A High-Efficient Significant Coefficient Scanning Algorithm for 3-D Embedded Wavelet Video Coding

Haohao Song*, Songyu Yu, Li Song, Hongkai Xiong
Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China 200030

ABSTRACT

3-D embedded wavelet video coding (3-D EWVC) algorithms become a vital scheme for state-of-the-art scalable video coding. A major objective in a progressive transmission scheme is to select the most important information which yields the largest distortion reduction to be transmitted first, so traditional 3-D EWVC algorithms scan coefficients according to bit-plane order. To significant bit information of the same bit-plane, however, these algorithms neglect the different effect of coefficients in different subbands to distortion. In this paper, we analyze different effect of significant information bits of the same bit-plane in different subbands to distortion and propose a high-efficient significant coefficient scanning algorithm. Experimental results of 3-D SPIHT and 3-D SPECK show that high-efficient significant coefficient scanning algorithm can improve traditional 3-D EWVC algorithms' ability of compression, and make reconstructed videos have higher PSNR and better visual effects in the same bit rate compared to original significant coefficient scanning algorithms respectively.

Keywords: Video Coding, Wavelet, Scalable, Scanning Algorithm

1. INTRODUCTION

Because different types of terminals and nodes in heterogeneous networks have different processing, memory and display capabilities, resolution, temporal, and SNR scalability of the coded video bit stream are needed essentially. It is achievable by realizing progressive transmission and embeddedness of information in decreasing order of its information content[1], which is an important characteristic of set partitioning and significance testing on hierarchical structures of transformed images[2]. This recognition has spawned more algorithms in image processing which include embedded zerotree wavelet (EZW)[1], set partitioning in hierarchical trees (SPIHT)[2], and set-partitioning embedded block coder (SPECK)[3]. Correspondingly, 3-D EWVC algorithms such as 3-D IEZW[4], 3-D SPIHT[5] and 3-D SPECK[6] become a vital scheme for state-of-the-art scalable video coding, which are very effective and computationally simple algorithms for video compression.

Note that, the motivation behind 3-D EWVC is to ensure that pixels with highest information content which yields the largest distortion reduction are coded and transmitted first. 3-D EWVC constructs partially ordering of the transformed video coefficients by magnitude with respect to a sequence of octavely decreasing thresholds, to advance the ordered bit plane transmission. So the bit-stream can be truncated at any point and can produce the corresponding reconstructions within all lower rate encodings.

Traditional 3-D EWVC algorithms scan coefficients based on bit-plane from high to low due to the coefficients with larger magnitude have a larger content of information. But when coding significant bit information of the same bit-plane, traditional scanning algorithm commonly considers their significance in decreasing of distortion is uniform. In traditional 3-D EWVC algorithms, bit-plane is generally scanned two times to finish the coding of significant bit information and outputs refinement information of known significant coefficients after all significance map and sign information are output. So it is very unfavorable in scanning significant bit information of same bit-plane.

In this paper, we analyze different effect of significant information bits of the same bit-plane in different subbands to distortion and propose a high-efficient significant coefficient scanning algorithm. Compared with original scanning

* songhaohao@sjtu.edu.cn; phone 86-21-62932341-1032

Visual Communications and Image Processing 2005, edited by Shipeng Li,
Fernando Pereira, Heung-Yeung Shum, Andrew G. Tescher, Proc. of SPIE Vol. 5960
(SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.633364

algorithm, the proposed scanning algorithm can improve coding performance apparently without the increasing of computational complexity.

The rest of the paper is organized as follows. Section 2 will present embedded information output order and rate-distortion analysis. The high-efficient significant coefficient scanning algorithm is addressed in Section 3. Section 4 contains experimental results. Section 5 summarizes the work and concludes the paper.

2. EMBEDDED INFORMATION OUTPUT ORDER AND RATE-DISTORTION ANALYSIS

3-D EWVC algorithms can output embedded bit-stream. After video sequences are transformed by 3-D wavelet transform (3-D WT), 3-D EWVC algorithms encode wavelet coefficients in respective coefficient organization format and set scanning mode. They usually think that the coefficients with larger magnitude should be transmitted first because they have a larger content of information. So in traditional 3-D EWVC algorithms, wavelet coefficients are scanned according to bit-plane order from high to low.

A bit-plane is generally scanned two times to shape three kinds of information: the results of significance checking for unknown pixels (significant map), the sign bits for newly identified significant pixels (sign information), and the refinement bits for significant pixels in the previous threshold (refinement information). In the first scanning (Sorting Pass), the unknown coefficients for significance in some organized set are scanned to output two kinds of information: (1) significance map of all coefficients in the set, (2) sign information of new significant coefficients. After all unknown coefficients against a corresponding binary representation threshold are checked, the second scanning (Refinement Pass) for magnitude refinement information of the known significant coefficients is performed to respectively output the bit information in the current bit-plane.

