AN ADAPTIVE THRESHOLDING TECHNIQUE FOR THE DETECTION OF ALL-ZEROS BLOCKS IN H.264

D.Wu, K.P. Lim, T.K. Chiew, J.Y. Tham, K.H. Goh

Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613

ABSTRACT

In H.264 video coding, there are a substantial number of 4x4 blocks becoming all-zeros after transformation and quantization. This is a waste of computational resources because these skipped blocks do not require forward transform and quantization. We proposed a very effective early detection of fast skipped block detections based on the theoretical derivation of H.264 integer transform and quantization. The experimental results show that the algorithm can detect 9.71%-43.35% more zero blocks than Yong's method.

***Index Terms*—** H.264, All-Zero-Blocks Detection, Fast Transform and Quantization

1. INTRODUCTION

In view of the high complexity of H.264 encoding, most research has looked into the fast INTER/INTRA mode selections [4-12] to reduce and simplify the RDO process. In our previous work, we used edge detection to determine the homogeneity of a macroblock and adaptively restrict selection of INTER modes [8-9]. In [12], Jing and Chau categorized macroblocks into two groups and code each group with a predefined set of INTER modes. More recently, Yong et. al [3] attempted to detect 4x4 all-zero-coefficient blocks (AZBs). The reason is that if AZBs can be pre-determined, the forward and backward transformation and quantization process can be skipped entirely and a higher computational savings can be achived. They have theoretically derived the sufficient condition for quantizing each coefficient to zero by comparing the SAD of the block-to-be-quantized with an adaptive threshold.

In this paper, we have improved their algorithm significantly with detection improvement ranging from 9.71% to 43.35%. This is achieved by theoretically deriving new sets of higher adaptive thresholds which is described in the next two sections.

2. H.264 TRANSFORM

The transformed of a 4x4 prediction error inputs X , in H.264 is given as follows [1]:

$$E = C \cdot X \cdot C^T \tag{0.1}$$

The approximated integer DCT transform matrix is:

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{pmatrix}$$

and the 4x4 input block is represented by the matrix X , where $X[i,j]$ represents the value at the ith row and jth column of X . The matrix transpose operation is denoted by T.

Let the transformed coefficient, $E[u,v]$ be the value of E at the uth row and vth column, then the matrix multiplication can be rewritten as follows:

$$E[u,v] = \sum_{i=0}^{3}\sum_{j=0}^{3}\left\{C[u,i]\cdot C[v,j]\cdot X[i,j]\right\} \tag{0.2}$$

Let the quantization parameter be Q_p ranging from 0 to 51, the quantized coefficient is therefore given as:

$$E_q(u,v,r,Q_p) = sign\left\{E[u,v]\right\} \times \frac{\left|E[u,v]\right| \cdot M\left[Q_p\%6, r\right] + f}{2^{15+\left(\frac{Q_p}{6}\right)}} \tag{0.3}$$

where % denotes the modular operator. f denotes the constant $\frac{2^{15+\frac{Q_p}{6}}}{6}$ in INTER mode and $\frac{2^{15+\frac{Q_p}{6}}}{3}$ in INTRA mode, respectively. The quantization coefficient $M\left[Q_p\%6, r\right]$ in (0.3) is predefined for each frequency as follows:

$$M = \begin{pmatrix} 5243 & 8066 & 13107 \\ 4660 & 7490 & 11916 \\ 4194 & 6554 & 10082 \\ 3647 & 5825 & 9362 \\ 3355 & 5243 & 8192 \\ 2893 & 4559 & 7282 \end{pmatrix} \tag{0.4}$$

where r = $2-(u\%2)-(v\%2)$.

From equation (0.3), the magnitude of the quantized coefficient is:

$$\left|E_q(u,v,r,Q_p)\right| = \frac{\left|E[u,v]\right| \cdot M\left[Q_p\%6, r\right] + f}{2^{15+\left(\frac{Q_p}{6}\right)}} \tag{0.5}$$

The absolute value of the coefficient $\left|E[u,v]\right|$ is limited to

$$
\begin{aligned}
\left|E[u,v]\right| &= \left|\sum_{i=0}^{3}\sum_{j=0}^{3}\left\{C[u,i]\cdot C[v,j]\cdot X[i,j]\right\}\right| \\
&\le \sum_{i=0}^{3}\sum_{j=0}^{3}\left\{\left|C[u,i]\cdot C[v,j]\right|\cdot\left|X[i,j]\right|\right\}
\end{aligned}
$$

(0.6)

