FPGA IMPLEMENTATION OF BIT PLANE ENTROPY ENCODER FOR 3-D DWT BASED VIDEO COMPRESSION

GOPIKA G NAIR, SABI S.

M. Tech. Scholar (Embedded Systems), ECE department, SBCE, Pattoor, Kerala, India, Email: gopikananthan@gmail.com

Assistant Professor, ECE department, SBCE, Pattoor, Kerala, India.

ABSTRACT

In video processing, energy efficient real time video compression and transmission finds numerous applications. In this paper entropy coding is performed for video compression based on three-dimensional discrete wavelet transform. Currently the most popular video coding approach is an extension of the H.264/SVC standard. It provides high compression efficiency due to motion compensation and inter-layer prediction. But, its implementation in a mobile device is a difficult task because of high computational complexity of motion estimation and inter-layer prediction. Therefore, in this paper a scalable video codec based on 3-D DWT and simple bit-plane entropy coding of wavelet subband matrices is proposed. The proposed entropy coder uses zero-run Levenstein coder and Adaptive binary range coder. For zero-run-length compression, Levenstein codes are used. Therefore, computation complexity of this coder is significantly less than bit-plane arithmetic coder in JPEG2000 standard and entropy encoder in 3-D SPIHT algorithm.

Index terms: Entropy coding; 3-D DWT; low complexity video coding.

1. INTRODUCTION

A video can be considered as a sequence of frames or images. Internet data gets doubled almost every year and reducing the storage space and increasing the speed of transmission has become the prior requirement and concern for multimedia communication. Therefore, for reducing the amount of information without any loss in quality, various video compression methods have been proposed and scalable video coding (SVC) is the most apt method for the compression of video streams as far as high bit error rates, packet losses and time varying bandwidth is concerned.

Currently the most popular scalable video coding approach is an extension of the H.264/SVC standard. It provides high compression efficiency and includes temporal, spatial and quality scalability. This standard also provides inter-layer prediction by exploiting the temporal redundancy as well as the redundancy between the layers. However the implementation of the H.264/SVC encoder in mobile devices, personal computers and other systems such as HD video surveillance, wireless home TV, mobile IPTV broadcast, etc is a very difficult task because of its high computational complexity.

The high computational complexity of H.264/SVC is due to the following reasons:

(1) A lot of computations are required for motion compensation and estimation even in the case of fast motion vector searching algorithms. High resolution video sequences requires high motion search radius for efficient coding which contradicts with low complexity coding case which requires as less as possible motion search radius. Therefore, the motion estimation can be not so efficient, especially taking into account that in some cases the compressed video sources does not contain significant motion.

(2) It requires backward looping in the encoding scheme in order to avoid resynchronisation with the decoder side and this loop includes inverse transform and quantization, de-blocking filter and other extra calculations.

So scalable video encoder based on three-dimensional discrete wavelet transform (3-D DWT) can be used as an alternative. Several 3-D DWT approaches were proposed during the last decade such as 3-D set partitioning in hierarchical trees, 3-D embedded subband coding with optimal truncation and so on. Currently three dimensional DWT schemes have a comparable rate-distortion performance with the H.264/SVC encoders due to intensive investigation of Motion compensated Temporal Filtering (MCTF). However, all these schemes require very high computational resources and they also require highly complex entropy encoder. The complexity of the 3-D DWT codec can be decreased by the development of various low complexity entropy encoding methods. As a result, the development of efficient low-complexity scalable video encoders is an important practical problem.

In this paper the proposed 3-D DWT codec uses two entropy encoding cores, Levenstein zero run coder and adaptive binary range coder. Levenstein zero-run coder is used for low entropy binary contexts and adaptive binary range coder (ABRC) is used for the remaining binary contexts, which in turn decreases the computational complexity without any significant rate-distortion performance degradation. An efficient rate-distortion criterion is also being proposed based on parent-child subbands tree. Using this criterion, the computational complexity of the proposed codec tends to the complexity of the 1-D temporal transform. More
and more 2-D subbands are skipped as the compression ratio increases and entropy coding and 2-D transforms are applied to limited subbands.

An overview of various 3-D DWT based video coding schemes are listed in section II, section III describes the basic subband entropy algorithm, low-complexity bit-plane entropy coding and parent-child skip criterion is introduced in section IV and finally conclusions are drawn in section V.

