

# Low complexity bit-plane entropy coding for 3-D DWT based video compression

E. Belyaev, K. Egiazarian and M. Gabbouj

Tampere University of Technology  
Korkeakoulunkatu 10, 33720 Tampere, Finland

## ABSTRACT

This paper is dedicated to entropy coding for scalable video compression based on three-dimensional discrete wavelet transform (3-D DWT). A new simple bit-plane entropy coding of wavelet subband matrices is proposed. Practical results show that 3-D DWT video codec with proposed entropy coding allows to increase the encoding speed 2-3 times for the same quality level in comparison with x.264 codec which is one of the fastest software implementation of H.264/AVC standard.

**Keywords:** Low-complexity video coding, 3-D DWT

## 1. INTRODUCTION

Energy efficient real-time video compression and transmission finds a numerous applications in video processing such as streaming, surveillance, conferencing and broadcasting. Because the throughput of these channels is less than raw video bit rate, the additional video compression on the transmitter side is required. Taking into account high bit error ratios, packet losses and time-varying bandwidth, the *scalable video coding* (SVC) is the most preferable compression method for video transmission.<sup>1</sup>

An extension of the H.264/SVC standard,<sup>2</sup> which is currently the most popular video coding approach, includes temporal, spatial and quality scalability and provides high compression efficiency due to motion compensation and inter-layer prediction exploiting the video source temporal redundancy and redundancy between different layers. But, because of high computational complexity of motion estimation and inter-layer prediction at the encoder side, implementation of H.264/SVC encoder in a mobile device is a difficult task.

As an alternative to highly complex H.264/SVC encoders, scalable video encoders based on three-dimensional discrete wavelet transform (3-D DWT) have been proposed. In<sup>3</sup> scalable Video Coding with 3-D Set Partitioning in Hierarchical Trees (3-D SPIHT) was introduced. But because high computational complexity of this algorithm development of low-complexity scalable video coders based on 3-D DWT is an important practical task.<sup>4,5</sup>

In this paper we present a scalable video codec based on 3-D DWT and low-complexity bit-plane entropy coding. First, in section 2 we describe the main idea of video compression scheme based on 3-D DWT, bit-plane context modeling and arithmetic coding. Then, in section 3, we analyze the statistical properties of binary sources corresponding to each context and introduce the low-complexity bit-plane coding based on this analysis. Finally, in section 4, we compare proposed video compression algorithm with x.264 codec which is one of the fastest software implementation of H.264/AVC standard and show that our algorithm can be more preferable for many applications where computational complexity plays a critical role.

---

Further author information: (Send correspondence to E. Belyaev)

E. Belyaev: E-mail: evgeny.belyaev@tut.fi

K. Egiazarian: E-mail: karen.egiazarian@tut.fi

M. Gabbouj: E-mail: moncef.gabbouj@tut.fi

## 2. VIDEO COMPRESSION BASED ON 3-D DWT AND ARITHMETIC CODING

Figure 1 shows the main scheme of the video coding based on 3-D DWT considered in our paper.

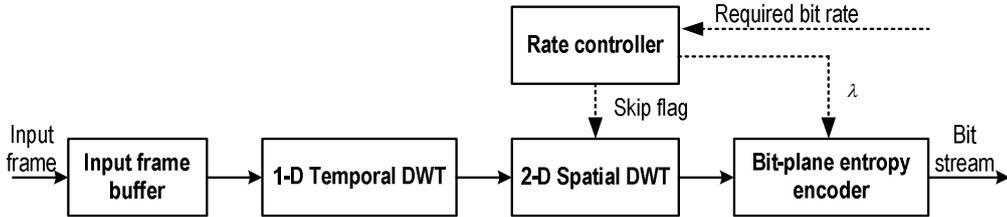


Figure 1. Video codec scheme based on three-dimensional discrete wavelet transform