Compression of the current bit-plane is finished when Refinement Pass is end. The encoding stops when some target stopping condition such as bit-rate is met. Otherwise, binary representation threshold is cut in half and the next bit-plane is encoded in the order as the current bit-plane.

It is apparent that the anterior information should have high energy than the posterior information with the scanning order from low subbands to high subbands during a Sorting Pass output, thereafter the corresponding refinement information is output only after all significance map and sign information in the current bit-plane are output. It hints that refinement information is less significant than significance map and sign information. Considering that the arbitrary truncated point of the bit-stream is possibly either on the verge of the boundary of bit-plane or within a bit-plane, especially in the point after Sorting Pass and before Refinement Pass. In Sorting Pass, when scanned set is checked as significant, significant map is output '1', otherwise, significant map is output '0'. Because sets are scanned from low subbands to high subbands, and magnitudes of high subbands coefficients is lower, a large number of '0's will be output in the latter part of Sorting Pass, especially in higher bit-plane. The large number of '0's is not helpful to the decoding reconstruction. Against different binary representation thresholds for 4 video sequences in 3-D SPIHT algorithm, Table 1 shows the percentage of significance map and sign information of unknown new significant coefficients from Sorting Pass output, and the percentage of refinement information of known significant coefficients whose refinement value are '1's from Refinement Pass output, in which, the percentage of significance map and sign information of unknown new significant coefficients from Sorting Pass output is signed as Percentage A and the percentage of refinement information of previous significant coefficients whose refinement value is '1' from Refinement Pass output is signed as Percentage B. The consequence comparison proves the consistence with our analysis.

It is deemed that the quality of reconstructed video be improved only when the decoder receives significance map and sign information of new significant coefficients or refinement information of known significant coefficients whose refinement value are '1's in current bit-plane. So it is very unfavorable that a large amount of useless significance map information of unknown insignificant coefficients and sets is output before refinement information of known significant coefficients whose refinement value are '1's in current bit-plane in traditional 3-D EWVC algorithms. Therefore, the rate-distortion optimization for an arbitrary truncated bit-stream point becomes an evitable issue for an applied progressive transmission. The problem is formulized to transmit the most valuable information as early as possible, by the constraint of reducing the distortion of the reconstructed video as most as possible. It promises a potential approximate optimality for the existing embeddedness generation derived from set partitioning and sorting.

Table 1. Against different binary representation thresholds for 4 video sequences in 3-D SPIHT algorithm

Bit-plane	Football		Mobile and Calendar		Foreman		Akiyo	
	A (%)	B (%)	A (%)	B (%)	A (%)	B (%)	A (%)	B (%)
Highest	0.30	0	2.96	0	4.76	0	5.09	0
Highest-1	3.63	0	4.15	8.99	4.32	22.76	4.13	2.58
Highest-2	9.43	17.65	6.97	31.46	4.57	32.23	3.81	38.63
Highest-3	15.75	33.17	17.55	37.88	11.27	39.66	5.45	37.98
Highest-4	23.04	37.04	23.17	28.78	20.77	43.70	14.66	38.11
Highest-5	27.29	28.80	27.47	27.69	30.26	34.56	23.97	33.71
Highest-6	32.59	28.94	31.96	30.53	30.26	35.35	30.48	37.46

3. HIGH-EFFICIENT SIGNIFICANT COEFFICIENT SCANNING ALGORITHM

Based on the embedded information output order and rate-distortion analysis of traditional 3-D EWVC algorithms, we propose a high-efficient significant coefficient scanning algorithm. The new scanning algorithm scans unknown coefficients, unknown sets, and known significant coefficients simultaneously according to significance of bit information both of different bit-planes and of same bit-plane. If the scanned coefficient or set with respect to embedded quantization-step in scanning procedure is marked as insignificant, its significance map information will be output, and if its significance map information is '1' and it is a coefficient not a set, its sign information will be output subsequently. If the scanned coefficient is known as significant, its refinement information will be output. The new scanning algorithm is capable of overcoming the disadvantage of traditional 3-D EWVC algorithms, achieving the scanning and information output of all unknown coefficients or sets and known significant coefficients in one time.

Being the same as traditional scanning algorithm of 3-D EWVC, the proposed scanning algorithm scans and encodes wavelet coefficients according to bit-plane order from high to low. But in same bit-plane, it is absolutely different from traditional scanning algorithm. The proposed scanning algorithm combines Sorting Pass with Refinement Pass, scans wavelet coefficients in the order from low subbands to high subbands in only one time to achieve the encoding of the current bit-plane. It can ensure that any encoded wavelet coefficient has higher information content than any uncoded wavelet coefficient when bit-stream is truncated at any point. When bit-stream is organized in the proposed scanning algorithm, a large number of '0's originally in latter part of Sorting Pass that is helpless in reconstructing video can be moved to the bit-stream end of the current bit-plane (originally before bit-stream of Refinement Pass). The new scanning algorithm satisfies demand above. Fig. 1 presents an example that compares the difference of the scanning and coding order of original scanning algorithm and the proposed scanning algorithm based on 2-D SPECK (3-D SPECK is similar with it).

