Telecommun Syst
DOI 10.1007/s11235-011-9656-4

VLC table prediction for CAVLC in H.264/AVC using correlation, statistics, and structural characteristics of mode information

Jin Heo · Yo-Sung Ho

Abstract The H.264/AVC video coding standard adopted context-based adaptive variable length coding (CAVLC) as an entropy coding tool. By combining adaptive variable length coding (VLC) with context modeling, we can achieve a high coding performance. However, CAVLC in H.264/AVC has the problem that VLC table prediction is not always accurate. In this paper, we propose a new VLC table prediction algorithm using the correlation between coding modes of the current and neighboring sub-blocks and the statistics of mode distribution in both intra and inter frames. In addition, we can further increase correctness of VLC table prediction considering the structural characteristics of mode information in inter frames. Experimental results show that the proposed algorithm increases correctness of VLC table prediction by 10.07% and reduces the bit rate by 1.21% on average without significant increment of encoding time.

Keywords H.264/AVC · CAVLC · VLC table prediction · Mode information

1 Introduction

The latest international video coding standard, H.264/AVC, was developed by the Joint Video Team (JVT) from the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group [1]. To achieve a higher compression efficiency, H.264/AVC has adopted several powerful coding techniques, such as variable block-size macroblock modes, multiple reference frames, sub-pixel motion estimation, various predictive direction modes in the intra macroblock mode, deblocking filter, integer discrete cosine transform (DCT), and efficient entropy coding techniques [2–4].

Context-based adaptive variable length coding (CAVLC) has two characteristics to improve coding efficiency: variable length coding (VLC) method and context-based adaptive (CA) method. The VLC method plays an important role in video coding using CAVLC. The principle idea of VLC is to minimize the average codeword length. Shorter codewords are assigned to frequently occurring data while longer codewords are assigned to less frequently occurring data. Another new important concept is the CA method. Previous video coding standards use a single VLC table to encode quantized transform coefficients. However, CAVLC in H.264/AVC uses a predicted VLC table among the four VLC tables by the characteristics of previously encoded syntax elements for quantized transform coefficients [5, 6].

Since the VLC tables are designed to match context conditions, coding efficiency of CAVLC is better than those of previous schemes which use a single VLC table. This means that efficiency of CAVLC depends on VLC table prediction. However, since VLC table prediction is not always accurate in the current CAVLC scheme, we cannot obtain a higher coding performance. Therefore, we need to enhance the accuracy of VLC table prediction.

In this paper, we propose a new VLC table prediction algorithm for CAVLC in H.264/AVC. Considering the relation between the VLC table and coding mode, we adopt two features; the correlation between coding modes of the current and neighboring sub-blocks, and the statistics of mode distribution to select a proper VLC table. Moreover, we further

J. Heo (✉) · Y.-S. Ho
Department of Information and Communications, Gwangju Institute of Science and Technology (GIST), 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Korea
e-mail: jinheo@gist.ac.kr

Y.-S. Ho
e-mail: hoyo@gist.ac.kr

Published online: 23 September 2011

3	7	-1	-2
9	7	2	0
8	-3	-5	-1
2	-2	1	0

Residual data in the sub-block

1	2	6	7
3	5	8	13
4	9	12	14
10	11	15	16

Zigzag scan order for the sub-block

Scanning Position		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Coefficient Level	Absolute Value	3	7	9	8	7	1	2	2	3	2	2	5	0	1	1	0
	Sign	+	+	+	+	+	-	-	+	-	+	-	-		-	+	

Reordered residual data according to scan order

Fig. 1 Zigzag scan order for a 4 × 4 sub-block

increase correctness of VLC table prediction to encode the current 4 × 4 sub-block by using the structural characteristics of mode information.

This paper is organized as follows. After we briefly introduce an overview of the CAVLC framework including five coding steps, we explain how to determine a VLC table for the current 4 × 4 sub-block in Sect. 2. In Sect. 3, we propose a new VLC table prediction algorithm for CAVLC. In Sect. 4, first, we compare the coding performance of the conventional CAVLC with that of the proposed coding technique. Second, we compare the complexity of the H.264/AVC and the proposed algorithm. Finally, we draw the conclusion in Sect. 5.

2 Context-based adaptive variable length coding in H.264/AVC

In this section, we briefly describe the CAVLC scheme in H.264/AVC and explain a VLC table prediction algorithm.

2.1 Overview of CAVLC

The entropy coding method of the baseline profile uses the zero-order Exponential Golomb (Exp-Golomb) code [7] and fixed length code (FLC) for all syntax elements with the exception of the residual data, which are coded using CAVLC. CAVLC is employed to encode residual data, zigzag scanned quantized transform coefficients, for a 4 × 4 sub-block. Figure 1 illustrates the zigzag scan order for the sub-block.

CAVLC was designed to take advantage of several characteristics of residual data in lossy coding: (1) after transform and quantization [8], sub-blocks typically contain many zeros, especially in high frequency regions; (2) the level of non-zero coefficients located in the highest frequency region tends to be as small as one; and (3) the level of non-zero coefficients tends to be larger toward the low frequency regions. Then, taking into consideration the above characteristics, CAVLC employs the syntax elements *coeff_token*, *trailing_ones_sign_flag*, *level_prefix*, *level_suffix*, *total_zeros*, and *run_before* to efficiently encode residual data. The encoding structure of CAVLC for a sub-block is depicted in Fig. 2.

