An Iterative Detection Aided Unequal Error Protection Wavelet Video Scheme Using Irregular Convolutional Codes

A. Q. Pham, J. Wang, L. L. Yang, and L. Hanzo
School of ECS, University of Southampton, SO17 1BJ, UK
Tel: +44-23-8059 3125, Fax: +44-23-8059 4508
Email:{apq03r,jw02r,lly,lh}@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract—A wavelet-based videophone scheme proposed, where the video bits are Unequal Error Protection (UEP) using Irregular Convolutional Codes (IRCCs). The proposed system uses Adaptive Arithmetic Coding (AAC) for encoding the motion vectors and individual wavelet subband coefficients. The turbo-equalized IRCC-aided videophone scheme is capable of attaining a near unimpaired video quality for channel Signal-to-Noise Ratios (SNRs) in excess of about 4.5dB over a five-path dispersive AWGN channel.

I. MOTIVATION

Under a number of idealistic conditions Shannon’s source and channel coding separation theorem [1] has shown that source and channel coding may be carried out separately. This has resulted in primarily independent development of source and channel codecs in the design of communication systems. However, in practice, the joint design of source and channel codecs constitutes a promising design in the context of practical finite-complexity, finite-delay speed [6] and video [5] systems. A well-proven approach to joint source-channel coder design is to cascade a source coder and a channel coder, and appropriately split the total overall bitrate between the source encoder and channel encoder. Typically Unequal Error Protection (UEP) is used for protecting the more vulnerable source bits.

In our system study, we invoke the wavelet video codec, proposed by the BBC Research and Development Department, as our source coder [15]. This source codec is capable of efficiently compressing video and yet achieving a high reconstructed video quality at a low bit-rate. As most video codecs, the wavelet codec invokes motion compensation for the successive video frames and Adaptive Arithmetic Coding (AAC) [8] of the motion vectors as well as of subbands’ data for the sake of achieving a high compression efficiency. However, the employment of AAC renders the compressed video bitstream sensitive to channel errors since the predictive coding inflicts error propagation upon the future frames as a share-ware design alternative to the Motion Picture’s Expert Group’s MPEG4 codec [9], which uses quite different design principles.

The allocation of UEP-based channel coding to the video bits is typically based on the codec’s bit-sensitivity analysis [5][6]. The family of Irregular Convolutional Codes (IRCCs) proposed by Tuchler and Hagenauer [11] has an innate ability to support UEP. This UEP property will be exploited in our proposed system for protecting the wavelet-compressed video stream, and powerful iterative detection will be used [7].

II. SOURCE CODING

A. NOTE

The basic philosophy of two-dimensional subband or wavelet-based video coding is that the video frame or the Motion Compensated Error Residual (MCER) frame is decomposed into a number of subbands containing the different-frequency representations of the original video or MCER frame. The decomposition to subbands is typically carried out in a number of consecutive steps, as outlined for example for a 10-subband video codec on page 629 of [5]. Briefly,
the first decomposition step seen in Fig 1 splits the frame into four 'quadrants', representing the High (H) and Low (L) frequency horizontal as well as vertical direction details in the frame. Hence the quadrant HH1 contains the high-frequency horizontal and vertical information found in the original full-band frame. By contrast, HL1 and LH1 represent the frame, where the high horizontal but low vertical and the low horizontal but high vertical frequency components have been retained. All these frame-quadrants appear to have quarter of the original frame’s size, since before the subband halving operation the video signal was decimated by a factor of two. As seen in Fig 1, the subjectively most important LL1 video segment is split further in two consecutive band-halving steps. This tree-structure based frequency band splitting allows the designer to selectively allocate the source encoding bits to those particular frequency bands, where they have the most beneficial subjective video quality improvements. The video codec used Daubechies’ so-called 9/7 biorthogonal filters in [4], where 9 and 7 denote the number of coefficients in high-pass and low-pass filters.

As in most existing video compression standards, each video frame is encoded in either intra-frame mode, where the original video frame is directly encoded, or in inter-frame mode, where the MCER generated with reference to the previous decoded frame is encoded. Fig 2 portrays the simplified block diagram of the wavelet video decoder.

In the **intra-frame mode**, the video frame is Discrete Wavelet Transformed (DWT) [16] and then quantised in the block diagram of the wavelet video decoder.

![Decoded Video Encoder](image)

**Fig. 2.** Block diagram of the Wavelet Video Encoder.

The bitstream of all video frames begins with a video frame header, as seen in Fig 3. The header is followed by the MV header, by the 30 subband DWT headers, MVs, and 30 subband DWT coefficients, as seen in Fig 3 (10 subbands for each of the Y, U, V components). The frame header commences with the current frame’s index (32 bits), followed by a number of other control parameters concerning whether or not a frame has to be skipped, the encoding mode (Intra, or Inter-frame), the number of reference frames used in MC, and the reference frame indices, etc. This is the most error-sensitive segment of the encoded video, since the entire video frame will be dropped, if the header is received in error.