First, group of frames (GOP) are accumulated in the input frame buffer and one-dimension DWT in temporal direction is applied with length  $N$ , which is a GOP size. Second, two-dimensional multilevel DWT for each frame is used. Then, each wavelet subband is independently compressed by using bit-plane entropy coding which uses context modeling similar to JPEG2000 standard<sup>6</sup> to split bit-planes in to a set of binary streams. Each binary stream is compressed by adaptive binary arithmetic coder from H.264/AVC standard<sup>7</sup> (M-coder) which is faster and more efficient than MQ-coder used in JPEG2000 standard.<sup>8</sup>

For the required bit rate achievement a *rate controller* chooses the Lagrange multiplier value  $\lambda$  for each wavelet subband by using one-pass rate control proposed in.<sup>9</sup> Entropy encoder starts to encode wavelet subband from the highest significant bit-plane  $i_{max}$  and use this  $\lambda$  value to calculate

$$\varphi_i = d_i + \lambda r_i, \quad (1)$$

where  $d_i$  and  $r_i$  are distortion and rate for current subband after compressing of bit-plane with number  $i$ . If  $\varphi_{i-1} > \varphi_i$  then entropy encoder stops and bit-planes  $i - 1, \dots, 0$  are not placed into the output bit stream. If  $\varphi_{i_{max}} > D$  for any subband in frame  $n = 1 \dots N$ , where  $D$  is a subband energy, then whole subband is skipped. In this case, the corresponding subbands in frames  $n + 1, \dots, N$  are skipped too and two-dimensional DWT for these subbands are not calculated.

## 3. PROPOSED BIT-PLANE ENTROPY ENCODING ALGORITHM

Figure 2 shows the binary memoryless entropy estimation for different binary contexts depending on Peak-to-noise ratio of luma component (Y-PSNR) achieved by video compression algorithm described in previous section.

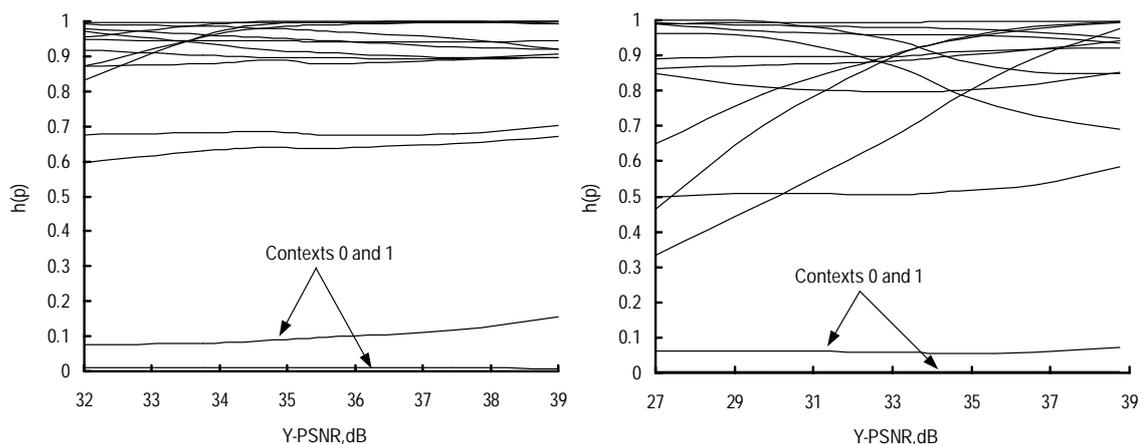


Figure 2. Entropy estimation in different contexts for “vassar.0“ (left) and “pedestrian area“ (right)

These dependences show that in the most of cases all binary contexts can be classified into the following two main groups: with entropy close to 1 and entropy close to 0. From the computational complexity point of view, arithmetic coding for the first group will be not efficient, because in this case it does renormalization part approximately for each input binary symbol<sup>10</sup> with a compression performance close to an uncompressed case. For the second group it is possible to achieve a comparable or better compression efficiency by using zero-run length coding approach.

Taking into account reasoning described above, we propose to remove an arithmetic coding stage and use the following simple bit-plane entropy encoding algorithm (see Fig. 3).

