Error Resilient Wavelet Video Transmission with Priority area Protection using Wyner-Ziv Coding

Zhuo Xue\textsuperscript{1}, K.K. Loo\textsuperscript{1}, J. Cosmas\textsuperscript{1}, P.Y. Yip\textsuperscript{2}
\textsuperscript{1} School of Engineering and Design, Brunel University
\textsuperscript{2} School of Computer Science, University of Hertfordshire
UNITED KINGDOM

Abstract: - In this paper, an efficient error resilience scheme aiming to give priority protection to certain area such as region of interest (ROI) in frame based on Wyner-Ziv (WZ) coding for wavelet based video transmission is proposed. By utilizing additional WZ coding, all the corresponding wavelet coefficients related to ROI will be specially protected during transmission on error prone channel and eventually ROI area can be better reconstructed. Comparing to FEC algorithm fully applied to all wavelet coefficients to realize protection, the proposed scheme has distinct advantages in saving bandwidth and robust transmission.

Key-Words: - Distributed Video Coding, Wyner-Ziv, Wavelet, ROI, Integer Wavelet Transform

1 Introduction
It is only until 1993 after Jussef Shapiro [1] firstly introduced embedded zero tree wavelet (EZW) to encode wavelet coefficients, the Discrete Wavelet Transform (DWT) begin to be widely used in image and video compression to exploit the spatial redundancy. Comparing to the traditional DCT transform, the DWT has distinct advantage in reducing the block artifacts especially at the low bit rate and inherited scalability characteristic. However, the zero tree coding used in EZW or SPIHT [2] make wavelet based video compression sensitive to error prone. Even a small error of wavelet coefficients in transmission could be propagated to influence whole image reconstruction quality. Therefore to design a robust video transmission techniques over heterogeneous and unreliable channels become the new challenge for the wavelet based video transmission. It is undoubted that the bandwidth will be timed if a FEC algorithm applied to protect all wavelet coefficients. However, in many practical applications such as medical image, video surveillance system, there exists within a frame that one or more regions of greater interest than others. Therefore by purposely protecting ROI area other than whole frame, namely give the priority protection to certain area in the frame, only few wavelet coefficients related to ROI area will be coded, which result in significant gain eventually in term of compression and bandwidth efficiency.

The Wyner-Ziv coding from lossy source coding theory has been recently adopted as error resilience method in video transmission. In this paper, a bandwidth efficient error resilience scheme based on WZ coding is proposed to protect ROI areas within the frame in order to trade off the bandwidth utilization and error resilience.

2 Related Work

2.1 ROI
ROI feature is supported in JPEG 2000 which adopted DWT to perform special compression, where ROI image can be coded with better quality than background. In general, two main kinds of methods are defined in [3, 4, 5], which are the general scaling based method and the maximum shift method. The principles of two methods are similar, in which after wavelet transform, the resulting coefficients not related to ROI will be scale down so that the ROI-associated bits are placed in the higher bit plane. During embedded encoding process, the bits in higher bit plane will be sent earlier than those bits in lower bit plane. To carry out this process, a ROI mask will be generated to indicate all the wavelet coefficients related to ROI. In this paper, the ROI mask generation method from [3] to identify the coefficients has been adopted. However, we won’t scale the ROI related coefficients as in JPEG 2000, but sent it to WZ codec for transmission.

2.2 Wyner-Ziv Coding
WZ coding is referred to lossy compression with side information at decoder. Details of WZ coding and application of distributed video coding can be referred to the literatures [6 7 8]. In [9], multiple embedded Wyner-Ziv video descriptions is first used in systematic lossy error protection for video waveform. In [10 11 12], by using the redundant slices and flexible macroblock ordering, the “SLEP” scheme is proposed for H.264/AVC, in which an additional WZ bit stream is transmitted to realize error resilience. In this paper, additional WZ bit stream will be used to protect ROI related wavelet coefficients. The related wavelet coefficients are encoded by turbo code and only the parity bits will be save into buffer to sent based on the request. At the decoder side, the reconstructed ROI related wavelet coefficients from wavelet video codec will be used as side information to help turbo code to perform decoding. The main advantage of using WZ coding is that bandwidth can be efficiently used for WZ stream. The required bandwidth is dynamically controlled by degree of error of side information, which is ROI related DWT coefficients passed through the error channel without any protection in normal wavelet video stream. If there are not too many errors happened in side information, then it can successfully decode the WZ stream with very small amount of parity bits. In contrast, if the side information has been corrupted badly during transmission, the bandwidth utilization will be increased by sending more parity bits for decoding and eventually recover the error of side information successfully.

