HYBRID VARIABLE LENGTH CODING FOR IMAGE AND VIDEO COMPRESSION

Dihong Tian, Wen H. Chen, and Pi Sheng Chang*

Cisco Systems, Inc.
{dtian, wenc, pschang}@cisco.com

Ghassan AlRegib and Russell M. Mersereau

Georgia Institute of Technology
{gregib, rmm}@ece.gatech.edu

ABSTRACT

The conventional run-level variable length coding (RL-VLC), commonly adopted in block-based image and video compression to code quantized transform coefficients, is not efficient in coding consecutive nonzero coefficients. To overcome the deficiency, hybrid variable length coding (HVLC) is proposed in this paper. HVLC takes advantage of the clustered nature of nonzero transform coefficients in the low-frequency (LF) region and the scattered nature of nonzero transform coefficients in the high-frequency (HF) region by employing two types of VLC schemes. A novel two-dimensional position and one-dimensional amplitude coding scheme is proposed to code the LF coefficients while RL-VLC or an equivalent VLC scheme is retained to code the HF coefficients. Experimental results show that HVLC greatly favors the coding of high-resolution, high-complexity scenes, while it preserves low computational complexity.

***Index Terms*—** Image coding, Video coding, Data compression.

1. INTRODUCTION

Image and video coding standards, e.g., JPEG [1], MPEG [2, 3] and H.26x [4, 5], commonly adopt a codec model that uses a block-based transform, quantization, and entropy coding. The spatial redundancy of an input image (or a motion-compensated residual frame in the video coding case) is reduced by applying a DCT or similar transform to the pixel (residual) samples and quantizing the results. The quantized coefficients are then compressed by an entropy coder, which removes statistical redundancy in the data. Variable length coding (VLC), for its efficiency and simplicity, is widely deployed for the entropy coding, particularly when the codec is desired to have low computational complexity.

In block-based transform coding, most of the significant transform coefficients concentrate in the upper-left corner of the block, designated as the "low-frequency" positions. In addition, the nonzero coefficients tend to cluster around the DC coefficient, in both the horizontal and vertical frequency directions. To allow efficient representation, the quantized transform coefficients are reordered to group together nonzero coefficients prior to entropy coding. A typical reordering path is the zigzag scan starting from the DC (top-left) coefficient toward the highest-frequency AC (bottom-right) coefficient of the transform block.

The output of the reordering process is an array that often contains one or more clusters of nonzero coefficients near the origin, followed by strings of zeros. Conventionally, the coefficient array is represented by a series of (run, level) pairs, where "run" indicates the number of zeros preceding a nonzero coefficient and "level" indicates the magnitude of the nonzero coefficient. The entropy encoder assigns one variable length codeword to each of the symbols[1], and optimal VLC tables are designed such that symbols appearing more often are encoded by shorter codewords and vice versa, thus resulting in a compressed bitstream.

The VLC based on the run-level representation, referred to as RL-VLC, is efficient in coding scattered nonzero coefficients. Nevertheless, it is inefficient in coding clustered nonzero coefficients, due to the fact that n separate codes are required to represent n consecutive nonzero coefficients, each of which has a run equal to zero. To overcome this deficiency, hybrid variable length coding (HVLC) is proposed in this paper[2]. HVLC takes advantage of the clustered nature of the quantized nonzero coefficients in the low-frequency (LF) region and the scattered nature of the quantized nonzero coefficients in the high-frequency (HF) region by employing two types of VLC schemes. In the LF region, runs of consecutive zero coefficients and runs of consecutive nonzero coefficients are coded as a pair by a two-dimensional VLC table, and the amplitudes of the nonzero coefficients are coded independently by a one-dimensional VLC table. In the HF region, RL-VLC or an equivalent scheme is retained to code the position and amplitude of each nonzero coefficient as a pair. Both coding schemes are optimized for their favorable coefficient patterns. Experimental results show that HVLC is a promis-

*This work was performed while the first author was with the Center for Signal and Image Processing at Georgia Institute of Technology.

