

Research Article

Transcoding-Based Error-Resilient Video Adaptation for 3G Wireless Networks

Sertac Eminsoy,¹ Safak Dogan,² and Ahmet M. Kondo²

¹ NEC Electronics (Europe) GmbH, Cygnus House, Sunrise Parkway, Linford Wood, Milton Keynes, Buckinghamshire MK14 6NP, UK

² I-Lab, Centre for Communication Systems Research (CCSR), University of Surrey, Guildford GU2 7XH, Surrey, UK

Received 2 October 2006; Revised 19 February 2007; Accepted 13 May 2007

Recommended by Ming-Ting Sun

Transcoding is an effective method to provide video adaptation for heterogeneous internetwork video access and communication environments, which require the tailoring (i.e., repurposing) of coded video properties to channel conditions, terminal capabilities, and user preferences. This paper presents a video transcoding system that is capable of applying a suite of error resilience tools on the input compressed video streams while controlling the output rates to provide robust communications over error-prone and bandwidth-limited 3G wireless networks. The transcoder is also designed to employ a new adaptive intra-refresh algorithm, which is responsive to the detected scene activity inherently embedded into the video content and the reported time-varying channel error conditions of the wireless network. Comprehensive computer simulations demonstrate significant improvements in the received video quality performances using the new transcoding architecture without an extra computational cost.

Copyright © 2007 Sertac Eminsoy et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The success of the Internet and mobile systems has motivated the development of various enhanced-capacity fixed and wireless networking technologies (e.g., 3G, WLAN, broadband Internet, etc.). The services supported by such networks helped to foster the vision of being connected at anywhere, anytime, and with any device for pervasive media applications. However, the coexistence of the different networking infrastructures and services has also led to an increased heterogeneity of compressed video communication systems and scenarios, in which a wide range of user-terminals with various capabilities access rich video content over a multitude of access networks with different characteristics. The mismatches between the content properties and several network and/or device-centric features, as well as diverse user preferences, call for efficient video delivery systems featuring effective video adaptation mechanisms [1]. In general, this concept has been addressed in literature with the theme of the universal multimedia access (UMA) [2, 3]. Several strategies have been developed for UMA, which are based on the multimedia content adaptation techniques using the context specifications and descriptions defined in Part 7: digital item adaptation (DIA) of the MPEG-21 standard [4–8].

An effective way of performing video adaptation is to utilise transcoding operations in the networks. Transcoding is particularly needed when compressed video streams traverse heterogeneous networks. In such cases, a number of content-specific properties of the coded video information require adaptation to new conditions imposed by the different networks and/or terminals to retain an acceptable level of video service quality. Network-based adaptation mechanisms can be employed at the edges or other strategic locations of different networks, using a fixed-location video adaptation gateway, node, or proxy as in conventional networking strategies [9–11]. Alternatively, video adaptation (i.e., transcoding) can be performed dynamically where and whenever needed using active networking technologies [12, 13].

This paper presents a comprehensive transcoding system to facilitate the efficient adaptation of video for both error resilience and rate control purposes in a scenario, where a high bit-rate compressed video stream, which can be accommodated over a high-bandwidth fixed network, is sent to a mobile terminal over a 3G wireless network. This is a typical heterogeneous video communications scenario that consists of two networks with asymmetric bandwidth and channel characteristics. Video transmission over the fixed

network is not subject to bit errors, thus can be assumed error-free. However, transmission over the wireless network is error-prone. The wireless channel effects of 3G networks are characterised by burst errors, which introduce noise into the transmitted video signal, and thus cause the decoder to misinterpret the received data. Such errors are caused by channel fading, which is defined as the aggregated effects of multipath, shadowing, intra/intercell interference, as well as the number and location of users in a cell [14–16]. The overall resulting effect of such error conditions is the significant deterioration of the received video quality. Therefore, in addition to the bit-rate matching requirement between the two networks, an error resilience adaptation mechanism is needed to provide robust video transmission and satisfactory levels of quality of service (QoS). For this purpose, a video transcoding architecture is developed, which can be deployed at intermediate points in the networks. It incorporates a rate control mechanism together with a suite of error resilience tools, which were originally designed for source coding methods, as described by the MPEG-4 video coding standard [17, 18]. However, the transcoder utilises them in harmony to add necessary amount of error robustness to coded nonresilient video streams prior to transmission over error-prone 3G wireless channels. Furthermore, a new scene activity and channel adaptive intra-refresh (SC-AIR) transcoding algorithm is proposed to enhance the added robustness with respect to the detected video scene activity and the reported time varying channel error conditions. The computer simulations performed demonstrate the effectiveness of the new transcoding architecture in terms of providing video adaptation for bit-rate controlled error resilience over error-prone 3G (i.e., W-CDMA) channels.