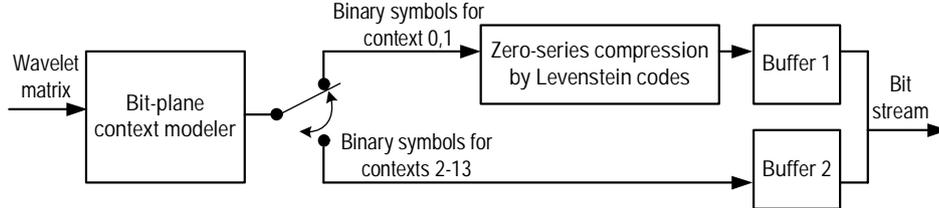


Figure 3. Proposed bit-plane entropy coding scheme

All binary symbols are separated into the following two groups. Binary sequence corresponding to contexts 0 and 1 is compressed by zero-run length coding and placed into the Buffer 1. Binary sequence corresponding to contexts 2–13 is placed into the Buffer 2 without any compression. For zero-run coding we propose to use Levenstein codes with prefix based on equal codes. It helps to avoid using look-up tables during coding process in comparison with the Huffman code that is more efficient for video codec implementation.

For compression of zero-run  $\underbrace{000\dots 0}_L 1$  value  $c$  is calculated:

$$c = \begin{cases} 0, & \text{if } L = 0 \\ \lfloor \log_2 L \rfloor + 1, & \text{if } L > 0. \end{cases} \quad (2)$$

Then the value  $L$  is represented by two parts:

$$\underbrace{\text{binary representation of } c}_{\lfloor \log_2(\log_2 L_{max}) \rfloor + 1 \text{ bits}} \quad \underbrace{\text{binary representation of } L - 2^{c-1}}_{c - 1 \text{ bits}}, \quad (3)$$

where  $L_{max}$  is the maximum possible zero-run value.

#### 4. PRACTICAL RESULTS

For our experiments, the proposed video coding algorithm based on three-dimensional discrete wavelet transform is compared with x.264 codec<sup>11</sup> which is one of the fastest software implementation of the H.264/AVC standard. Proposed codec was run with GOP size  $N = 8$  with Haar wavelet transform in temporal direction and 5/3 spatial wavelet transform at three-levels decomposition. x.264 codec was run in very low complexity mode with intra-frames period 8. In both cases codecs were run in constant bit rate mode and speed results were achieved without any assemblers, threads and other program optimization techniques\*.

Practical results were obtained for the typical for video surveillance test video sequences “vassar\_0“ and “ballroom\_0“<sup>12</sup> with 640×480 resolution, and “pedestrian area“ and “rush hour“<sup>13</sup> with 1920×1080 resolution. Figures 4–7 show the rate-distortion-complexity comparison for considered codec’s. In our work computation complexity is defined as a number of frames which can be encoded in one second by using fixed processor architecture. In the case of vassar\_0 and ballroom\_0 we use Intel Atom CPU N270 1.6GHz, in the case of “pedestrian area“ and “rush hour“ we use Intel Core 2 DUO CPU 3.0GHz.

\*Command line example for x.264:

x264.exe -output vassar\_0.264 vassar\_0.avs -preset ultrafast -keyint 8 -bitrate 1000 -no-asm -threads 1

Our results demonstrate that bit-plane entropy coding algorithm can be significantly simplified by removing of arithmetic coder and using Levenstein codes for zero-run compression for low-entropy contexts and uncompressed approach for high-entropy contexts. In the most cases, 3-D DWT video codec with the proposed entropy coding allows to increase the encoding speed 2-3 times for the same Y-PSNR level in comparison with the one of the fastest software implementation of a H.264/AVC standard. Taking into account that a scalable extension of the H.264/SVC standard is more complex than single-layer H.264/AVC compression, our scalable video compression scheme can be more preferable in many applications such as video surveillance, mobile TV and other where computational complexity plays a critical role.

## ACKNOWLEDGMENTS

This work was supported by the Academy of Finland (project no. 213462, Finnish Program for Centers of Excellence in Research 2006-2011) and by the project of NSFC International Young Scientists (project no. 61150110166).