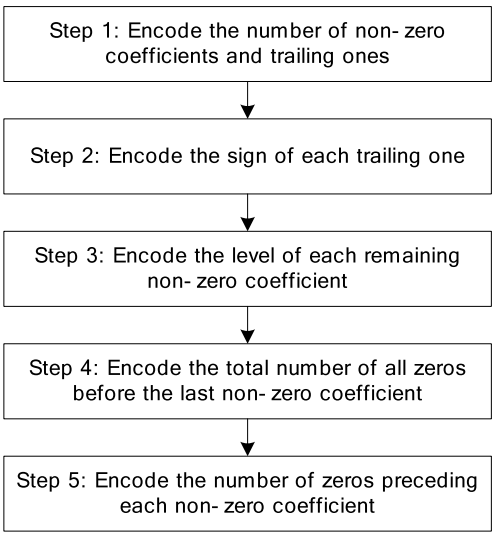


Fig. 2 Flowchart of CAVLC

The detailed coding procedure of CAVLC is as follows:

In Step 1, we encode both the total number of non-zero coefficients and the number of trailing ones as a combination, using a selected VLC table among the four VLC tables (three variable length code tables (*Num-VLC0*, *Num-VLC1*, and *Num-VLC2*) and one fixed length code table (*FLC*)) based on the number of non-zero coefficients in the neigh-

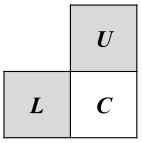


Fig. 3 Upper (U) and left (L) sub-blocks of the current (C) sub-block

boring sub-blocks. A trailing one is one of up to three consecutive non-zero coefficients at the end of a scan of non-zero coefficients having an absolute value equal to 1. If there are more than three trailing ones, only the last three are treated as trailing ones, with any others being coded as normal coefficients.

In Step 2, since trailing ones are all equal to ± 1, they only need the sign specification. Thus, the sign of each trailing one is encoded with a single bit codeword ('+' = 0, '−' = 1) in reverse order.

In Step 3, we encode the level (sign and magnitude) of each remaining non-zero coefficient in the current 4×4 sub-block in reverse order. Each absolute level value is encoded by a selected Lev-VLC table from seven predefined Lev-VLC tables (*Lev-VLC0* to *Lev-VLC6*), with selection of Lev-VLC table dependent on the magnitude of each recently encoded level.

In Step 4, we encode the number of all zeros preceding the highest frequency non-zero coefficient of each 4×4 sub-block.

In Step 5, the number of consecutive zeros preceding each non-zero coefficient is encoded in reverse order.

2.2 VLC table prediction scheme

At the first step, there are four VLC tables used to encode both the total number of non-zero coefficients and the number of trailing ones in each 4×4 sub-block. Selection of the VLC table depends on the number of non-zero coefficients in the current 4×4 sub-block which is predicted using the number of non-zero coefficients in the upper and left sub-blocks. Figure 3 shows the upper and left sub-blocks (U and L) of the current sub-block (C). The size of each sub-block is 4×4.

When both upper and left sub-blocks are available, the number of predicted non-zero coefficients for the current sub-block is calculated by

$$N_C = round(N_U + N_L)/2 \quad (1)$$

where N_C represents the number of predicted non-zero coefficients for the current sub-block. N_U and N_L are the number of non-zero coefficients in the upper and left sub-blocks, respectively. If only the upper sub-block (U) is available, $N_C = N_U$. If only the left sub-block (L) is available, $N_C = N_L$. If neither is available, N_C is set to zero. Using the parameter N_C, we choose an appropriate VLC table from Table 1 for encoding both the total number of non-zero coefficients and the number of trailing ones in one 4×4 sub-block.

Table 1 Choice of VLC table

N_C	VLC table
0, 1	*Num-VLC0*
2, 3	*Num-VLC1*
4, 5, 6, 7	*Num-VLC2*
8 or above	*FLC*

