AN ENHANCED ENTROPY CODING SCHEME FOR INTERFRAME WAVELET

Han-Kuang Hsu, Hsiang-Cheh Huang, and Hsueh-Ming Hang

Department of Electronics Engineering,
National Chiao Tung University, Hsinchu, Taiwan, R.O.C.
hmhang@mail.nctu.edu.tw

ABSTRACT[1]

An enhanced entropy coding scheme is incorporated into the interframe wavelet coding architecture in this paper. Interframe wavelet coding has the advantage of SNR, temporal, and spatial scalability, and is a potential candidate for the on-going MPEG-21 scalable video coding (SVC) standard. Motion-Compensated Temporal Filtering (MCTF) and Wavelet Transform Coding are two most essential components in the interframe wavelet coding architecture. The arithmetic entropy subsystem is an indispensable element in Wavelet Transform Coding. It produces the final output bitrate. In this paper, we modify the entropy coding syntax/scheme originally specified in the MPEG SVC Core Experiment (CE) reference software. We observe some bit savings of this technique in our simulations based on the conditions specified by the MPEG core experiments; however, the full potential of this technique is yet to be further explored.

1. INTRODUCTION

Video compression is an essential element in multimedia applications. Conventional video coding systems, including MPEG-1, MPEG-2, H.261 and H.263 international standards, employ the so-called *hybrid coding* structure. In these schemes, the reconstructed previous frame is used to predict the current frame after motion compensation.

The on-going MPEG-21 Scalable Video Coding (SVC) standard employs a new approach different from the hybrid coding structure, Motion-Compensated Temporal Filtering (MCTF) with Wavelet Transform Coding, to achieve SNR, spatial, and temporal scalability. Ohm first proposed a motion-compensated t+2D coding structure [1], as shown in Figure 1 [2]. The major difference between the hybrid coding and the t+2D coding is that the latter does not contain the closed-loop (interframe) DPCM. In addition, the t+2D coding scheme fulfills for the scalable video coding requirements. One of the improved and highly efficient realizations of this concept is the interframe wavelet video coder proposed by Woods and his co-workers [2]. This scheme is called Motion Compensated Temporal Filtering – Embedded Zero Block Coding (MCTF-EZBC or MC-EZBC). Its architecture is shown in Figure 2 [4][6]. Essentially, the same basic structure was adopted by the MPEG committee in March 2004 as the first reference model of SVC.

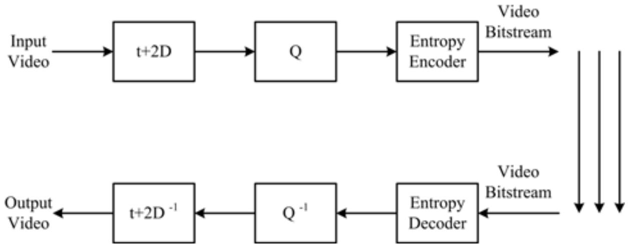

Figure 1. Block diagram of t+2D transform coding system.

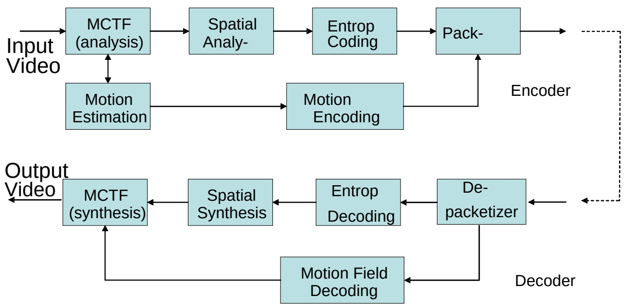

Figure 2. The interframe wavelet video coder.

In this paper, we focus on improving the entropy coding scheme in the aforementioned interframe wavelet coding structure. As illustrated in Figure 1 and Figure 2, no matter how the motion compensation is performed in SVC, entropy coding is a must to further reduce the bits in the output bitstream.

