Error-Resilient Video Transcoding for Robust Internetwork Communications Using GPRS

Safak Dogan, Student Member, IEEE, Akin Cellatoglu, Student Member, IEEE, Mustafa Uyguroglu, Member, IEEE, Abdul H. Sadka and Ahmet M. Kondoz, Member, IEEE

Abstract—A novel fully comprehensive mobile video communications system is proposed in this paper. This system exploits the useful rate management features of the video transcoders and combines them with error resilience for transmissions of coded video streams over general packet radio service (GPRS) mobile-access networks. The error-resilient video transcoding operation takes place at a centralised point, referred to as a video proxy, which provides the necessary output transmission rates with the required amount of robustness. With the use of this proposed algorithm, error resilience can be added to an already compressed video stream at an intermediate stage at the edge of two or more different networks through two resilience schemes, namely the adaptive intra refresh (AIR) and feedback control signalling (FCS) methods. Both resilience tools impose an output rate increase which can also be prevented with the proposed novel technique in this paper. Thus, an error-resilient video transcoding scheme is presented to give robust video outputs at near target transmission rates that only require the same number of GPRS timeslots as the non-resilient schemes. Moreover, an ultimate robustness is also accomplished with the combination of the two resilience algorithms at the video proxy. Extensive computer simulations demonstrate the effectiveness of the proposed system.

Index Terms—Mobile video communications, error-resilient video proxy, GPRS mobile-access networks, MPEG-4 video standard, video transcoding

I. INTRODUCTION

As opposed to the conventional source-driven resilient transmissions, recent research is focusing on the addition of resilience to the video data where or whenever it is needed. Bearing this in mind, error resilience can also be introduced into an already encoded video stream at an intermediate stage. This particular stage where the addition of error resilience to the video stream takes place can simply be the video proxy at the edge of two or more networks [1], [2], as depicted in Fig. 1. The video proxy comprises a video transcoder or a set of transcoders that provides the necessary bit rate management between different networks.

Therefore, bandwidth bottleneck problems can be resolved dynamically during media transmissions rather than by signalling back to communication sources. This evidently enables faster system responses and more efficient congestion control techniques with the utilisation of the useful features of the video transcoders [3]. However, it should be noted that increased intelligence of network proxies/gateways or nodes in such a way might render the entire networking infrastructure quite fragile due to added overall networking complexity and dynamic behaviour.

S. Dogan, A. Cellatoglu, A.H. Sadka, and A.M. Kondoz are with the Centre for Communication Systems Research (CCSR), University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom (e-mails: S.Dogan, A.Cellatoglu, A.Sadka, A.Kondoz@eim.surrey.ac.uk). M. Uyguroglu is with the Advanced Technologies Research and Development Institute, Eastern Mediterranean University, Gazi Magusa, North Cyprus (e-mail: mustafa.uyguroglu@emu.edu.tr).

Fig. 1. A GPRS networking scenario with an error-resilient video proxy.

In addition to the rate management skills of video transcoders, a further need for the error-resilient handling of the transcoded video stream may arise over mobile-access networks, such as GPRS. The nature of the GPRS channels imposes quite bursty error characteristics causing deep fades of the signal strength caused mainly by the co-channel interference and the multipath effects. Due to this fact, the video transmission will greatly be affected over the GPRS channels resulting in perturbed images with significantly reduced quality of service (QoS) levels. Thus, during the access via GPRS, video proxies will play an important role not only matching the transmission rates to the user requirements, but also providing the necessary protection for the transcoded video streams prior to their transmissions.

The proxy interconnects a relatively low bit-error-rate (BER) and high bandwidth network, such as the integrated services digital network (ISDN) and/or the public-switched telephone network (PSTN), to a relatively high BER and low bandwidth network, like the mobile-wireless network, as illustrated in Fig. 1. The output bit rate from the proxy can be adjusted by monitoring the occupancy of frame
buffers within the network monitoring module situated at the end of the video transcoding block. The state of these buffers varies according to the channel bandwidth conditions. The amount of resilience added to the video data can also be controlled by monitoring the proxy output rate and the change in error conditions of the network. This is accomplished by the means of feedback signalling, as also shown in Fig. 1.

By moving the error resilience support from the source encoder to the video proxy, a more rapid and dynamic way of error-handling at the edge of different networks is achieved. This paper focuses on the combination of two particular resilience schemes, namely the AIR and FCS methods, whilst preserving the transmission rate management features of the video transcoders. In this way, the destructive error effects of GPRS on the transcoded video streams are believed to be alleviated with the added resilience. This is due to the fact that both error resilience tools aim at the provision of prevention mechanisms against temporal error propagation effects caused by error-prone transmissions over GPRS. Thus, the primary objectives of such a scheme are envisaged as to increase the robustness of transcoded streams to transmission errors of mobile channels whilst meeting the bandwidth requirements of such networks, user preferences and client-device capabilities.

The rest of the paper is organised as follows. Section II gives a brief introductory background on video transcoding and the two resilience techniques used. An overview of the GPRS networks is presented in Section III. Section IV describes the resilient video transcoding architecture and Section V demonstrates the experiments and computer simulation results. Section VI presents further discussions of the simulation results. Finally, Section VII concludes the paper.

II. ERROR-RESILIENT VIDEO TRANSCODING

A. Video Transcoding Background

The frequent variations in the network conditions and constraints, such as the congestion characteristics, forced the necessary adaptations to these changes to take place dynamically at a centralised point at the edge of two or more networks. This specific location is referred to as a video proxy, as depicted in Fig. 1. Such a device enables faster network responses whilst maintaining the user video encoders and decoders free of unnecessary complexities normally incurred by the scalability features [4]. Moreover, a video proxy facilitates a seamless and transparent interconnection of various heterogeneous networks. A video proxy can consist of a single or a group of video transcoders operating together to establish such interconnectivity [3], [5].