[1] We refer to the VLC scheme adopted in H.263 [4] for its simplicity, where each (run, level) pair is encoded jointly with a binary symbol that indicates if the last nonzero coefficient in the block has been reached. H.264 [5], as today's most advanced video coding standard, employs a more sophisticated VLC scheme which is specifically designed for small transform blocks (4×4 or 2×2). In this paper, we mainly focus on larger transform blocks such as 8×8 blocks.

[2] The intellectual property in this paper is patented.

ing scheme for coding high-resolution, high-complexity scenes, which is the trend of future image and video applications.

The subsequent sections are organized as follows. In Section 2 we provide an overview of the hybrid variable length coding and introduce the concept of a breakpoint, which divides a block of coefficients into LF and HF regions. The proposed coding scheme for LF coefficients is described in detail in Section 3. We present experimental results in Section 4 and conclude the paper in Section 5.

2. HYBRID VARIABLE LENGTH CODING

Figure 1 provides an illustration on the coefficient distribution of an 8×8 transform block along a pre-defined reordering path such as zigzag, where the nonzero coefficients are statistically more clustered in the low-frequency (LF) region while more scattered in the high-frequency (HF) region.

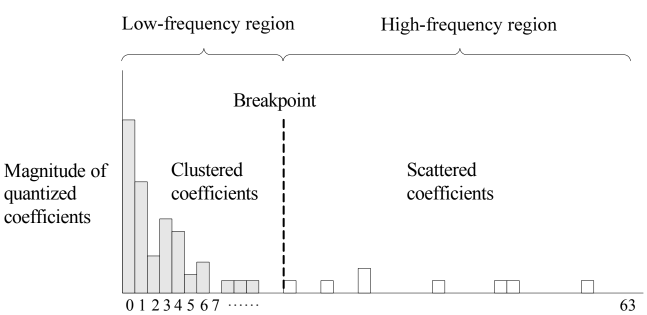

Fig. 1. Coefficient scan of an 8×8 block along a pre-defined path, e.g., zigzag.

In HVLC, a breakpoint along the coefficient scan path is first defined, as shown in Figure 1. The coefficients below and above the breakpoint are considered as low-frequency (LF) and high-frequency (HF) coefficients, respectively. A new coding scheme is developed for the LF coefficients, which will be detailed in the next section. The conventional RL-VLC or an equivalent scheme is used for the HF coefficients. A block diagram of HVLC is shown in Figure 2.

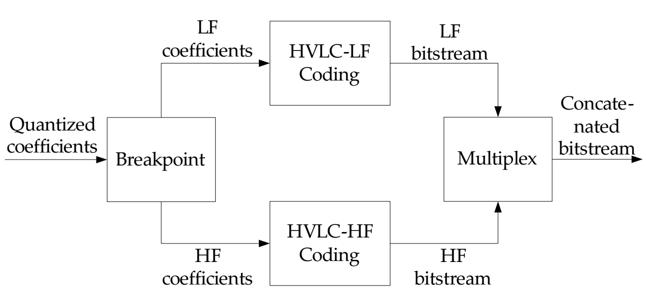

Fig. 2. The block diagram of HVLC.

The breakpoint, as mentioned above, is a coefficient index that divides the coefficient sequence into LF and HF regions along the reordering path. The breakpoint must be known to the decoder to properly decode the coefficients. On the one hand, given the LF and HF coding schemes, there exists an optimal breakpoint within each block of coefficients, which results in the minimum number of bits in coding the block. On the other hand, depending on the coefficient distribution, this optimal breakpoint varies among blocks and therefore needs to be included in the bitstream, which introduces a considerable overhead. For simplicity, in this paper we consider a constant breakpoint for all blocks that have the same quantization parameter. Furthermore, instead of choosing the breakpoint arbitrarily from the coefficient indices, we choose the breakpoint such that the resulting LF (and consequentially HF) region includes the same number of coefficients in both the horizontal and vertical frequency directions. For example, in the zigzag matrix shown in Figure 3, several candidates of the breakpoints are $N = 5, 9, 14, 20, \ldots$. Each of the candidate breakpoints divides the matrix into an upper-left (LF) and a lower-right (HF) region, both of which are symmetric along the diagonal of the matrix. In other words, they include the equal numbers of coefficients in the horizontal and vertical frequency directions. Given the quantization parameter, the breakpoint of a particular block will be selected from the candidate breakpoints. A table that maps the breakpoint with the quantization parameter can be constructed offline and stored at both the encoder and decoder. Therefore no overhead is needed to code the breakpoint into the bitstream .