By inspection, the maximum value of $\left|C[u,i]\cdot C[v,j]\right|$ is 4 and (0.6) becomes:

$$
\begin{aligned}
\left|E[u,v]\right| &\le 4\cdot\sum_{i=0}^{3}\sum_{j=0}^{3}\left|X[i,j]\right| \\
&\le 4\cdot SAD
\end{aligned}
\qquad (0.7)
$$

SAD is the sum of absolute difference of the prediction errors in the 4x4 block. Using equation (0.5) and (0.7), we have the following inequality:

$$
\left|E_q(u,v,r,Q_p)\right| \le \frac{4\cdot SAD\cdot M\left[Q_p\%6,r\right]+f}{2^{15+\left(\frac{Q_p}{6}\right)}} \qquad (0.8)
$$

If $\left|E_q(u,v,r,Q_p)\right|$ is less than 1, it implies that the quantized coefficient is zero. Therefore, the sufficient condition that the coefficients become zero is therefore the right hand side of equation (0.8) less than 1 i.e. [3]:

$$
\frac{4\cdot SAD\cdot M\left[Q_p\%6,r\right]+f}{2^{15+\left(\frac{Q_p}{6}\right)}} < 1, \quad \forall u,v \qquad (0.9)
$$

$$
SAD < \frac{2^{15+\left(\frac{Q_p}{6}\right)}-f}{4\cdot M\left[Q_p\%6,r\right]}, \quad \forall u,v
$$

We let the right hand side of the equation be the threshold $T\left(Q_p,r\right)$.

In summary, from the equations derived, the quantized coefficients in 4x4 block will be all-zeros if the 4x4 block SAD is less than the threshold value $T\left(Q_p,r\right)$.

3. PROPOSED APPROACH

In order to improve the algorithm, we seek to find a higher threshold value so that more all-zeros block can be detected. As can be seen from equation (0.6),(0.7) and (0.9), the smaller the upper bound of $\left|C[u,i]\cdot C[v,j]\right|$, the bigger the threshold in (0.9).

In [3], three set of quantized coefficients are analyzed to derive three corresponding threshold values. The quantized coefficients are grouped by $r = 0,1$ or 2 used in H.264 periodic table, the corresponding position of the quantized coefficients are:

$$
r = \begin{cases}
0, & \text{for } [u,v]\in\left\{[1,1],[1,3],[3,1],[3,3]\right\} \\
1, & \text{for } [u,v]\in\left\{\begin{matrix}[0,1],[0,3],[1,0],[1,2],\\ [2,1],[2,3],[3,0],[3,2]\end{matrix}\right\} \\
2, & \text{for } [u,v]\in\left\{[0,0],[0,2],[2,0],[2,2]\right\}
\end{cases}
$$

It has been shown [3] that the maximum value for $\left|C[u,i]\cdot C[v,j]\right|$ derived for $r = 0,1$ and 2 are $C(r) = 4,2$ and 1 respectively. This also corresponds to the threshold $T(Q_p,r) = \frac{2^{15+\left(\frac{Q_p}{6}\right)}-f}{C(r)\cdot M\left[Q_p\%6,r\right]}$. So, $SAD < T(Q_p,r)$ would imply that the quantized coefficients belonging to r will be all-zeros. Also, it has been pointed out [3] that $T(Q_p,0) < T(Q_p,1) < T(Q_p,2)$, therefore, as long as $SAD < T(Q_p,0)$, all the coefficients will be zeros and this is the method similar to [2].

In this paper, we derived a separate threshold value for the positions at $r = 0$, for $[u,v]\in\left\{[1,1],[1,3],[3,1],[3,3]\right\}$ and $r = 1$, for $[u,v]\in\left\{\begin{matrix}[0,1],[0,3],[1,0],[1,2],\\ [2,1],[2,3],[3,0],[3,2]\end{matrix}\right\}$

3.1 Conditions for quantized coefficients to be all zeros for r=0

For position $[u,v]=[1,1]$, expanding equation (0.2), we have

$$
\begin{aligned}
E[1,1] &= \sum_{i=0}^{3}\sum_{j=0}^{3}\left\{C[1,i]\cdot C[1,j]\cdot X[i,j]\right\} \\
&= (2)\cdot(2)\cdot X[0,0]+(2)\cdot(1)\cdot X[0,1]+(2)\cdot(-1)\cdot X[0,2]+(2)\cdot(-2)\cdot X[0,3] \\
&\quad +(1)\cdot(2)\cdot X[1,0]+\ldots
\end{aligned}
$$