The motivation behind our scheme is the observation of clusters of “1”-bits on the wavelet coefficient bitplanes. Thus, we develop our entropy coding based on the quadtree concept. At the end, we compare the performance between our proposed scheme and that in [7] and [8], and show a somewhat better performance of the proposed scheme.

This paper is organized as follows. In Sec. 2, we outline the motivation behind our proposed scheme. The conventional 3D EBCOT technique is summarized in

[1] This work is supported in parts by National Science Council (Taiwan, ROC) under Grant NSC 92-2219-E-009-003 and OES, Industrial Technology Research Institute (Taiwan, ROC) under Grant 93B-05.

Sec. 3. In Sec. 4, we describe the coding process of new entropy coding scheme by modifying the existing 3D EBCOT. The changes on CE software for integrating the proposed algorithm are described in Sec. 5. Simulation results are shown in Sec. 6, in which we compare the results with the existing scheme. We conclude this paper in Sec. 7.

2. MOTIVATION

In the SVC core experiment one software, the 3D EBCOT entropy coding procedure is applied after MCTF and spatial transform [7][8]. We observe that high energy wavelet coefficients often cluster together [10][9]. In order to save coding bits, we propose a modified coding procedure as described in Section 3. Essentially we construct another layer that records the bitplane locations of the Significant Bits (SB) of all coefficients. We observe bit savings of this technique in our simulation; however, the full potential of this technique is yet to be further explored.

3. 3D EMBEDDED BLOCK CODING WITH OPTIMAL TRUNCATION SCHEME

In the MPEG SVC core experiment reference software [8], the coefficients are coded by the 3D Embedded Block Coding with Optimal Truncation (3D EBCOT) process after the temporal and spatial subband transform. Each subband, generated by temporal and spatial transforms, is divided into 3D codeblocks, which is coded independently. Next, the entropy coding module is applied to these codeblocks. It encodes each bitplane sequentially using context-based arithmetic coding.

Three coding operations are employed to encode the samples in a bitplane: [8] [8]

- **Zero Coding (ZC)**: When a sample is not yet significant in the previous bitplane, this primitive operation is utilized to code the new information of the sample. The definition of "significance" is described in Sec. 4.
- **Sign Coding (SC)**: Once a sample becomes significant in the current bitplane, Sign Coding operation is performed to encode the sign of the sample. Sign Coding also utilizes an adaptive context-based arithmetic coder to compress the sign symbols.
- **Magnitude Refinement (MR)**: Magnitude Refinement is employed to encode the new information of a sample, which has already become significant in the previous bitplane.

For each bitplane, the EBCOT coding procedure consists of three distinct passes, applied in turn. The three passes are:

- **Significant Propagation (SP) pass:** In this pass, samples which are not yet significant but have significant neighbor sample(s) are processed.
- **Magnitude Refinement (MR) pass:** Significant samples are coded in this pass.
- **Normalization pass:** During this pass, those samples which are not yet coded in SP and MR passes (insignificant samples) are coded. So zero coding and sign coding primitives are applied in this pass.

4. PROPOSED ENTROPY CODING SCHEME

We propose the so-called SB-reach method in this Section. In the 3D EBCOT described in Sec. 3 (also in the core experiment software [8]), all wavelet coefficients are initially "insignificant". A coefficient becomes "significant" when its non-zero bit is first found. The first non-zero bit will thus be called Significant Bit (SB) (of a coefficient). For each bitplane, we construct another binary bitplane – so-called SB-reach plane. As shown in Figure 3, a single sample in the SB-reach plane represents a square *mapping block* of *n* by *n* coefficients. The size of the SB-reach plane thus decreases as its representing *mapping block* becomes larger. The binary sample on an SB-reach plane is set to 1, if its square mapping block contains one or more significant coefficients. On the other hand, if the binary sample on the SB-reach plane is 0, it means that all its associated bits in the coefficient bit plane are zero.