2.3 ROI mask generation

The ROI mask is a bitplane indicating a set of wavelet coefficients whose exact transmission is sufficient for the receiver in order to reconstruct the desired region perfectly. The details of ROI mask derivation should refer to [3]. In this paper, the mask is a matrix which was initialized to zero with same size of the frame. The mask is derived following the same steps as the forward transform by tracing the inverse transform backwards. The position of all the coefficients used to reconstruct the pixel in ROI area will be marked in mask. However, it should be mentioned here that in order to get the complete reversible transform, the integer wavelet transform based on lifting scheme is used here. During the transform, at each decomposition level, the mask will indicate which coefficients are needed exactly at this level so that the inverse transform will reproduce the ROI related coefficients in previous level exactly. The ROI mask matrix is grown slowly following the forward transform until whole transform is finished. According to this mask, the ROI related coefficients will be picked out and sent to WZ codec for transmission. A typical example definition of ROI mask is show below

\[
M(x,y) = \begin{cases} 
1, & \text{the coefficient located at } (x,y) \text{ is related to ROI and therefore need to be picked out and protected} \\
0, & \text{the coefficient located at } (x,y) \text{ is not related to ROI therefore can be neglected}
\end{cases}
\]

3 Proposed Architecture

Fig.1 illustrated the concept of the proposed error resilience scheme based on Wyner-Ziv coding. The input signal is compressed and sent by normal wavelet based video transmission on an error prone channel without protection. The wavelet based video transmission adopted generic video prediction structure, which employs DWT, motion estimation, and entropy coding to realize spatial, temporal and data compression. The ROI area is manually defined before DWT decomposition. The ROI mask which indicates the corresponding coefficients in wavelet domain for reconstruct the ROI is generated after DWT decomposition. By using ROI mask, wavelet coefficients related to ROI will be progressively scanned and sent to WZ codec to transmit. It should be mentioned that we only protect ROI area of I frame, by which the corresponding ROI area in P or B frames will also be better reconstructed since I frame is used as reference frame to build P or B frame. Note that ROI mask still need to be sent to decoder via normal channel in order to rebuild the side information. Each ROI related coefficient will be uniformly quantized and sent by bit plane. The turbo encoder used here worked with RCPT [13] to realize dynamic rate control. The generated parity bits are saved into buffer and wait for transmission, while all systematic bits are discarded. A small fraction of parity bits will be sent to decoder first for initial decoding. At the turbo decoder, by using received ROI mask, the error corrupted wavelet coefficients of ROI region will be used as side information to help turbo decoder to perform decoding. An ideal error detection at decoder is assumed. If the BER of whole bit plane is over than 10^{-3}, we consider that the majority of ROI related coefficients haven’t been
correctly recovered. The decoder will send the request for more parity bits to encoder via feedback channel. This process continued repeatedly until whole bit plane is successfully decoded. This kind of ‘decode and request’ process greatly adapt the channel errors automatically and can dynamically control the bit rate of WZ stream and efficiently utilized the bandwidth. After all bit planes are decoded, the whole symbol will be dequantized and all the ROI related wavelet coefficients will be rebuilt, which later than replaced the error corrupted ROI coefficients. IDWT is then performed and the whole picture is rebuilt subsequently.