The proposed scanning algorithm doesn't need any context-template and can easily be used in traditional 3-D EWVC by properly adjusting their coding structure to improve their ability of compression without the increasing of computational complexity. In order to explain new scanning algorithm clearly and directly, we present it in outline below.

Algorithm:
1) *Initialization*: output the number of the highest bit-plane, initialize sets used in corresponding algorithm.
2) *Combining Pass*: In the order of low subbands to high subbands
   2.1) if the scanned coefficient is known as significant, output its refinement information.
   2.2) if the scanned coefficient or set with respect to embedded quantization-step in scanning procedure is marked as insignificant, output its significance map information; and if it is significant and it is a coefficient not a set, output its sign information and set it as significant coefficient.
3) *Quantization-step update*: decrement quantization-step by 1 and go to Step 2).

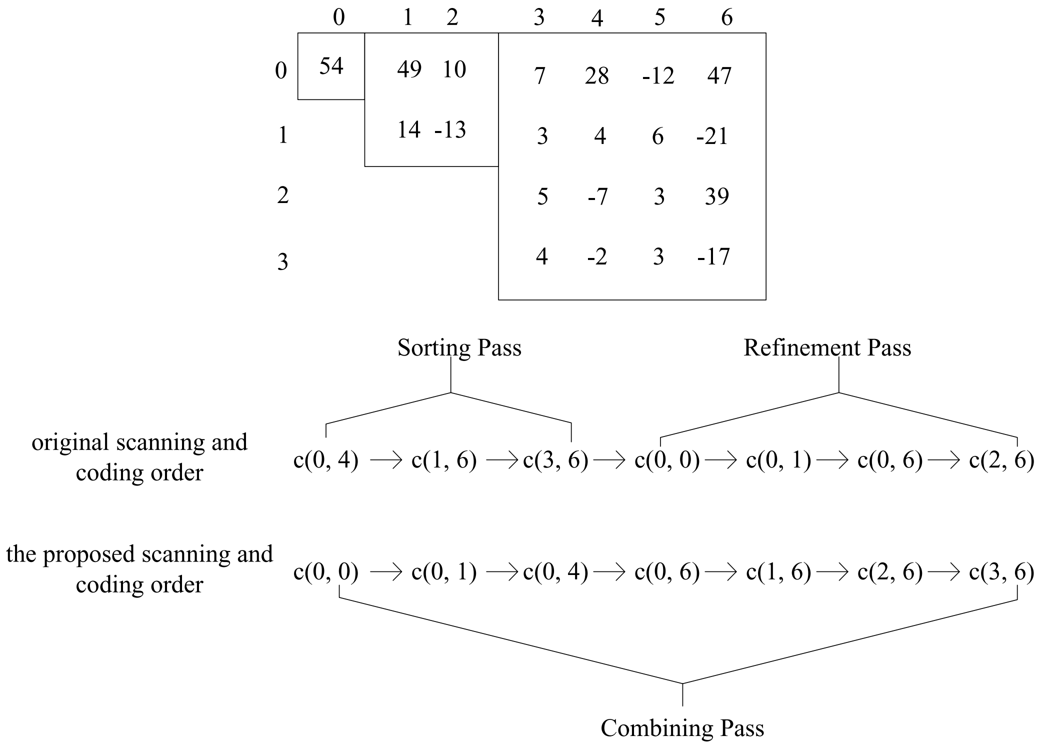

Figure 1. The comparison of the scanning and coding order of original scanning algorithm and the proposed scanning algorithm based on 2-D SPECK (c(i, j) is the coefficient of (i, j), Threshold =16)

4. EXPERIMENTAL RESULTS

In order to validate the high efficiency of the proposed significant coefficient scanning algorithm, extensive experiments have been carried out on several standard test sequences. We used it in two kinds of 3-D EWVC algorithms as 3-D SPIHT and 3-D SPECK and run on 4 gray level (8 bit/pixel) CIF (352×288) sequence 'Football', 'Mobile and Calendar', 'Foreman' and 'Akiyo' sampled at 30 frames per seconds at test bit rates of 760 kbps (0.3 bit/pixel) and 2.53 Mbps (1.0 bits/pixel). We used the 9/7 biorthogonal wavelet filters[7] separately in all dimensions. The same filtering operation is performed in both temporal and spatial domain with reflection extensions both at each frame boundary and at the boundary of each 16 frames as a GOP. The output bit stream isn't further compressed with an arithmetic coder because we only put emphasis on researching the high efficiency of new scanning algorithm. 48 frames (3 GOPs) were coded for all sequences. Quality of reconstruction is measured by peak signal to noise ratio (PSNR) defined as $PSNR = 10\log_{10}(\frac{255^2}{MSE})\,dB$, where MSE denotes the mean squared-error between the original and reconstructed frame.