Rearranging and listing the terms with coefficients $(2)\cdot(2),(2)\cdot(-2),(2)\cdot(-2)$ and $(-2)\cdot(2)$ first, we have

$$
\begin{aligned}
E[1,1] &= \sum_{i=0}^{3}\sum_{j=0}^{3}\left\{C[1,i]\cdot C[1,j]\cdot X[i,j]\right\} \\
&= (2)\cdot(2)\cdot X[0,0]+(2)\cdot(-2)\cdot X[0,3]+(-2)\cdot(2)\cdot X[3,0]+(-2)\cdot(-2)\cdot X[3,3]+\ldots \\
&= (2)\cdot(2)\cdot\{X[0,0]+X[3,3]-X[0,3]-X[3,0]\}+\ldots \\
&= 2\cdot\underbrace{\{X[0,0]+X[3,3]-X[0,3]-X[3,0]\}}_{\lambda_{1,1}}+\underbrace{(2)\cdot\{X[0,0]+X[3,3]-X[0,3]-X[3,0]\}+\ldots}_{s_{1,1}}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad \left|E[1,1]\right| &= \left|2\cdot\lambda_{1,1}+s_{1,1}\right| \\
&\le 2\cdot\left|\lambda_{1,1}\right|+\left|s_{1,1}\right|
\end{aligned}
$$

Note that the modulus of the coefficients in the terms of $s_{1,1}$ are all ≤ 2. We therefore have the following inequality:

$$
\begin{aligned}
E[1,1] &\le 2\cdot\left|\lambda_{1,1}\right|+\sum_{i=0}^{3}\sum_{j=0}^{3}2\cdot\left|X[i,j]\right| \\
&\le 2\cdot\left|\lambda_{1,1}\right|+2\cdot SAD
\end{aligned}
$$

The upper bound of quantized coefficient at [1,1] will be:

$$\left|E_q(1,1,r,Q_p)\right| \le \frac{\left(2\cdot|\lambda_{1,1}|+2\cdot SAD\right)\cdot M\left[Q_p\%6,r\right]+f}{2^{15+\left(\frac{Q_p}{6}\right)}}$$

And the necessary condition for quantized coefficient at [1,1] to be zero will be:

$$\frac{\left(2\cdot|\lambda_{1,1}|+2\cdot SAD\right)\cdot M\left[Q_p\%6,r\right]+f}{2^{15+\left(\frac{Q_p}{6}\right)}}<1$$

$$\Rightarrow SAD < \frac{2^{15+\left(\frac{Q_p}{6}\right)}-f}{2\cdot M\left[Q_p\%6,r\right]}-|\lambda_{1,1}|$$

$$\Rightarrow SAD < 2\cdot T(Q_p,0)-|\lambda_{1,1}|$$

With the same analysis, for coefficient at [1,3],

$$\left|E[1,3]\right| \le 2\cdot|\underbrace{\{X[0,2]+X[3,1]-X[0,1]-X[3,2]\}}_{\lambda_{1,3}}|+2\cdot SAD$$

The necessary condition for coefficient at [1,3] to be zero is thus $SAD < 2\cdot T(Q_p,0)-|\lambda_{1,3}|$.

For coefficient at [3,1],

$$\left|E[3,1]\right| \le 2\cdot|\underbrace{\{X[2,0]+X[1,3]-X[1,0]-X[2,3]\}}_{\lambda_{3,1}}|+2\cdot SAD$$

The necessary condition for coefficient at [3,1] to be zero is thus $SAD < 2\cdot T(Q_p,0)-|\lambda_{3,1}|$.

For coefficient at [3,3],

$$\left|E[3,3]\right| \le 2\cdot|\underbrace{\{X[1,1]+X[2,2]-X[1,2]-X[2,1]\}}_{\lambda_{3,3}}|+2\cdot SAD$$

The necessary condition for coefficient at [3,3] to be zero is thus $SAD < 2\cdot T(Q_p,0)-|\lambda_{3,3}|$.