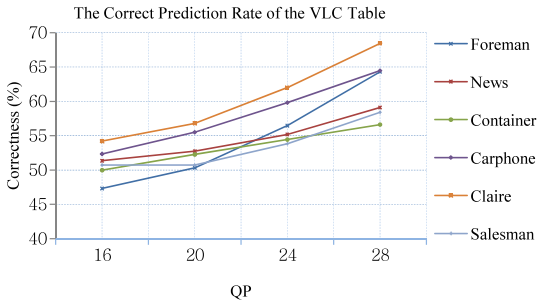


Fig. 4 Correct prediction rate of VLC table in H.264/AVC CAVLC

2.3 Problems of CAVLC

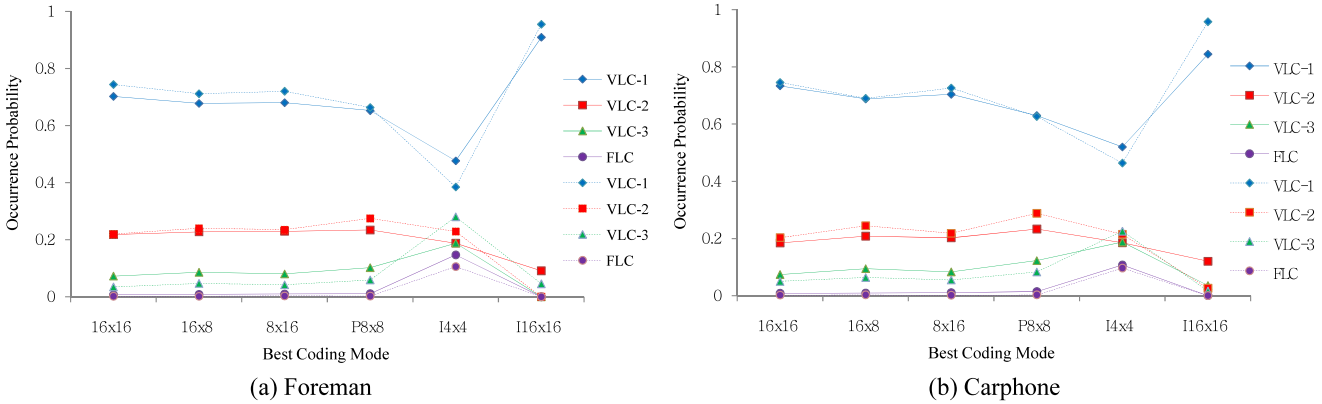
Unlike VLC algorithms in previous video coding standards, CAVLC selects one VLC table from the four possible VLC tables adaptively according to the values of previous syntax elements. Since VLC tables are context dependent, coding efficiency of CAVLC is better than those of other schemes which use a single VLC table in previous video coding standards.

However, H.264/AVC has a drawback that correctness of VLC table prediction is low. Figure 4 shows the correctness of VLC table prediction according to the four quantization parameters (QPs). As shown in Fig. 4, the correct prediction rate of VLC table for six test sequences (Foreman, News, Container, Carphone, Claire, and Salesman) is about 55% on average. Consequently, the optimal VLC table is not used for encoding both the total number of non-zero coefficients and the number of trailing ones in one 4×4 sub-block. Therefore, it reduces coding efficiency.

3 Proposed VLC table prediction algorithm

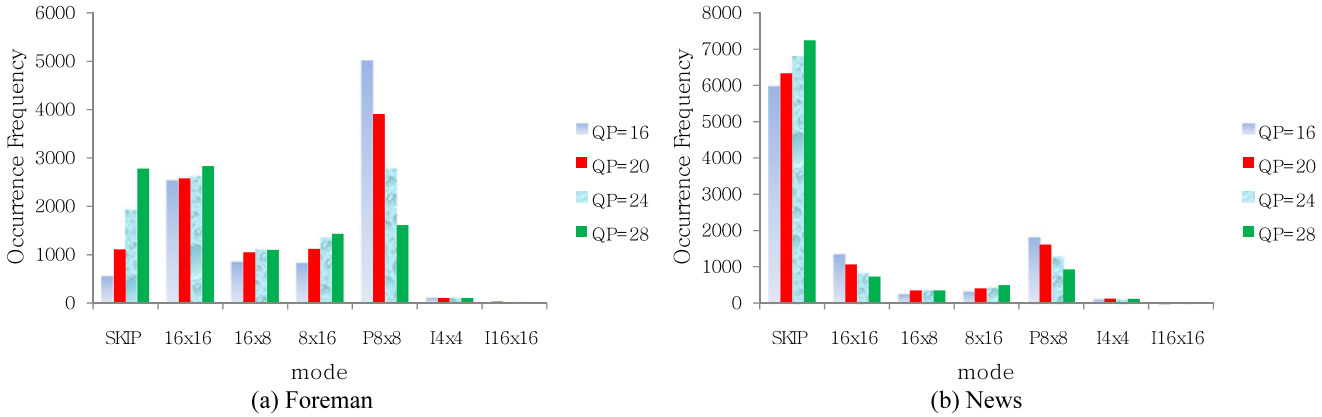
3.1 Observation of VLC table according to the mode

For a given macroblock, H.264/AVC chooses the best coding mode from seven different potential prediction modes: *SKIP*, 16×16, 16×8, 8×16, $P8 \times 8$, *Intra*4×4, and *Intra*16×16. It also uses a rate-distortion optimization



(a) Foreman (b) Carphone

Fig. 5 Comparison of the occurrence probability of each VLC table according to the best coding mode for $QP = 24$



(a) Foreman (b) News

Fig. 6 Distribution of the best coding mode according to QP

(RDO) algorithm [9] to choose the best coding mode for one macroblock.

Figure 5 shows the occurrence probabilities of VLC tables for the six modes: $16 \times 16, 16 \times 8, 8 \times 16$, $P8 \times 8$, *Intra*4×4, and *Intra*16×16. We perform preliminary experiments with 100 frames of Foreman and Carphone sequences in the QCIF format. The coding structure of preliminary experiments is IPPP$\cdots$P (only first frame is encoded in intra coding and the other frames are encoded in inter coding). In Fig. 5, solid lines and dotted lines indicate the ideal VLC table that has 100% correctness of VLC table prediction and the predicted VLC table that is used in the original encoding process, respectively. From Fig. 5, we can observe that the occurrence probabilities of VLC tables are different according to the best coding mode. Moreover, there is a difference between the occurrence probability of the ideal VLC table and the occurrence probability of the predicted VLC table at each best coding mode.