Video transcoding is a method which makes the interoperability of different multimedia networks possible. Therefore, the objective of video transcoding consists of changing the format, size, transmission rate and/or syntax of an incoming compressed video stream without fully decoding and re-encoding the video information. Thus, a high transfer rate, high resolution compressed video stream can be converted into lower rates and resolutions whilst also complying with the syntax requirements. As a result, the complexity, processing power and the delay incurred by this process are minimised whilst achieving improved QoS levels [3], [6]-[12].

B. Resilience Tools

In this paper, error resilience is provided by both the AIR and FCS methods. AIR is a method whereby the error propagation within a video stream is prevented temporally by the use of a pre-determined number of intra (I) refresh macroblocks (MBs). The scheme works in an adaptive way to enhance and protect the visual quality of fast motion portions of a video stream. The definition and the detailed operation of AIR are discussed in Annex-E.1.5 of the MPEG-4 visual standard [13], [14]. On the other hand, the FCS algorithm is an adoption of Annex-N: reference picture selection mode of the H.263+ standard which relies on a back channel signal from the decoder to inform the source coder of the lost or the properly delivered video frames [15], [16]. Thus, this particular feedback signal helps the transmitter adapt its encoding scheme according to the varying channel conditions and/or constraints. In this way, the reference picture selection and the long-term prediction operations are accomplished by the source encoder.

In most cases whereby a video stream is susceptible to transmission errors, re-synchronisation of the end-decoder with the received video data is a significant operation to achieve an acceptable level of quality. Maintaining synchronisation is typically performed with the help of re-synchronisation words in a video stream. In this research work, this particular accomplishment was also inevitable at the very end-receivers for a successful decoding operation as the source coding MPEG-4 simulation software was operated without the use of any error resilience options [17]. This is due to the fact that the aim of the proposed transcoding algorithm here is to insert the necessary amount of resilience with the most adequate method at an intermediate stage during the GPRS transmission of a compressed video stream. Thus, such an operation allows the video source to be free of the extra burdens imposed by the resilient source coding techniques. Moreover, the choice of the two resilience tools retains compatibility with standard MPEG-4 decoders, which is an imperative feature of a transcoder.

III. OVERVIEW OF GPRS SYSTEMS

GPRS [18] is a new non-voice value added service that allows information to be sent and received across a mobile telephone network. It is an end-to-end mobile packet communication system which makes use of the same radio architecture as global system for mobile (GSM) communications [18], [19]. GPRS is also the name for an international packet-switched networking standard in GSM systems, initiated and developed by the European Telecommunication Standards Institute (ETSI).
GPRS involves overlaying a packet-based air interface on the existing circuit-switched GSM network. This gives the user an option to use a packet-oriented data service. A new set of logical channels has been defined for GPRS traffic as opposed to the circuit-switched networks where all the signalling and information transfers make use of one channel only. This set includes control channels and packet data traffic channels. A physical channel allocated for GPRS traffic is called a packet data channel (PDCH). The PDCH consists of a multi-frame pattern that runs on timeslots assigned to GPRS [20], [21]. Thus, the GPRS data is transmitted over the PDCH and is protected by four different channel protection schemes: CS1, CS2, CS3 and CS4 [22]. The channel coding is used to protect the transmitted data packets against transmission errors. CS1-3 use convolutional codes and block check sequences of varying strengths, so as to produce different rates. CS1-3 are based on a 1/2 rate convolutional codes, which is punctured to obtain approximate rates 1/2, 2/3 and 3/4, respectively. On the other hand, CS4 is uncoded whereby it only provides error detection functionality [20], [23]. Each of the four channel protection schemes is assigned a maximum of eight timeslots [18], [24]. The coding schemes and resulting bit rates per one timeslot are described in Table I.

<table>
<thead>
<tr>
<th>Coding Scheme</th>
<th>Convolutional Code Rate</th>
<th>Payload per Block [bits]</th>
<th>User Bit Rate [kbit/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1</td>
<td>1/2</td>
<td>181</td>
<td>9.05</td>
</tr>
<tr>
<td>CS2</td>
<td>~2/3</td>
<td>268</td>
<td>13.4</td>
</tr>
<tr>
<td>CS3</td>
<td>~3/4</td>
<td>312</td>
<td>15.6</td>
</tr>
<tr>
<td>CS4</td>
<td>1</td>
<td>428</td>
<td>21.4</td>
</tr>
</tbody>
</table>

The choice of one of the four coding schemes for the coding of PDCHs depends on the quality of the channel. Under very bad conditions, a very reliable CS1 may be used and a data rate of 9.05 kbit/s per GPRS timeslot can be obtained. Under good channel conditions, data can be transmitted without convolutional coding and a transport rate of 21.4 kbit/s per timeslot can be achieved. Hence, with the use of eight slots of this channel protection scheme, namely CS4, a maximum data rate of 171.2 kbit/s can be obtained in theory. This is significantly faster than the data transmission speeds possible over today’s fixed telecommunication networks and the current circuit-switched data services on GSM networks. Thus, GPRS promises to fully enable the use of new applications on the move with the increased communication speeds. However, in practice, multiple users share the timeslots, and hence a much lower bit rate is available to an individual user [25], [26].

