# Error-resilient Scheme for Wavelet Video Codec using Automatic ROI Detection and Wyner-Ziv coding over Packet Erasure Channel

Z. Xue, K.K. Loo, J. Cosmas, M. Tun, L.F. Feng, and P.Y. Yip

***Abstract*—The error-resilient for video transmission over the Internet in which regarded as the packet erasure channel is always a tough task and has gained lots of attentions. The main contradictory problem lies between error-resilient and bandwidth usage. Additional redundant data has to be added to achieve robust transmission which leads to huge bandwidth usage. In this paper, an error-resilient scheme called Wyner-Ziv Error-Resilient (WZER) based on a receiver driven layered Wyner-Ziv (WZ) coding framework is proposed. The WZER purposely emphasizes on the protection of the Region of Interest (ROI) area in the frame thus to achieve the better tradeoff between the bandwidth usage and error-resilience. WZER is designed to work for the scenario of wavelet based video coding over packet erasure channel, where several techniques including automatic ROI detection, ROI mask generation, Rate distortion optimization (RDO) quantization, WZ coding with layer design, and packet level Low Density Parity Check (LDPC) code are used. The performances of the proposed WZER are simulated based on average PSNR of luminance, perceptual reconstruction and bandwidth usage and compared with normal Forward Error Correction (FEC) full protection scheme and no protection scheme. The results show the advantages of the proposed WZER over traditional FEC protection, especially in the aspects of the recovery of the subject area and bandwidth efficiency.**

***Index Terms*—Wyer-Ziv coding, Error-resilient, ROI, Wavelet video coding**

## I. Introduction

With the rapid growth of internet and modern communications technology, the strong demand of various multimedia applications and services has promoted the research of robust transmission of compressed image and video data. In recent years, the design of robust video transmission techniques over heterogeneous and unreliable channels has been an active research area. However, error control in image/video communications is proved to be a very tough task. In most popular video coding standards, compressed video streams are very susceptible to transmission errors due to the spatial-temporal prediction coding structure and Variable Length Coding (VLC) at the source coder. A single incorrect recovered pixel in a reference frame can lead to more errors in the following reconstructed frames. Therefore, in order to deliver the compressed bitstream over an error-prone network, the video has to be coded in a resilient format to combat the channel errors. In [1], Wang et al. generalized most error-resilient techniques and categorized them into three groups: Firstly, the coder works at the source and channel to make the bitstream more error-resilient such as adopting Redundant Slices and the Flexible Macro-Block Ordering (FMO) in H.264/AVC [2] or inserting some periodic macroblocks during transmission as introduced in [3], or applying FEC mechanism to the video stream; Secondly, the error concealment method is applied at decoder side and such algorithm can be found in [4-6]; Thirdly, the interaction between the encoder part and decoder is introduced so that the encoder can adapt the operation based on loss information provided by the decoder. A feedback channel working with ARQ is usually combined for this algorithm.

In most video coding standards, the Discrete Cosine Transform (DCT) is widely used to realize spatial compression. However, the Discrete Wavelet Transform (DWT) emerging later has become increasingly popular in last decades since Shapiro [7] and Said [8] introduced the Embedded Zerotree Wavelet (EZW) and Set Partitioning In Hierarchical Trees(SPIHT) to efficiently code the wavelet coefficients. Later in [9] and [10], 3D-EZW and 3D-SPIHT were proposed with superior performance. DWT has distinct advantages in reducing the block artifacts especially at the low bit rate and inherited scalability characteristic comparing to the DCT transform. With these advantages of DWT, the wavelet based video coding combining the DWT transform and block based predictive coding has been proposed in various applications. However, in order to widen the applications of such wavelet based video coding scenario, error-resilient ability of video stream has to be considered as well. Comparing to the DCT based video coding algorithms, DWT based video coding will experience more problems due to the use of EZW or SPIHT. Not only it will suffer the same problems resulted by predictive

Manuscript received in April 2009. Revised in May 2010

Z. Xue and J. Cosmas are with the School of Engineering and Design, Brunel University, UB83PH, UK (email: Zhuo.Xue@hotmail.com, John.Cosmas@brunel.ac.uk).

K.K. Loo is with the School of Engineering and Information Sciences, Middlesex University, NW4 4BT, UK (email: J.Loo@mdx.ac.uk)

M. Tun is with Dialogic Corp., Quebec, H4M 2V9, Canada (email: Myo.Tun@dialogic.com)

L.F. Feng is with the University of Science and Technology Beijing, China (email: lifang.feng@gmail.com).

P.Y. Yip is with the School of Computer Sciences, University of Hertfordshire, AL10 9AB, UK (email: p.y.a.yip@herts.ac.uk).

coding and VLC entropy coding as discussed above, but also it will face more serious situation than the DCT, because EZW or SPIHT also will produce code words with variable length and similarly a single bit of error will result in the loss of synchronization. Furthermore, the influence of error could propagate to the whole frame in video reconstruction.

## II. Related work

There have been several approaches to realize the error-resilience of wavelet video coding over error-prone channel. In [11], Creusere proposed the basic idea of realizing the error-resilience by partitioning the wavelet transform coefficients into groups, which will be processed individually by embedded encoder therefore a bit error happened in one group would not affect others. The similar idea can also be found in [12] where the block based coefficient partitioning was proposed.

However, the above methods that only add the error-resilient format to bitstream at source encoder cannot make the stream robust enough for transmission over error-prone channel. This has lead to the research that considering the application of FEC to the video bitstream to against channel errors. In [13-15], the sub-bitstream is protected by different code rate of Rate Compatible Punctured Convolutional (RCPC) coding according to the importance of the content of bitstream, thus to achieve the different protection. Later Kim et al. extended this idea to working with 3D-SPIHT [16], where the ARQ is also adopted which constrains the algorithm to be applied in real-time application. In 2002, Cho et al. extended the ideas from [11] to 3D-SPIHT and use same RCPC as in [16] to protect resulted packets [17]. The work has been further explored in [18, 19], where the error concealment mechanism and RCPC in conjunction with CRC were added. In [20], Tun et al. proposed the similar error-resilient algorithm to DIRAC [21], which is considered as the most mature wavelet video codec so far, by extending the partition method of wavelet transform coefficients to motion compensated residual frame and protecting each packet equally by RCPC and Turbo Code with cyclic redundancy check (CRC). These works suggested that FEC has to be adopted in one way or another in order to protect the wavelet-based compressed bitstream. However, the application of FEC inevitably costs huge bandwidth and the error resilience is highly depended on the coding rate.

The Wyner-Ziv (WZ) coding from distributed source coding theory [22-24] has been recently adopted as an option for error-resilience in video transmission. It has been shown that using additional WZ bitstream in a systematic lossy error protection framework can provide competitive error-resilience comparing to FEC, especially having advantage in limiting the quality degradation and error propagation with less consumption of bandwidth. In [25,26], Arron et al. first reported the result of applying WZ coding for error-resilient video transmission. The basic idea is that the video signal compressed by MEPG-2 is transmitted over an error-prone channel without any protection. A supplementary stream which is a low rate representation of the transmitted video sequence through coarse quantization is generated using WZ encoding. The received error-prone MPEG-2 bitstream is used as side information to decode the WZ bitstream. The decoder combines the error-prone side information and the WZ description to yield an improved decoded video signal. The work has been further improved in [27] by composing the WZ stream with coarsely quantized prediction error from MEPG-2 compression and applying RS codes in the WZ codec. The algorithm later was named as Systematic Lossy Error Protection (SLEP) in Rane's later work [28, 29], where further improvement have been proposed including using H.264/AVC to generate the lossy systematic bitstream and using coarse description of redundant slices for WZ bitstream etc. Based on Rane's work in [29], Baccichet et al. [30] introduced the Flexible Macroblock Ordering (FMO) in SLEP to coarsely quantized the region of interest in the frame thus to improve the subject quality. In [31], the multiple WZ bitstreams containing embedded video descriptions was proposed to better exploit the trade-off between error-resilience and the residual distortion from coarse quantization in the WZ codec. The similar work can also be found in [32-35].