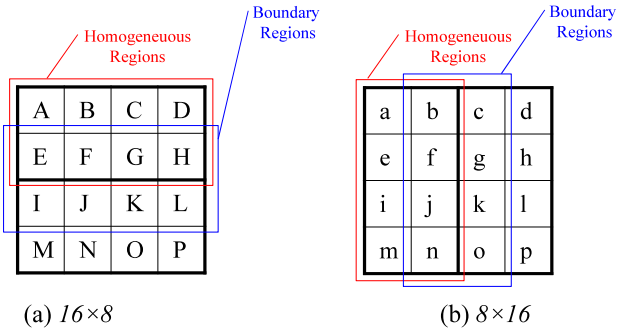
Figure 6 shows the distribution of the best coding mode for the seven modes according to the four QP values. We perform preliminary experiments with 100 frames. The coding structure is IPPP$\cdots$P (only first frame is encoded in intra coding and the other frames are encoded in inter coding). In Foreman sequence, the most popular coding mode is $P8 \times 8$ in the low QP value. However, as the QP value is increased, the occurrence frequencies of both *SKIP* and 16×16 are increased. In News sequence, the most popular coding mode is *SKIP* in all QP values. It is found that, in general, except for *SKIP*, the occurrence frequencies of 16×16 and $P8 \times 8$ are higher than those of other modes. Therefore, if 16×16 or $P8 \times 8$ is occurred in the neighboring sub-blocks, the VLC table for the current sub-block is likely to be predicted to be the VLC table of the neighboring sub-block with 16×16 or $P8 \times 8$. Using the statistics of mode distribution, we can determine a proper VLC table for the current 4×4 sub-block.

3.2 VLC table prediction algorithm for intra frame

The available coding modes for a given macroblock in I-slice include *Intra*4×4 and *Intra*16×16. In I-slice coding, we develop some conditions for proper VLC table prediction. In Table 2, BM_C, BM_U, and BM_L indicate the best coding mode for the current (C), upper (U), and left (L) sub-blocks, respectively (refer to Fig. 3). N_C indicates the

Table 2 VLC table prediction conditions for $BM_C = BM_U = BM_L$ and $BM_C = (BM_U \text{ or } BM_L)$ in intra frame

Condition	Current 4 × 4 sub-block mode	N_C
$BM_C = BM_U = BM_L$	*Intra*4 × 4	round$(N_U + N_L)/2$
	*Intra*16 × 16	
$BM_C = (BM_U \text{ or } BM_L)$	*Intra*4 × 4	$N_{Intra4\times4}$
	*Intra*16 × 16	$N_{Intra16\times16}$



(a) *16×8* (b) *8×16*

Fig. 7 Two inter prediction modes

number of non-zero coefficients in the current 4 × 4 sub-block. $N_{Intra4\times4}$ and $N_{Intra16\times16}$ represent the number of non-zero coefficients in the 4 × 4 sub-block with *Intra*4 × 4 and *Intra*16 × 16, respectively.

In the first condition in Table 2, we can predict N_C using (1) regardless of the current 4 × 4 sub-block mode because BM_C, BM_U, and BM_L are the same. In the second condition, if the current 4 × 4 sub-block mode is *Intra*4 × 4, the probability that N_C is similar to $N_{Intra4\times4}$ is high because the occurrence frequency of *Intra*4 × 4 is higher than that of *Intra*16 × 16 as shown in Fig. 6. Therefore, in this case, we determine $N_{Intra4\times4}$ for N_C. If the current 4 × 4 sub-block mode is *Intra*16 × 16, the probability that the current 4 × 4 sub-block is included in one macroblock with *Intra*16 × 16 is high. This means that the characteristics of all 4 × 4 sub-blocks in one macroblock with *Intra*16 × 16 are similar each other. Therefore, in this case, N_C is set to $N_{Intra16\times16}$.

3.3 VLC table prediction algorithm for inter frame

H.264/AVC supports all seven modes for a given macroblock in P-slice. However, the occurrence frequencies of *Intra*4 × 4 and *Intra*16 × 16 in inter frames are relatively small against the occurrence frequencies of all other five modes, such as *SKIP*, 16 × 16, 16 × 8, 8 × 16, and *P*8 × 8, as shown in Fig. 6. Thus, we do not consider *Intra*4 × 4 and *Intra*16 × 16 in inter frames.

Figure 7 shows two inter prediction modes, 16 × 8 and 8 × 16. Label 'A'–'P' and 'a'–'p' represent a 4 × 4 sub-block within one macroblock in 16 × 8 and 8 × 16, respectively. We define two homogeneous regions and one boundary region in 16 × 8 and 8 × 16, respectively. In 16 × 8, one homogeneous region is composed of eight sub-blocks; 'A', 'B', 'C', 'D', 'E', 'F', 'G', and 'H' and the other homogeneous region is composed of eight sub-blocks; 'I', 'J', 'K', 'L', 'M', 'N', 'O', and 'P'. In 8 × 16, each homogeneous region is composed of eight sub-blocks; 'a', 'b', 'e', 'f', 'i', 'j', 'm', and 'n' and 'c', 'd', 'g', 'h', 'k', 'l', 'o', and 'p', respectively. Each boundary region is composed of eight sub-block; 'E', 'F', 'G', 'H', 'I', 'J', 'K', and 'L' in 16 × 8 and 'b', 'c', 'f', 'g', 'j', 'k', 'n', and 'o' in 8 × 16.