### IV. ERROR-RESILIENT VIDEO TRANSCODER ARCHITECTURE

In this paper, the video transcoding has further been exploited to add error resilience to the transcoded data in addition to the rate management characteristics. For this purpose, the transcoding system has been modified, as illustrated in Fig. 2. Referring to this figure, the video transcoder reduces the incoming bit rate whilst adding resilience to the transcoded video data simultaneously. The rate reduction algorithm provides drift-free transcoding qualities with refined motion vectors (MVs) [3], [7], [27], [28]. Furthermore, the increase in the output rate due to the addition of resilience is compensated for using an adaptive transcoding operation. The overall resilience is provided with the use of AIR and FCS algorithms, details of which were discussed in Section II. Both AIR and FCS can work independently as well as together in combined harmony depending on the choice of “error resilience decision block” which reflects the necessary action required against the varying channel conditions, as indicated by the relevant feedback signal. Since both the AIR and FCS methods increase the overall transmission rate, the video transcoder adaptively transforms the bit rate as required by the congested or bandwidth-limited network(s). The rate regulation is simply carried out by the adaptation of the quantisation parameter (QP) to the newly required conditions. During transcoding, an increase in QP results in a bit rate reduction whilst a decrease gives faster transcoder output rates.

Fig. 2. The error-resilient video transcoder architecture.

Adaptive operation of the video transcoder is maintained by two primary feedback control mechanisms:

1. The first control system comprises feedback signals which contain up-to-date information directly related to the output channel conditions, such as BER, carrier-to-interference (C/I) ratio, delay, lost/received video frames, etc. Relying on the received feedback data, one or both of the two error resilience schemes, namely the AIR and FCS blocks of Fig. 2, make(s) an attempt to insert the necessary robustness to the transcoded data within the drift-correction loop, which constitutes the core transcoding mechanism. The decision of which resilience block(s) to be employed is dynamically accomplished by the received control feedback data, comprising transmission channel characteristics. This decision is a logical operation conducted by the resilience decision block which relies on the back channel data reporting the status of the destination network. Such particular information is gathered at the network monitoring module prior to its conveyance back to the two resilience and the decision blocks. With or without the use of error robustness algorithms with respect to the varying channel
conditions, transcoding is performed via customary drift correction and MV refinement operations.

(i) For increased BER (decreased C/I) conditions, the AIR block acts as the major resilience tool to stop the potential error accumulation effects resulting from transmission errors. This particular operation of the video transcoder regulates the output bit rate whilst also introducing improved robustness to the video stream particularly for high motion areas [13]. Since high motion areas are more susceptible to channel bit errors, these particular portions of the video stream are transcoded to I-MBs rather than inter predictive (P) MBs. I-MBs hence do not require motion compensation, and therefore a potential error accumulation is prevented with added resilience. In addition to processing the high motion data in an error-resilient way, the transcoder also encodes these particular portions of the video sequence with an increased number of I-MBs whilst compensating for the resulting increased bit rates. The compensation for the increase in bit rate is performed by increasing the value of QP.

(ii) On the other hand, for entirely lost video frames during error-prone transmissions, the video transcoder is designed in a fashion to receive any kind of transmission feedback signal, such as an acknowledgement (ACK), non-acknowledgement (NACK) or both, from the end-receiver. Depending on the received return signal from the end-user with its associated latency, the video transcoder adapts its transcoding scheme according to the reported channel conditions. According to the feedback signal obtained from the receiver end, the video transcoder can judge which video frames are not correctly received and/or lost during transmission. Consequently, the currently transcoded frame is predicted using the last acknowledged stored video frame in the transcoder buffer [15]. Thus, a certain degree of error resilience is inserted by referring to the most recent error-free video frame in the transcoder buffer, hence resulting in a better QoS. The addition of robustness is accompanied by the regulation of the increased transmission rate due to the FCN algorithm. The error propagation effects can be minimised at a much earlier point at the edge of different networks rather than waiting for the ACK/NACK messages to arrive at the source end. Moreover, this kind of a video transcoder operation can also produce the necessary robust output to counteract the detrimental impacts of video frame drops resulting from network congestions.

(iii) Lastly, for extreme channel conditions whereby not only do high BERs (low C/I) persist, but also full frame losses exist, then the combined AIR-FCS operation is performed as a result of the error resilience decision block. Consequently, the significant effects of channel bit errors coupled with severe frame losses are mitigated.

2. The second feedback control mechanism comprises adaptive rate transcoding. This scheme requires a feedback signalling method for the control of the output bit rate from the video transcoder, as shown in Fig. 2. The feedback signal is originated from the output video frame buffer within the network monitoring module which constantly monitors the flow conditions. In case of an underflow, it returns a signal to the transcoder seeking an increase in the output rate. On the other hand, the rate reduction is flagged back to the transcoder in case of an overflow. Thus, a straightforward rate controlling scheme is established for a congestion control or a bandwidth bottleneck resolution with the use of variable quantisation.

V. COMPUTER SIMULATIONS AND ANALYSIS OF RESULTS

In this section, the proposed error-resilient video transcoding algorithm is tested with three different experiments. Prior to the further discussion of each test however, a brief description of the simulations and test set-up, which is common for the whole three test models, is given herein. The test sequences chosen for the simulations were encoded, transcoded and decoded in compliance with the MPEG-4 standard with the use of the unrestricted MVs and the advance prediction modes. The frame rates, frame sizes and the operation modes were set to 25 fr/s, quarter common intermediate format (QCIF: 176×144 pixels) and I-P-P-P-... layout for the video clips, respectively. Each set of experiments is accompanied by both objective and subjective test results. The objective measurements indicate a quality performance averaged over the results of 10 different simulations run with 10 different random seeds. The remaining simulation parameters, which are specific to individual experiments, are separately described in the following sub-sections.

