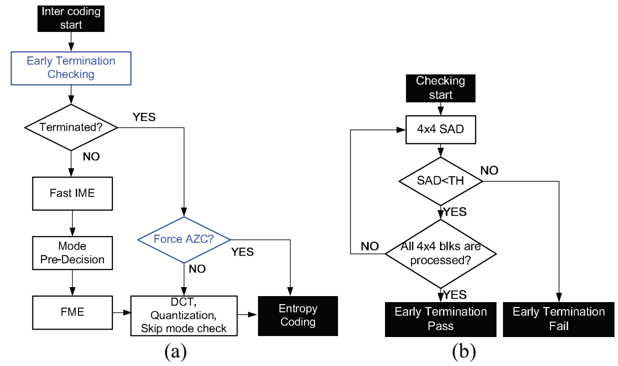


Fig. 2. (a) Proposed fast inter prediction algorithm with early termination scheme and inter mode pre-decision. (b) The detail decision flow of proposed early termination scheme.

$$\mathrm{SAD} < \frac{16\times \mathrm{TH}_{\mathrm{TR}}[\mathrm{QP}][0][0]}{9.47\sqrt{2}\cdot n} \equiv \mathrm{TH}_{\mathrm{Laplacian}} \tag{4a}$$

In (4a), the threshold value $\mathrm{TH}_{\mathrm{Laplacian}}$ is the function of QP and parameter n. It means that threshold value $\mathrm{TH}_{\mathrm{Laplacian}}$ will raises with the increasing of the quantization step. The parameter n controls the probability that transformed coefficients will be quantized to zero. For $n = 3$, the probability of Laplacian distribution that a value will fall within $(-3\sigma, 3\sigma)$ is about 99%. When $n = 2$, then the probability of this coefficient being zero after quantization will be reduced to 94%. In the same way, the threshold values based on the normal distribution and the GGD model ($\alpha = 1.5$) are shown in (4b) and (4c) respectively.

$$\mathrm{SAD} < \frac{16\times \mathrm{TH}_{\mathrm{TR}}[\mathrm{QP}][0][0]}{9.47\sqrt{\frac{2}{\pi}}\cdot n} \equiv \mathrm{TH}_{\mathrm{Normal}} \tag{4b}$$

$$\mathrm{SAD} < \frac{16\times \mathrm{TH}_{\mathrm{TR}}[\mathrm{QP}][0][0]}{9.47\times 1.303127236\times n} \equiv \mathrm{TH}_{\mathrm{GGD}} \tag{4c}$$

3. Model-Based Early Termination Scheme

Fig. 2(a) shows the data flow of the proposed fast inter prediction with the early termination scheme and inter mode pre-decision. When the H.264/AVC inter-prediction coding process enables, early termination checking scheme will be performed at first. If the checking is negative, IME and FME processing are executed as the regular flow. The checking result is positive, all inter prediction process are terminated, and the motion vector will be set as the (0, 0) or MVP. The motion compensated residual data will be fed into intra frame coding engine. Here, the AZC flag is also utilized for the intra frame coding engine to decide the residual coding flow. When AZC flag is true, all transform quantized coefficients of the motion-compensated residual MB force to zeros, and the residual coding flow is also terminated to reduce the power consumption.

The detail early termination checking flow is shown in Fig. 2(b). Owing to the 4×4 integer transform adopted in H.264, each MB is divided into 16 4×4 blocks, and then compare each SAD value with predefined threshold. If any of these sixteen SAD values is larger than the threshold value, the early termination checking is aborted to assure all 4×4 blocks of the all-zero blocks.

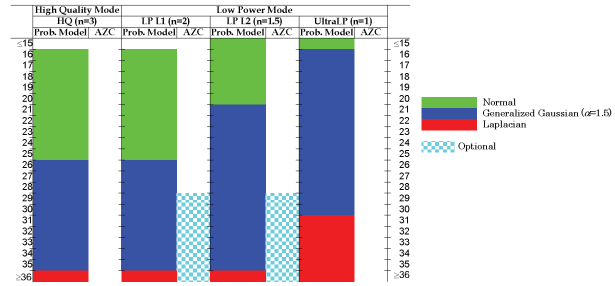


Fig. 3. Proposed hybrid model based on generalized Gaussian distribution (GGD).