In order to evaluate the influence of vertical and horizontal boundaries on the selection of a correct VLC table in 16 × 8 and 8 × 16, we compare the selection probabilities of the same VLC tables in the homogeneous and the boundary regions. First, in order to calculate the selection probability of the same VLC table in 16 × 8, we compare the VLC tables of (A, E), (B, F), (C, G), and (D, H) and the VLC tables of (I, M), (J, N), (K, O), and (L, P) in the homogeneous regions. Moreover, we compare the VLC tables of (E, I), (F, J), (G, K), and (H, L) for calculating the selection probability of the same VLC table in the boundary region. In 8 × 16, we compare the VLC tables of (a, b), (e, f), (i, j), and (m, n) and the VLC tables of (c, d), (g, h), (k, l), and (o, p) in the homogeneous regions. In order to calculate the selection probability of the same VLC table in the boundary region, we compare the VLC tables of (b, c), (f, g), (j, k), and (n, o).

In Table 3, we can observe that the probability that the same VLC table is selected in the homogeneous regions is higher than the probability that the same VLC table is selected in the boundary region in both 16 × 8 and 8 × 16. Therefore, we can further increase correctness of VLC table prediction considering these structural characteristics of mode information in 16 × 8 and 8 × 16.

Table 4 shows three VLC table prediction conditions for $BM_C = BM_U = BM_L$, $BM_C = BM_U$, and $BM_C = BM_L$ in inter frames. In Table 4, $N_{Homogeneous\_Region}$ represents the number of non-zero coefficients of the 4 × 4 sub-block in the homogeneous region. In these conditions, we classify the current 4 × 4 sub-block into two different types; one includes 16 × 8 or 8 × 16 and the other includes 16 × 16 or *P*8 × 8.

In the first condition in Table 4, if the current 4 × 4 sub-block mode is 16 × 8 or 8 × 16, we check whether the neighboring sub-block (upper sub-block or left sub-block) is included in the boundary region or not. If either neighboring

Table 3 Comparison of probability of the same VLC table selection in 16×8 and 8×16

Sequence	QP	16×8		8×16	
		Homogeneous region (%)	Boundary region (%)	Homogeneous region (%)	Boundary region (%)
Foreman	20	41.10	36.83	47.36	44.10
	28	57.60	55.09	57.52	56.41
News	20	54.58	44.66	45.81	37.79
	28	47.96	45.30	54.30	53.05

Table 4 VLC table prediction conditions for $BM_C = BM_U = BM_L$, $BM_C = BM_U$, and $BM_C = BM_L$ in inter frame

Condition	Current 4×4 sub-block mode	Whether the neighboring sub-block is included in the boundary region?	N_C
$BM_C = BM_U = BM_L$	16×8 or 8×16	Yes	$N_{Homogeneous\_Region}$
		No	$round(N_U + N_L)/2$
	16×16 or $P8 \times 8$	–	$round(N_U + N_L)/2$
$BM_C = BM_U$	16×8	Yes	Refer to Table 5
	16×8 or 8×16	No	N_U
	16×16 or $P8 \times 8$	–	N_U
$BM_C = BM_L$	8×16	Yes	Refer to Table 5
	16×8 or 8×16	No	N_L
	16×16 or $P8 \times 8$	–	N_L

Table 5 VLC table prediction conditions for $BM_C \neq BM_U \neq BM_L$ in inter frame

Current and neighboring sub-blocks mode	Comparison of an occurrence frequency	N_C
SKIP, 16×16, $P8 \times 8$	$C_{16\times16} < C_{P8\times8}$	$N_{P8\times8}$
	$C_{16\times16} > C_{P8\times8}$	$N_{16\times16}$
SKIP, 16×16, (16×8 or 8×16)	–	$N_{16\times16}$
16×16, 16×8, 8×16	–	$N_{16\times16}$
16×16, (16×8 or 8×16), $P8 \times 8$	$C_{16\times16} > C_{P8\times8}$	$N_{16\times16}$
	$C_{16\times16} < C_{P8\times8}$	$N_{P8\times8}$

sub-block is included in the boundary region, we can select $N_{Homogeneous\_Region}$ for N_C. Otherwise N_C is determined using (1). If the current 4×4 sub-block mode is 16×16 or $P8 \times 8$, we can directly determine N_C using (1).

In the second condition, if the current 4×4 sub-block mode is 16×8, we check whether the upper sub-block is included in the boundary region or not. If the upper sub-block is included in the boundary region, three sub-blocks (the current, upper and left sub-blocks) have all different characteristics. In this case, we determine N_C using Table 5. Next, we will explain this condition in more detail. If the current 4×4 sub-block mode is 16×8 or 8×16 and the upper sub-block is not included in the boundary region, we can select N_U for N_C. In fact, since the boundary region is a vertical line in 8×16, we do not need to check whether the upper sub-block is included in the boundary region or not in 8×16. If the current 4×4 sub-block mode is 16×16 or $P8 \times 8$, we can directly select N_U for N_C.