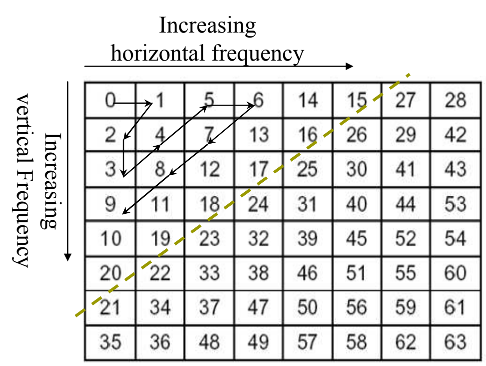

Fig. 3. The zigzag coefficient matrix of an 8×8 transform block and the example of candidate breakpoint.

3. PROPOSED LF CODING

There are essentially two types of information that need to be represented in coding a sequence of quantized transform coefficients: the positions of the nonzero coefficients and their corresponding amplitudes. The run-level coding represents the position and amplitude of each nonzero coefficient individually. Alternatively, for a cluster of nonzero coefficients, their positions can be represented altogether by two quantities: the run of zeros that precede the nonzero cluster and

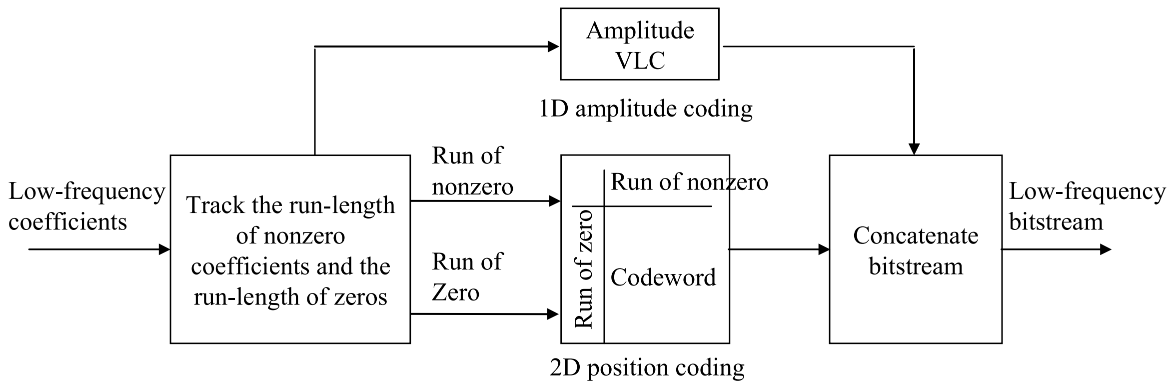

Fig. 4. The block diagram of the proposed 2DP1DA coding for LF coefficients.

the run of nonzero coefficients within the cluster. Their amplitudes can be represented separately. Based on such a representation, a two-dimensional position and one-dimensional amplitude coding scheme is developed, which is described below. For the ease of presentation, this coding scheme is referred to as 2DP1DA hereinafter.

3.1. The 2DP1DA Coding

As its name suggests, the 2DP1DA coding consists of two components: a position coding component that encodes the run of consecutive zero coefficients and the run of consecutive nonzero coefficients as a pair using a two-dimensional VLC table, and an amplitude coding component that encodes the amplitudes of the nonzero coefficients using an independent, one-dimensional VLC table. Figure 4 shows the block diagram of 2DP1DA in coding LF coefficients.