A. Transcoding with AIR over GPRS

1) Experiments and Results

The robust transcoding performance was tested over a GPRS channel simulator which was genuinely designed and implemented within the Centre for Communication Systems Research (CCSR). In terms of error effects, the characterisation of a GPRS channel is modelled as a bursty error-prone transmission environment where fairly big chunks of the transmitted data become highly susceptible to the detrimental error impacts [22], [29]. This kind of errors corrupts the conveyed information more significantly than random error effects as far as QoS is concerned. This impact particularly destroys the video communication data since even a single bit error, in the form of a bit loss or an inversion, leads to a serious synchronisation problem or a rapidly increasing and spreading error propagation within the transmitted video sequence. Thus, the error propagation has to be stopped and the synchronisation has to be resumed during the transmission of the video data.

In this sub-section, two different 200-frame video sequences were tested over the GPRS channel model. The two test sequences were deliberately chosen to comprise two different motion activity natures: “Mother & Daughter” and “Foreman” with moderate and high activity scenes, respectively. The original bit rate of the “Mother & Daughter” sequence prior to the transcoding operation was 70.553 kbit/s on average, giving an average PSNR level of 36.047 dB. This sequence was later transcoded down to an average rate of 25.818 kbit/s with a PSNR level of 32.683 dB. Similarly, the “Foreman” sequence was transcoded...
from an average rate of 87.403 kbit/s with a PSNR level of 33.582 dB down to 46.835 kbit/s on average with a PSNR level of 30.029 dB. Thus, the rate reductions applied on “Mother & Daughter” and “Foreman” were 63.5% and 46.5%, respectively. Moreover, the MV refinement window sizes were set to ±2 pixels and ±5 pixels for “Mother & Daughter” and “Foreman”, respectively.

### Table II

<table>
<thead>
<tr>
<th>Timeslots</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1</td>
<td>6.8</td>
<td>13.6</td>
<td>20.4</td>
<td>27.2</td>
<td>34.0</td>
<td>40.8</td>
<td>47.6</td>
<td>54.4</td>
</tr>
<tr>
<td>CS2</td>
<td>10.5</td>
<td>21.0</td>
<td>31.5</td>
<td>42.0</td>
<td>52.5</td>
<td>63.0</td>
<td>73.5</td>
<td>84.0</td>
</tr>
<tr>
<td>CS3</td>
<td>12.2</td>
<td>24.4</td>
<td>36.6</td>
<td>48.8</td>
<td>61.0</td>
<td>73.2</td>
<td>85.4</td>
<td>97.6</td>
</tr>
<tr>
<td>CS4</td>
<td>17.2</td>
<td>34.4</td>
<td>51.6</td>
<td>68.8</td>
<td>86.0</td>
<td>103.2</td>
<td>120.4</td>
<td>137.6</td>
</tr>
</tbody>
</table>

The bit rate reductions were essential to enable the video streams to transport over the GPRS channels in such a typical video communication scenario, as depicted in Fig. 1. As the last column of Table I clearly indicates, the amount of user data for the transport over GPRS is strictly limited depending on the selected channel protection scheme. However, the timeslotting feature of GPRS can overcome this kind of limitation to some extent. Nevertheless, despite the multi-slotting feature, GPRS rates are still far too low for video communications if frame droppings are not employed. Therefore, a successful error-resilient video transcoding for transmission rate reduction is necessary prior to the GPRS network transport. Multiple slots can be used to further increase the user bit rate as multiples of the base transmission rate, as depicted in Table II. In this table, the first column is illustrated with a shaded pattern as to describe that following slots are multiples of the data rates given in this first column. Although this particular table seems to indicate different user data rates for different channel protection schemes from the figures given in Table I, there is indeed not any kind of mismatches between these particular two tables. This is only due to the fact that actual raw application level user rates for the user applications are given in Table II whereas Table I also comprises the added overheads on the physical link level. Naturally, Table I user rates are slightly higher than those of Table II. However, it has to be denoted that the raw data rates presented in Table II were obtained from a series of video transmissions over GPRS with various test sequences; they do not constitute a part of the GPRS standard. During the GPRS simulations presented in this paper, this particular table, namely Table II, guided the selection of the transcoded raw user video rates and channel protection schemes from the figures given in Table I, respectively.

### Table III

Average PSNR and Bit Rate Values Against Different C/I Ratios

<table>
<thead>
<tr>
<th>Scheme</th>
<th>C/I=7</th>
<th>C/I=9</th>
<th>C/I=12</th>
<th>C/I=15</th>
<th>C/I=18</th>
<th>Rate [kbit/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>200-frame “Mother &amp; Daughter”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>err-free</td>
<td>32.683</td>
<td>25.818</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>err-prn</td>
<td>19.781</td>
<td>24.106</td>
<td>30.550</td>
<td>32.683</td>
<td>N/A</td>
<td>25.818</td>
</tr>
<tr>
<td>err-relnt</td>
<td>21.811</td>
<td>27.256</td>
<td>31.538</td>
<td>32.683</td>
<td>N/A</td>
<td>27.000</td>
</tr>
<tr>
<td>CS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>err-prn</td>
<td>15.595</td>
<td>18.036</td>
<td>22.280</td>
<td>28.389</td>
<td>31.742</td>
<td>25.818</td>
</tr>
<tr>
<td>err-relnt</td>
<td>16.449</td>
<td>19.154</td>
<td>25.302</td>
<td>30.597</td>
<td>31.980</td>
<td>27.000</td>
</tr>
<tr>
<td>CS3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AIR is provided in these simulations as the major error resilience tool on the transcoded video streams. Thus, all the simulations were initiated with a pre-determined number of I-MBs which was set to be a maximum of 3 MBs per frame. However, it should also be indicated that the number of intra (I) refresh MBs vary with the motion activity and the output transcoded transmission rate variations in an adaptive way, details of which were discussed in the preceding section.