The third condition is quite similar to the second condition. The unique difference between them is that the sub-block that has the same mode with the current sub-block is not the upper sub-block but the left sub-block. Therefore, we can apply the rule for the second condition to the third condition except that we check whether the left sub-block is included in the boundary region or not in the third condition.

In Table 5, $C_{16\times16}$ and $C_{P8\times8}$ represent the cumulative occurrence frequencies of 16×16 and $P8 \times 8$, respectively. $N_{16\times16}$ and $N_{P8\times8}$ represent the numbers of non-zero coefficients in the 4×4 sub-block with 16×16 and $P8 \times 8$,

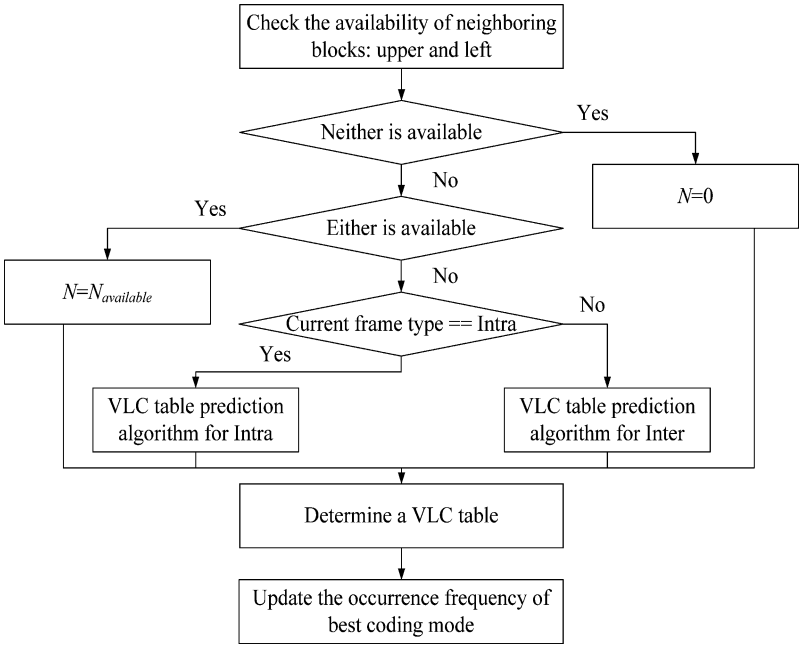


Fig. 8 Flowchart of the proposed algorithm

respectively. When the current sub-block mode is *SKIP*, no motion or residual information is encoded in the current 4×4 sub-block. Since there is not any encoded information in the current 4×4 sub-block, the number of non-zero coefficient is set to zero. Therefore, we do not consider the cumulative occurrence frequency for *SKIP*, in $BM_C \neq BM_U \neq BM_L$ condition.

In the first condition, we compare $C_{16\times16}$ with $C_{P8\times8}$. If $C_{16\times16} < C_{P8\times8}$, we can directly select $N_{P8\times8}$ for N_C. Otherwise N_C is determined as $N_{16\times16}$.

In the second condition, since the characteristics of 16×8 and 8×16 are similar to each other as shown in Fig. 5 and Fig. 6, we can consider these two different modes as the same one mode. Besides, we can observe that $C_{16\times16}$ is generally higher than $C_{16\times8}$ or $C_{8\times16}$ from Fig. 6. Therefore, we can directly select $N_{16\times16}$ for N_C.

In the third and the last conditions, we know that $C_{16\times8}$ or $C_{8\times16}$ is generally lower than $C_{16\times16}$ or $C_{P8\times8}$. Therefore, we can directly determine $N_{16\times16}$ for N_C in the third condition. In the last condition, first we compare $C_{16\times16}$ with $C_{P8\times8}$ and then determine N_C based on the comparison result of $C_{16\times16}$ and $C_{P8\times8}$.

3.4 New VLC table prediction algorithm

In this section, we propose a new VLC table prediction algorithm. The proposed algorithm depends on the correlation of coding mode, the statistics of mode distribution, and the structural characteristics of mode information. Figure 8 shows the flowchart of the proposed VLC table prediction algorithm.

4 Experimental results and analysis

In order to evaluate the performance of the proposed algorithm, we encoded first 100 frames from six test video sequences in the QCIF format. JM 11.0 [10] was used to conduct experiments. We used the baseline profile. In motion estimation, one reference frame is enabled with the maximum search range ±16. The coding structure is IPPP$\cdots$P (only first frame is encoded in intra coding and the other frames are encoded in inter coding). We tested for various QPs (16, 20, 24, and 28).