2. OVERVIEW OF 3-D VIDEO CODING SCHEMES

Wavelet based video coding schemes can be classified into four major classes in literature. The first class treats the video sequence as a volume and performs 3-D DWT followed by entropy coding of the resulting subbands. The second class performs first spatial discrete wavelet transforms followed by in-band Motion Compensated Temporal Filtering (MCTF) and encoding of the resulting spatio-temporal subbands. Third class of video coding schemes apply MCTF on the original image data and then transform and encode the remaining using a critically-sampled wavelet transform. Final class is an extension of the second and third class which is a generic spatio-temporal wavelet decomposition structure alternating spatial transforms, MCTF and subsequent spatial transform. The necessity to calculate the spatial wavelet transform for each frame, even for video sequences which have group of frames (GOF) with low motion or without any motion, is the main disadvantage of the second class schemes. The third class and the fourth class schemes require motion compensation and motion estimation, which will introduce high computational complexity. Therefore for low complexity video coding we consider the first class of wavelet based video coding schemes and use a simplified 3-D extension of the JPEG 2000 standard as the basic video coder. The adopted scheme is shown below in figure 1.

For video compression based on 3-D DWT first, a group of frames (GOF) of length N are accumulated in the input frame buffer and then one-dimensional multilevel DWT of length N (where N is the GOF size) is applied in the temporal direction. All the frames in the GOF are processed. For each frame, from low-frequency to high-frequency, the spatial subbands are processed. Then, two-dimensional multilevel DWT is performed for each frame and each wavelet subband is independently compressed by using bit-plane entropy coding to split bit-planes in to a set of binary streams. Each binary stream is compressed by adaptive binary arithmetic coder (which is also known as the M-coder) from H.264/AVC standard, which is faster and more efficient as compared to the MQ-coder used in JPEG2000 standard.

3. BASIC SUBBAND ENTROPY ENCODER

A. Basic subband entropy algorithm

Input: Subband \( \{ x[i,j] \}, \lambda \)

1: \( n_{\text{max}} \leftarrow \left\lfloor \log_2 \left( \frac{\text{max}(x[i,j])}{(i,j)} \right) \right\rfloor \)

2: \( D \leftarrow \sum (i,j) x[i,j] \)

3: arithmetic_encode(\( n_{\text{max}} \))

4: \( \Psi \leftarrow D \)

5: for \( n = n_{\text{max}}, \ldots, 0 \) do

6: for \( i = 0..S_n - 1 \) do

7: for \( j = 0..S_w - 1 \) do

8: if \( |x[i,j]| \& 2^n \neq 0 \) then

9: \( \text{bit} \leftarrow 1 \)

10: else

11: \( \text{bit} \leftarrow 0 \)

12: end if

13: if \( s[i,j] = 0 \) then

14: \( H \leftarrow s[i-1,j] + s[i+1,j] \)

15: \( V \leftarrow s[i,j-1] + s[i,j+1] \)

16: \( D \leftarrow s[i-1,j-1] + s[i+1,j+1] \)

17: \( c \leftarrow \text{context_encode(bit,c)} \)

18: arithmetic_encode(bit,c)

19: if \( \text{bit} = 1 \) then

20: \( s[i,j] \leftarrow 1 \)

21: arithmetic_encode(sign(x[i,j]),11)

22: end if

23: else

24: arithmetic_encode(bit,9+\( f[i,j] \))

25: \( f[i,j] \leftarrow 1 \)

26: end if

27: \( D \leftarrow D - ((\{x[i,j]\} \& 2^{n+1} - 1))^2 \)

28: \( D \leftarrow D - ((\{x[i,j]\} \& 2^n - 1))^2 \)

29: end for
30: end for

31: $R \leftarrow \text{arithmetic\_get\_bit\_stream\_size()}$

32: if $\psi < D + \lambda R$ then

33: stop encoding

34: else

35: $\psi \leftarrow D + \lambda R$

36: end if

37: end for

Bit-plane entropy coding splits the bit-planes into a set of binary sources. Then each binary source is compressed by an adaptive binary arithmetic coder. First $x[i,j]$ is defined as the value of the coefficient with coordinates $(i,j)$ in a subband of size $S_w \times S_h$, significance flag is $s[i,j]$, $f[i,j]$ is a flag which shows that the current bit is first after the significant bit. First, the number of significant bit-planes in subband $n_{max}$ is calculated and encoded, subband distortion $D$ is calculated and flags $s[i,j]$ and $f[i,j]$ are set to zero. From the highest to the lowest bit plane each coefficient $x[i,j]$ is processed.

For each coefficient the current bit value bit in bitplane $n$ is determined. If $x[i,j]$ is not significant, then the number of significant neighbours $H,V$ and $D$ are calculated. The context number $c$ is determined as it is shown in Table 1 and bit value is compressed by an adaptive binary arithmetic coder corresponding to the context number $c$ and $x[i,j]$ becomes significant if the bit value becomes equal to 1. In this case the $s[i,j]$ is modified and the coefficient sign is compressed.