The organisation of the paper is as follows. Section 2 highlights the need for rate-controlled error-resilient video adaptation, and describes the developed transcoding architecture in detail. Section 3 introduces the proposed visual scene activity and channel adaptive error-resilient transcoding mechanism (i.e., SC-AIR-based transcoding). Section 4 presents the computer simulation results and provides elaborate discussions on the resilient video adaptation performance. Finally, Section 5 outlines the concluding remarks.

2. TRANSCODING-BASED VIDEO ADAPTATION

Noisy channel conditions in wireless networks introduce errors in the compressed video streams, which are manifested as degradations in the perceived quality of the decoded video. Due to the predictive coding techniques applied in video compression, errors are likely to propagate into the future frames, making video communications perceptually unacceptable. To mitigate such effects, MPEG-4 has adopted a number of error resilience tools that make the video streams more robust against error-prone transmission over wireless channels [17, 18].

However, there exist numerous video applications all of which may not be optimised for error-resilient transmission. This is due to the fact that most error resilience tools reduce the coding efficiency, demand more bandwidth, and in-

crease the encoding and decoding complexities. Therefore, users should be provided with error-resilient video adaptation by the intermediate video adaptation nodes when and where necessary. The problems addressed above hence call for error-resilient transcoding units to be deployed in the networks.

Video transcoding is the process of converting the format of an input coded video content into another format [19–21]. It is primarily used to adapt the bit-rates, temporal, and spatial resolutions of the incoming compressed video streams, as well as to provide syntax translations between different video coding standards. Moreover, video transcoding with error resilience properties is particularly popular due to the fixed/wireless internetworking interoperability issues, and hence has been extensively addressed in literature [22–29].

In general, transcoders are designed to operate either in the discrete cosine transform domain (DCT domain) or in the pixel domain. Working in the pixel domain offers higher flexibility in terms of executing various different transcoding services simultaneously (i.e., controlling the output rate while inserting necessary amount of error robustness). The transcoding architecture presented in this paper is based on a cascaded pixel domain transcoder (CPDT) formation [19–21]. Typically, CPDT operation is computationally more expensive than the DCT domain architectures. Consequently, the reduction of the complexity while providing a highly functional and efficient architecture has been the main driving force behind many research activities in this area.

Figure 1 shows the block diagram of the developed transcoding architecture in this work. In its operation, the input compressed video is decoded down to the pixel domain, and then partial reencoding is performed on the decoded video in the required adapted format. Partial reencoding is utilised to reduce the inherent complexity of the CPDT operation. It is an operation, whereby the original frame headers are reused, and the time consuming reordering of the video frames is avoided. More significantly, motion compensation is performed reusing the original motion vectors of the input video stream. This is because, previous research has shown that methods like computationally complex motion reestimation and refinement can introduce considerable delay to the transcoder's operation, and yield only a minor video quality gain in the case of transmission over error-prone wireless channels [30]. In return, the overall effect is the significantly reduced computational complexity and processing time during transcoding.

The transcoder is compatible with the MPEG-4 video coding standard [17]. It is designed to perform both rate control (RC) and error resilience (ER) operations on the input video streams. RC is generally required to provide better utilisation of the bandwidth, rate matching between heterogeneous networks, and fair treatment to congestion responsive applications. In the context of this paper, RC is employed to smooth out the fluctuations in the throughput and convert input high bit-rate streams into lower rates [25]. As a result, users with diverse terminal capabilities,