For the performance comparison between H.264/AVC CAVLC and our proposed algorithm, we used delta VLC table prediction (Δ VLCTP), bit saving (BS), and delta encoding time (Δ T) as shown in (2), (3), and (4).

$$\Delta\,\text{VLCTP} = \text{VLCTP}_{\text{Proposed}} - \text{VLCTP}_{\text{H.264/AVC}}\ (\%). \quad (2)$$

$$\text{BS} = (\text{Bitrate}_{\text{Proposed}} - \text{Bitrate}_{\text{H.264/AVC}}) / \text{Bitrate}_{\text{H.264/AVC}} \times 100\ (\%). \quad (3)$$

$$\Delta\,\text{T} = (\text{Encoding time}_{\text{Proposed}} - \text{Encoding time}_{\text{H.264/AVC}}) / \text{Encoding time}_{\text{H.264/AVC}} \times 100\ (\%). \quad (4)$$

Table 6 shows the performance of the proposed algorithm. The proposed algorithm achieved 7.81$\sim$13.36% correctness of VLC table prediction and 0.64$\sim$1.61% bit saving with a slight increase in complexity. The additional complexity caused by the proposed algorithm occupies only minority portion (about 0.58%) of the total encoding time. From Table 6, we found that the proposed algorithm works more effectively on high $+\Delta$ VLCTP sequences, such as

Table 6 Comparison in the performance measures

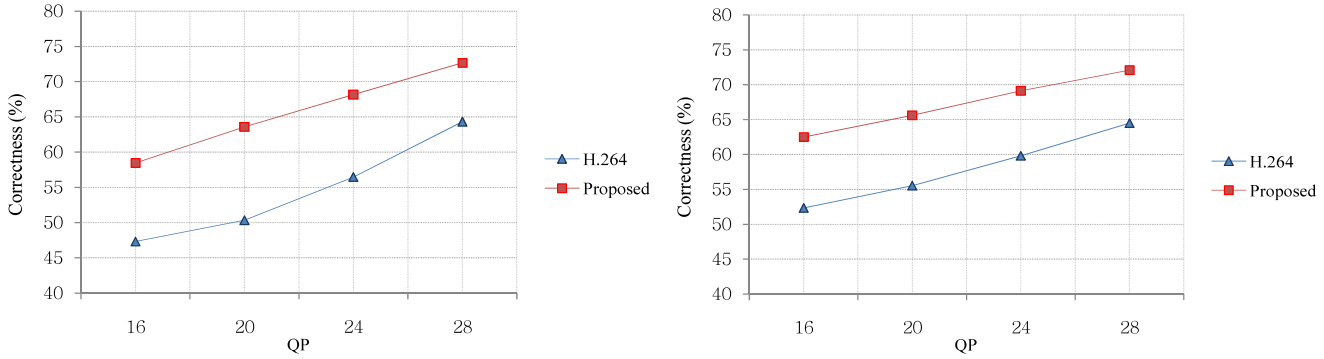
Test sequence	QP	H.264/AVC			Proposed			Δ VLCTP (%)	Δ T (%)	BS (%)
		Correctness of VLC table (%)	Complexity (second)	Bit rate (kbps)	Correctness of VLC table (%)	Complexity (second)	Bit rate (kbps)			
Foreman	16	47.32	114.76	698.42	58.45	115.35	688.38	+11.13	+0.51	−1.44
	20	50.32	105.93	402.89	63.57	106.62	396.40	+13.36	+0.65	−1.61
	24	56.47	98.70	233.38	68.15	99.01	230.35	+11.88	+0.31	−1.30
	28	64.31	93.23	136.95	72.67	93.64	135.80	+8.58	+0.44	−0.84
News	16	51.37	106.86	312.44	60.90	107.59	308.72	+9.53	+0.68	−1.19
	20	52.76	101.21	199.45	62.48	101.99	196.98	+9.72	+0.77	−1.24
	24	55.19	96.27	125.34	65.40	96.92	123.74	+10.21	+0.68	−1.28
	28	59.12	92.02	76.05	67.71	92.55	75.36	+8.59	+0.58	−0.91
Container	16	49.99	108.14	342.80	60.25	108.59	338.04	+10.60	+0.42	−1.39
	20	52.27	100.86	174.57	64.64	101.23	171.78	+12.58	+0.37	−1.59
	24	54.45	94.90	83.74	64.96	95.56	82.56	+10.88	+0.70	−1.41
	28	56.63	90.44	40.10	65.09	90.96	39.68	+8.76	+0.57	−1.06
Carphone	16	51.58	106.76	590.59	60.79	107.84	583.38	+9.21	+1.01	−1.22
	20	54.93	99.92	340.15	65.85	100.73	335.35	+10.92	+0.81	−1.41
	24	59.62	94.59	195.15	69.61	94.95	192.71	+9.99	+0.38	−1.25
	28	63.12	89.97	106.05	71.33	90.58	105.13	+8.21	+0.68	−0.91
Claire	16	52.34	94.02	182.41	62.50	94.77	180.24	+10.16	+0.80	−1.19
	20	55.51	90.02	102.60	65.60	90.57	101.33	+10.45	+0.61	−1.24
	24	59.81	86.78	58.35	69.11	87.39	57.89	+9.64	+0.70	−0.82
	28	64.48	84.00	32.21	72.08	84.66	32.00	+7.81	+0.79	−0.64
Salesman	16	49.41	112.53	287.02	58.44	113.13	283.35	+9.03	+0.53	−1.28
	20	52.79	105.36	164.31	63.15	105.66	162.09	+10.36	+0.28	−1.35
	24	57.83	98.42	96.99	68.49	98.89	95.65	+10.66	+0.48	−1.38
	28	63.51	92.98	56.88	72.85	93.22	56.41	+9.34	+0.26	−0.83
Average	16							+9.94	+0.66	−1.29
	20							+11.23	+0.58	−1.41
	24							+10.54	+0.54	−1.24
	28							+8.55	+0.55	−0.87