![Graph](image-url)
200-frame “Foreman”, MV refinement window size: 25 pixels

<table>
<thead>
<tr>
<th>C/I [dB]</th>
<th>Scheme BER</th>
<th>Scheme BER</th>
<th>Scheme BER</th>
<th>Scheme BER</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>S 2.704e-3</td>
<td>CS 2.047e-2</td>
<td>C</td>
<td>5.144e-2</td>
</tr>
<tr>
<td>12</td>
<td>S 2.956e-4</td>
<td>S 3.323e-3</td>
<td></td>
<td>1.360e-2</td>
</tr>
<tr>
<td>15</td>
<td>I 0.00000</td>
<td>2 4.440e-4</td>
<td></td>
<td>1.968e-3</td>
</tr>
<tr>
<td>18</td>
<td>I 0.00000</td>
<td>2 4.206e-4</td>
<td></td>
<td>1.968e-3</td>
</tr>
</tbody>
</table>

Fig. 4. Subjective results of the 200th frames of “Foreman” for a particular seed at C/I = 12 dB. (a) Error-free direct enc/dec at high rate. (b) CS1, non-resilient error-prone. (c) CS2, non-resilient error-prone. (d) CS3, non-resilient error-prone. (e) Error-free. (f) CS1, error-resilient. (g) CS2, error-resilient. (h) CS3, error-resilient sequences transcoded down to the lower rate.

Simulation results are depicted in Figs. 3-4 for both objective and subjective comparisons. All the results presented in this sub-section comprise the simulations using three different channel protection schemes as the fourth scheme (CS4) is not practically feasible for video applications [30]. Thus, the results demonstrate the non-resilient error-prone and error-resilient transcoding applications along with the results of the error-free sequences for comparative referencing purposes.

Table IV demonstrates that the error-resilient “Mother & Daughter” sequence performed slightly better than the error-resilient “Foreman” sequence for all the three CS conditions. This outcome implies that the high motion activity of “Foreman” might have imposed a limitation over the performance improvement especially during the significantly perturbed transmission conditions. As demonstrated, the degradation in quality is quite distinguishable since the distortion effects of bursty errors are more critical to the error-sensitive video data. Particularly, the objective video qualities of “Mother & Daughter” with CS3 and “Foreman” with CS2 and CS3 at C/I = 7 dB (PSNR: below 15 dB) are unacceptable, as depicted in Fig. 3. At this very low C/I ratio, the sole 3-MB AIR resilience method did not perform satisfactorily for either of the test video clips. In addition, the error-resilient “Foreman” sequence also presented similar low quality results for CS2 and CS3 at C/I = 9 dB. However, this is not the case for “Mother & Daughter” at C/I = 9 dB as the error sensitivity of the high motion activity of “Foreman” has a major contribution to the QoS loss in error-prone conditions.

TABLE V

<table>
<thead>
<tr>
<th>Scheme BER</th>
<th>Scheme BER</th>
<th>Scheme BER</th>
<th>Scheme BER</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>S 2.704e-3</td>
<td>CS 2.047e-2</td>
<td>C</td>
</tr>
<tr>
<td>12</td>
<td>S 2.956e-4</td>
<td>S 3.323e-3</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>I 0.00000</td>
<td>2 4.440e-4</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>I 0.00000</td>
<td>2 4.206e-4</td>
<td></td>
</tr>
</tbody>
</table>

Thus, a combination of suitable error resilience tools is recommended at these particularly very low C/I ratios over GPRS. On the other hand, the AIR method presented quite satisfactory performance improvements at various other C/I ratios and with different CS schemes, as seen in Table V.
and Fig. 3. Naturally, for low BERs, or high C/I ratios (e.g. C/I = 18 dB), quality improvement features of the error resilience methods are limited. The experimental BERs versus different C/I ratios can be seen in Table IV. It is clear from these tables that as the C/I decreases, the BER increases. Moreover, the BER also increases for one particular C/I for different CS conditions, CS1 having the lowest BERs and CS3 bearing the highest.

Finally, Fig. 4 illustrates the GPRS channel effects on the non-resilient and error-resilient transcoded video qualities. This figure depicts the 200th frames of the “Foreman” video clip for three different CSs at C/I = 12 dB. The figure clearly shows the perceptual improvement in the video service quality performance with error-resilient transcoding. This significant improvement was achieved at near target bit rates despite the transmission rate increase incurred by the AIR method. Such rate management was accomplished by the rate reduction features of the video transcoder. Hence, the error resiliency was introduced to the compressed video streams at an intermediate level at the expense of merely 1 kbit/s and 0.1 kbit/s growths for “Mother & Daughter” and “Foreman”, respectively. The obtained near target bit rates were 27 kbit/s on average for “Mother & Daughter” and 47 kbit/s on average for “Foreman”. These particular rates allowed the former to be transmitted over 4 CS1, 3 CS2 or CS3 timeslots and the latter to be conveyed over 7 CS1, 5 CS2 or 4 CS3 timeslots via the GPRS access network.

B. Transcoding with FCS over GPRS

1) Experiments and Results

The FCS experiments were designed to simulate the effects of full frame losses and the FCS resilience operation at various ACK/NACK reception delay conditions. The different transcoded video performances were tested for the back channel signal reception times of up to 480 msec, which coincide with the duration of 12 transcoded video frames at the frame rate of 25 fr/s. The maximum delay was deliberately set to 12 frames to investigate the effects of significantly long delays of the ACK/NACK signal over a GPRS mobile-access network. This particular end-decoder-to-transcoder delay is assumed to be ~450 msec (11.25 video frames at 25 fr/s), in line with phase-1 of the initial GPRS standard [18]. Thus, the experimental set-up was built in such a way that a loss of a GPRS radio packet is reported back to the video proxy from a receiving end-terminal in 450 msec, as depicted in Fig. 5. In a real-life GPRS scenario however, the round-trip end-to-end latency may be much longer. Here, this delay refers to the time elapsed whilst waiting for an ACK or NACK to arrive back at the video proxy. Meanwhile, the transcoder keeps on processing the input video frames at the frame rate of 25 fr/s, and hence the proxy continues transmitting the transcoded video frames in GPRS radio packets. The assumption made here for the GPRS access network experiments is that one video frame fits into one GPRS radio packet prior to transmission. Therefore, the loss of a GPRS packet is directly related to the loss of a video frame for a simplified simulation model. However, on a few occasions during the tests, two consecutive video frame losses were also experienced which were assumed to fit in one GPRS radio packet.