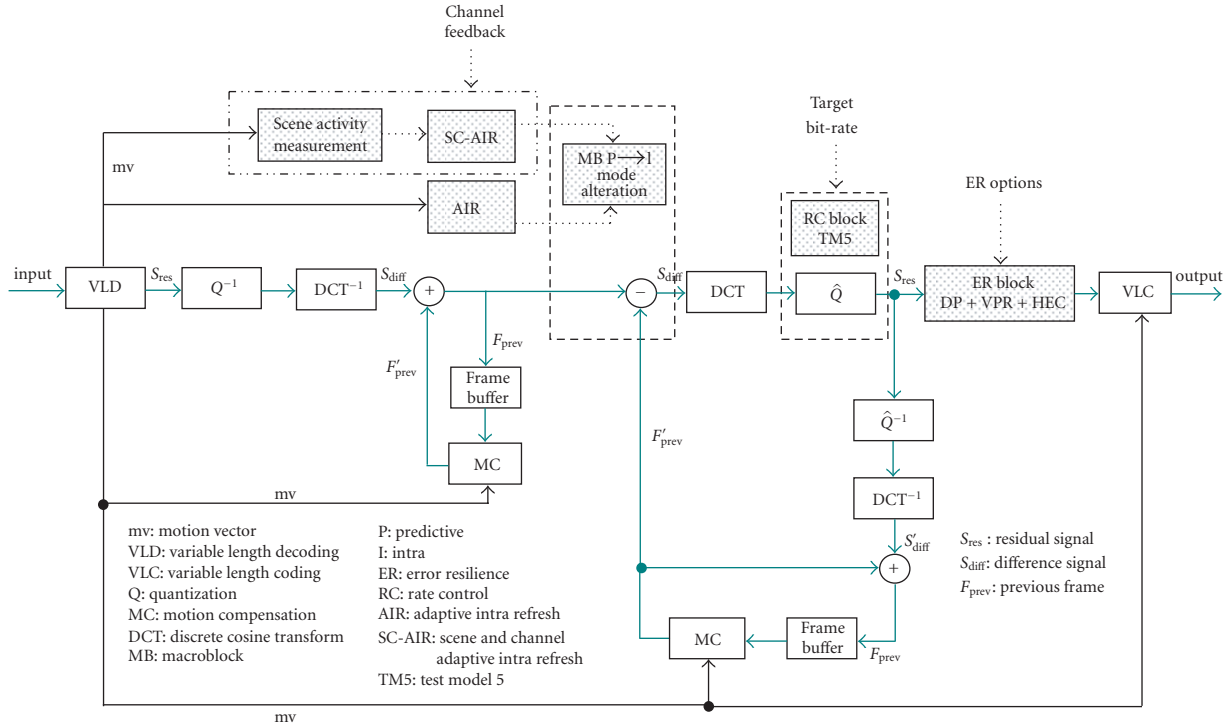


FIGURE 1: Block diagram of the error resilience and rate control capable video transcoding architecture.

personal requirements, and service level agreements can be accommodated. The mechanism employed for RC is a macroblock-based rate control algorithm, known as test model 5 (TM5) [31]. As shown in Figure 1, the RC block regulates the quantisation step size during partial reencoding, so as to compress the input bit-stream at the target bit-rate.

The ER tools are essential in alleviating the artefacts resulting from the transmission errors. However, considering that transcoding needs to be performed with minimal latency, error resilience tools are required to be computationally lightweight algorithms. Therefore, the developed transcoding architecture utilises computationally simple, but effective ER tools for mitigating the error effects in the perceived video due to the transmission errors that are introduced by the 3G wireless channel (i.e., W-CDMA channel). These ER tools are data partitioning (DP), video packet resynchronisation (VPR), header extension code (HEC), adaptive intra-refresh (AIR), and the novel scene activity and channel adaptive intra-refresh (SC-AIR) algorithm. While DP, VPR, and HEC are applied directly on the quantised video information, AIR or SC-AIR algorithms are executed prior to the partial reencoding process in order to alter the macroblock mode decisions extracted from the decoding operation. Depending on the operator's choice, either AIR or SC-AIR algorithms can be chosen as the intra-update mechanism. As illustrated in Figure 1, both algorithms make use of the input motion vector information to decide which macroblocks to update.

3. SCENE ACTIVITY AND CHANNEL ADAPTIVE INTRA-REFRESH (SC-AIR) TRANSCODING

The standard AIR algorithm uses a fixed and predetermined number of intra-macroblocks per frame for refreshing the frames of a video stream. This means that a specific AIR rate is determined before encoding or transcoding, and not altered throughout the encoding/transcoding process. In addition, the AIR algorithm lacks a mechanism for determining the optimum intra-refresh rate and providing adaptation to the spatiotemporal characteristics of the video stream with respect to varying channel conditions. For this reason, the effectiveness of AIR is significantly undermined.

In attempt to improve video communications over error-prone channels, various adaptive intra-refresh mechanisms have been presented in literature. A rate-distortion-based method was proposed by Côté and Kossentini, which measured the degradation in quality associated with the effects of losing individual macroblocks, and encoded certain blocks in intra-mode accordingly [32]. Although this was shown to be an effective method, it involved complex computations, and thus can be unsuitable for low-latency or real-time video applications. Similarly, Liao and Villasenor presented an intra-refresh mechanism that was based on error-sensitivity metrics [33]. This mechanism was implemented at the encoder and modelled the transmission medium as a random error channel with a specific bit-error-rate (BER). Another intra-refresh approach was employed by Reyes et al. as a tool for providing temporal resilience in a rate-distortion-based