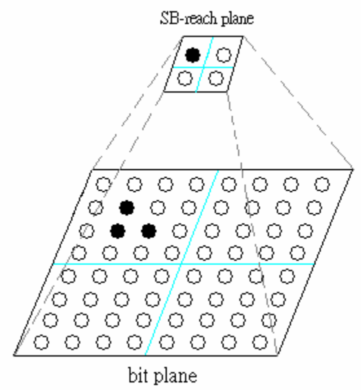

Figure 3. One binary sample on the SB-reach plane is associated with 4x4 mapping block bits.

In this modified coding process, we first construct all the SB-reach bit planes up to the selected "SB-reach depth", as illustrated in Figure 4. Each SB-reach bitplane is associated with one bit plane of the original coefficients. We first encode an SB-reach bit plane before encoding its associated coefficient bit plane using the core experiment software (CES) procedure. In encoding an SB-reach plane, we perform the Significant Propagation pass and the Normalization pass following the scanning order in CES. If a sample is classified significant in a previous SB-reach plane, it must be a "1" bit in the current SB-reach plane and thus is not coded.

After coding one SB-reach plane, we code its associated coefficient bitplane. The coefficients on the bit plane are not coded, if its corresponding SB-reach plane bit is zero (insignificant). If a bit on the SB-reach plane is 1, and then its associated coefficient mapping block bits are coded in the order shown in Figure 5. We perform three coding passes as the original CES does on these coefficient bits. One example is given in Figure 6 to illustrate the procedure in our proposed algorithm. The “coding bits” in this figure are the bits (samples) to be coded by the context-based arithmetic coder.

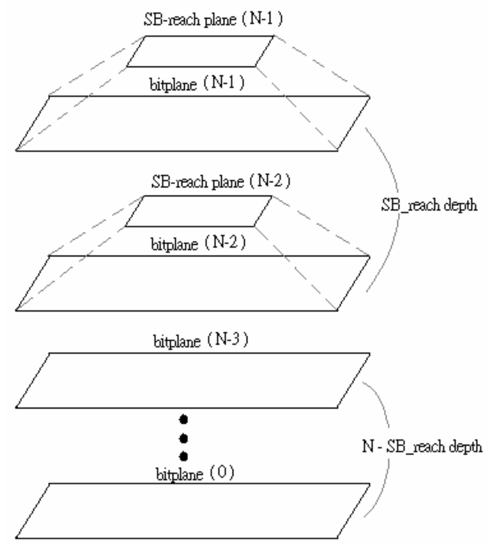

Figure 4. Illustration of “SB-reach depth”.

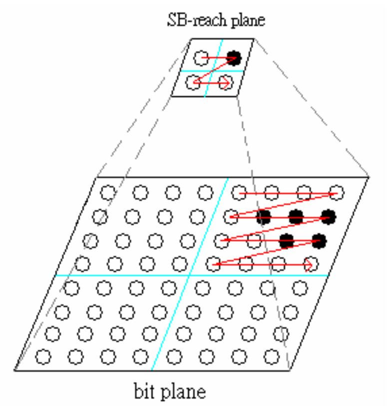

Figure 5. The encoding process of an SB-reach plane and its associated bitplane.

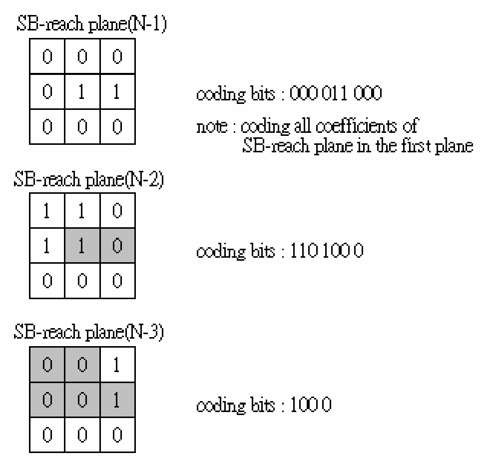

Figure 6. An example of coding steps with our algorithm.

With the method described above, we try all combinations of mapping block size and SB-reach depth, and we then compare the resulting coded bits of all combinations. The best combination of mapping block size and SB-reach depth is retained and coded.