There is a trade-off between error-resilience and bandwidth usage. It is always necessary to design such a system to find the balance between these two. In many practical application areas such as medical image, video surveillance system etc., there exists one or more regions of greater interest than others within a frame. Therefore, for this kind of applications, it is not necessary to waste the bandwidth to protect the whole frame since the end users only concern the quality of Region Of Interest (ROI) area. Hence it is sensible to give higher priority to those ROI areas than other areas during transmission. Technically speaking, it is feasible to purposely protect the ROI area other than the whole frame. By scarifying the quality of some unimportant area, the ROI can have better reconstruction. Since only a small bitstream will be generated additionally for ROI area protection comparing to the whole frame protection, huge bandwidth cost can be saved and the video quality of ROI area can also be improved significantly. By this way the overall system gain can be obtained in term of compression and bandwidth usage. This idea is referred as content based video coding in [5].

In our previous work [36], we proposed a bandwidth efficient error-resilient algorithm for wavelet video coding based on the WZ protection. In the proposed architecture, the video signal is compressed by a generic wavelet video coding to compose the systematic lossy stream and sent through error-prone channel without FEC protection. Meanwhile a supplementary WZ stream which only contains the description of ROI area in wavelet domain will be sent to the decoder for error protection. The ROI area is predefined and the AWGN channel is adopted in the paper. In this paper, we further extend the previous work in following aspects: firstly an automatic ROI detection method is proposed to generate ROI area automatically which gives much more flexibility than the method that define the ROI area in advance manually. Secondly,

we use the maximum shift method adopted in JPEG 2000 [37], RDO quantization [20], and Rate Compatible Punctured Turbo (RCPT) coding [38] to encode the ROI related wavelet coefficients to compose the WZ stream. Thirdly, the proposed scheme will be working on the packet loss channel. The generated WZ stream will be divided into packets for transmission. Moreover, the packet level Low Density Parity Check (LDPC) codec will be applied to protect WZ packets via adding redundant parity packets. After the LDPC decoding, the decoder will combine the received lossy systematic stream and WZ stream to yield the protected ROI related coefficients which promise higher output quality in ROI area. The trade-off between bandwidth usage and error correction ability is optimized in the way that using limited bandwidth to protect the most important area (ROI) that the receiver concerns most. Finally, in order to satisfy the various requirements from heterogeneous groups with various bandwidth conditions and make the scheme compatible in the application like video multicast, the WZER is designed based on a receiver-driven layered protection framework, which enable the receiver choose the best size of WZ stream via joining different layers based on available bandwidth.

The remainder of this paper is organized as follows: Section III, we describe the details of the proposed automatic ROI detection algorithm and the generation method of ROI mask. The concept of receiver driven layered WZ coding for error-resilience is described in Section IV. The details of applying packet level LDPC codec will be explained in Section V. The whole system architecture and operation is introduced in Section VI. Section VII will reveal the simulation results and shown the performance of proposed scheme. Finally, the paper is concluded in Section VIII.

## III. Automatic ROI detection and ROI mask generation

ROI area usually is predefined in many applications, however, this is based on the assumption that the encoder has the knowledge of video content thus can manually define the ROI area in advance. In many applications, the video content is not predicable therefore it is hard to define the ROI area in advance.

In this paper, an automatic ROI detection method which can be considered as a simple version of video segmentation method from [39, 40] is proposed, by which the subjective area with high motion in the frame can be detected automatically and is to be defined as ROI area. This prediction is accurate in most circumstances, especially for the video with comparatively static background. The proposed algorithm consists of three processes namely frame difference mask definition, ROI generation and $D_{th}$ decision as described in the following sub-sections.

### A. Frame difference mask definition

Denote the length of a group of Frame (GOP) as *n*, the frame difference between current frame and previous frame is calculated as

$$f_d(x,y,t) = \left| f(x,y,t+1) - f(x,y,t) \right| \tag{1}$$

where $f(x,y,t)$ is the representation of the frame data in which $x$, $y$ denote the coordinates and $t$ is frame index with $1 \le t \le n\text{-}1$. $f_d$ denotes the frame difference. With $f_d$, the frame difference mask, denoted by $f_{dm}$ can be calculated as:

$$f_{dm}(x,y,t) = \begin{cases} 1, & \text{if } f_d(x,y,t) \ge D_{th} \\ 0, & \text{if } f_d(x,y,t) < D_{th} \end{cases} \tag{2}$$

Note that the parameter $D_{th}$ need to be set in advance by using $f_d(x, y, 1)$. The exact calculation is shown in sub-section *C*. Pixels marked by $f_{dm}$ are considered as "moving pixels".

### B. ROI generation

According to $f_{dm}$, pixels moving for a long time in the adjacent frames are defined as ROI pixels. The procedure of ROI generation can be shown as:

$$f_n(x,y,n-1) = \sum_{t=1}^{n-1} f_{dm}(x,y,t) \tag{3}$$

$$ROI(x,y) = \begin{cases} 1, & \text{if } f_n(x,y,n-1) \ge F_{th} \\ 0, & \text{if } f_n(x,y,n-1) < F_{th} \end{cases} \tag{4}$$

where $f_n$ stores the number of times that a pixel moved in a whole GOP and *ROI* denotes the region of interest. The initial value of $f_n$ and *ROI* are all set to "0." The parameter $F_{th}$ is a manually predefined constant which presents the sensitivity of the ROI area detection. A pixel can be viewed as ROI pixel only if it moves for more than $F_{th}$ times.

### C. $D_{th}$ decision

This section reveal the process to derive the parameter $D_{th}$ used in (2), which is based on two steps: Gaussianity Test and $D_{th}$ output.

#### 1) Gaussianity Test

The frame difference of background part follows Gaussian distribution and those of ROI area are normally not. The reason is that for video with comparatively static background, the frame difference values between two adjacent frames vary slightly in the part of background (most are zero value plus camera noise) and significantly in the part of foreground (ROI) area. The Gaussianity test [39, 40] can be used to indicate if a group of values is Gaussian distributed or not, with which we can roughly distinguish the background and ROI area. First, the frame difference $f_d(x,y,t)$ is divided into many blocks (assume that the size of the block is $M \times N$). The Gaussianity test is then applied to each block to examine if the frame differences in the block are distributed in Gaussian or not. The block distributed in Gaussian is deemed to belong to background, and non-Gaussian block belongs to the foreground. The Gaussianity test can be shown as

$$F_{d\,k}(t) = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} \left( f_d(m,n,t) \right)^k \tag{5}$$

$$\begin{aligned} &H(F_{d1}, F_{d2}, F_{d3}, F_{d4}) \\ &= F_{d3} + F_{d4} - 3F_{d1}(F_{d2} - F_{d1}^2) - 3F_{d2}^2 - F_{d1}^3 - 2F_{d1}^4 \end{aligned} \tag{6}$$

where ***H*** is the Gaussianity test function and it is defined as follows: when $H(F_{d1}, F_{d2}, F_{d3}, F_{d4}) < G_{th}$, the current block is Gaussian distributed. Vice versa, the block is Non-Gaussian when $H(F_{d1}, F_{d2}, F_{d3}, F_{d4}) \geq G_{th}$.

The parameter $G_{th}$ here can be set as a constant "1" because the values of ***H*** of foreground blocks and background blocks are dramatically different. As mentioned above, the Gaussianity test can only roughly distinguish background parts and foreground parts. To get more precise distinguishing, we need the next step to find the optimum value of $D_{th}$.