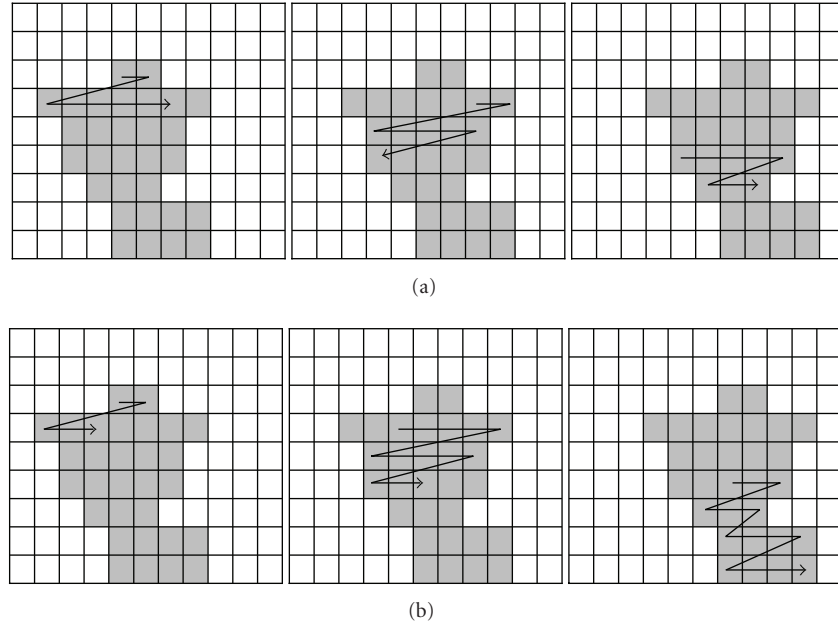


FIGURE 2: Intra-refresh of motion-active macroblocks: (a) operation of the standard AIR algorithm with a fixed number of macroblocks per frame; (b) operation of the SC-AIR algorithm with a variable intra-refresh rate.

error resilience scheme [22]. In this approach, the intra-refresh resilience was altered with respect to the output bit-rate and the BER of the channel. Mean-square-error (MSE) measurements were used to calculate the distortion introduced due to the lost macroblocks. Based on MSE measurements, the intra-refresh mechanism altered the number of intra coded blocks in every frame to provide optimal resilience. This algorithm involved complex computations, and hence can also be regarded unsuitable for low-latency video communications. Stuhlmüller et al. used a similar approach, where intra-refresh was based on a slightly modified version recommended in H.263 standard [34]. Another adaptive intra-refresh algorithm was proposed by Chiou et al. that required the encoder to extract some distortion information offline before the transmission, which was later used by a two-pass error-resilient video transcoder to decide on a prioritised intra-refresh strategy [26]. This idea was then developed into a more comprehensive error-resilient transcoder, which adaptively varied the intra-refresh rate according to the video content and communication channel's packet-loss rate to protect the most important macroblocks against packet losses over wireless networks [29, 35]. Moreover, a profit tracing scheme has recently been proposed by Chen et al. [36], so as to further improve the efficiency of intra-refresh allocation to macroblocks in the content-aware intra-refresh method developed in [29, 35]. A more practical intra-refresh method was also introduced by Worrall et al. [37]. In this approach, the intra-refresh mechanism at the encoder was based on a simple motion activity analysis of each frame and different GPRS channel conditions.

In this section, we introduce a video transcoding method using a new intra-refresh algorithm, namely SC-AIR, which

is adaptive to changing channel conditions and source specific characteristics (e.g., motion-based scene activity) of the video stream. The operation of the algorithm is similar to that of the standard AIR algorithm, where a motion map is formed to mark the location of the motion-active macroblocks in every frame and intra code (i.e., refresh) them sequentially [17]. This means that the motion-active macroblocks are first determined, and a number of them are refreshed starting from the first one until the end of the admissible intra-refresh rate (i.e., the number of macroblocks to be refreshed) in a predictive frame. In the subsequent predictive frame, the refresh algorithm continues from the point where it was left in the previous frame. This process continues in every predictive frame until the whole of the motion-active macroblocks are fully refreshed, and then the algorithm goes back to the first active macroblock to start the process again. However, as depicted in Figure 2, the SC-AIR algorithm computes the optimal intra-refresh rate for each video frame in contrast to standard AIR, where the intra-refresh rate is a predetermined fixed number of macroblocks throughout its operation. The information about the activity levels of a video scene is determined by examining different aspects of the motion-active macroblocks. This information is then coupled with the instantaneous channel condition factor to decide on the optimum number of intra-refresh blocks that is required to obtain the best possible performance. The operation of this algorithm has comparable computational complexity to the standard AIR operation, and thus it can be suitable for low-latency or real-time video communications.