For our early termination scheme, one design bottleneck is how to determine the threshold value in Fig. 2(b). According to above discussion, the residual data can be modeled by GGD model with different shape parameters. In the light of three shape parameters, three corresponding threshold values are utilized in terms of the quantization step and the encoding status. In order to get the best choice for the trade-off between video quality and power consumption, a hybrid model based on GGD model is proposed in Fig. 3. For the different quantization steps, different models are exploited. In order to ensure the accuracy, it is necessary to employ the pessimistic model as the conservative estimation. For example, we utilize the GGD model with $\alpha = 1.5$ instead of the normal distribution for QP = 15–20 at ultra low power (UltraLP) mode.

4. Performance Analysis

The R-D performance comparison among our early termination scheme with hybrid model and the JM12.2 reference software is presented in Fig. 4. In this experiment, Akiyo and Stefan CIF video sequences are exploited. Moreover, the detail performance comparison of the △PSNR, △bitrate, and S% of the proposed early termination scheme is also presented in Table II. In this table, S% stands for the detection rate of the early termination scheme. The large S% value represents more computation of motion estimation can be eliminated, and more decreasing of the power consumption. There are four power-aware modes, as shown in Fig. 3 and Table II, can be selected to achieve different levels of the low power requirements. When we select the HQ mode, the good RD performance can be reserved and 10% – 30% MBs will be terminated. The termination rate will increase when we select other low power mode. When the UltraLP mode is selected, more than 20% MBs will not perform inter prediction processing at low energy situation. For the low motion sequence or video surveillance like Akiyo sequence, even the S% values are larger than 50% as shown in Table II.

TABLE II

RD PERFORMANCE COMPARISON BETWEEN PROPOSED FAST INTER PREDICTION AND JM12.2 REFERENCE SOFTWARE CODE

	QP	JM12.2		Early Termination Scheme											
				HQ (n=3)			LP L1 (n=2)			LP L2 (n=1.5)			UltraLP (n=1)		
		PSNR	Bitrate (kb/sec)	△PSNR	△bitrate	S%	△PSNR	△bitrate	S%	△PSNR	△bitrate	S%	△PSNR	△bitrate	S%
STEFAN	24	38.546	2389.693	-0.005	0.31%	3.95%	-0.023	0.52%	12.19%	-0.011	0.61%	8.50%	-0.048	0.98%	15.55%
	26	37.000	1834.074	-0.004	0.19%	1.42%	-0.009	0.70%	6.99%	-0.019	1.00%	12.71%	-0.060	1.94%	17.30%
	28	35.501	1395.492	-0.008	0.65%	3.02%	-0.017	1.10%	10.46%	-0.031	1.50%	14.38%	-0.080	2.70%	18.31%
	30	33.819	1026.660	-0.003	0.43%	5.19%	-0.012	1.56%	12.49%	-0.029	2.02%	15.58%	-0.070	3.34%	19.07%
	32	32.282	755.658	-0.002	1.28%	9.39%	-0.015	1.93%	14.89%	-0.037	2.26%	17.16%	-0.068	4.09%	19.84%
	34	30.878	568.286	-0.005	2.49%	11.71%	-0.023	3.05%	15.83%	-0.040	3.88%	17.92%	-0.072	6.72%	21.19%
	36	29.383	421.750	-0.008	3.19%	12.79%	-0.011	4.00%	16.39%	-0.034	5.26%	18.40%	-0.090	10.17%	23.14%
AKIYO	24	42.261	207.882	-0.006	0.07%	14.47%	-0.005	0.22%	55.24%	0.002	0.13%	35.92%	-0.014	0.10%	67.12%
	26	41.144	156.068	-0.004	0.25%	3.00%	0.009	0.16%	24.22%	-0.001	0.05%	47.90%	-0.035	0.33%	70.53%
	28	40.002	120.600	0.002	0.24%	9.35%	0.006	0.23%	32.46%	-0.007	0.14%	51.96%	-0.050	0.51%	72.74%
	30	38.700	92.678	0.008	-0.03%	13.82%	0.007	-0.14%	35.09%	0.005	-0.002%	52.94%	-0.039	1.11%	74.57%
	32	37.493	72.854	0.017	-0.50%	21.45%	0.002	-0.71%	43.48%	-0.005	-0.73%	60.02%	-0.062	0.93%	75.10%
	34	36.345	57.776	0.012	0.30%	27.57%	0.017	0.51%	46.39%	-0.017	0.61%	61.42%	-0.081	2.86%	78.27%
	36	35.126	46.080	0.003	0.51%	28.31%	-0.011	0.31%	47.75%	-0.030	0.71%	60.91%	-0.115	3.82%	81.53%