*2) $D_{th}$ output*

Considering the digitizing effect of digital systems, $D_{th}$ can be one of fame difference values of background range from 0~255. Since the frame difference of background part is Gaussian distributed, and the probability distribution of the absolute value of frame difference of background in digital domain should be (7):

$$P_k = \begin{cases} \int_0^{0.5} \frac{2}{\sigma\sqrt{2\pi}} e^{(-x^2/2\sigma^2)} dx, & \text{for } k=0 \\ \int_{k-0.5}^{k+0.5} \frac{2}{\sigma\sqrt{2\pi}} e^{(-x^2/2\sigma^2)} dx, & \text{for } k=1,2,...,255 \end{cases} \tag{7}$$

As defined in (2) and (4), the pixel located at $(x_1,y_1)$ in the frame is defined as background eventually only when frame differences of the pixel $(x_1,y_1)$ in a GOP which are $f_d(x_1,y_1,1)$, $f_d(x_1,y_1,2)$…$f_d(x_1,y_1,n\text{-}1)$ must be smaller than $D_{th}$ for $(n\text{-}1\text{-}F_{th})$ times. Assume $D_{th}$ value is k(0~255) for a GOP of frame difference, and the optimal value of $D_{th}$ threshold should be its expected value which can be obtained by the following:

$$E(D_{th}) = \sum_{k=0}^{255} kP(D_{th}=k) = \sum_{k=0}^{255} k\left(P_k\right)^{F_{th}}$$
$$P(D_{th}=k) = \left(P_k\right)^{F_{th}} \tag{8}$$

The standard deviation $\sigma$ can be calculated by (9). For more details of derivation of $D_{th,}$ readers are advised to refer to [39], [40] where deep analysis and elaborations are given.

$$\sigma = \sqrt{\sum_{pixels \in backgroundblocks} {f_d}^2} \tag{9}$$

Fig. 1 shows the example of detected ROI areas generated by different threshold $F_{th}$ for Akiyo sequence. As introduced in (4), only pixels moved more than $F_{th}$ times that will be considered as the pixels in ROI area. Hence, the value of threshold $F_{th}$ reflects the sensitivity of motion detection. It can be observed in the figure that the area with low motion will be neglected during detection if high $F_{th}$ is set and vice versa. On the other hand, $F_{th}$ can be used as a parameter to control the size of ROI area thus generate the different number of ROI related coefficients, which directly determine the size of WZ bitstream later.

After detecting the ROI area, a ROI mask will be defined to mark ROI related wavelet coefficients in wavelet domain so that the positions of the coefficients can be tracked and then coefficients are protected via WZ codec. ROI is a standard feature supported in JPEG 2000 which also adopted DWT to perform spatial compression. In JPEG 2000, ROI image can be coded with better quality than background. In general, two main kinds of methods are defined in [37, 41, 42], which are the general scaling based method and the maximum shift method. The principles of these two methods are similar, in which after wavelet transform, the resulting coefficients not related to ROI will be scaled 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 key component is the ROI mask that will be generated to indicate the positions of all the wavelet coefficients related to ROI. In this paper, the ROI mask generation method from [42] to identify the coefficients is adopted.


Fig. 1. ROI area detection with different $F_{th}$

The ROI mask is a bit plane indicating a set of wavelet coefficients in which the exact transmission is sufficient for the receiver in order to reconstruct the desired region perfectly. The details of ROI mask derivation can be found in [42]. A brief introduction will be given in this paper. The mask is a matrix which was initialized to zero with same size of the frame. Following the same steps as the forward transform, the mask is derived by tracing the inverse transform backwards. The positions of all the coefficients used to reconstruct the pixels in ROI area will be marked in the mask. In order to get the complete reversible transform, the integer wavelet transform based on lifting scheme has to be used.

During the transform, at each decomposition level, the mask will indicate the coefficients which are needed at this level so that the inverse transform will reproduce the ROI related coefficients in previous level exactly. The ROI mask matrix is growing 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 expressed below:

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

Fig. 2 shows a ROI mask generated using the automatic ROI detection with threshold $F_{th}$=2. The coefficients highlighted in non-black area are the ROI related coefficients and need to be specially protected by WZ steam.

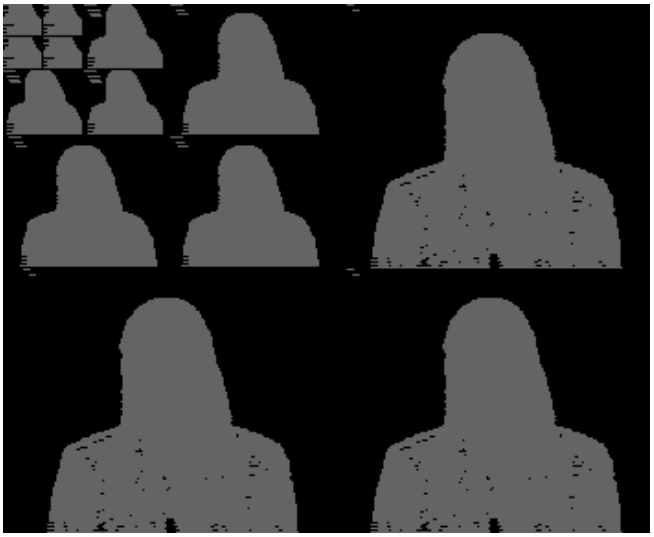


Fig. 2. Example of ROI mask generated using the automatic ROI detection with threshold $F_{th}$=2

## IV. Receiver driven layered Wyner-Ziv coding

Wyner-Ziv coding refers to distributed lossy source coding with side information at the decoder. It is suggested in [43, 44] that efficient compression still can be achieved if two statistical dependent sources *X* and *Y* are separately coded and jointly decoded. In Wyner-Ziv coding, *X* is coded without knowing of *Y*, and the decoder conditionally decode *X* with *Y*. In practical Wyner-Ziv coding applications, this usually is achieved via following method. The Wyner-Ziv source *X* is coded by certain type of channel codec (such as turbo code, LDPC etc), but the systematic bits will be discarded after the encoding and only the generated parity bits will be sent. The decoder need to first generate the side information *Y*, which is considered as channel corrupted or estimated version of *X*. Then by combining the *Y* and the received parity bits, *X* is estimated and decoded at decoder side. In the WZ error-resilient scenarios introduced in [25, 26], the side information is extracted out from the main systematic system passed through the channel, which can be regarded as the error description of encoded WZ information. Similarly, in this paper, the WZ stream contains the description of ROI in wavelet domain. The corresponding side information is located via using the ROI mask from systematic stream. With the side information and received parity bits, the WZ stream can be decoded. The advantages of applying WZ coding algorithm in our error-resilient framework are two folds. Firstly, only the ROI area is protected, hence the size of WZ steam is very small and the bandwidth usage is small. Secondly, since we do not send the systematic bits after encoding, and the number of parity bits sent is based on the coding rate chosen, the bandwidth usage is very much reduced as compared to a normal FEC.

Moreover, the proposed error-resilient scheme is based on a receiver-driven layered protection framework, namely, the receiver determines which layer can be combined in order to get the suitable size of WZ stream within the available bandwidth. This type of layer framework is typically applied in video multicasting. More specific design about such framework can be found in work [45-48]. Basically, in the proposed scenario, there are three factors directly affecting the size of WZ stream, which are the $F_{th}$ threshold value, the number of subbands of ROI related coefficients need to be protected and the encoding rate of turbo codec in WZ codec. According to this, the whole layered structure is designed and shown in Fig. 3. Firstly, the base layer is called $F_{th}$ layer, in which $F_{th}$ can be set to determine the size of ROI area. The second layer is called subband layer, which decides the number of subbands of ROI related coefficients should be protected by WZ codec. More coefficient subbands involved in WZ stream would lead to more picture details recovered in ROI area. The third layered is called parity layer, in which different coding rates (or different puncturing) are selected to generate different size of parity bits set in WZ codec.

Fig. 3 shows the structure of layer design, the receiver chooses the best option at each layer thus to determine a suitable size of WZ stream to achieve corresponding error-resilience performance within the available bandwidth.