The first stage of the SC-AIR algorithm involves extracting the activity information of each frame by means of a set

of functions, which analyse the scene-activity-related information from the input video scene. The second step involves the modulation of the activity information with a channel condition function, which represents the signal-to-noise ratio experienced at the downlink W-CDMA channel. The outcome of this operation determines the optimum number of intra-refresh blocks required for a frame. The complete SC-AIR algorithm is formulated as

$$\Omega(t, j) = \beta(t) \cdot \text{IRR}(j), \quad (1)$$

where Ω represents the number of macroblocks that need to be refreshed in the motion map of frame j (as described in Figure 2(b)), for any channel condition at time t . $\text{IRR}(j)$ stands for the intra-refresh rate determined from the scene activity analysis and $\beta(t)$ is derived from the instantaneous channel condition. The detailed explanations of these functions are given in the following subsections.

3.1. Activity measurement

In motion compensation-based video coders, the sensitivity to error can be related to the amount of motion within a scene [37]. Motion is defined by the motion vectors, and the activity is associated with the motion. Based on these assumptions, it can be claimed that as the amount of activity increases in a video scene, an increased number of intra-refresh blocks may be required to prevent the temporal error propagation. Thus, activity measurements can be used for developing an adaptation mechanism to counter the adverse effects of changing channel conditions on the perceived video quality. The activity measurement technique presented here forms the core of the SC-AIR transcoding algorithm. It reveals the required number of intra-refresh blocks for every predictive frame. In the development of this algorithm, various standard and in-house produced video test sequences with different scene activity levels were used. However, due to space constraints, the discussions of the algorithm presented in this paper are limited to the two standard test sequences, namely, “Foreman” and “Students”.

The SC-AIR algorithm is composed of a number of functions which represent the different aspects of the activity in a video scene. The algorithm performs the primary activity analysis using a function named the normalised activity index (NAI). In addition, a number of supplementary functions are also used to shape the output obtained by the NAI analysis. These functions are namely the motion macroblock factor and range index. The shaped NAI output is used to determine the optimum number of intra-refresh blocks needed for a frame.

Activity index (AI) is a function, which computes the cumulative magnitude of all motion vectors within a frame. NAI is the normalised variation of the AI function with respect to the number of macroblocks within a frame. This normalisation is required, so that the NAI measurements of different video sequences can be comparable with each other. If the NAI value is high, it is likely that the frame is a part of a highly active scene, which may indicate the need for more frequent intra-refresh (e.g., as in the “Foreman” se-

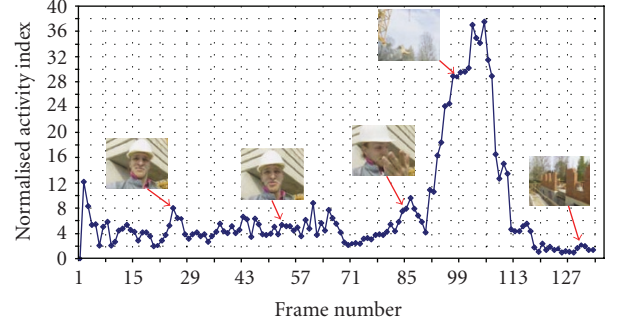


FIGURE 3: NAI computation for the “Foreman” stream.

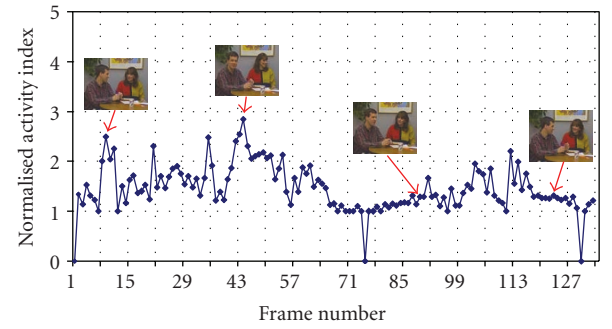


FIGURE 4: NAI computation for the “Students” stream.

quence) than a low-motion scene (e.g., as in the “Students” sequence). The NAI function can be written as

$$\text{NAI}(j) = \frac{\sum_{n=1}^{\tau} |mv_j(n)|}{i(j)}, \quad (2)$$

where $i(j)$ is the number of motion-active macroblocks in frame number j , n is the macroblock number, τ represents the total number of macroblocks in a frame (e.g., 99 for quarter common intermediate format: QCIF), and $mv_j(n)$ is the motion vector (both in x and y directions) of the n th macroblock in the j th frame. The NAI computations for the “Foreman” and “Students” sequences are depicted in Figures 3 and 4, respectively. As can be seen from these figures, the NAI is able to indicate the activity-level characteristics of both video sequences.

