A Parallel Context Model for Level Information in CABAC

Min Gao, Xiaopeng Fan, Qiang Wang, Debin Zhao
Department of Computer Science and Technology
Harbin Institute of Technology
Harbin, China
mgao.hit.vilab@gmail.com,{fxp, dbzhao}@hit.edu.cn

Wen Gao
School of Electronic Engineering and Computer Science
Peking University
Beijing, China
wgao@jdl.ac.cn

***Abstract*—Context-adaptive binary arithmetic coding (CABAC) is one of the most time-consuming modules in H.264/AVC decoder. A potential way to accelerate CABAC is by parallelization. However, the context modeling process for level information in CABAC is highly serial in nature and can not be parallelized in the coefficient level. In order to improve the throughput of CABAC, in this paper we present a parallel context model for level information. The key feature of the model is to use the total number of the significant coefficients and the scanned position of the current significant coefficient in the quantized transform coefficient block as the context information. Since the context information is independent of the previously decoded significant coefficients, parallel decoding in coefficient level is achieved. In experiments, the proposed context model achieves the similar compression efficiency as the CABAC.**

I. Introduction

The processes in context adaptive binary arithmetic coding (CABAC) consist of the following three elementary steps: (1) binarization: the non-binary valued syntax elements need to be mapped into a sequence of binary decisions, so-called bins, by the variable length coding method. (2) Context model: in the context model of CABAC, the selection of the context is determined by the previously decoded significant coefficients, and the probability used for coding the next bin is updated according to the value of the current coded bin. (3) Binary arithmetic coding: the arithmetic coding is based on the interval subdivision principle. The range, value and offset used to determine the interval, on which the coded bin value is identified uniquely, is updated in a serial fashion.

From the above description, the process of CABAC is serial in nature. That will make CABAC engine run at extremely high frequencies in order to decode high bit rate video bit-stream and sometimes may be not feasible in the worst case.

To improve the decoding throughput of CABAC, a lot of parallel entropy coding techniques are proposed. An N-bins/cycle coding (NBAC) was introduced in [2], which decodes N bins in one cycle to achieve N-fold improvement in throughput. PIPE/V2V schemes were proposed in [3] and [4], in which a parallel bin decoding scheme is employed. The NBAC/PIPE/V2V schemes reduce the serial dependency in binary arithmetic coding (BAC) block.

The context model process may be the bottle-neck of CABAC, when increasing the throughput of BAC block. So some parallel context models are proposed. Some parallel context models for coefficient sign, significant coefficients, and significant map are proposed in [5].And a coding order that groups the significant map and the last significant map according to their contexts is proposed in [6]. Besides, a context model having the trait of parallel orientation for variable length coding is proposed in [7].

However, the context modeling process for level information in CABAC is still serial, for the selection of the context for the current bin is determined by the previously coded significant coefficients.

In this paper, we reexamine the CABAC scheme and developed a new context model for level information in CABAC, such that the context modeling of each level in a block is parallel. The proposed context technique, in terms of coding efficiency, is similar to the original CABAC scheme in H.264/AVC.

II. Overview of The Context Model for Significant Coefficient in CABAC

A. Decoding Process of The Syntax Element in CABAC

In CABAC [1], there are three syntax elements. The significant map is a binary array indicating whether or not each coefficient is significant. The last significant map is a binary array indicating whether or not each significant coefficient is the last significant one in the block. The level information is the value of each significant coefficient containing the absolute value and the sign information.

Usually, the decoding process for syntax elements in CABAC is as follows. The significant map in the current scanned position i is decoded to judge whether the coefficient in the scanned position i is a significant coefficient or not. If it is a significant coefficient, the last significant map is decoded to judge whether the significant coefficient is the last one in the transform quantized block. After that, if the coefficient is a significant coefficient, the level information is decoded to get its value.

When decoding the level information, the absolute value of the significant coefficient is a non-binary valued syntax element, which needs to be mapped into a sequence of the binary decisions by the binarization process UEG0 [1].

B. *Context Model for Significant Coefficient in CABAC*

In CABAC, the decoding of each syntax element requires the context information. The context information in CABAC gives the knowledge about the statistic information of the syntax element. Different syntax elements have different context information, and the context information is updated during the decoding process.

The context model for the significant map and last significant map are both the scanned position, such as a coefficient *coeff[i]* scanned at position *i*, the context model of its significant map and last significant map are determined as follows:

$$\chi_{SIG}(coeff[i]) = \chi_{LAST}(coeff[i]) = i \tag{1}$$

From (1), we can see that there is no dependency between any two significant coefficients in the context modeling.