The set of objective and subjective results for the frame loss and the remedial FCS experiments are demonstrated in Figs. 6-7 and Table VI. This particular set comprises the simulation results of the 150-frame “Suzie” and 200-frame “Foreman” video test sequences. Objective results include the average PSNR variations against the various delay conditions for the ACK/NACK reception. Table VI presents the detailed quality levels and the changes in bit rates imposed by the added resilience. The subjective results illustrate the last frames of “Foreman” with different frame delays for the resilient and non-resilient cases as well as the error-free ones provided for reference.

![Fig. 5. Feedback signal delay over GPRS.](image)

![Fig. 6. Objective results of (a) “Suzie” and (b) “Foreman” for the average PSNR variations against various feedback signalling delay times.](image)
resilient. (e) Non-resilient error-prone. (f) 4-frame delay resilient. (g) 8-frame delay resilient. (h) 12-frame delay resilient.

Fig. 7. Subjective results of the 200th frames of “Foreman”. (a) Error-free. (b) 2-frame delay resilient. (c) 6-frame delay resilient. (d) 10-frame delay resilient. (e) Non-resilient error-prone. (f) 4-frame delay resilient. (g) 8-frame delay resilient. (h) 12-frame delay resilient.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Av. Bit Rate [kbit/s]</th>
<th>Av. PSNR [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>150-frame “Suzie”</strong> MV refinement window size: 22 pixels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>direct enc/dec 4th high bit rate</td>
<td>172.685</td>
<td>40.523</td>
</tr>
<tr>
<td>transcd err-free</td>
<td>46.353</td>
<td>35.446</td>
</tr>
<tr>
<td>transcd non-resilient err-prn</td>
<td>46.353</td>
<td>32.931</td>
</tr>
<tr>
<td>transcd 2-frame delay err-rslnt</td>
<td>47.207</td>
<td>31.749</td>
</tr>
<tr>
<td>transcd 4-frame delay err-rslnt</td>
<td>47.233</td>
<td>31.781</td>
</tr>
<tr>
<td>transcd 6-frame delay err-rslnt</td>
<td>47.305</td>
<td>31.757</td>
</tr>
<tr>
<td>transcd 8-frame delay err-rslnt</td>
<td>47.521</td>
<td>31.643</td>
</tr>
<tr>
<td>transcd 10-frame delay err-rslnt</td>
<td>47.912</td>
<td>31.434</td>
</tr>
<tr>
<td>transcd 12-frame delay err-rslnt</td>
<td>48.410</td>
<td>31.715</td>
</tr>
<tr>
<td><strong>200-frame “Foreman”</strong> MV refinement window size: 25 pixels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>direct enc/dec 4th high bit rate</td>
<td>87.403</td>
<td>33.582</td>
</tr>
<tr>
<td>transcd err-free</td>
<td>46.835</td>
<td>30.029</td>
</tr>
<tr>
<td>transcd non-resilient err-prn</td>
<td>46.835</td>
<td>27.752</td>
</tr>
<tr>
<td>transcd 2-frame delay err-rslnt</td>
<td>47.168</td>
<td>28.099</td>
</tr>
<tr>
<td>transcd 4-frame delay err-rslnt</td>
<td>47.036</td>
<td>27.923</td>
</tr>
<tr>
<td>transcd 6-frame delay err-rslnt</td>
<td>46.636</td>
<td>28.034</td>
</tr>
<tr>
<td>transcd 8-frame delay err-rslnt</td>
<td>47.411</td>
<td>28.055</td>
</tr>
<tr>
<td>transcd 10-frame delay err-rslnt</td>
<td>47.117</td>
<td>27.989</td>
</tr>
<tr>
<td>transcd 12-frame delay err-rslnt</td>
<td>47.018</td>
<td>28.021</td>
</tr>
</tbody>
</table>

Furthermore, during the FCS simulations, a 5% of the transmitted video frames were randomly lost. The 5% frame loss case is a typical packet loss rate for GPRS CS2 code at C/I = 12 dB condition [30] which was also chosen as the operating point for the error-resilient video transcoding tests over GPRS.

2) Analysis of the Results

The 5% frame loss experiments presented varying quality levels with and without the FCS resilience algorithm. The effective resilience performance has been shown to rely on the motion activity of the video clip, making it sequence-dependent, as depicted in Fig. 6. Therefore, the performance improvement with the FCS scheme has been demonstrated as ~2.4 dB for the “Suzie” sequence whilst the results of the “Foreman” sequence showed quality enhancement of ~0.4 dB at most. This is due to the fact that the particular loss of the very high motion activity frames in the middle of the “Suzie” sequence caused significant quality losses, as seen in Fig. 6(a) and Table VI. Evidently, the FCS scheme performed much better in this particular case as the long-term temporal referencing with feedback signalling achieved an enhancement in the perceptual quality. However, the increase in the latency of the feedback signal decreased the degree of quality improvement, as observed from Table VI.