Nevertheless, the NAI computation on its own is insufficient in terms of defining the accurate levels of activity in a video scene. This is because the output of the NAI function is directly proportional with the motion vector sizes and inversely proportional with the number of motion macroblocks in a frame. As a result, if macroblocks with relatively small motion vectors are dominant in a particular frame, the NAI parameter will yield a small value, indicating low activity in that scene. This is the case where more frequent refreshing of the scene is required although the NAI result gives low values (e.g., the last 2 seconds of the “Foreman” sequence). Alternatively, if there is relatively small number of motion-active macroblocks but with relatively large motion vectors,

then the NAI value will yield a very large value, indicating an extreme activity in the scene. In this case, a high intra-refresh rate will be chosen, which may lead to compromise in the compression efficiency and degradation in the perceptual quality.

The empirical studies performed at the development stage have revealed that the inefficiency of the NAI function in determining the accurate levels of activity in a video scene can be compensated using a function called the motion macroblock factor (δ). The product of the NAI and the $\delta(j)$ functions determines the activity. The $\delta(j)$ function is computed for every frame, and is related to the ratio of the number of motion macroblocks to the total number of macroblocks in a frame. The use of this function can alleviate the anomalies in the NAI decision (as in the last 2 seconds of the “Foreman” sequence). The representation of the $\delta(j)$ function is given in

$$\delta(j) = \alpha \left(1 + \frac{i(j)}{\tau} \right), \quad (3)$$

where,

$$\alpha = \begin{cases} 0.75 & \text{if } R(j) \leq 5, \\ 1 & \text{if } 5 < R(j) \leq 10, \\ 2 & \text{if } 10 < R(j) \leq 20, \\ 3 & \text{if } 20 < R(j) \leq 30, \\ 4 & \text{if } 30 < R(j) \leq 40, \\ 5 & \text{if } 40 < R(j) \leq 50, \\ 7 & \text{if } R(j) > 50, \end{cases} \quad (4)$$

α is the scaling coefficient, whose value depends on the computed range index (R) function for every frame. The $R(j)$ function is given in

$$R(j) = \frac{i(j)}{\text{NAI}(j)}. \quad (5)$$

$R(j)$ stands for the ratio between $i(j)$ and the $\text{NAI}(j)$, and is calculated for every predictive frame. The scaling coefficient of the δ function is directly proportional with the outcome of $R(j)$ function. In other words, $R(j)$ is useful in scaling $\delta(j)$, which shapes the NAI output in a way that an optimum number of refresh blocks are chosen. As seen in (4), $R(j)$ defines seven different ranges for the scaling coefficient of the $\delta(j)$ function. These ranges and the weights of $\delta(j)$ were determined experimentally by observing the effects of the W-CDMA channel errors on the video quality with respect to the number of intra-refresh blocks enabled by the proposed algorithm.

In effect, the ranges defined by the $R(j)$ function correspond to seven different levels of confidence on the NAI decision. As $R(j)$ increases, the confidence on the NAI result decreases. Hence, as the $R(j)$ results in higher values, $\delta(j)$ function needs to be scaled with a higher coefficient. Consequently, these ranges enable a differentiated refreshing priority between frames with different motion characteristics. For instance, if a frame contains macroblocks with higher average

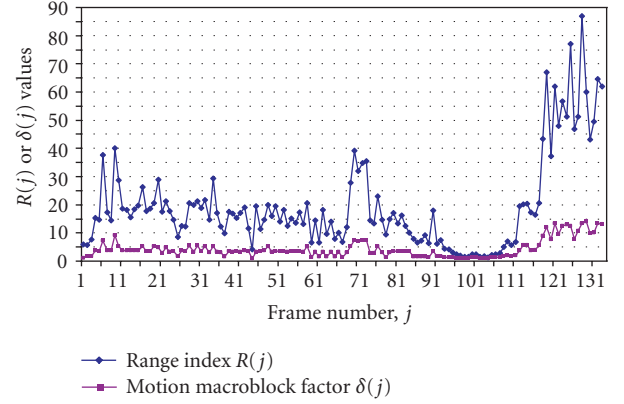


FIGURE 5: Motion macroblock factor and range index analysis for the “Foreman” stream.

motion vector magnitude than a frame with the same number of motion macroblocks but smaller average motion vector magnitude, then the proposed algorithm will always set a higher refresh rate. In other words, $R(j)$ function maintains a balance between the influence of the number of motion-active macroblocks and their average motion vector magnitude in determining the optimum intra-refresh rate.