The MV header and subband header information have the next level of importance, since motion compensation and/or wavelet decomposition cannot be performed without it. The MV data and the DWT coefficients transmitted at the end of the frame are the least error sensitive. The error sensitivity of these bits will be discussed in Section III.

**III. BIT SENSITIVITY STUDY**

In order to provide robust source-matched error protection for the wavelet-based video stream, it was subjected to bit sensitivity investigations by systematically corrupting all of

<table>
<thead>
<tr>
<th>PH</th>
<th>MH</th>
<th>$S_1YH$</th>
<th>$S_2YH$</th>
<th>MV</th>
<th>$S_1YD$</th>
<th>$S_2YD$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PH</strong> Frame header</td>
<td><strong>MH</strong> Motion vector header</td>
<td><strong>$S_1YH$</strong> Subband1’s header for component Y (DC Subband)</td>
<td><strong>$S_2YH$</strong> Subband1’s header for component V</td>
<td><strong>MV</strong> Motion vectors</td>
<td><strong>$S_1YD$</strong> Subband1’s data for component Y</td>
<td><strong>$S_2YD$</strong> Subband1’s data for component V</td>
</tr>
</tbody>
</table>

**Fig. 3.** The structure of the encoded video frame.
its bits in the encoded video frame and then evaluating the average PSNR degradation inflicted. The simulations were carried out using a 150-frame duration, 144x176-pixel QCIF-resolution Miss America video clip. The PSNR degradations were recorded for each of the parameters under the following conditions:

1) The effects of corruption on the wavelet video header parameters were somewhat difficult to estimate, because these parameters contain vital information related to the entire video sequence, and to a particular video frame. The decoder cannot even commence its decoding operations, when those parameters are corrupted.

2) When for example the sensitivity of bit 1 of an intra-frame coded picture was investigated, this bit was consistently corrupted in each occurrence of an intra-frame coded picture, while keeping all other bits of the same frame intact.

More details on the bit allocation used within a frame are provided in Table I, where $Y$, $U$, $V$ represent the luminance and two colour difference components of the video frame. An overview of the objective Peak Signal to Noise Ratio (PSNR) degradations inflicted by the systematic bit corruption events upon the MVs and subband DWT coefficients is given in Fig 4. Observe in this figure that

1) The bit sensitivity is typically higher at the beginning of a subband and then reduces towards the end of the subband, which is because an error at the beginning of a AAC segment corrupts more bits than one near the end of it.

2) The subband’s bit sensitivity also depends on the particular subband’s frequency. More specifically at the same position of a low frequency subband, the bit sensitivity is higher and vice versa.

3) The number of subband DWT coefficients is an other factor influencing the sensitivity of a subband, since the so-called context-based AAC continuously updates based on the past received symbols for the sake of decoding the remaining part of the subband concerned consequently the rest of the entire subband data is affected. The same characteristic behaviour can also be observed for the MVs.

The results of Fig 5 and Fig 6 are also interesting to analyze. Since all video frames in a Group Of Picture (GOP) directly or indirectly use the intra-coded frame as a reference frame, the corruption of the intra-frame results in error propagation to subsequent inter-coded frames. The PSNR degradation seen in the subsequent frames is almost as high as the degradation in the corrupted inter-frame, as seen in Fig 5. Fig 6 shows that similar trends are also valid for the corruption of P-frame in GOP.

<table>
<thead>
<tr>
<th>Subband</th>
<th>Y comp</th>
<th>U comp</th>
<th>V comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC subband</td>
<td>1888</td>
<td>116</td>
<td>502</td>
</tr>
<tr>
<td>Subband 9</td>
<td>466</td>
<td>28</td>
<td>109</td>
</tr>
<tr>
<td>Subband 8</td>
<td>1049</td>
<td>145</td>
<td>230</td>
</tr>
<tr>
<td>Subband 7</td>
<td>1208</td>
<td>110</td>
<td>273</td>
</tr>
<tr>
<td>Subband 6</td>
<td>490</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td>Subband 5</td>
<td>1038</td>
<td>60</td>
<td>104</td>
</tr>
<tr>
<td>Subband 4</td>
<td>1504</td>
<td>0</td>
<td>334</td>
</tr>
<tr>
<td>Subband 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Subband 2</td>
<td>215</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Subband 1</td>
<td>302</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE I
THE NUMBER OF SUBBAND BITS IN THE FIRST FRAME OF THE MISS AMERICA QCIF VIDEO SEQUENCE USING INTRA-FRAME MODE.

Fig. 4. Average PSNR degradation due to the corruption of the Y component for the Miss America QCIF video sequence encoded at 30 frame/s using intra-frame mode.
The performance of the wavelet codec was investigated, when communicating over dispersive AWGN channels. In order to reduce the channel induced Inter-Symbol Interference (ISI), powerful turbo equalization schemes were employed. The system model is depicted in Fig. 7. At the transmitter, the input video data is first compressed by the wavelet encoder and then packetized into fixed-length packets before being fed into the channel encoder. The channel encoder consists of an outer encoder and a rate-1 precoder [18], which are separated by the interleaver Π of Fig 7. The outer code adds redundancy for the sake of providing protection for the video data, while the precoder renders the channel recursive so that reduced bit error rates [18]. On the other hand, at the receiver a turbo equalizer is invoked, which consists of an inner channel equalizer and an outer channel decoder. Both of them are capable of processing soft inputs and generating soft outputs, which are exchanged between the two components in a number of consecutive iterations for the sake of improving the attainable decoding/detection performance. At the last iteration hard decisions are made, and the resultant bit stream is depacketized and fed into the wavelet decoder for the sake of reconstructing the video sequence.