In the 2D position coding, a run of nonzero coefficients in fact implies that the following coefficient is a zero-value coefficient (otherwise it would have been counted into the nonzero cluster and a larger cluster would have be obtained). Therefore, each run of zero-value coefficients can be reduced by 1 before it is coded, with the exception of the first run at the beginning of a block. This results in a shorter codeword to encode the positions of LF coefficients.

Similar to H.263 [4], the indication of the last nonzero coefficient in the block is achieved by coding each pair of zero and nonzero runs jointly with a "last" symbol. Finally, each symbol coded by 2DP1DA has the following form:

$$\begin{cases} (R_z, R_n, last), & \text{for the first symbol in a block,} \\ (R_z - 1, R_n, last), & \text{otherwise,} \end{cases}$$

where R_z and R_n denote the run of zeros and the run of nonzeros, respectively.

3.2. Soft Breakpoint

The definition of 2DP1DA suggests that for a given breakpoint, coefficients following the breakpoint in the scanned sequence may be counted into the run of zeros or the run of nonzeros and therefore be doubly coded by both the LF and HF coding schemes. To avoid using two codewords for one coefficient around the breakpoint, the breakpoint is extended beyond the LF region to the last coefficient coded by the LF coding scheme. This extended breakpoint is termed a *soft* breakpoint and denoted by N_s in the paper.

To illustrate our proposed coding scheme, consider the following coefficient sequence:

Index:	1	2	3	4	5	6	7	8	9	10	11	$\cdots$
Coeff:	2	3	2	0	0	1	-2	1	0	0	-1	$\cdots$

All the remaining coefficients in the sequence are assumed to be zeros, and a constant breakpoint, $N = 6$, is provided. Using 2DP1DA for LF coding and RL-VLC for HF coding, the coefficient sequence will be coded as

$$C_P(0,3,0)\, C_A(2)\, S(0)\, C_A(3)\, S(0)\, C_A(2)\, S(0) C_P(1,3,0)$$
$$C_A(1)\, S(0)\, C_A(2)\, S(-1)\, C_A(1)\, S(0)\, C_{RL}(1,1,1)\, S(1)$$

In the codeword stream above, C_P, C_A denote the position and amplitude codewords, respectively, for 2DP1DA, and C_{RL} denotes the codeword of RL-VLC. S indicates the positive or negative sign of a nonzero amplitude. Note that the soft breakpoint for this example is $N_s = 9$, i.e., the ending position of the second 2DP1DA symbol.

4. EXPERIMENTAL RESULTS

We report preliminary test results of HVLC with an H.263 video codec. H.263 deploys an 8×8 DCT transform, and the quantized transform coefficients are encoded as (run, level, last) triples where "last" indicates the end of the block. We incorporate HVLC into the codec by replacing the sole run-level based entropy coding with HVLC, i.e., 2DP1DA for LF coefficients and RL-VLC for HF coefficients.

The test video sequences in our experiments consist of three resolutions: QCIF (176x144), CIF (352x288), and 4CIF (704x576), with two test sequences for each resolution. Each sequence has 300 frames with a playback rate at 30 frames per second (fps). We enforce one frame to be coded as an INTRA

Table 1. Bit-rate results for QP = 6 and N = 20

Sequence	Resolu-tion	Overall			INTRA			INTER		
		H.263	HVLC		H.263	HVLC		H.263	HVLC	
Carphone	QCIF	268.56	258.94	3.58%	66.30	57.52	13.24%	202.25	201.42	0.41%
Pingpong	QCIF	273.21	256.90	5.97%	73.37	63.85	12.98%	199.84	193.05	4.40%
Foreman	CIF	970.34	956.12	1.47%	224.72	203.19	9.58%	745.61	752.93	-0.98%
Stefan	CIF	2713.26	2460.26	9.32%	405.22	334.37	17.48%	2308.04	2125.89	7.89%
Soccer	4CIF	4187.94	3987.95	4.78%	667.95	629.49	5.76%	3519.99	3358.46	4.59%
Harbour	4CIF	5657.80	5367.56	5.13%	1027.14	871.97	15.11%	4630.66	4495.59	2.92%