5. CHANGES ON CE SOFTWARE FOR INTEGRATING THE PROPOSED ALGORITHM

On the top of the core experiment software, we changed some syntax and decoding procedure as follows. We add the SB-reach plane architecture to the original 3D EBCOT. The information for the mapping block size, the SB-reach plane depth, and the SB-reach planes is added to the original syntax as shown in Figure 7.

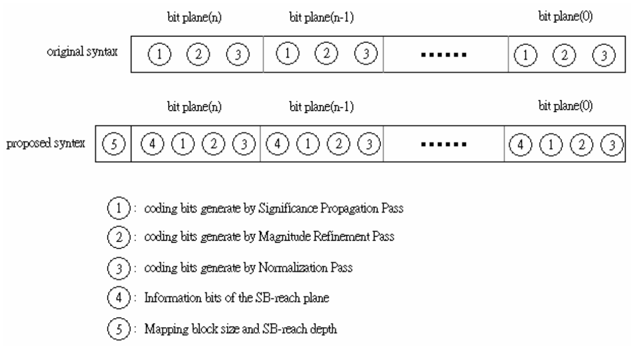

Figure 7. Changes between the original and the proposed syntax.

Here are some definitions in the newly added terms.

- Mapping_blk_size: The mapping block size information is defined in Figure 8. Bit pattern “00” = size 4x4; “01” = 8x8; “10” = 16x16; and “11” = 32x32.
- SB-reach depth: The depth of SB-reach planes. Bit pattern “00” = depth 2; “01” = depth 3; “10” = depth 4; and “11” = depth 5.
- SB_plane: Record SB-reach bits of the corresponding bitplane.

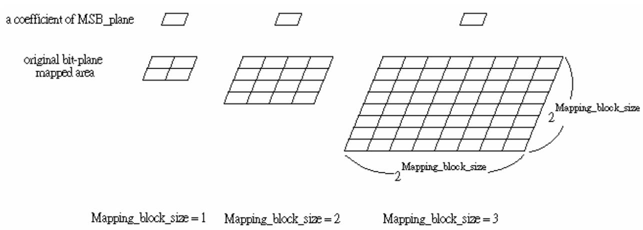

Figure 8. Square mapping block size.

6. SIMULATION RESULTS

We evaluate the performance of our algorithm by measuring the bitrate savings between the proposed algorithm and the core experiment software. We follow the core experiment (CE) specifications to conduct a series of experiments and to test the effectiveness of the proposed algorithm. Eight sequences are tested, namely, CREW, HARBOR, SOCCER, CITY, BUS, FOOTBALL, FOREMAN, and MOBILE, under different spatial, temporal and bitrate test points. Spatial resolutions are QCIF, CIF, and 4CIF, temporal resolutions are 15, 30 and 60 frame/sec, and bitrates vary from 96 kbit/sec to 3 Mbit/sec. The objective image quality, or the PSNR values, are almost the same between our results and the results from the core experiment software. Besides, the subjective qualities are almost identical. Therefore, we compare the resulting bits generated by our algorithm and those by the CE reference software. Some savings in bits with our algorithm are observed.

In Table I to Table III, the bitrate savings are expressed in percentage. In these tables, each entry is the total bitrate saving accumulated from the 1st bitplane to the current one. For example, the cumulative biplane 2 means the total bits saved for the 1st, 2nd and 3rd bitplanes together. The positive numbers denote bitrate savings, while the negative numbers mean bitrate loss. The LL, LH, HL, and HH bands in these tables are the spatial subbands of all spatial resolutions accumulated.

Table I. Bitrate savings (in percentage) for the FOREMAN and BUS sequences of the H frames at temporal levels 1 and 2.

Cumulative bitplane	FOREMAN				BUS			
	LL	LH	HL	HH	LL	LH	HL	HH
2	0.22%	0.17%	0.27%	0.18%	-1.86%	-0.63%	-0.59%	-0.17%
3	0.67%	0.45%	0.51%	0.37%	0.36%	0.51%	0.30%	0.45%
4	0.46%	0.25%	0.28%	0.23%	0.18%	0.23%	0.17%	0.22%

Table II. Bitrate savings (in percentage) of the H frames at temporal levels 3 and 4.