If $x[i,j]$ was already significant, then the current bit is compressed depending on the position after a significant bit and the subband distortion $D$ is recalculated. The bit stream size $R$ of the subband is determined after encoding of each bit plane and the Lagrange sum $D+\lambda R$ is also determined. If the Lagrange sum $D+\lambda R$ for the bit plane $n$ is greater than the same sum for the bit plane $n+1$, then the bit stream is truncated at the bit plane $n+1$. Encoding is continued otherwise.

<table>
<thead>
<tr>
<th>Table 1: Bit plane context model for wavelet subband</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL and LH subbands</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

4. PROPOSED LOW-COMPLEXITY ENTROPY CODER

4.1 COMBINED ENTROPY CODING

The adaptive binary range coder (ABRC) has less computational complexity due to use of bytes as an output bit stream element and byte renormalization and better compression efficiency due to the assignment of a specific window length selected according to the statistical properties of the corresponding binary source. By the replacement of the M-coder by the ABRC, the computational complexity of the bit-plane entropy coder can be reduced. Using the ABRC the encoding speed of 3-D DWT coder can be increased from 20% to 30%. The properties of different binary contexts are considered in order to further decrease the computational complexity. Figure 2 shows the typical fraction of the context $c=0$ in all input binary data depending on Peak signal-to-noise ratio of the luma component (Y-PSNR) for the entropy coder given in Section III. Typical binary memory less entropy estimation for different binary contexts is shown in figure 3.

![Figure 2: Fraction of context 0 in all input binary data](image)
Figure 3: Binary entropy estimation for different contexts

The context $c=0$ has more than 88% fraction in all binary input data for any Y-PSNR. Binary entropy of the source which corresponds to the context $c=0$ does not exceeds 0.1. This means that the source can be efficiently compressed by zero-run length coding. A combined bit-plane entropy encoding algorithm is thus proposed as shown in Figure 4. The binary sequence corresponding to the context $c=0$ is compressed by the zero-run length coding and placed into Buffer 1 and other binary sequences are compressed by the ABRC and placed into the Buffer 2. Levenstein codes are used for zero run coding with a prefix based on equal codes and this avoids the usage of look-up tables as compared to Huffman codes during the coding process.

For compression of zero-run $000 \ldots 01$

Value $m$ is calculated as follows:

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

Then the value is represented in two parts, $\lfloor \log_2 (\log_2 L_{\text{max}}) \rfloor + 1$ bits for binary representation of $m$ and 2) $m-1$ bits for binary representation of $L - 2^{m-1}$, where $L_{\text{max}}$ is the maximum possible zero-run value. The proposed combined-coder increases the encoding speed of 3-D DWT coder from 77% to 85% on average in comparison with the one using the M-coder.

Figure 4: Proposed combined entropy coder

4.2 SKIPPING OF 2-D WAVELET TRANSFORM AND ENTROPY ENCODING

From algorithm 1, it follows that if

$$D (n_{\text{max}}) + AR (n_{\text{max}}) > \sum_{(i,j)} x^2[i,j] \quad (2)$$

Where $D (n_{\text{max}})$ is the subband distortion and $R (n_{\text{max}})$ is the bit stream size after highest significant bit plane encoding. Computational resources are wasted by the encoder in calculating 2-D DWT and for forming the subband bit stream which is not included in the output bit stream. Therefore a criterion, parent child tree based subband skip criterion, is proposed which can guarantee with a high probability that the current subband will be skipped before processing of this subband. The value of the wavelet coefficient in the child subband is correlated with the corresponding coefficient in the parent subband. The subbands are processed from low-frequency to high frequency temporal subbands and from low-frequency to high frequency spatial subbands. Then a coding assumption is made that if for any subband, inequality (2) holds, the same inequality also holds then for all temporal-spatial child-subbands. In this case, all child-subbands will not be processed and hence spatial transform calculation and entropy coding steps are skipped.

Figure 5: Example of parent-child skipping tree

Frames after temporal wavelet decomposition are L3, H2, H1 and H0 where L3 is the low-frequency frame. If in the frame H2 the subband HH0 is skipped then the corresponding child-subbands HH0 & HH1 in the frame H1, HH0 & HH1 in the frame H0 frame are also skipped without any processing.

Figure 6: Simulation output of bit plane entropy encoding
5. CONCLUSION

In this paper a real-time scalable video codec based on 3-D DWT with low-complexity bit-plane entropy coding is presented. Simulations showed that the proposed codec exhibits a lower computational complexity compared to the most efficient software implementation of H.264/AVC with comparable rate-distortion performance.

REFERENCES