On the contrary, Fig. 6(b) shows that this experimental observation is valid only for “Suzie” which has an overall moderate motion activity in most video frames with the exception of a few frames in the middle of the sequence. Due to the inherent high motion and scene activity associated with “Foreman”, the quality improvement for this particular sequence with the FCS was not as significant as for the “Suzie” sequence. Nevertheless, the resilient transcoding results demonstrated better qualities than the non-resilient error-prone transcoding result, as seen in Fig. 6 and Table VI. As opposed to the “Suzie” results, “Foreman” presented a varying error-resilient transcoding performance due to the variation of the feedback signal time delay. The reason is that as the waiting time latency for the reception of the back channel signal increases, the lack of correlation between the reference and the current frames causes more intra (I) mode transcoded MBs, due to the significant amount of scene changes, which in turn increases the output rate whilst also improving the resilient transcoding quality. This is mostly perceived in the results at very high delay values, such as 12-frame delays, since the FCS scheme in this case was unable to handle the vast lack of correlations between the transcoded pictures and the reference ones. Such behaviour was observed to be sequence-dependent. The increase in the output rate is further reduced with the transcoder.

Moreover, the subjective results, as seen in Fig. 7, also depict the effects of the delay on the transcoded video quality. Generally, the results have shown that the FCS algorithm gives limited improvements on the picture quality compared to the source coding FCS resilience method. This is mainly due to the fact that small MV refinement window sizes put a limitation on the quality improvement of the motion active scenes in the error-resilient mode. Furthermore, resilience over an already reduced quality video (due to the re-quantisation process at the video transcoder) results in smaller improvements than those achieved by source coding resilience techniques.

Table VI gives detailed output rate values. These results show that in most cases, the bit rate increases as the latency for the feedback signal reception increases. This is due to the lack of correlation between the long-term reference and the current video frames. However, the rate increase can easily be managed with a straightforward adaptive rate reduction algorithm which operates at the resilient video transcoder, as presented here.

C. Transcoding with Combined AIR and FCS over GPRS

1) Experiments and Results

In these experiments, AIR was also employed for the video transcoding performance tests in addition to the FCS algorithm. This achievement was established to provide the transcoded video streams with the ultimate resilience prior to transmissions over fairly high BER GPRS networks. Hence, these particular experiments show the novel
The transcoding with combined AIR and FCS simulation results have been illustrated in Figs. 8-10 and Table VII for the 150-frame “Suzie”, 150-frame “Salesman” and 200-frame “Foreman” sequences. “Suzie”, “Salesman” and “Foreman” were transcoded from 78.118 kbit/s (37.560 dB), 89.700 kbit/s (37.383 dB) and 87.403 kbit/s (33.582 dB) down to 29.908 kbit/s (33.471 dB), 43.630 kbit/s (34.991 dB) and 46.835 kbit/s (30.029 dB), respectively. The MV refinement window sizes were pre-set as ±4, ±2 and ±5 pixels in the same order as above.

2) Analysis of the Results
The FCS results have proved that even the 12-frame delay resilience cases (480 msec at 25 fr/s) performed well above the non-resilient video communication qualities. Hence, the motivation obtained from these results led us to apply and test this particular scheme as a complementary resilience method to the AIR algorithm over the GPRS networks when frame droppings are inevitable. The results of these tests have been demonstrated in Figs. 8-10 and Table VII. It has been shown that the transcoding with combined resilience achieved superior quality levels against the non-resilient schemes over the GPRS channels with frame losses. In these figures, the corresponding performance improvements have been observed to be 4 dB, 4 dB and 2.5 dB on average for “Suzie”, “Salesman” and “Foreman”, respectively. Moreover, “Suzie” results have been presented in such a way that the quality gains of the combined method of AIR and FCS are compared against the AIR and FCS only resilience results at similar conditions, BER ≅ 3.3e-3, over the same GPRS channel at C/I = 12 dB with CS2. The results of these particular experiments, shown in Fig. 8 and Table VII, demonstrate 1 dB, 2 dB and 4 dB quality improvements in favour for the FCS only, the AIR only and the combined methods, respectively, compared to the non-resilient scheme. The associated quality improvements were achieved with the minimal output bit rate growths with the use of the rate management features of the video transcoder. The bit rate increases due to the use of the combined resilience methods were reported to be ~2.5 kbit/s, ~0.3 kbit/s and ~3 kbit/s on average for “Suzie”, “Salesman” and “Foreman”, respectively. These increases in bit rates are so little that the GPRS timeslots required for the transfers of the resilient and non-resilient data are exactly the same.

Figs. 9-10 present the subjective results obtained for the 150-frame “Suzie” and 200-frame “Foreman” sequences,
respectively. The reason for choosing the 67th frames of “Suzie” is that these particular frames show the effects of frame losses in a high motion region of the sequence. In this way, a more lucid comparison of the frame loss effects and the quality improvements obtained with the combined resilience algorithms can be demonstrated.

VI. DISCUSSIONS

The AIR transcoding performance has been tested over an error-prone GPRS channel model. These tests have shown that the GPRS error effects on the transcoded video quality were quite detrimental. This is due to the fact that the error-sensitive video data is significantly vulnerable to the loss of long bursts of visual information. Hence, interleaving of data prior to its transmission at the video proxy is believed to improve the QoS in error-prone conditions as this will randomise the burstiness of errors. The inherent GPRS channel interleaving and protection schemes, namely CS1-3, provide a certain degree of protection against transmission errors by means of convolutional coding. However, for video communications, these built-in schemes have been demonstrated to be practically inefficient at higher BER levels. The simulations have shown that as the protection schemes of the different GPRS channels got weaker, the BER increased significantly. This increase in BER at low C/I ratios, such as C/I = 7, 9 dB, notably degraded the perceptual quality of video communications. At these low C/I ratios, the resilience provided only by the AIR algorithm was not very satisfactory. This hinted at the necessity of additional protection/resilience schemes. Conversely, at moderate and high C/I ratios, even a 3-MB AIR method gave quite satisfactory results compared to the non-resilient ones. During the addition of AIR to the compressed video data, the transcoder produced video streams which required the same number of GPRS timeslots to be transmitted as the non-resilient ones. During the experiments, it has also been demonstrated that the detrimental effects of transmission errors and the remedial effects of AIR varied with the change in the motion activity within the test sequences. The higher the motion activity, the less robust the video stream to errors.