## REFERENCES

- [1] Gallant, M. and Kossentini, F., "Rate-distortion optimized layered coding with unequal error protection for robust internet video," *IEEE Transactions on Circuits and Systems for Video Technology* **11**, 357–372 (2001).
- [2] "Advanced video coding for generic audiovisual services," *ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC)* (2009).
- [3] Kim, B., Xiong, Z., and Pearlman, W., "Low bit-rate scalable video coding with 3-d set partitioning in hierarchical trees (3-d spiht)," *IEEE Transactions on Circuits and Systems for Video Technology* **10**, 1374–1378 (2000).
- [4] Moinuddin, A., Khan, E., and Ghanbari, M., "Reduced complexity embedded 3-d wavelet video coding," *International Symposium on Telecommunications* (2008).
- [5] Lopez, O., Martinez-Rach, M., Piol, P., Malumbres, M., and Oliver, J., "A fast 3d-dwt video encoder with reduced memory usage suitable for iptv," *IEEE International Conference on Multimedia and Expo* (2008).
- [6] "Jpeg 2000 image coding system: Core coding system, itu-t recommendation t.800 and iso/iec 15444-1," *ITU-T and ISO/IEC JTC 1* (2000).
- [7] Marpe, D., Schwarz, H., and Wiegand, T., "Context-based adaptive binary arithmetic coding in the h.264/avc video compression standard," *IEEE Transactions on Circuits and Systems for Video Technology* **7**, 620–636 (2003).
- [8] Marpe, D. and Wiegand, T., "A highly efficient multiplication-free binary arithmetic coder and its application in video coding," *IEEE International Conference on Image Processing* (2003).
- [9] Belyaev, E., "Low bit rate video coding based on three-dimensional discrete pseudo cosine transform," *International Conference on Ultra Modern Telecommunications* (2010).
- [10] Belyaev, E., Veselov, A., Turlikov, A., and Liu, K., "Complexity analysis of adaptive binary arithmetic coding software implementations," *The 11th International Conference on Next Generation Wired/Wireless Advanced Networking* (2011).
- [11] "x.264 video codec," <http://x264.nl/>.
- [12] "Mvc test sequences," <http://www.merl.com/pub/avetro/mvc-testseq/orig-yuv/>.
- [13] "Xiph.org test media," <http://media.xiph.org/video/derf/>.

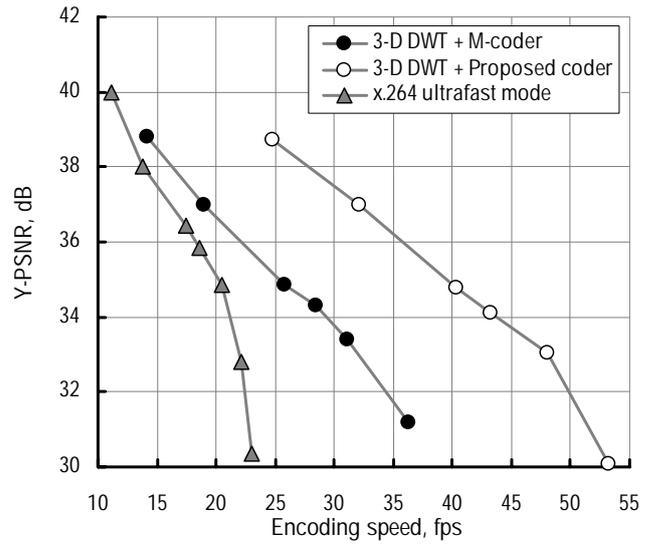
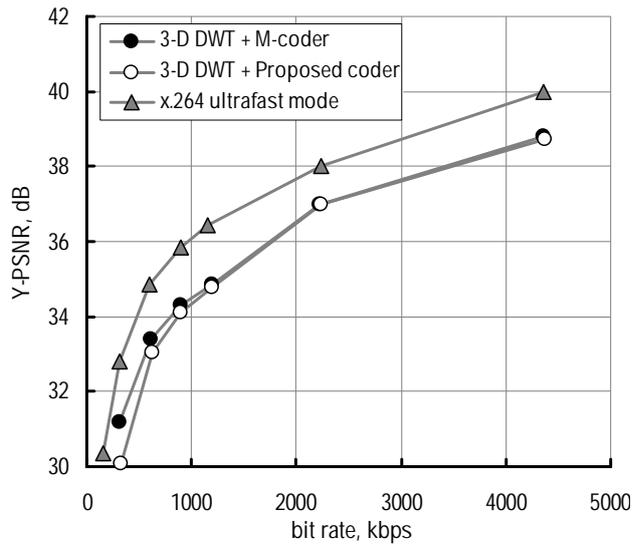