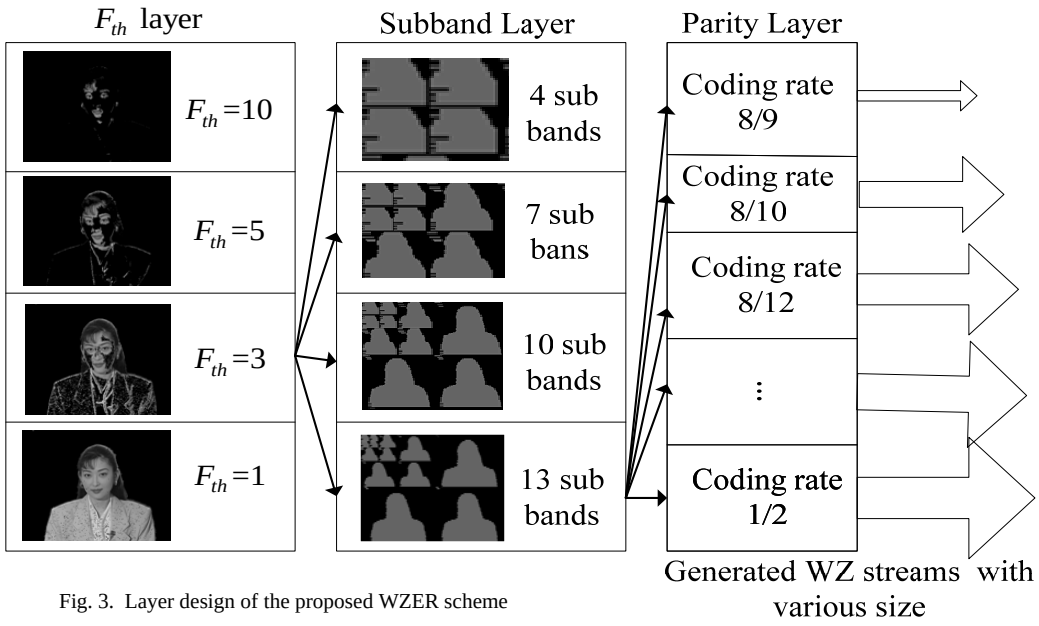



Fig. 3. Layer design of the proposed WZER scheme

## V. Packet Level LDPC Codec Protection

There has been a lot of research around the packet level FEC protection in order to against the packet loss [49-51]. In this paper, an advanced packet level LDPC-triangle codec [52], which recently is adopted as standard of Internet Engineering Task Force (IETF), is used to protect the generated WZ over packet erasure channel. The LDPC codec used is a new powerful FEC codec that can survive in the packet loss channel with loss rate nearly up to 30% with code rate of 2/3.

In this paper, the LDPC codec is used to protect the generated WZ packets and 2/3 code rate is sufficient to guarantee the WZ packet delivered at decoder with error free under the packet erasure channel even with maximum packet loss of 30%. The LDPC codec will be used as full FEC scheme to protect the whole systematic stream and it is compared with the WZER in the simulation. The details of FEC scheme using packet level LDPC-triangle codec will be introduced in the simulation part.

## VI. System Architecture and Operations

Fig. 4 illustrates the architecture of WZER. The system is composed by two streams, a main wavelet compressed systematic stream and an additional WZ stream to provide error-resilience. Firstly, the input signal is passed through main wavelet video codec and compressed. The encoder adopts a generic video prediction structure, DWT, and entropy coding to realize temporal, spatial and data compression. Same as in most popular video compression standard, the GOP mode is also introduced where there are three types of frames: I, P and B frame which are coded by intra frame coding and interframe coding respectively. The fast integer DWT with lifting scheme is applied to perform wavelet transform in order to get the perfect reconstruction in IDWT. The compressed bitstream consisting of all bit steams from all subbands is firstly interleaved and divided into packets to transmit on the packet erasure channel without any protection. Meanwhile at the encoder, the video frames of a GOP are first sent to video buffer, the proposed automatic ROI detection method will be applied to generate the ROI area for current GOP and each GOP will only be assigned with one ROI in order to keep it updated during encoding. After the ROI area is defined, a ROI mask is generated after DWT decomposition, in which the positions of ROI related wavelet coefficients are marked. The ROI related coefficients in each subband will be individually scanned and uniformly quantized. The rate distortion optimization (RDO) quantization is performed for each subband, in which the best quantizer for current subband will be chosen by minimizing the Lagrangian combination of rate and distortion. The generated quantized symbols from each subband will be binarized into bitstreams, which then are multiplexed into one serial bitstream.

The multiplexing is performed in the way that the lower frequency subbands are placed first then higher frequency bands. The purpose is to make the stream robust to the channel loss so that if the video bitstream is truncated at any time during transmission, the end user still can use currently received bitstream to realize the partial error-resilience. The multiplexed stream is fed into turbo encoder for encoding, which is implemented as RCPT [38] with the ability to dynamically rate control of WZ stream thus to achieve different error protection. After the encoding, the systematic bits are discarded and only the parity bits are sent as WZ stream. The WZ stream is divided into the packets to transmit. In order to combat with the packet loss in the channel and guarantee the WZ packet delivered with error free, the packet level LDPC code with code rate of 2/3 introduced in Section V is applied to protect the WZ packet. The WZ packets and the generated LDPC parity packets will be multiplexed with packets from main systematic stream for transmission.

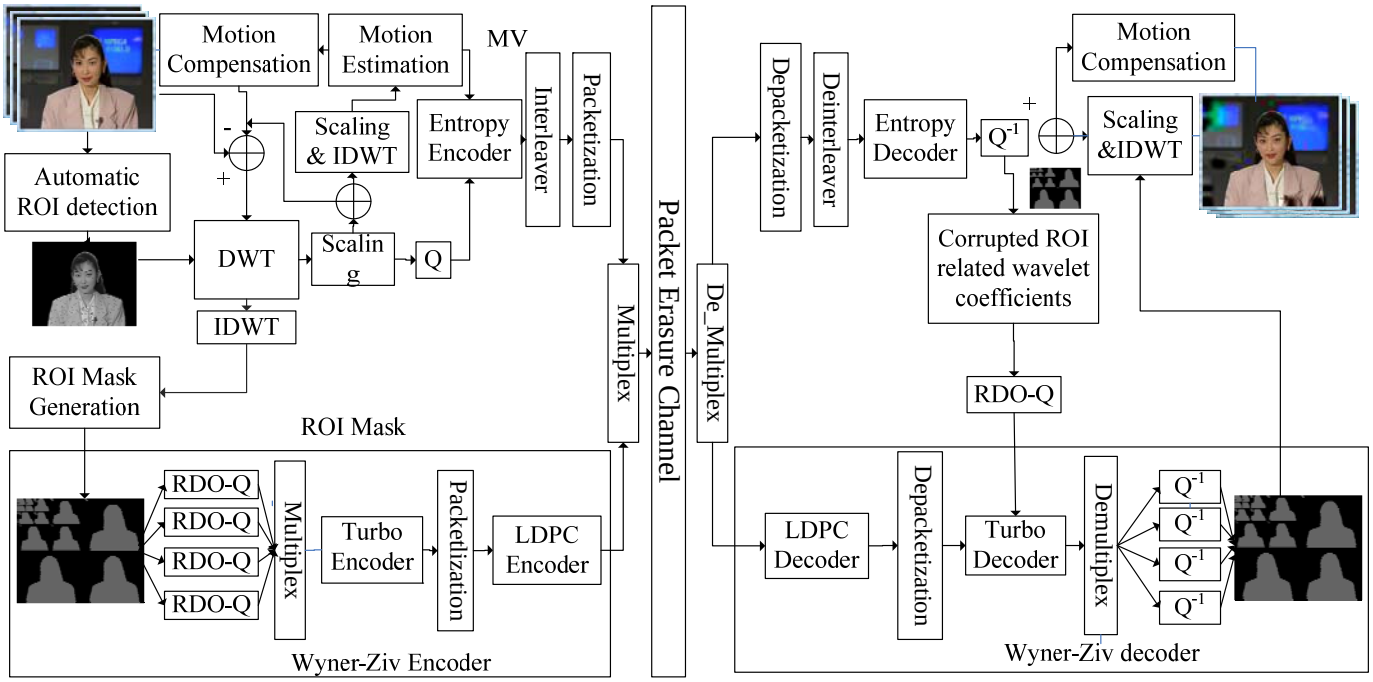



Fig. 4. The proposed WZER for wavelet video coding over packet erasure channel

It should be specially mentioned here that for each GOP, only the ROI area of I frame is protected; nevertheless the corresponding ROI area in P or B frames can also be reconstructed since the I frame is used as reference frame to build P or B frame. Note that ROI mask and quantization parameter of each subband still need to be sent to decoder via normal channel in order to rebuild the side information. Since the ROI mask only count for a negligible size of data stream, we assume this mask can be perfectly received at decoder. In practice, this can be achieved by encapsulating to the header of the main stream etc. At decoder, the received packets are regrouped into main systematic packets, WZ packets and LDPC parity packets. The Depacketization and LDPC decoding process are performed and the main systematic stream and WZ stream are obtained again. At the turbo decoder, by using received ROI mask, the error corrupted wavelet coefficients of ROI region are marked out from the main systematic stream and used as side information to help the turbo decoder to perform WZ decoding. The error-prone wavelet coefficients in the same wavelet decomposition level are quantized by the same procedure as in the encoder and play the role of channel corrupted systematic bits to help turbo decoding with received parity bits from WZ stream. After turbo decoding, all symbols for the current level are de-quantized and the ROI related wavelet coefficients are rebuilt, which later then replaced the error corrupted ROI coefficients in systematic stream. The IDWT is performed afterwards and the whole picture with better ROI reconstruction is rebuilt.