Figure 5 shows the computed $R(j)$ and $\delta(j)$ functions for the “Foreman” stream. As seen from this figure, the value of the $R(j)$ function falls into its highest range during the last 2 seconds (i.e., between frame numbers 118 and 134) of the “Foreman” sequence, indicating the lowest confidence on the NAI output. That is to say, NAI function indicates low “activity” if the number of motion macroblocks is high, but the magnitude of their cumulative motion vectors is low. Thus, an appropriate α value should be used in the $\delta(j)$ function to compensate for the deficiency of the NAI function. Similarly, in the region that was indicated as the highest activity region by the NAI function (i.e., between the 9th and 11th seconds), the output of the $R(j)$ function falls into its lowest range. In this way, over-refreshing and consequently degradation in the compression efficiency are prevented. Other than accounting for these extreme cases, $R(j)$ function is utilised to assign an accurate weight to the $\delta(j)$ function, such that the best possible intra-refresh strategy can be applied.

On the other hand, the low-motion nature of the “Students” stream is also reflected on the output of its $R(j)$ function. As depicted in Figure 6, the output values of this function here are far less variable than in the case with the “Foreman” stream. In contrast, the $R(j)$ values for this video stream are occasionally below 5, which indicates that there are only a few motion-active macroblocks, and hence the confidence on the NAI analysis is high.

Having introduced all the functions which play a part in the activity analysis of the input video streams, the intra-refresh rate (IRR) required for a given frame can be calculated using the function given in

$$\text{IRR}(j) = \delta(j) \cdot \text{NAI}(j). \quad (6)$$

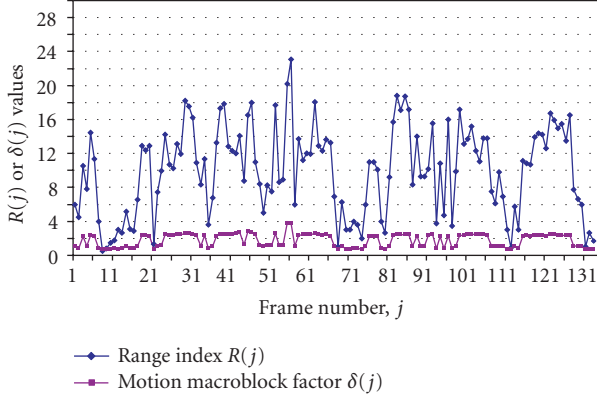


FIGURE 6: Motion macroblock factor and range index analysis for the “Students” stream.

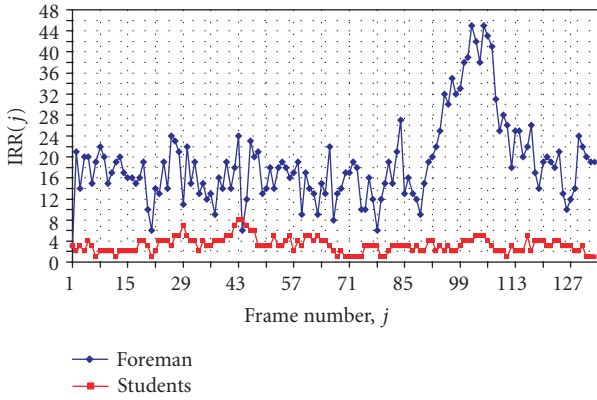


FIGURE 7: The intra-refresh rate computed for the “Foreman” and “Students” streams.

The results of the computation of this function for both the “Foreman” and “Students” streams are shown in Figure 7. As can be observed from this figure, the activity measurement functions are effective in differentiating between the activity levels of these two streams and assign particular number of intra-refresh macroblocks for each frame accordingly.

3.2. Channel factor

For wireless video communications, the varying channel conditions should also be considered in determining the optimum number of intra-refresh blocks to be used. In general, a faster refresh rate is required as the channel conditions worsen, while a limited number of intra-blocks will suffice when the channel conditions become better. The information on the instantaneous downlink bit-energy-to-noise ratio (E_b/N_o) value of the W-CDMA channel is available to the base station. In the proposed algorithm, this information is thus assumed to be fed back to the network node that performs the error-resilient video adaptation (i.e., transcoding). In 3G systems, such feedback can be made available to the

video adaptation node in less than 250 milliseconds of delay [38].