Additionally, the FCS transcoding performance has also been tested and the video quality has been demonstrated to improve by a couple of dBs in most cases. The effect of an ACK/NACK delay has been observed to vary depending on the motion activities in the test sequences. In these tests, it has been shown that the increasing feedback delay also increased the bit rate and affected the video quality depending on the correlation of the video information between the reference and the current video frames. Similarly, the increase in bit rate was also easily managed here with the rate and error resilience control feedback loops of the video transcoder.

Furthermore, a combination of the AIR and FCS resilience methods has also been demonstrated. This combination has been shown to achieve superior transcoding qualities to the non-resilient video qualities at near target output bit rates, requiring the same number of GPRS timeslots. During these particular experiments, a 5% video frame loss was also considered in addition to the inherently error-prone GPRS transmission model at C/I = 12 dB, using the CS2 protection scheme. The tests were repeated for several video sequences and similar results were obtained with 2.5~4 dB enhancements in error-prone environments.

Finally, the reason why combined AIR-FCS method performed better than either alone is that whilst AIR compensated for the quality degradation caused by the GPRS channel bit errors, FCS also mitigated the effects of the full video frame losses. As it can be recalled from the set-up of the particular experiments, the video transmission over error-prone (CS2 C/I = 12 dB) GPRS channel was coupled with a 5% frame loss effect. Thus, AIR alone was only able to alleviate the GPRS bit error propagation effects within the received media stream whereas sole FCS could merely mitigate the temporal artefacts resulting from accumulation of errors due to full frame losses. Therefore, during the design of the proposed transcoding algorithm, it was envisaged to successfully stop the quality damaging error propagation effects with the use of a combined AIR-FCS method at the error-resilient video transcoder.

VII. CONCLUSION

An intermediate stage error resilience addition to an already compressed and transmitted video stream has been discussed in this paper. For this purpose, a video transcoder has been exploited to produce an error-resilient and standards-compliant output. The resilience was achieved with the use of separate and combined AIR and FCS techniques during the transcoding operations. The trade-off of both resilience schemes, namely the undesired inherent output bit rate increase due to their operations, was easily overcome and resolved by employing an adaptive rate transcoding scheme. Thus, a more efficient adoption of the resilience algorithms could be accomplished with output rates fairly close to the requirements. The adaptive operation of the combined rate and error resilience control feedback loops produced output rates at near target bit rates whilst injecting the necessary amount of robustness to pre-compressed video streams. Numerous experiments gave superior transcoding performances over the error-prone GPRS channels to the non-resilient video qualities.

Since this paper has presented an incorporation of the error resilience schemes into the video transcoding algorithm, it consequently shows another objective of the video transcoders: the provision of error resilience to compressed video streams. Thus, it can be said that the next generation video proxies will carry most of the burdens of the networks allowing the source encoders and end-decoders to stay free of complex resilience or rate regulation tasks.

REFERENCES

resilient MPEG-4 video transcoder for bit rate regulation,” in Proc. 

transcoder for multimedia traffic planning,” in Proc. IEEE 2nd 
Int. Conf. 3G Mobile Commun. Technol., IEEE conf. pub. no. 477, 

networks,” IEEE J. Select. Areas Commun., vol. 7, no. 5, pp. 771-

Iwakita and F. Watanabe, “Video transcoding proxy for 3G wireless 
mobile internet access,” IEEE Commun. Magazine, vol. 38, no. 10, 

coded video for transmission at lower bit rates,” in Proc. IEEE Int.
Conf. Acoustics, Speech, and Signal Processing, vol. 4, Atlanta, GA, 

video coding,” IEEE Trans. Consumer Electron., vol. 44, no. 1, 

minimum buffer requirement for compressed MPEG-2 bitstream,” 

[9] G. Keesman, R. Hellinghuizen, F. Hoeksema and G. Heideman,
“Transcoding of MPEG bitstreams,” Signal Processing: Image 

Technol., vol. 6, no. 2, pp. 191-199, April 1996.

4/H.263 video transcoder for interoperability of heterogeneous 

3845, Boston, MA, Sept. 1999, pp. 164-175.

[13] Information technology-Generic coding of audio-visual objects-
Vancouver meeting, July 1999.

[14] K. Stuhlmuller, N. Farber, M. Link and B. Girod, “Analysis of 
video transmission over lossy channels,” IEEE J. Select. Areas 


standard,” IEEE Commun. Magazine, vol. 36, no. 6, pp. 112-119, 
June 1998.

[18] Overall description of the GPRS radio interface stage 2, 
ETSI/SMG, GSM 03.64, Version 5.2.0, 1998.

the new GSM phase 2+ general packet radio service,” IEEE 

general packet radio service GPRS: Architecture, protocols, and air 

1997.

[22] Digital cellular telecommunications system (GSM radio access 

[23] P. Larsson and H. Olafsson, “Performance evaluation of 
different frequency reuse patterns and channel coding schemes in 
GPRS,” in Proc. IEEE 48th Vehicular Technol. Conf., vol. 1, 
Ottawa, Canada, May 1998, pp. 139-143.