## VII. Simulation Results

In this section, the performance of WZER is demonstrated. DIRAC [20, 21], is used as wavelet video codec to generate the main systematic stream. Turbo code composed by two identical constituent convolution encoder of rate 1/2 with constraint length of 4 and with polynomial generator of (13,11) is adopted in WZ codec. The puncturing period of RCPT is set to 8, which provides various code rates of (8/9,8/10,8/11,….1/3 etc) for WZ stream. Two CIF sequences "Silent" and "Akiyo" have been tested during simulation. In DIRAC, the wavelet transform filters used are the Daubechies (9,7) filter with lifting scheme to perform fast integer wavelet transform for 4 levels, thus total number of subbands is 13. GOP size is 36 with structure of $L_1L_3L_3L_2L_3L_3L_2$ (DIRAC definition, similar as IBBPBBP structure). the quality factor is set to '7' and the function of coefficient partition is enabled with '33' format [15], with which Silent and Akiyo are compressed with bit rate of 246.5kbps and 218.1kbps respectively. The frame rates for both sequences are 13fps therefore the frame rate actually for WZ codec is only 0.3fps, since the WZER only deal with L1 frame (I frame). Other parameters for DIRAC keep default value.

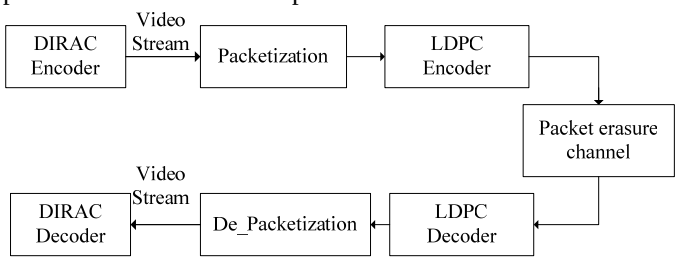



Fig. 5. Structure of algorithm using FEC to protect whole video stream

The rate of WZER can be varied and influenced by each option in each layer. In following sections, we will show the error-resilient performance of WZER with different

combinations of coding choices. For example, we can choose $F_{th}$=5 in order to protect ROI related coefficients of all 13 subbands with the LDPC coding rate of 8/14 and 8/16 for WZ stream. Otherwise, the coding rate and number of subbands need to be protected can be fixed, but varying the value $F_{th}$.

The Table I and Table II shows the performance of two WZER coding configurations compared with FEC scheme as shown in Fig. 5 (that uses LDPC codec with coding rate 8/10 or 8/12) and No protection scheme. The individual occupied bandwidth for each scheme is listed in Table I and II. Further comparisons include average PSNR performance of Y video component (PSNR-Y), and reconstructed picture quality and bandwidth usage, are revealed in following sub-sections.

TABLE I ERROR RESILIENT SCHEMES FOR AKIYO

| Scheme type | ANC | ANB | CR | FR (fps) | RBS (kbps) | ABR (kbps) | TAB (kbps) |
|---|---|---|---|---|---|---|---|
| No Protection | N/A | N/A | N/A | 13 | 218.1 | N/A | 0 |
| WZER $F_{th}$=5 | 12805 | 70428 | 8/14 | 0.3 | 15.5 | 7.8 | 23.3 |
| WZER $F_{th}$=5 | 12805 | 70428 | 8/16 | 0.3 | 20.6 | 10.3 | 30.9 |
| WZER $F_{th}$=7 | 10290 | 80185 | 8/14 | 0.3 | 11.8 | 5.9 | 17.7 |
| WZER $F_{th}$=7 | 10290 | 80185 | 8/16 | 0.3 | 15.6 | 7.8 | 23.4 |
| FEC | N/A | N/A | 8/10 | 13 | 272.7 | N/A | 54.6 |
| FEC | N/A | N/A | 8/12 | 13 | 327.1 | N/A | 109 |

TABLE II ERROR RESILIENT SCHEMES FOR SILENT

| Scheme type | ANC | ANB | CR | FR (fps) | RBS (kbps) | ABR (kbps) | TAB (kbps) |
|---|---|---|---|---|---|---|---|
| No Protection | N/A | N/A | N/A | 13 | 246.5 | N/A | 0 |
| WZER $F_{th}$=3 | 34196 | 174401 | 8/14 | 0.3 | 38.3 | 19.2 | 57.5 |
| WZER $F_{th}$=3 | 34196 | 174401 | 8/16 | 0.3 | 51 | 25.5 | 76.5 |
| FEC | N/A | N/A | 8/10 | 13 | 308.1 | N/A | 61.6 |
| FEC | N/A | N/A | 8/12 | 13 | 368.8 | N/A | 123.2 |

**ANC=A**verage **n**umber of **R**OI related coefficients protected per I frame
**ANB=A**verage **n**umber of **b**its after RDO quantization per I frame
**CR**= **C**oding **r**ate
**FR**= **F**rame **r**ate
**RBS=Resulted b**andwidth by **s**cheme.
**ABR=A**pproximate **b**andwidth caused by **LDPC** coding (coding rate 8/12) to protect WZ packets in simulation
**TAB**=**T**otal **a**dditional **b**andwidth required for error-resilience
The formulation to calculate bandwidth usage of WZ stream (namely RBS for WZ) is:

$$RBS_{WZ} = \frac{ANB * \left(\frac{1}{CR} - 1\right) * FR}{1024}$$

**N/A**=**N**ot **a**pplicable
Note: The choice of transformation layer is 13 for all proposed WZER schemes in the table

### *A. PSNR performance*

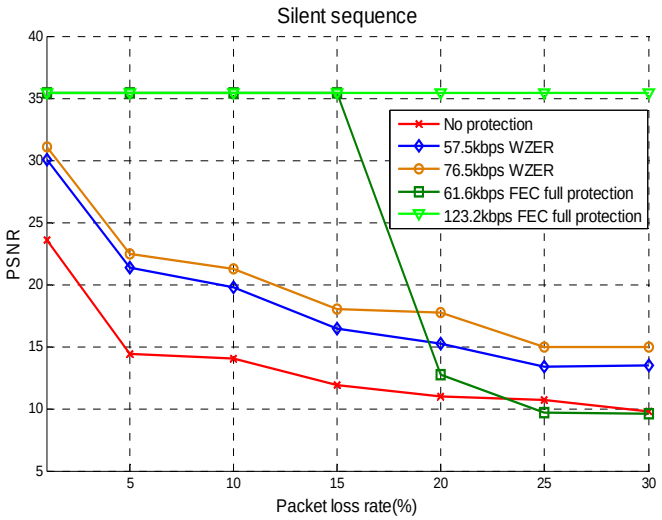
Fig. 6 and 7 show the PSNR-Y performance of "Silent" and "Akiyo" CIF sequences protected by different error resilient schemes over packet erasure channel. From the figures, it is observed that the packet drop has severely corrupted the video stream with no protection and a very low PSNR is observed generally.

The video stream protected by the full FEC with coding rate of 8/10 and 8/12 has superior PSNR performance up to 15% of packet loss. In particular, the FEC with code rate of 8/12 survived in the packet loss rate up to 30%. However, both the full FEC schemes occupy a rather high bandwidth that are 54.6kbps and 109kbps w.r.t. the code rate of 8/10 and 8/12, in the case of Silent . With 54.6kbps (half the bandwidth of 8/12), the 8/10 scheme has inferior performance than that of 8/12 and the PSNR dropped dramatically after 15% packet loss. The PSNR result is worst off at packet loss over 20% as compared to WZER scheme.