3.2 Conditions for quantized coefficients to be all zeros for r=1

To futher improved capturing the all-zeros blocks, we futher analyzed the coefficients when r=1. The corresponding positions are

$$[u,v]\in\left\{\begin{matrix}[0,1],[0,3],[1,0],[1,2],\\ [2,1],[2,3],[3,0],[3,2]\end{matrix}\right\}$$

For coefficients at [1,0] and [1,2], expanding equation (0.2) we have:

$$\begin{aligned}
E &= \sum_{i=0}^{3}\sum_{j=0}^{3}\{C[1,i]\cdot C[v,j]\cdot X[i,j]\}\\
&= \sum_{j=0}^{3}\{C[1,0]\cdot C[v,j]\cdot X[0,j]\}+\sum_{j=0}^{3}\{C[1,1]\cdot C[v,j]\cdot X[1,j]\}\\
&+\sum_{j=0}^{3}\{C[1,2]\cdot C[v,j]\cdot X[2,j]\}+\sum_{j=0}^{3}\{C[1,3]\cdot C[v,j]\cdot X[3,j]\}\\
&= \sum_{j=0}^{3}\{(2)\cdot C[v,j]\cdot X[0,j]\}+\sum_{j=0}^{3}\{(1)\cdot C[v,j]\cdot X[1,j]\}\\
&+\sum_{j=0}^{3}\{(-1)\cdot C[v,j]\cdot X[2,j]\}+\sum_{j=0}^{3}\{(-2)\cdot C[v,j]\cdot X[3,j]\}
\end{aligned}$$

Note that $\forall j$, $\left|C[v,j]\right|=1$ for v=0 or 2 in this case. Therefore,

$$E \le 2\cdot SAD-\left(\sum_{j=0}^{3}\left|X[1,j]\right|+\sum_{j=0}^{3}\left|X[2,j]\right|\right)$$

Let $H(m,n)=\sum_{j=0}^{3}\left|X[m,j]\right|+\sum_{j=0}^{3}\left|X[n,j]\right|$ and

$$V(m,n)=\sum_{i=0}^{3}\left|X[i,m]\right|+\sum_{i=0}^{3}\left|X[i,n]\right|$$

$$\Rightarrow E \le 2\cdot SAD-H(1,2)$$

From equation (0.8), the necessary condition for the coefficients at [1,0] and [1,2] to be zero is thus,

$$\frac{\left(2\cdot SAD-H(1,2)\right)\cdot M\left[Q_p\%6,r\right]+f}{2^{15+\left(\frac{Q_p}{6}\right)}}<1$$

$$\Rightarrow SAD < \frac{2^{15+\left(\frac{Q_p}{6}\right)}-f}{2\cdot M\left[Q_p\%6,r\right]}+\frac{H(1,2)}{2}$$

$$\Rightarrow SAD < T(Q_p,1)+\frac{H(1,2)}{2}$$

Similarly, the necessary condition for the coefficients at [3,0] and [3,2] to be zero is $SAD < T(Q_p,1)+\frac{H(0,3)}{2}$

The necessary condition for the coefficients at [0,1] and [2,1] to be zero is $SAD < T(Q_p,1)+\frac{V(1,2)}{2}$

The necessary condition for the coefficients at [0,3] and [2,3] to be zero is $SAD < T(Q_p,1)+\frac{V(0,3)}{2}$

From these analysis, we derived an adaptive threshold method that make use of the higher threshold value at around $T(Q_p,2)$, to detect all-zeros coefficients as follows:

IF SAD < $T(Q_p,2)$
  IF SAD < $T(Q_p,0)$
    BREAK => SKIP block and proceed to process next block
  ELSEIF $SAD \ge 2\cdot T(Q_p,0)-\lambda_{1,1}$ || $SAD \ge 2\cdot T(Q_p,0)-\lambda_{1,3}$
    Perform forward transform and quantization
  ELSEIF $SAD \ge 2\cdot T(Q_p,0)-\lambda_{3,1}$ || $SAD \ge 2\cdot T(Q_p,0)-\lambda_{3,3}$
    Perform forward transform and quantization
  ELSEIF SAD < $T(Q_p,1)$
    BREAK => SKIP block and proceed to process next blk
  ELSEIF $SAD \ge T(Q_p,1)+\frac{H(0,3)}{2}$ || $SAD \ge T(Q_p,1)+\frac{H(1,2)}{2}$
    Perform forward transform and quantization
  ELSEIF $SAD \ge T(Q_p,1)+\frac{V(0,3)}{2}$ || $SAD \ge T(Q_p,1)+\frac{V(1,2)}{2}$
    Perform forward transform and quantization
  ELSE
    BREAK => SKIP block and proceed to process next block
ELSE
  Perform forward transform and quantization

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed algorithm, simulation was performed. For ease of comparison, H.264 JM6.1 encoder is used as in Yong's method [3]. Several QCIF (176x144) sequences and one CIF (352x288) were tested in this simulation. Qp values 28, 32, 36, 40 are used. Table I shows the comparisons of the number of all-zero blocks (AZBs) detected by our proposed algorithm and Yong's algorithm. INC denotes for percentage improvement. The CS denotes the computational saving, which is the percentage saving in total calculations required for the proposed algorithm compared to Yong's method. It shows that despite of slight overhead, the proposed method achieves constantly computational saving compared to Yong's method. The computational saving can be as high as 9.67%. It demonstrates that the proposed algorithm eliminates all-zero blocks more effectively.