4 Experimental Results

In this section, the performance of proposed error resilience scheme is demonstrated. We investigate the scenario of wavelet based video transmission over wireless link. Typically, the AWGN channel has been adopted as type of error-prone channel. Turbo code composed by two identical constituent convolutional encoder of rate 1/2 with constraint length of 4 and with polynomial generator of (13,11) is adopted in this architecture. The puncturing period of RCPT is set to 8, which provided various code rates of (8/9,8/10,8/11,....1/3 etc). The wavelet transform filters used here are the Daubechies (9,7) filter set with lifting scheme to perform fast integer wavelet transform. The frame rate is set to 12.5fps, GOP size is 60. Note that since proposed scheme only deal with I frame, therefore the frame rate actually sent into WZ codec is only 0.2fps. 256 levels Quantization is performed for ROI related coefficient to provide lossless quantization. Three QCIF sequences “Foreman”, ”Akiyo”, and “Carphone” have been tested during simulation. Moreover, to simplify the analysis, we manually defined ROI area in rectangular shape.

Fig.2, 3, and 4 illustrated PSNR performance of 3 sequences under the different channel condition with different error protection schemes. We test each sequence with different degree of protection, which are no protection, ROI protection, and full protection (protect whole frame by turbo code with rate 1/2). It
can be clearly seen that without any error resilience scheme, the video sequence is badly corroded by channel, therefore showed very poor PSNR performance. Conversely, as expected, a fully protected scheme showed the best PSNR performance, but obviously it requires huge bandwidth to keep error resilience ability. The proposed ROI protection scheme has improved system PSNR performance up to 4.9 dB, with only compensation of little additional bandwidth. (the peak area at 61st frame is due to the a new I frame is reconstructed).

Most importantly, with ROI protection, the ROI area, which is considered more interesting and important area than background can be better reconstructed. Fig. 5, 6 and 7 showed the perceptual quality comparison for one frame of three sequence. The perceptual quality of the reconstructed image without any protection is too bad to accept. However, with ROI protection, if we only concern about the quality of ROI area, circulated in red rectangular area, then perceptual quality of ROI area is certainly acceptable. Moreover, with the increase in the accuracy of ROI related coefficients, not only the ROI area is better reconstructed, but also other areas which would use ROI related DWT coefficients for reconstruction will also be improved in term of perceptual quality.

Table 1 analysis the actual bandwidth utilization for each scheme illustrated in figure 2, 3 and 4. It is obvious that full protected scheme occupied twice bandwidth of the original stream to maintain higher error correction ability. WZ stream bit rate, influenced by ROI size, WZ frame rate, quantization level and final decoding code rate is little bandwidth usage comparing to two other methods. However, this little bandwidth usage can efficiently protect the ROI and provide the decent output video quality in ROI area. Especially compare to fully protected stream, the bandwidth could save near to 50% and have been efficiently utilized.

5 Conclusion

In this paper, we propose a bandwidth efficient error resilience scheme based on WZ coding to give priority protection to ROI area for wavelet based video transmission on error prone channel. By introducing additional WZ coding with negligible complexity, the wavelet based video system can be expected much more robust during transmission. Furthermore, comparing to the FEC algorithm fully applied to whole wavelet coefficients; the introduction of ROI coding and WZ coding has efficiently used the limited bandwidth and provide better video quality in the interesting area in frame.

References:


Figure 7. Perceptual Quality Comparison of 213th frame of Carphone sequence

a) No protection  
b) ROI protected  
c) Fully protected

Table I. Bandwidth unitization comparison.

<table>
<thead>
<tr>
<th></th>
<th>ROI size</th>
<th>WZ stream bit rate</th>
<th>NO protection bit rate</th>
<th>ROI protection bit rate</th>
<th>Fully protection bit rate</th>
<th>Bandwidth saved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreman</td>
<td>71x77</td>
<td>9.1kbps</td>
<td>152.5kbps</td>
<td>152.5*2</td>
<td>305kbps</td>
<td>305-161.6 = 143.4kbps</td>
</tr>
<tr>
<td>Carphone</td>
<td>63x92</td>
<td>9.6kbps</td>
<td>165.9kbps</td>
<td>175.5*2</td>
<td>351kbps</td>
<td>351-175.5 = 175.5kbps</td>
</tr>
<tr>
<td>Akiyo</td>
<td>48x72</td>
<td>3.6kbps</td>
<td>67.3kbps</td>
<td>67.3*2</td>
<td>134.6kbps</td>
<td>134.6-70.9 = 63.7kbps</td>
</tr>
</tbody>
</table>