Figure 4. Rate-distortion-complexity comparison for test video sequence "vassar\_0"

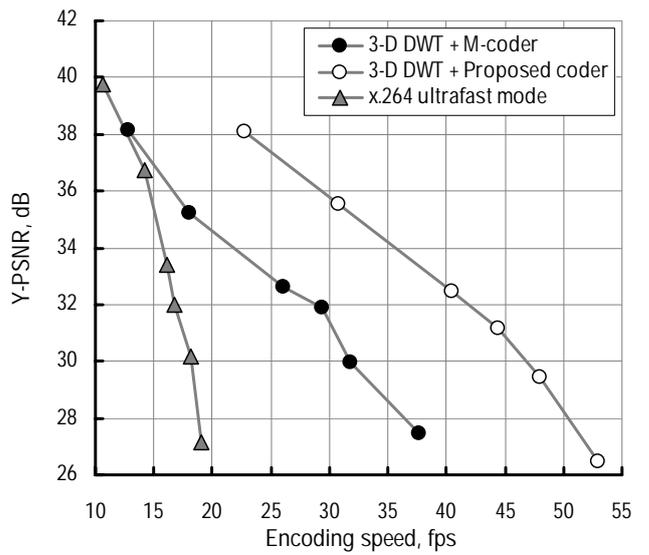
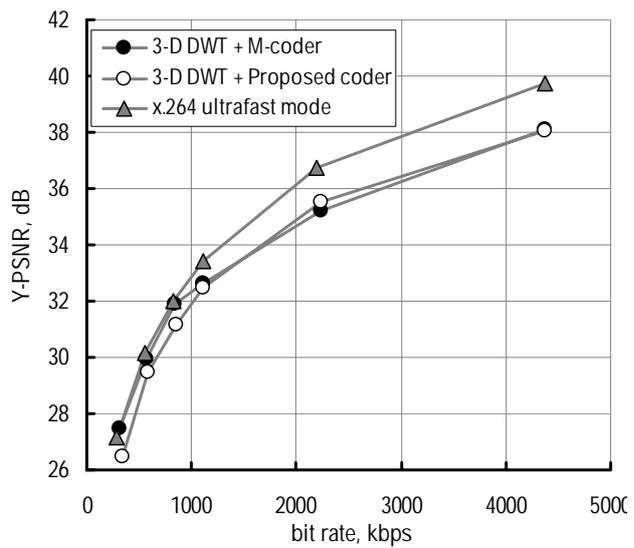


Figure 5. Rate-distortion-complexity comparison for test video sequence "ballroom\_0"