Foreman and Container. Table 6 also shows that correctness of VLC table prediction depends on QP. Moreover, we could observe that the proposed algorithm is more effective in high bit rate (when QP value is small). This means that VLC table prediction in low bit rate is generally higher than VLC table prediction in high bit rate in the conventional CAVLC. Therefore, this result shows that there is more room for increasing correctness of VLC table prediction in high bit rate. We already confirmed this fact in Fig. 4.

Figure 9 illustrates the correctness curves for Foreman and Claire sequences. Correctness curves for Foreman and Claire sequences represent the best case and the worst case among experimental results, respectively. From Fig. 9, we can observe that correctness curves of the proposed algorithm are better than those of H.264/AVC CAVLC. This achieves coding gain of CAVLC.

5 Conclusion

In this paper, we proposed a new VLC table prediction algorithm for CAVLC in H.264/AVC. Considering the correlation of coding mode between the current and neighboring sub-blocks and the statistics of mode distribution based on the relation between the VLC table and coding mode, we developed conditions for a proper VLC table prediction in both intra and inter frames. Moreover, we can further increase the correct VLC table prediction rate using the structural characteristics of mode information. Experimental results show



(a) Foreman (b) Claire

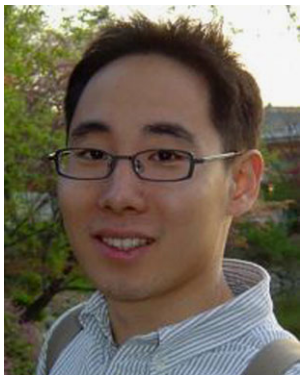
Fig. 9 Correctness curves

that the proposed VLC table prediction algorithm increases the correctness of VLC table prediction by 10.07% and reduces the bit rate by 1.21% on average, compared to CAVLC in H.264/AVC.

Acknowledgements This research was supported by The Ministry of Knowledge Economy (MKE), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) [NIPA-2011-(C1090-1111-0003)].

References

1. Joint Video Team of ITU-T and ISO/IEC JTC 1 (2003). *Draft ITU-T recommendation and final draft international standard of joint video specification* (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), Doc. JVT-G050, March 2003.
2. Wiegand, T., Sullivan, G. J., Bjontegaard, G., & Luthra, A. (2003). Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *13*(7), 560–576.
3. Sullivan, G. J., & Wiegand, T. (2005). Video compression—from concepts to the H.264/AVC standard. *Proceeding of the IEEE*, *1*, 18–31.
4. Luthra, A., Sullivan, G. J., & Wiegand, T. (2003). Introduction to the special issue on the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *13*(7), 557–559.
5. Richardson, I. E. G. (2003). *H.264 and MPEG-4 video compression—video coding for next-generation multimedia*. New York: Wiley.
6. Bjontegaard, G., & Lillevold, K. (2002). *Context-adaptive VLC (CVLC) coding of coefficients*. JVT Document JVT-C028, May 2002.
7. Sayood, K. (2003). *Lossless compression handbook*. San Diego: Academic Press.
8. Malvar, H., Hallapuro, A., Karczewicz, M., & Kerofsky, L. (2003). Low-complexity transform and quantization in H.264/AVC. *IEEE Transactions on Circuits and Systems for Video Technology*, *13*(7), 598–603.
9. Sullivan, G. J., & Wiegand, T. (1998). Rate-distortion optimization for video compression. *IEEE Signal Processing Magazine*, *15*, 74–90.
10. JVT Reference Software JM 11.0, available online at: http://iphome.hhi.de/suehring/tml/download/jm_old/jm11.0.zip.



Jin Heo received the B.S. degree in Electrical Engineering from Kwang-Woon University, Korea, in 2004 and the M.S. degree in Information and Communication Engineering from Gwangju Institute of Science and Technology (GIST), Korea, in 2006. He is currently pursuing the Ph.D. degree from the Department of Information and Communications at GIST, Korea. His research interests include digital image and video coding, multiview video coding, and high efficiency video coding technologies.

Yo-Sung Ho received both B.S. and M.S. degrees in Electronic Engineering from Seoul National University, Korea, in 1981 and 1983, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of California, Santa Barbara, in 1990. He joined the Electronics and Telecommunications Research Institute (ETRI), Korea, in 1983. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, NY, where he was involved in development of the advanced digital high-definition television (AD-HDTV) system. In 1993, he rejoined the Technical Staff of ETRI and was involved in development of the Korea direct broadcast satellite (DBS) digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology (GIST), where he is currently a professor in the Department of Information and Communications. His research interests include digital image and video coding, image analysis and image restoration, advanced coding techniques, digital video and audio broadcasting, 3D television, and realistic broadcasting.