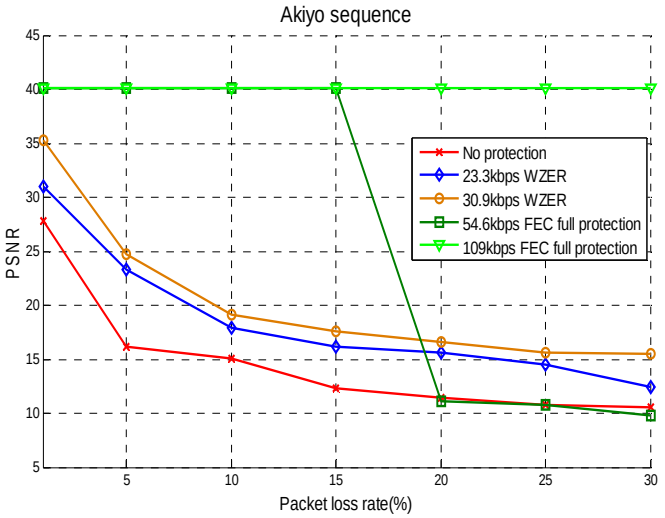
In WZER, the ROI of I frame is protected by LDPC with code rate of either 8/14 or 8/16 which delivered as WZ packets. From the figures, it is observed that the video stream with WZER delivers a better PSNR than the video stream with no protection generally. It is also observed that full FEC does not necessary delivers good performance over the packet erasure channel; for example, the FEC with 61.6kbps delivered a lower PSNR than the WZER scheme at packet loss over 20%.

Despite low PSNR at packet loss up to 20%, the quality of the reconstructed picture of WZER scheme at the decoder is still very well maintained especially in the ROI. The rationale is all down to the trade-off between bandwidth and picture quality. It is worth to mention here that low PSNR in WZER scheme does not necessary ending up with low picture quality.

The next sub-section evaluates the quality of the reconstructed picture at the decoder side.

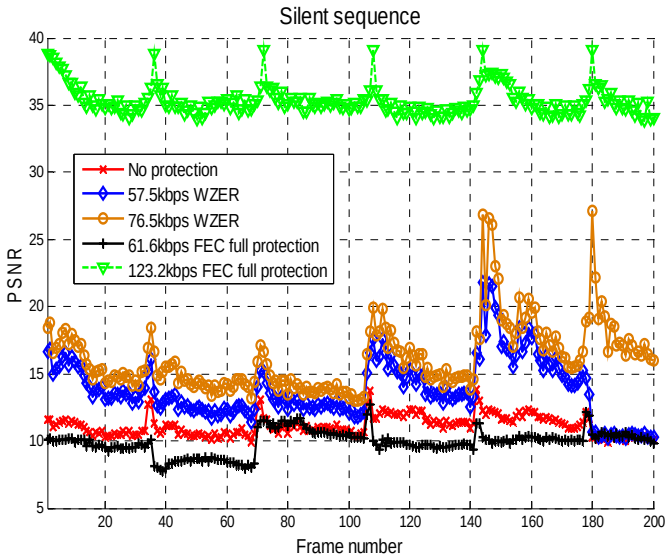



(a) Silent

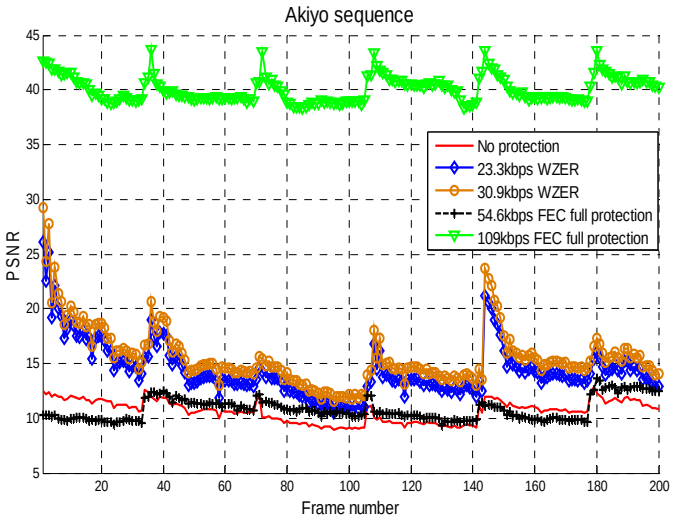



(b) Akiyo

Fig. 6. PSNR performance vs. packet loss rate (13 subbands protected)




a) Silent




b) Akiyo

Fig. 7. PSNR performance vs. frame number (Packet loss rate =25%; 13 subbands protected**)**

### *B. Picture quality*

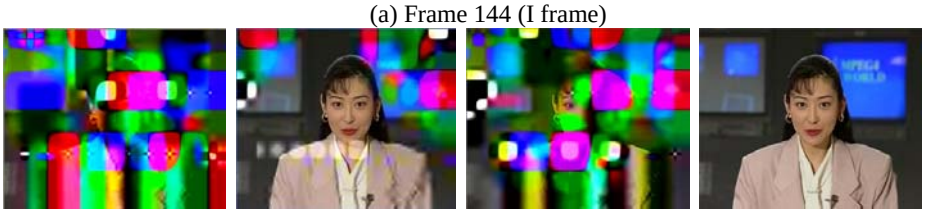
It is known that PSNR does not reflect the perceive quality of video. Fig. 8 show the picture quality of I, P and B frames of Akiyo CIF sequence. The picture quality comparison has clearly shown the advantage of WZER. On the basis of a simple assessment on the Fig. 8 (a), it can be seen that the 54.6kbps FEC unable to correct errors occurred in the frame at 25% packet loss. With FEC, many errors occurred at background are corrected but the ROI area is still not recovered. This is not acceptable for end users who actually concern the quality of ROI area much more than the rest areas. In other words, the redundant parity bits sent in FEC actually are wasting in protecting those background areas, which are not the interest of end users. In order to correct more errors occurred in ROI area, more redundant bits are need. In this case, 109kbps FEC would be enough to decode all the errors in the frame. However, larger bandwidth is inevitably required.

Conversely, the 30.9kbps WZER has fully utilized the parity bits to protect the ROI area, thus has higher quality in the ROI area than that of 54.6kbps FEC. In Fig. 8 (a), I frame that was partially protected by the 30.9kbps WZER has a comparable quality with the 109kbps FEC. Furthermore, it can be observed that not only the ROI area has significant improvement, but also the adjacent area of ROI area has also been improved to certain extent.

Even though the WZER only protects I frame, from Fig. 8 (b) and (c), it is observed that the output quality of ROI area in P and B frames are correspondingly improved following the enhancement of I frame. This proved the fact that the protection of I frame could lead to a good recovery in the GOP since that the I frame is the reference frame to construct the adjacent P and B frames. However, some errors (white dots) still can be spotted inside ROI area of P and B frames. These errors are caused by the errors of MV and residual, which are not covered by WZER thus cannot be corrected.

No proteciton 30.9kbps WZER 54.6kpbs FEC 109kbps FEC

(a) Frame 144 (I frame)



No proteciton 30.9kbps WZER 54.6kpbs FEC 109kbps FEC

(b) Frame 146( B frame)

No proteciton 30.9kbps WZER 54.6kpbs FEC 109kbps FEC

(c) Frame 150 (P frame)

Fig. 8. Picture quality of WZER and FEC with Akiyo CIF sequence
($F_{th}$=5; 13 subbands protected, Packet loss rate =25%)

### *C. Bandwidth utilization and computation complexity analysis*

In the suggested applications, the perceive picture quality is more concerned than the PSNR. The average PSNR gain can be significantly dropped because of bad quality of areas outside the ROI, which is not necessary to be protected. However, the traditional FEC algorithm cannot distinguish which part of the bitstream should be protected therefore part of bandwidth is actually wasted for the protection of unnecessary area.

As in the proposed WZER scheme, the most important area (ROI) is marked and a reasonably low bandwidth is efficiently utilized to protect it. The significant improvement in ROI area of the frame can be observed in the picture quality comparison presented in the previous section. Therefore, given the condition that end users only concern the ROI area quality in the frame, the WZER scheme actually outperform FEC scheme with less bandwidth requirement. Take the case of Akiyo sequence as an example, the 30.9kbps WZER gives a more satisfying perceive quality than the 54.6kbps FEC as shown in Fig. 8. Besides that the 30.9kbps WZER scheme only occupies around 28.3% of bandwidth. Higher $F_{th}$ could lead to further reduction in bandwidth but it will offer different error-resilience performance, which is discussed in the next sub-section.

In term of the computation complexity, WZER and FEC certainly require more computation than no protection scheme. For FEC scheme, the additional computation only originated from LDPC encoding and decoding. For WZER scheme, the ROI mask generation, RDO quantization, WZ coding and LDPC coding, are all contributed to the increase of the computation.