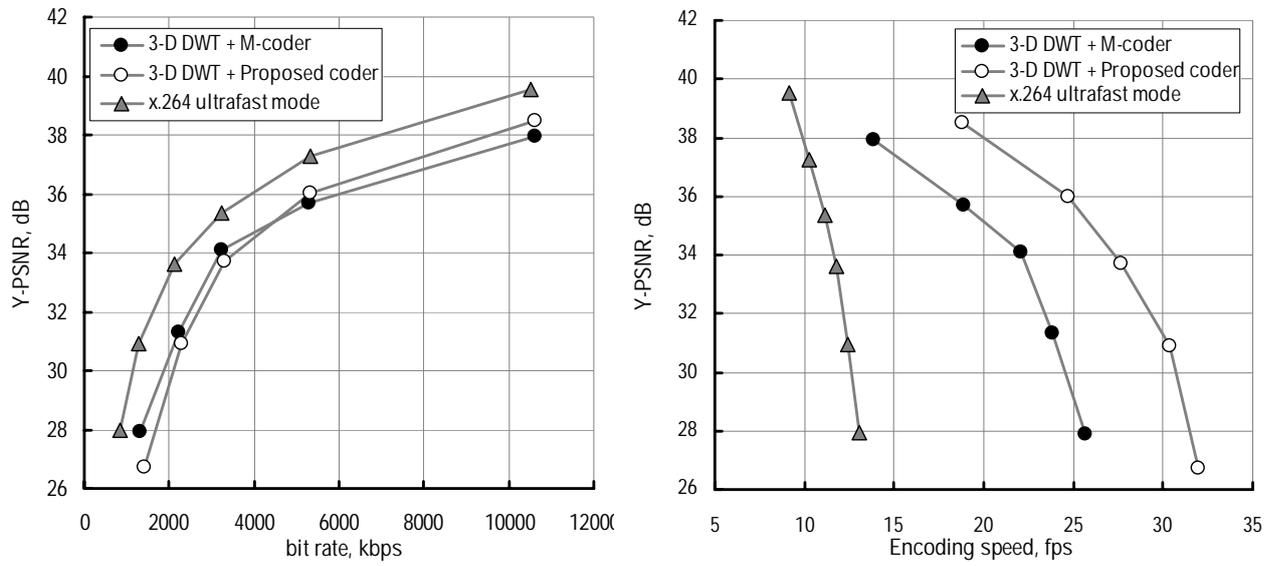


Figure 6. Rate-distortion-complexity comparison for test video sequence “pedestrian area“

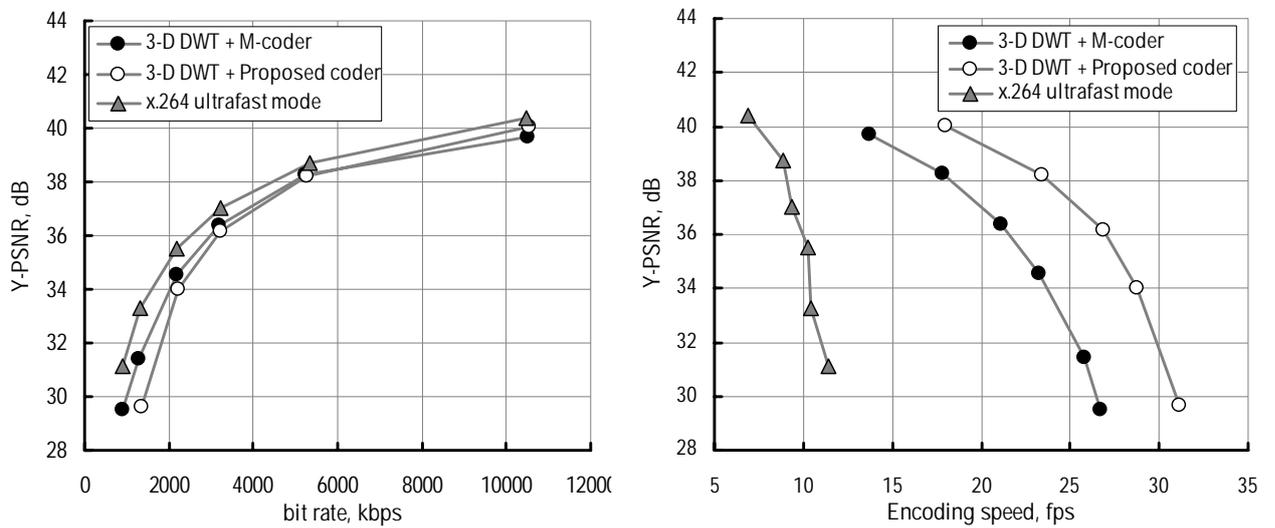


Figure 7. Rate-distortion-complexity comparison for test video sequence “rush hour“