The context models for the level information include the model for the absolute value and the model for the sign information.

For the absolute value, after binarization, bins with bin index between 0 and 13 are coded by two adequate context models.

The context of the first bin with bin index equal to 0 is determined by (2).

$$\chi_{Abscoeff}(i, bin\_index = 0) = \begin{cases} 4 & ,if\ NumLgt1(i)>0 \\ \max(3, NumT1(i)), otherwise \end{cases} \tag{2}$$

And the context of the bins with bin index between 1 and 13 is determined by (3).

$$\chi_{Abscoeff}(i, bin\_index) = 5 + \max(4, NumLgt1(i)) \tag{3}$$

For all bins with bin index greater than 13 as well as sign information, the bypass coding is used, in which the probability is 0.5.

In the two context models just described above, *NumT1(i)* denotes the accumulated number of previously decoded levels with absolute value equal to 1, and *NumLgt1(i)* denotes the accumulated number of previously decoded levels with absolute value greater than 1. The two counters *NumT1(i)* and *NumLgt1(i)* are related to the current scan position *i* in the reverse scan order.

Thus the processes of updating the *NumT1(i)* and *NumLgt1(i)* bring about the context dependency between levels, which make it difficult to decode the levels in parallel.

Based on the observation, we proposed a new context model that breaks the context dependency between levels.

III. THE PROPOSED CONTEXT MODEL FOR LEVEL INFORMATION

It is known that the positions adjacent to the up-left corner in the coefficient block transformed by DCT usually have larger magnitude coefficients dual to the characteristic of energy compact of DCT transform, and in the description above about the context model in CABAC, the scanned position is used as the context of the significant map and last significant map. Thus the scanned position can be used as the context indicator. But it is not enough. In our statistics, we find the probability distribution of the level is different even in the same scanned position when the total number of the significant coefficients is different, as shown in Fig.1.

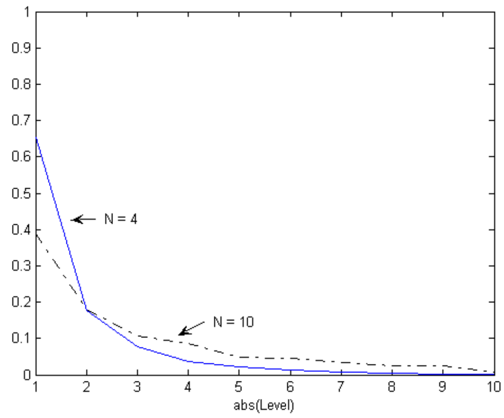

Figure.1 Probability distribution of the absolute value of the significant coefficient at position 11 when the total number of significant coefficients N = 4 and N =10 in Football Qcif video.

Therefore, the combination of the total number of the significant coefficients and the scanned position is used as the context to indicate the probability distribution of the level.

The input of the binary arithmetic coding is the probability of the bins generated by the binarization, so we use the scanned position and the total number of the significant coefficients as the primary context model to get the probability distribution of the level, and then use the bin index as the secondary context to estimate the probability of bins.

A. *Primary Context*

The primary context is defined on the random variable $C_L(Pos_i, N)$, which is defined as follows:

$$C_L(Pos_i, N) = Pos_i + (N-1) \times B \tag{4}$$

Where Pos_i represents the scanned position of the significant coefficients in a DCT coefficient block (here the scanned position is numbered in zig-zag scanning order), and $0 \le Pos_i < B$. N represents the total number of the significant coefficients in a DCT coefficient block and $1 \le N \le B$, and B represents the number of the coefficients including the zero coefficients in a DCT coefficient block.

Equation (4) indicates that the range of $C_L(Pos_i, N)$ is between 0 and B^2-1 for a DCT coefficient block with size equal to B. Take 4×4 block as an example, $C_L(Pos_i, N)$ can have 256 different values. To balance the model accuracy against the model cost and avoid the context dilution problem, we quantize the context by using the distance D_{AB} between the probability distribution of the level under different contexts defined by (5).

$$D_{AB} = \sum_l (p_A(L=l \mid Pos_A, N_A) - p_B(L=l \mid Pos_B, N_B))^2 \qquad (5)$$

Where $p(L=l \mid Pos, N)$ represents the probability of the level L equal to l under the context (Pos, N).

In our implementation, the number of the primary context is reduced to four according to the statistics on the typical video sequences by using the k-means cluster algorithm.

B. Secondary Context

To estimate the probability of the bins, the secondary context is designed under each primary context.