Although WZER needs to go through more processes and looks more complicated than FEC, the computation of WZER scheme is lower because the information that WZER dealt with is much less than that for FEC. Firstly, the WZER only dealt with I frame. Secondly, WZER only dealt with ROI part of I frame. The full FEC schemes not only need to encode the whole I frame but residual and motion vectors as well. Take the Akiyo CIF sequence as an example, for WZER with $F_{th}$=5, as seen in Table I, there are only 12805 coefficients per I frame need to dealt as compared to the full FEC scheme needs to code 352x288=101376 coefficients which is nearly ten times more than the WZER.scheme. Moreover, the RDO quantization and Turbo coding in WZER can be very fast with such a small number of coefficients to deal with. However, the computational complexity required by the WZER varies according to the value of $F_{th}$. The $F_{th}$ cannot be set too small which will significantly increases the number of coefficients to be processed and the computational complexity will be increased exponentially.

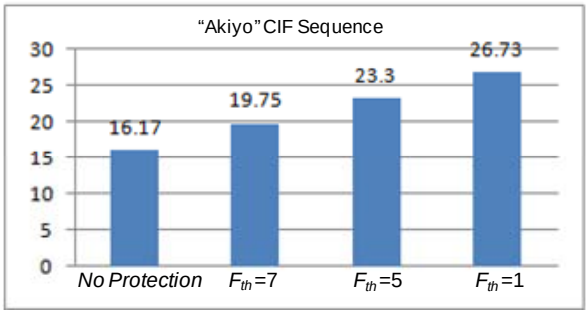
### *D. PSNR and picture quality over different $F_{th}$*

As analyzed in the previous sub-section, $F_{th}$ value directly controls the size of ROI area, which results in the different size of WZ stream eventually. Fig. 9(a) shows the average PSNR gain of different WZ stream. Smaller $F_{th}$ gives larger bit rate of the WZ stream, and hence higher PSNR gain can be expected. In Fig. 9(b), it is observed that higher $F_{th}$ results smaller ROI

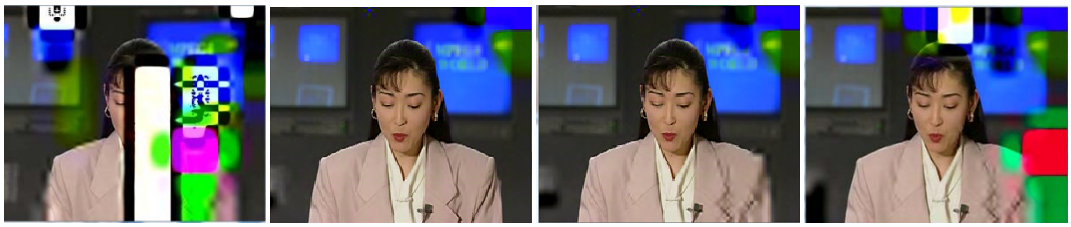
area required to be protected. In the case of $F_{th}$=1, the basic shape of Akiyo has been nicely protected. When it changes to $F_{th}$=5, some blur areas can be spotted around the right shoulder part. The protected ROI area is shrunk into only the face part of Akiyo for the case of $F_{th}$=7, hence other part of the frame can be erroneous after undergoing the packet erasure channel.

### *E. Picture quality over different number of subbands*

The number of subbands of ROI related coefficients is another key factor that can influence the size of WZ stream. More subbands (bigger ROI area) for WZER protection would require higher bit rate of WZ stream. In the proposed layered framework, the end user can choose the number of subbands that they want to receive judged by the quality resolution level they are satisfied with. For example, as shown in Fig. 10, the 7 subbands approximately can satisfy the requirement of application like video surveillance system etc. But for application in medical image etc, more than 10 subbands probably are needed. Moreover, the WZ stream is composed in the way that the lower frequency subbands (more important) are first to be sent out. This gives the algorithm advantage to combat with the channel loss. If the WZ stream is truncated during the transmission, the decoder can use the currently received subbands to partially recover ROI area with its best.




a) PSNR performance of WZER scheme with different $F_{th}$

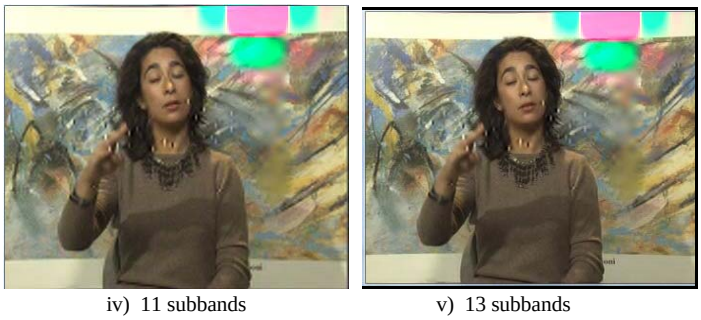


i) No protection ii) $F_{th}$=1 iii) $F_{th}$=5 iv) $F_{th}$=7

b) Picture quality of WZER scheme with different $F_{th}$

Fig. 9. PSNR performance and picture quality of WZER scheme with different $F_{th}$
(WZ Coding rate=8/14, 13 subbands protected, Packet loss rate 5%, $F_{th}$=1,5,7, Frame: 108)

i) No protection ii) 4 subbands iii) 7 subbands



iv) 11 subbands v) 13 subbands

Fig. 10. Picture quality of WZER with different number of subbands of ROI realted coefficients received (Packet loss rate=5%, WZER coding rate= 8/14 ;*Fth*=3 , the 36th Frame)

## VIII. Conclusion and Future Work

In this paper, an efficient error-resilient scheme called WZER based on a receiver driven layered WZ coding framework is proposed for wavelet video transmission. The WZER purposely protects the ROI area in the frame, which is important for some applications. The proposed WZER detects the ROI area automatically by using an automatic ROI detection, in which a parameter $F_{th}$ is used to control the sensitivity of motion detection and control the size of ROI and the number of ROI related wavelet coefficients. The ROI related coefficients is coded by WZ codec. In order to combat packet erasure channel, the generated WZ stream is further protected by packet level LDPC, where a group of parity packets are added in to help the delivery of WZ packets. The ROI area is protected in such way that the decoder able to use the recovered ROI related wavelet coefficients to replace the corresponding error-prone wavelet coefficients. The proposed WZER scheme utilizes reasonably low bandwidth to protect ROI area in the frame. The simulation results revealed that WZER is capable to deliver a satisfactory perceive picture quality under harsh packet loss channel condition despite of low bandwidth stream of WZ parity bits. Furthermore, because of the perfect reconstruction of ROI related coefficients, the adjacent areas in the frame are benefited in the recovery due to the property of IDWT in the picture reconstruction. The multiplexing way of WZ stream makes the scheme more robust in the packet erasure channel, where the decoder can recover from the partially received stream if the truncation occurred to the bitstream during transmission. Generally, the WZER receiver driven framework is suitable for multicast application, where receivers from heterogeneous group with various bandwidth availability can be satisfied. The WZER scheme has several advantages over the full FEC scheme in the aspects of bandwidth efficiency, ROI protection, computational complexity and the tradeoff between bandwidth and error-resilience etc. However, since the WZER does not protect the residual and MV of to ROI area, the quality of P and B frames will be dramatically affected especially whenwhen packet loss rate goes higher than 30%. The best WZER solution should cover the protection of the residual and MV related to ROI area in order to survive in the more error-prone channel. This remains as future work in our research.

## References


[1] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol. 17, pp. 61-82, 2000.

[2] B. Katz, S. Greenberg, N. Yarkoni, N. Blaunstien, and R. Giladi, "New error-resilient scheme based on FMO and dynamic redundant slices allocation for wireless video transmission," *IEEE Trans. Broadcasting*, vol. 53, pp. 308-319, 2007.

[3] J. Zheng and L. P. Chau, "Error-resilient coding of H.264 based on periodic macroblock," *IEEE Trans. Broadcasting*, vol. 52, pp. 223-229, 2006.

[4] J. W. Suh and Y. S. Ho, "Error concealment techniques for digital TV," *IEEE Trans. Broadcasting*, vol. 48, pp. 299-306, 2002.

[5] L. Tang, "Combined and iterative form of spatial and temporal error concealment for video signals," *IEEE Trans. Broadcasting*, vol. 52, pp. 356-361, 2006.

[6] S. C. Huang and S. Y. Kuo, "Optimization of hybridized error concealment for H.264," *IEEE Trans. Broadcasting*, vol. 54, pp. 499-516, 2008.