Table 2. Bit-rate results for QP = 18 and N = 9

Sequence	Resolu-tion	Overall			INTRA			INTER		
		H.263	HVLC		H.263	HVLC		H.263	HVLC	
Carphone	QCIF	69.54	68.67	1.25%	26.49	25.17	4.98%	43.06	43.50	-1.02%
Pingpong	QCIF	80.22	80.36	-0.17%	29.45	28.61	2.85%	50.77	51.75	-1.93%
Foreman	CIF	271.12	267.59	1.30%	88.22	85.12	3.51%	182.90	182.47	0.24%
Stefan	CIF	761.32	747.52	1.81%	158.89	149.11	6.21%	602.43	598.41	0.67%
Soccer	4CIF	1404.79	1374.59	2.15%	269.36	263.41	2.21%	1135.43	1111.19	2.13%
Harbour	4CIF	1514.11	1441.18	4.82%	451.98	409.08	9.49%	1062.13	1032.10	2.83%

frame in every 15 frames. The VLC tables are constructed based on the measured statistics of symbols and are generated separately for INTRA and INTER coding modes.

Table 1 and Table 2 present the bit-rate results for coding the test sequences at two constant quantization parameters: QP = 6 and QP = 18. The corresponding breakpoints for the two quantization parameters are $N = 20$ and $N = 9$, respectively. For each sequence, we provide both the overall bit-rate results and the respective statistics for INTRA- and INTER-coded blocks. All the bit-rate results are in kbits/sec. A positive percentage in the tables indicates the bit-rate reduction achieved by HVLC, while a negative percentage indicates that the bit rate is increased by HVLC.

From the results, it can be seen that HVLC outperforms H.263 considerably in the lower QP case, and especially for blocks that are coded in the INTRA mode. This is not difficult to understand, as more nonzero coefficients present when a small quantization parameter is used and/or a block is coded as INTRA. The most significant improvement is observed for the STEFAN and HARBOUR sequences. In Table 1, for example, the INTRA bit rates of STEFAN and HARBOUR are reduced by more than 15%. The two sequences present large motion and therefore contain more clustered nonzero coefficients in the low-frequency region.

5. CONCLUSIONS AND FUTURE WORK

The proposal of the hybrid variable length coding (HVLC) is inspired by the observation that in block-based image and video coding, nonzero transform coefficients are statistically more clustered in the low-frequency (LF) region while scattered in the high-frequency (HF) region. Our investigations have shown that HVLC is specifically suitable for coding high-resolution, high-complexity scenes, which is the trend of future image and video applications.

The proposed LF coding scheme, named 2DP1DA, encodes the position and amplitude of LF coefficients separately. Our ongoing research effort is devoted to developing integrated position and amplitude coding schemes. Also, in this paper we used a fixed breakpoint with a constant quantization parameter. In general, the breakpoint should be determined based on the coding parameters of a block or, more intelligently, the coefficient distribution of the block. A context adaptive mechanism to automatically determine variable breakpoints is our another research direction. The performance of the refined HVLC scheme will be evaluated in comparison with the 2D-VLC schemes used in the newest video codec such as H.264 under an 8×8 transform condition [5] and China's AVS standard [6].

6. REFERENCES

[1] ISO/IEC 10918-1, *The JPEG still picture compression standard*, 1994.

[2] ISO/IEC 13818, *Generic coding of moving pictures and associated audio (MPEG-2)*, 1995.

[3] ISO/IEC 14496-2, *Generic coding of audio-visual objects – Part 2: Visual*, 1999.

[4] Draft ITU-T Recommendation H.263, *Video coding for low bitrate communication*, 1996.

[5] ISO/IEC 14496-10, *Draft of Version 4 of H.264/AVC*, 2005.

[6] Q. Wang, D.-B. Zhao, and W. Gao, "Context-based 2D-VLC entropy coder in AVS video coding standard," *J. Comput. Sci. Technol.*, vol. 21, no. 3, pp. 315–322, 2006.