Combined with the binarization process of UEG0, we can calculate the probability of the bin $p_{bin}(j)$ with its bin index j using (6), if we know the probability distribution of level $p(L \mid Pos, N)$ under some context. Then the probability of Least Probable Symbol (LPS) $p_{LPS}(j)$ and the value of Most Probable Symbol (MPS) $val_{MPS}(j)$ can be inferred from $p_{bin}(j)$ by (7) and (8).

$$p_{bin}(j) = \frac{p(L=j \mid Pos, N)}{1-\sum_{k=0}^{j-1} p(L=k \mid Pos, N)} \qquad (6)$$

$$p_{LPS}(j) = \min(p_{LPS}(j), 1-p_{LPS}(j)) \qquad (7)$$

$$val_{MPS}(j) = \begin{cases} 1 & , p_{bin}(j) \le 0.5 \\ 0 & , otherwise \end{cases} \qquad (8)$$

Thus we can use the bin index as the secondary context, and the context model process of the bin generated by the significant coefficient at the scanned position i can be described as follows: firstly, we get the probability distribution of level information according to the scanned position i and the number of the significant coefficients N in the quantized transform coefficient block, then combined with the bin index j of the current bin, we can get the probability of the LPS and the value of the MPS of the current bin.

C. Analysis of The Parallelism

The proposed context model breaks the context dependency between the successive significant coefficients, thus the context modeling process for the levels can be implemented in parallel. This subsection quantitatively analyzes the data throughput by the number of the basic operations used in the original and proposed decoding schemes for the level.

The decoding scheme of the significant map and last significant map is the same as CABAC, in the following we mainly focus on the decoding of the level.

We assume that there are N significant coefficients in a quantized transform coefficient block. And the absolute value of the significant coefficient is denoted by L. We also assume that there is only one binary arithmetic decoder; it means that the binary decoding process for all bins is serial.

TABLE I: BASIC OPERATIONS OF THE DECODING PROCESS

index	Basic operation	Time cost
1	Add/Subtract	1
2	Shift	1
3	Table lookup	1
4	Compare (ie. a>b)	1
5	Logical and (&)	1
6	Decode one bit	3

TABLE II: THE NUMBER OF BASIC OPERATIONS IN THE DECODING PROCESS FOR A LEVEL

	Module	Basic operations					
		1	2	3	4	5	6
original	Context	1		1|2	1|2		
	Decoder		*L*	*L*		*L*	*L*
proposed	Context			1			
	Decoder		*L*	*L*		*L*	*L*

Table I list all the basic operations in the original and proposed decoding process for the level information, and the time cost of each basic operation. The basic operation of decoding one bit consists of two subtractions to update the range of the interval and the state of the arithmetic decoder and a table look-up to update the probability for the next bin.

Table II lists the modules of the original and proposed schemes, and gives the number of the basic operations to decode a level in each module. If $L > 1$, there are 2 context modeling processes in the original decoding scheme.

Thus we can calculate the time cost to decode a block of the original scheme $T_{original}$ and the proposed scheme $T_{proposed}$ according to the Table I and Table II by (10) and (11).

$$Y = \begin{cases} 1, If\ L > 1 \\ 0, otherwise \end{cases} \qquad (9)$$

$$T_{original} = 3\times N + 2\times\sum_{k=1}^{N}(Y_k) + 6\times\sum_{k=1}^{N} L_k \qquad (10)$$

$$T_{proposed} = 1 + 6\times\sum_{k=1}^{N} L_k \qquad (11)$$

Since the data throughput *TP* is inversely proportional to the time cost, the data throughput improvement Δ_{TP} [%] of the proposed one relative to the original is calculated by (12).

$$\Delta_{TP} = \frac{T_{original} - T_{proposed}}{T_{proposed}} \quad (12)$$

IV. EXPERIMENTAL RESULTS

In this section, we present the coding performance and the data throughput improvement of the proposed context model relative to the original CABAC.

The experiment is done at JM17.2 under the main configuration file. The frame structure is IBBPBBP mode and *Transform8x8Mode* is disabled. The value of QP for B pictures is set to QP (I/P) + 2, where the value of QP for I, P pictures is set to 16, 20, 24, and 28. The test sequences include *Football*, *Foreman*, and *News* in QCIF format, and *Bus*, *Football*, and *Tempete* in CIF format, and *City*, *Harbour* and *Ice* in 4CIF format. In these test sequences, *News* in QCIF format, *Tempete* in CIF format and *Ice* in 4CIF format are not in the training set.