Cumulative bitplane	FOREMAN				BUS			
	LL	LH	HL	HH	LL	LH	HL	HH
2	1.04%	1.04%	1.11%	0.91%	-0.71%	-0.32%	-0.49%	0.10%
3	1.47%	1.26%	1.21%	1.20%	0.53%	0.61%	0.29%	0.68%
4	0.81%	0.66%	0.59%	0.83%	0.27%	0.30%	0.14%	0.37%

Table III. Bitrate savings (in percentage) at the bottommost temporal level.

Cumulative bitplane	FOREMAN				BUS			
	LL	LH	HL	HH	LL	LH	HL	HH
2	-0.05%	0.91%	0.39%	0.97%	0.31%	0.28%	-0.06%	0.05%
3	0.13%	0.13%	0.99%	1.03%	0.24%	0.47%	0.24%	0.61%
4	0.14%	0.80%	0.67%	0.68%	0.15%	0.25%	0.10%	0.30%

As shown in the simulation results regarding to the output bitrates, our algorithm performs somewhat better than the CE software. In general, we gain more at the cumulative biplane 3. Particularly, the HH bands at higher temporal levels perform better. Even better results may be obtained by selecting good context and probability models for arithmetic coding. Also, we should tune carefully the parameter values in our algorithm.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we propose an enhanced entropy coding scheme to further increase the compression efficiency of the interframe wavelet coding algorithm. We modify the entropy coding unit by adding an extra SB-reach layer. Several test conditions specified by the core experiment are tested. So far, our proposed algorithm has somewhat better performance at low- to mid-bitrates comparing to the MPEG Core Experiment (CE) reference software. Further parameter tuning should provide better results, and the full potential of this technique is yet to be further explored.

8. REFERENCES

[1] J.-R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 559–571, Sep. 1994.

[2] S.-T. Hsiang and J. W. Woods, "Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank," *Signal Processing: Image Communications*, vol. 16, pp. 705–724, May 2001.

[3] T. Rusert, et al., *Recent improvements to MC-EZBC*, ISO/IEC/JTC1 SC29/WG11 doc. M9232, Dec. 2002.

[4] S. S. Tsai, *Motion information scalability for interframe wavelet video coding*, MS thesis, National Chiao Tung University, Hsinchu, Taiwan, R.O.C., Jun. 2003.

[5] C. Y. Tsai, H. K. Hsu, H.-C. Huang, H.-M. Hang, and G. Z. Wu, "Enhanced motion estimation for interframe wavelet video coding," *IEEE Int'l Conf. Image Processing*, 2004.

[6] C. Y. Tsai, H. K. Hsu, H.-M. Hang, and T. Chiang, "A scalable video coding scheme based on interframe wavelet technique," ISO/IEC/JTC1 SC29/WG11 MPEG2004/M10569/S08, Mar. 2004.

[7] J. Xu, Z. Xiong, S. Li, and Y. Q. Zhang, "Three-dimensional embedded subband coding with optimized truncation (3-D ESCOT)," *Applied and Computational Harmonic Analysis*, vol. 10, pp.290–315, May 2001.

[8] J. Xu, R. Xiong, B. Feng, G. Sulliman, M. C. Lee, F. Wu, and S. Li, "3D sub-band video coding using barbell lifting," ISO/IEC/JTC1 SC29/WG11 MPEG2004/ M10569/S05, Mar. 2004.

[9] W. H. Peng, T. Chiang and H.-M. Hang, "Context-based binary arithmetic coding for fine granuality scalability," *Proc. Seventh Int'l Symp. Signal Processing and Its Applications*, vol. 3, pp.105–108, 2003

[10] D. Taubman, "EBCOT (Embedded Block Coding with Optimized Truncation) A complete reference", ISO/IEC JTC1/SC29/WG1 N988, Sep.1998.