[7] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3462, 1993.

[8] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 6, pp. 243-250, 1996.

[9] Y. Chen and W. A. Pearlman, "Three-dimensional subband coding of video using the zero-tree method," in *Proc. of The International Society for Optical Engineering*, vol. 2727, pp. 1302-1312, 1996.

[10] B.-J. Kim and W. A. Pearlman, "Embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT)," in *Proc. of Data Compression Conference*, pp. 251-260, 1997.

[11] C. D. Creusere, "Robust image coding using the embedded zerotree wavelet algorithm," in *Proc. of Data Compression Conference*, pp.432, 1996.

[12] W.T. Tan and A. Zakhor, "Resilient compression of video for transmission over the Internet," in *Conference Record of the Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 243-247, 1998.

[13] P. Greg Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Signal Processing Letters*, vol. 4, pp. 189-191, 1997.

[14] A. A. Alatan, M. Zhao, and A. N. Akansu, "Unequal error protection of SPIHT encoded image bit streams," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 814-818, 2000.

[15] H. Man, R. L. De Queiroz, and M. J. T. Smith, "Three-dimensional subband coding techniques for wireless video communications," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 12, pp. 386-397, 2002.

[16] Z. Xiong, B.-J. Kim, and W. A. Pearlman, "Progressive video coding for noisy channels," in *IEEE Int.Conf. Image Processing*, vol. 1, pp. 334-337, 1998.

[17] [17] S. Cho and W. A. Pearlman, "A full-featured, error-resilient, scalable wavelet video codec based on the set partitioning in hierarchical trees (SPIHT) algorithm," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 12, pp. 157-171, 2002.

[18] S. Cho and W. A. Pearlman, "Error resilient video coding with improved 3-D SPIHT and error concealment," in *Proc. of The International Society for Optical Engineering*, pp. 125-136, 2003.

[19] S. Cho and W. A. Pearlman, "Multilayered protection of embedded video bitstreams over binary symmetric and packet erasure channels," *Journal of Visual Communication and Image Representation*, vol. 16, pp. 359-378, 2005.

[20] M. Tun, K. K. Loo, and J. Cosmas, "Error-resilient performance of dirac video codec over packet-erasure channel," *IEEE Trans. Broadcasting*, vol. 53, pp. 649-659, 2007.

[21] "Dirac Software," available at http://diracvideo.org/.

[22] R. Puri and K. Ramchandran, "PRISM: A "reversed" multimedia coding paradigm," in *IEEE Int.Conf. Image Processing*, pp. 617-620, 2003.

[23] R. Puri, A. Majumdar, and K. Ramchandran, "PRISM: A video coding paradigm with motion estimation at the decoder," *IEEE Trans. Image Processing*, vol. 16, pp. 2436-2448, 2007.

[24] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," in *Proc. of the IEEE*, vol. 93, pp. 71-83, 2005.

[25] A. Aaron, S. Rane, D. Rebollo-Monedero, and B. Girod, "Systematic lossy forward error protection for video waveforms," in *IEEE Int. Conf. Image Processing*, pp. 609-612, 2003.

[26] A. Aaron, S. Rane, R. Zhang, and B. Girod, "Wyner-Ziv coding for video: Applications to compression and error resilience," in *Proc. of IEEE Data Compression Conference*, pp. 93-102, 2003.

[27] S. Rane, A. Aaron, and B. Girod, "Systematic lossy forward error protection for error-resilient digital video broadcasting - A Wyner-Ziv

coding approach," in *Proc. of Int.Conf. Image Processing*, ICIP, 2004, pp. 3101-3104.
[28] S. Rane and B. Girod, "Systematic lossy error protection versus layered coding with unequal error protection," in *Proc. of The International Society for Optical Engineering*, 2005, pp. 663-671.
[29] S. Rane and B. Girod, "Systematic lossy error protection of video based on H.264/AVC redundant slices," in *Proc. of The International Society for Optical Engineering*, 2006.
[30] P. Baccichet, S. Rane, and B. Girod, "Systematic lossy error protection based on H.264/AVC redundant slices and flexible macroblock ordering," *Journal of Zhejiang University: Science*, vol. 7, pp. 900-909, 2006.
[31] S. Rane, A. Aaron, and B. Girod, "Error-resilient video transmission using multiple embedded Wyner-Ziv descriptions," in *Proc. of Int. Conf. Image Processing*, pp. 666-669, 2005.
[32] A. Sehgal, A. Jagmohan, and N. Ahuja, "Wyner-Ziv Coding of Video: An Error-Resilient Compression Framework," *IEEE Trans. Multimedia*, vol. 6, pp. 249-258, 2004.
[33] M. Wu, D. Song, and C. W. Chen, "Multiple description distributed image coding with side information for mobile wireless transmission," in *Proc. of The International Society for Optical Engineering*, pp. 572-582, 2005.
[34] Q. Xu, V. Stankovic, and Z. Xiong, "Layered Wyner-Ziv video coding for transmission over unreliable channels," *Signal Processing*, vol. 86, pp. 3212-3225, 2006.
[35] O. Crave, C. Guillemot, B. Pesquet-Popescu, and C. Tillier, "Distributed temporal multiple description coding for robust video transmission," *Eurasip Journal on Wireless Communications and Networking*, , 2008.
[36] Z. Xue, K. K. Loo, and J. Cosmas, "Bandwidth efficient error resilience scheme for wavelet based video transmission," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2008*, 2008.
[37] D. Nister and C. Christopoulos, "Lossless region of interest coding," *Signal Processing*, vol. 78, pp. 1-17, 1999.
[38] D. N. Rowitch and L. B. Milstein, "On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo (RCPT) codes," *IEEE Trans. Communications*, vol. 48, pp. 948-959, 2000.
[39] M. N. Gurcan, Yardimci, Yasemin, Cetin, A.Enis, "Influence function based Gaussianity tests for detection of microcalcifications in mammogram images," in *IEEE Int. Conf. Image Processing*, vol. 3, pp. 407-411, 1999.
[40] S. Y. Chien, Y. W. Huang, B. Y. Hsieh, S. Y. Ma, and L. G. Chen, "Fast video segmentation algorithm with shadow cancellation, global motion compensation, and adaptive threshold techniques," *IEEE Trans. Multimedia*, vol. 6, pp. 732-748, 2004.
[41] E. Atsumi and N. Farvardin, "Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees," in *IEEE Int. Conf. Image Processing*, pp.87-91, 1998.
[42] C. Christopoulos, J. Askelof, and M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard," *IEEE Signal Processing Letters*, vol. 7, pp. 247-249, 2000.
[43] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Information Theory*, vol. IT-19, pp. 471-480, 1973.
[44] A. D. Wyner and J. Ziv, "Rate-Distortion function for source coding with side information at the decoder," *IEEE Trans. Information Theory*, vol. IT-22, pp. 1-10, 1976.
[45] S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven layered multicast," *Computer Communication Review*, vol. 26, pp. 117-130, 1996.
[46] T. Zhang and Y. Xu, "Unequal packet loss protection for layered video transmission," *IEEE Trans. Broadcasting*, vol. 45, pp. 243-252, 1999.
[47] P. A. Chou, A. E. Mohr, A. Wang, and S. Mehrotra, "Error control for receiver-driven layered multicast of audio and video," *IEEE Trans. Multimedia*, vol. 3, pp. 108-122, 2001.
[48] J. Wang and K. Ramchandran, "Receiver-driven multicast over wireless with distributed source coding and FEC," in *Proc. of the 2007 International Wireless Communications and Mobile Computing Conference*, pp. 600-605, 2007.
[49] C. P. Ho and C. J. Tsai, "Content-adaptive packetization and streaming of wavelet video over IP networks," *Eurasip Journal on Image and Video Processing*, vol. 2007, 2007.
[50] M. Wu, S. S. Karande, and H. Radha, "Network-embedded FEC for optimum throughput of multicast packet video," *Signal Processing: Image Communication*, vol. 20, pp. 728-742, 2005.
[51] J. Kim, R. M. Mersereau, and Y. Altunbasak, "Error-resilient image and video transmission over the Internet using unequal error protection," *IEEE Trans. Image Processing*, vol. 12, pp. 121-131, 2003.
[52] "LDPC-Staircase code," Available at: http://planete-bcast.inrialpes.fr/