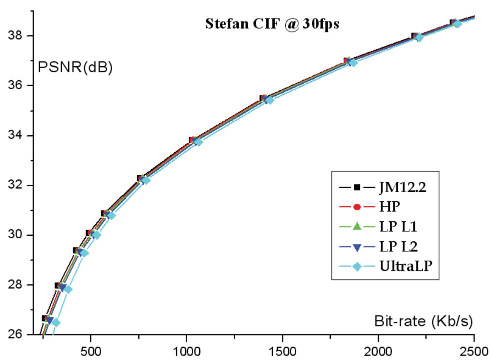


Fig. 4. RD performance analysis of the proposed fast inter prediction by using Stefan CIF sequence

5. CONCLUSION

In this paper, a model-based early termination scheme of the inter prediction is proposed for power-aware H.264 encoder. We propose the early termination scheme to early detect the homogeneous or stationary blocks to early terminate the H.264 inter prediction procedure. The proposed early termination scheme is based on the statistical modeling of the motion compensated residual data. According to the statistical GGD model with different shape parameters, we develop a power-aware adaptive mechanism with multiple thresholds. According to the experimental results, our model-based early termination scheme for the H.264 inter prediction not only preserves fine RD performance but also eliminates the unnecessary operation in both IME stage and FME stage to realize the low-power and power-aware H.264 encoder system.

6. REFERENCE

[1] Y.-H. Chen, T.-C. Chen, and L.-G. Chen, “Power-scalable algorithm and reconfigurable macro-block pipeline architecture of H.264 encoder for mobile application,” in *Proc. IEEE Int. Conf. Multimedia and Expo*, pp. 281-284, July 2006.

[2] T.-C. Chen, Y.-H. Chen, C.-Y. Tsai, and L.-G. Chen, “Low power and power aware fractional motion estimation of H.264/AVC for mobile applications,” in *Proc. IEEE Symp. Circuits Syst.*, pp. 5531-5334, May 2006.

[3] C.-L. Su, W.-S. Wang, Y.-L. Chen, Y. Li, C.-W. Chen, J.-I. Guo, S.-Y. Tseng, “Low complexity high quality fractional motion estimation algorithm and architecture design for H.264/AVC,” in *Proc. IEEE Asia-Pacific Conf. Circuits Syst.*, pp. 678-581, Dec. 2006.

[4] Y.-K. Lin, C.-C. Lin, T.-Y. Guo, T.-S. Chang, “A hardware-efficient H.264/AVC motion estimation design for high definiation video,” *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 55, no. 6, pp. 1126-1135, July 2008.

[5] D. Wu, F. Pan, K. P. Lim, S. Wu, Z. G. Li, X. Lin, S. Rahardja, and C. C. Ko, “Fast intermode decision in H.264/AVC video coding,“ *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 953-958, July 2005.

[6] C. S. Kannangara, I. E. G. Richardson, M. Bystrom, J. R. Solera, Y. Zhao, A. MacLennan, and R. Cooney, “Low-complexity skip prediction for H.264 through Lagrangian cost estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, pp. 202-208, Feb. 2006.

[7] C.-C. Lin, Y.-K.Lin, and T.-S. Chang, “Hardware efficient skip mode detection for H.264/AVC,” in *Proc. IEEE Int. Conf. Consum. Electron.*, pp. 1-6, March 2008.

[8] I.-M. Pao and M.-T. Sun, “Modeling DCT coefficients for fast video encoding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 608-616, June 1999

[9] H. Wang, and S. Kwong, “Hybrid model to detect zero quantized DCT coefficients in H.264,” *IEEE Trans. Multimedia*, vol. 9, no. 4, pp. 728-735, June 2007.

[10] Y.-W Huang, B.-Y. Hsieh, S.-Y. Chien, S.-Y. Ma, and L.-G. Chen, “Analysis and complexity reduction of multiple reference frames motion estimation in H.264/AVC,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 4, pp. 507-522, April 2006.