Table 2 and Table 3 show that average PSNR results with high-efficient significant coefficient scanning algorithm in 3-D SIPHT are 0-1.1 dB better than original 3-D SIPHT, and in 3-D SPECK are 0.3-1.1 dB better than original 3-D SPECK, respectively. Fig. 2 and Fig. 3 present the visual quality comparison of 3-D SPIHT and 3-D SPECK of our proposed scanning algorithm and original 3-D SPIHT and 3-D SPECK, respectively. The better subjective quality can be acquired by using the proposed high-efficient significant coefficient scanning algorithm in traditional 3-D EWVC algorithms apparently.

Table 2. 3-D SPIHT Coding results (average PSNR in dB)

Sequence	Rate (bpp)	3-D SPIHT with the proposed scanning algorithm (dB)	Original 3-D SPIHT (dB)
Football	0.3	26.6	26.5
	1.0	34.7	33.6
Mobile and Calendar	0.3	23.3	22.9
	1.0	29.2	28.9
Foreman	0.3	33.7	33.1
	1.0	39.8	39.8
Akiyo	0.3	44.5	43.8
	1.0	51.5	50.8

Table 3. 3-D SPECK Coding results (average PSNR in dB)

Sequence	Rate (bpp)	3-D SPECK with the proposed scanning algorithm (dB)	Original 3-D SPECK (dB)
Football	0.3	27.0	26.3
	1.0	34.3	33.3
Mobile and Calendar	0.3	23.9	23.6
	1.0	30.1	28.8
Foreman	0.3	34.4	33.6
	1.0	40.9	39.8
Akiyo	0.3	45.6	44.5
	1.0	52.6	52.0

(a) Reconstructed frame in original 3-D SPIHT

(b) Reconstructed frame in 3-D SPIHT with the proposed scanning algorithm

Figure 2. Same reconstructed frame at 1.0 bpp average rate in 3-D SPIHT for *football* sequence.

(a) Reconstructed frame in original 3-D SPECK

(b) Reconstructed frame in 3-D SPECK with the proposed scanning algorithm

Figure 3. Same reconstructed frame at 0.3 bpp average rate in 3-D SPECK for *foreman* sequence

5. CONCLUSION

In this paper, we propose a high-efficient significant coefficient scanning algorithm. It can scan unknown coefficients, unknown sets, and known significant coefficients in one time and make three kinds of information (significance map, sign information and refinement information) be output simultaneously according to R-D theory instead of original scanning algorithm that scans a bit-plane two times to finish the coding of current bit-plane and outputs refinement information of known significant coefficients at last. Experimental results of 3-D SPIHT and 3-D SPECK show that high-efficient significant coefficient scanning algorithm can improve traditional 3-D EWVC algorithms' ability of compression, and make reconstructed videos have higher PSNR and better visual effects in the same bit rate compared to original significant coefficient scanning algorithms respectively.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of RFDP No.20040248047 and Shanghai NSF No.04ZR14082.

REFERENCES

1. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficient," *IEEE Trans. on Signal Processing*, vol. 41, pp. 3445-3462, Dec. 1993.
2. A. Said and W. A. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 6, pp. 243-250, June 1996.
3. W. A. Pearlman, A. Islam, N. Nagaraj and A. Said, "Efficient, low-complexity image coding with a set-partitioning embedded block coder," *IEEE Trans. Circuits Syst. Video Technol.* vol. 14, no. 11, pp. 1219-1235, Nov. 2004.
4. Y. Chen, and W. A. Pearlman, "Three-Dimensional Subband Coding of Video Using the Zerotree Method", *Visual Communications and Image Processing '96*, Proc. SPIE 2727, pp. 1302-1312, March. 1996.
5. B. J. Kim, and W.A. Pearlman, "An Embedded Wavelet Video Coder Using Three-Dimensional Set Partitioning in Hierarchical Trees", *IEEE Data Compression Conference*, pp. 251-260, March 1997.
6. Z. Lu, and W. A. Pearlman, "Wavelet Video Coding of Video Object by Object-Based SPECK Algorithm", *Picture Coding Symposium (PCS-2001)*, Seoul, Korea, pp. 413-416, 25-27 April, 2001.
7. M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transformation". *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 205-220, 1992.