To be specific, the proposed scheme employs an IRCC scheme as the outer channel code, which was designed to match the channel characteristics as well as to provide UEP for the video data. An IRCC is constructed from a set of subcodes [10], which have different coding rates and are generated using puncturing from the same mother convolutional code. Each subcode encodes a fraction of the input video bits while maintaining the overall coding rate at a specified value. For more details on the design of IRCCs, please refer to [13], [10], where the same IRCC was used, except that the coding rate ratios of the subcodes were different. The design method proposed in [10] was used here for optimizing the IRCCs for the specific dispersive channel considered and for the wavelet video codec invoked. As our benchmark scheme, a classic maximum free-distance Non-Systematic Convolutional (NSC) code was employed as the outer channel code, which provides Equal Error Protection (EEP) for the video bits.

V. Simulation Results

The achievable system performance was evaluated for $K = 2000$ bits of video bits per packet, resulting in an interleaver length of $L = 4000$ bits at a total average IRRC code-rate of 0.5. A five-path dispersive channel

$$h[n] = 0.227\delta[n] + 0.46\delta[n - 1] + 0.688\delta[n - 2]$$
$$+ 0.46\delta[n - 3] + 0.227\delta[n - 4]$$

imposing severe ISI was selected from [14] for our investigations. Again the outer IRCC has an average coding rate of 0.5 and a constraint length of 5. For the proposed UEP scheme, the IRCC consists of a set of subcodes having coding rates of $[0.35 0.45 0.5 0.55 0.6 0.8]$, which encode $[34\% 1.3\% 4.8\% 22\% 6.5\% 31\%]$ of the input data bits, respectively. For our outer NSC benchmark scheme the...
generator polynomials of $g_0 = 1 + D + D^2 + D^4$, and $g_1 = 1 + D^3 + D^4$ were used. For both schemes, the rate-1 precoder having a generator polynomial of $g_0 = 1 + D$ was invoked, and both the channel equalizer and the channel decoder employed the Maximum A Posteriori (MAP) algorithm [17]. In all simulations, BPSK modulation and an AWGN channel were assumed.

The PSNR performances of both the UEP and EEP system are depicted in Fig. 8. It can be seen that the UEP system attains a near error-free PSNR transmission quality in excess of $E_b/N_0 = 2.1$ dB. Although at this point the overall BER is about $2 \times 10^{-3}$, the most sensitive video bits protected by the strongest subcodes are almost error-free. The EEP system reaches the same BER at an $E_b/N_0$ of about 1.7 dB, but the PSNR attained is significant worse ($< 14$ dB). It is also worth noting that for the UEP system, the residual BER becomes so low that it has a negligible effect on the PSNR performance.

![Fig. 8. Comparison of the achievable video PSNR using both EEP and UEP. All system parameters are summarized in Table II](image)

TABLE II
THE PARAMETERS OF THE WAVELET-BASED VIDEOPHONE SCHEME.

<table>
<thead>
<tr>
<th>The system parameters</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>average frame size</td>
<td>2200 bits</td>
</tr>
<tr>
<td>Packet size</td>
<td>2000 bits</td>
</tr>
<tr>
<td>Packet header size</td>
<td>14 bits</td>
</tr>
<tr>
<td>Code rate</td>
<td>1/2</td>
</tr>
<tr>
<td>Interleaver length</td>
<td>4000</td>
</tr>
<tr>
<td>No. decoder iterations</td>
<td>10</td>
</tr>
<tr>
<td>Outer NSC polynomials</td>
<td>$g_0 = 1 + D + D^2 + D^4$, $g_1 = 1 + D^3 + D^4$</td>
</tr>
<tr>
<td>Constraint length</td>
<td>5</td>
</tr>
<tr>
<td>Precoder polynomials</td>
<td>$g_0 = 1 + D$</td>
</tr>
</tbody>
</table>

VI. CONCLUSIONS AND FUTURE WORK

The BBC wavelet-based video codec invoked constitutes the state-of-the art compression schemes, and it achieves a remarkably high reconstructed video quality at a low bit-rate. However, the compressed video sequence is sensitive to channel errors. When the video frame header, the motion vector header or subband header are corrupted, the decoder loses synchronization with the encoder, and hence UEP was used. The iteratively detected and turbo-equalized IRCC-aided receiver attained an iteration gain of 2 dB over the 5-path dispersive AWGN of Equation 1 when using a 4000-bit random interleaver. Our future research will consider dispersive fading channels, multilevel modulation and improved AAC schemes.

REFERENCES