5. CONLUSIONS

In this paper, we proposed a very effective early detection algorithm of fast skipped blocks. It is based on theoretical analyses and derivation of H.264 integer transform and quantization. From the simulation results, it is shown that the proposed algorithm can detect 9.71%-43.35%more zero blocks than Yong's method. Despite of slight overhead, computation complexity is still lower. In addition, the proposed algorithm does not cause any degradation of the quality.

6. REFERENCES

[1] Iain E. G. Richardson, "H.264 and MPEG-4 Video Compression", John Wiley and Sons, September 2003.

[2] L. A. Sousa, "General method of eliminating redundant computations in video coding," Electronics Letter, Feb 2000

[3] Y. H. Moon, G. Y. Kim, and J. H. Kim, "An Improved Early Detection Algorithm for All-Zero Blocks in H.264 Video Encoding," IEEE Trans. On Circuits and System on Video Technology, Vol. 15, No. 8, pp. 1053-1057, Aug 2005

[4] K. P. Lim, G. Sullivan and T. Wiegand, "Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods", ISO/IEC JTC1/SC29/WG11, JVT-R095, Jan 2006.

[5] P. Yin, H. Y. Cheong, A. Tourapis and J. Boyce, "Fast mode decision and motion estimation for JVT/H264", IEEE ICIP 2003, pp. 853-856 (III).

[6] C. Grecos and M. Yang, "Fast inter mode prediction for P slices in the H.264 video coding standard,"
IEEE Trans. on Broadcasting, vol. 51, no. 2, Jun. 2005.

[7] A. Yu, K. Ngan, and G. Martin, "Efficient intra- and inter-mode selection algorithms for H.264/AVC,"special issue of Journal of Visual Communication and Image Representation, vol. 17, no. 2, pp. 322-343., Apr. 2006.

[8] 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. on Circuits and Systems for Video Technology, vol. 15, no. 7, pp. 953-958, Jul. 2005.

[9] K. P. Lim et. al., "Fast inter mode selection," ISO/IEC JTC1/SC29/WG11, JVT- I020, Sep. 2003.

[10] J. Lee and B Jeon, "Fast mode decision for H.264," in Proc. of IEEE International Conference on Multimedia and Expo (ICME) 2004, vol. 2, pp. 27-30, Jun. 2004.

[11] I. Choi, J. Lee and B Jeon, "Efficient coding mode decision in MPEG-4 part -10AVC/H.264 main profile," in Proc. of International Conference on Image Processing (ICIP) 2004, vol. 2, 1141-1145, Oct. 2004.

[12] X. Jing and L. Chau, "Fast approach for H.264 inter mode decision," IEE Electronics Letters, vol. 40,
no. 17, Aug. 2004.

[13] C. S. Kannangara, I. E.G. Richardson, M. Bystrom, J. R. Solera, Yafan Zhao, A. MacLennan and R. Cooney, "Low-complexity skip prediction for H.264 through Lagrangian cost estimation," IEEE Trans. on Circuits and Systems for Video Technology, vol. 16, no. 2, pp. 202-208, Feb. 2006.

		Our method	Yong's method	INC	CS
Sequences	QP	AZB	AZB	(%)	(%)
Silent (QCIF)	28	24865	17346	43.35	2.84
	32	39909	28573	39.67	4.43
	36	63323	48368	30.92	6.59
	40	90373	72401	24.82	9.67
Foreman (QCIF)	28	43808	32719	33.89	4.67
	32	62025	50996	21.63	5.05
	36	78622	67917	15.76	5.32
	40	96299	83561	15.24	7.82
Container (QCIF)	28	60137	45198	33.05	6.21
	32	78042	65159	19.77	5.72
	36	92832	83431	11.27	4.85
	40	103070	93945	9.71	5.47
news (QCIF)	28	51683	43679	18.32	3.27
	32	64362	55284	16.42	4.17
	36	77776	67991	14.39	5.15
	40	92973	83250	11.68	5.50
Mobile (CIF)	28	130028	100982	28.76	2.79
	32	168810	142929	18.11	2.48
	36	210995	180054	17.18	3.15
	40	271927	230478	17.98	4.81

Table 1. AZB detection rate and computational saving (CS) of the proposed method compared to Yong's method.