A. Coding Performance

TABLE III: CODING PREFORMANCE

Seq	BD-PSNR[dB]	BD-Rate[%]
Football@QCIF	0.012	-0.135
Foreman@QCIF	-0.006	0.112
News@QCIF	0.012	-0.173
Bus@CIF	-0.005	0.092
Football@CIF	0.017	-0.271
Tempete@CIF	-0.022	0.383
City@4CIF	-0.007	0.202
Harbour @4CIF	0.019	-0.380
Ice@4CIF	0.003	-0.091
Average	0.0026	-0.029

From table III, it can be seen that the proposed context model has the same coding efficiency as the original one. The average bit rate (BD-Bit Rate) is decreased 0.029% relative to the original one.

B. Data Throughput Improvement

From the statistics of the test sequences, we get the probability distribution of the total number of the significant coefficients $p(N)$ and the conditional probability of the level under the total number of the significant coefficients $p(L=l\mid N)$, the mean of the $T_{proposed}$, $T_{original}$ can be calculated by (13) and (14) .

$$E(T_{proposed}) = 1 + 6\sum_{N=1}^{16}\sum_{l} Nlp(N)p(L=l\mid N) \quad (13)$$

$$\begin{aligned} E(T_{original}) = 3\sum_{N=1}^{16} Np(N) + 2\sum_{N=1}^{16}\sum_{y=0}^{1} Nyp(N)p(Y=y\mid N) \\ + 6\sum_{N=1}^{16}\sum_{l} lNp(N)p(L=l\mid N) \end{aligned} \quad (14)$$

$$Avg(\Delta_{TP}) = \frac{E(T_{original}) - E(T_{proposed})}{E(T_{proposed})} \quad (15)$$

Thus, the average data throughput improvement to decode a quantized transform block under each QP can be calculated by (15), as shown in table IV.

TABLE IV: DATA THROUGHPUT IMPROVEMENT [%] TO DECODE A BLOCK

Seq \ QP	16	20	24	28
Football@QCIF	23.8	27.9	31.6	34.5
Foreman@QCIF	27.2	30.0	32.7	34.9
News@QCIF	18.4	21.5	24.2	27.4
Bus@CIF	27.5	29.9	31.8	33.8
Football@CIF	26.8	30.0	32.6	34.6
Tempete@CIF	28.5	31.2	32.8	34.7
City@4CIF	35.5	35.5	34.9	35.8
Harbour @4CIF	30.9	33.7	34.6	35.2
Ice@4CIF	31.93	32.4	32.5	34.2
Average	27.847	30.233	31.967	33.900

V. CONCLUSION

The paper presents a new context model for level in CABAC, which can parallelize the context modeling process for level. From the experimental results, the new context model has the similar compression efficiency with CABAC and the average data throughput improvement to decode a quantized transform block is between 27.847% and 33.900% relative to the original one under the assumption that there is only one binary decoder. If there is more than one binary decoder, the more data throughput improvement is achieved.

VI. ACKNOWLEDGEMENTS

This work was partly supported by the National Science Foundation of China, 60736043 and the Major State Basic Research Development Program of China, 973 Program 2009CB320905。

VII. REFERENCES

[1] Detlev Marpe, HeikoSchwarz, and Thomas Wiegand, “Context-Based adaptive binary arithmetic coding in H.264/AVC video compression standard,” IEEE Transaction on circuits and systems for video technology, vol. 13, pp. 620-636, July 2003.

[2] V. Sze and M. Budagavi, “Parallel CABAC,” ITU-T SG 16/Q.6 Doc.COM16-C334. Geneva, Switzerland, April 2008.

[3] M. Winken et. al, “Video coding technology proposal by Fraunhofer HHI,” JCTVC-A116, Joint Collaborative Team on Video Coding meeting. Dresden, Germany, Apirl 2010.

[4] D. He, G. Korodi, G. Martin-Cocher, E. h. Yang, X. Yu and J. Zan, “Video coding technology proposal by Research Motion,” JCTVC-A120, Joint Collaborative Team on Video Coding meeting. Dresden, Germany, Apirl 2010.

[5] Madhukar Budagavi and Mehmet Umut Demircin, “Parallel context processing techniques for high coding efficiency entropy coding in HEVC,” JCTVC-B088, Joint Collaborative Team on Video Coding meeting. Geneva, Switzerland, July 2010.

[6] Dake He, Gergely Korodi and Gaelle Martin-Cocher, “Improving throughput for V2V coding in HEVC,” JCTVC-B036, Joint Collaborative Team on Video Coding meeting. Geneva, Switzerland, July 2010.

[7] Q. Wang, D. Zhao, W, Gao and S. Ma, “High efficiency context-based variable length coding with parallel orientation,” Proc. PCM’05. Jeju Island, Korea, Nov 2005