As presented in (7), the channel factor $\beta(t)$ is a coefficient, whose values were determined experimentally for various channel conditions (i.e., E_b/N_o) in W-CDMA. $\beta(t)$ is a time-dependent function which implies that channel conditions may vary over the time. In the exhaustive number of experiments conducted, only those error patterns corresponding to E_b/N_o ranging from 6 dB to 10 dB were used. Considering the performance figures presented in [39], it can be argued that it is not possible to conceive acceptable quality video communications at E_b/N_o rates below 6 dB, and the intra-refresh application thus becomes ineffective. On the contrary, the effects of errors at E_b/N_o rates above 10 dB reach to saturation, and the channel condition factor can remain the same. As the channel condition worsens, a more aggressive intra-refresh rate should be applied to the video streams in order to limit the increased error propagation into the future frames during transcoding. Conversely, when the channel conditions start to improve (i.e., $E_b/N_o \geq 9$ dB), intra-refresh should be made less frequent in order not to compromise the coding efficiency by the introduction of unnecessary number intra macroblocks.

Having tested the output of the $IRR(j)$ function with various video streams and under different channel conditions, the following $\beta(t)$ values were found to provide the optimum efficiency for the SC-AIR algorithm operating at the following specified conditions:

$$\beta(t) = \begin{cases} 1.8 & \text{if } \frac{E_b}{N_o} = 6 \text{ dB}, \\ 1.5 & \text{if } \frac{E_b}{N_o} = 7 \text{ dB}, \\ 1.0 & \text{if } \frac{E_b}{N_o} = 8 \text{ dB}, \\ 0.9 & \text{if } \frac{E_b}{N_o} = 9 \text{ dB}, \\ 0.35 & \text{if } \frac{E_b}{N_o} = 10 \text{ dB}. \end{cases} \quad (7)$$

The difference between consecutive E_b/N_o values was chosen to be 1 dB. According to the test results given in [39], 1 dB difference in E_b/N_o value represents a noticeable change in the channel conditions. Hence for the SC-AIR algorithm, any intermediate E_b/N_o figure can be rounded to its closest integer value.

4. ADAPTATION PERFORMANCE EVALUATION BY COMPUTER SIMULATIONS

The advantages of employing an error-resilient transcoder (i.e., using HEC, VPR, DP, and AIR) in an EDGE network were demonstrated in our earlier work [25]. In this section, we discuss the transcoding performance using these tools over the 3G network in Section 4.2 while also presenting the performance evaluation of our proposed SC-AIR transcoding algorithm in Section 4.3. Therefore, a series of experiments were conducted using the simulation scenario

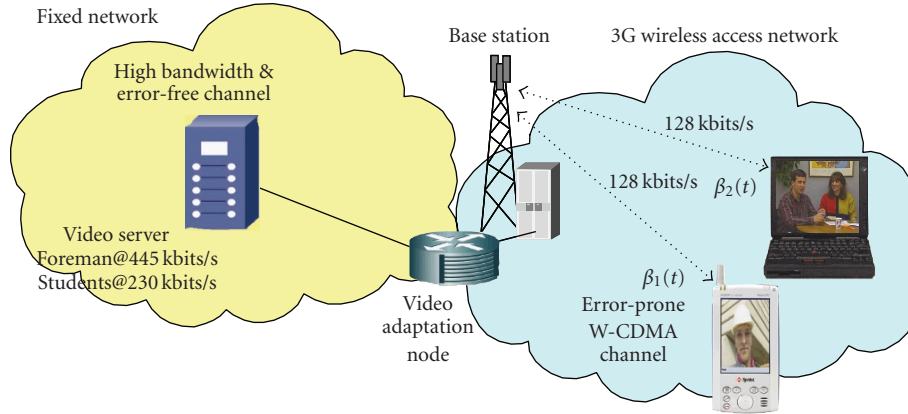


FIGURE 8: Experimentation scenario for the error-resilient video adaptation by transcoding.

shown in Figure 8. In this scenario, the video source lies in the fixed network and transmits live or stored video content to subscribers in the 3G network, which uses W-CDMA technology in its radio access network. The video server sends the video stream through one of the video adaptation nodes, which performs the necessary video adaptation for heterogeneous media transmission. This node operates the transcoder, and is assumed to be aware of the user profile, requirements, and channel conditions. Consequently, it is capable of modifying the incoming bit-stream to the required format. Due to their location, speed, and various interference sources, mobile users experience time varying channel conditions, as represented by $\beta(t)$.

4.1. Simulation setup

In the experiments, two different test sequences were deliberately chosen to comprise two different motion activity properties: “Students” and “Foreman” with low and high activity scenes, respectively. The video server depicted in Figure 8 employs an MPEG-4 video encoder, and is set to produce a single base-layer stream without any rate control (with a fixed quantisation step size at $Q_p = 2$). The server produces QCIF-sized video streams encoded at 10frames/s, and with I-P-P-P... layout for both video sequences. The duration of the encoding operation was limited to 13.4 seconds (i.e., 134 frames) for both video sequences. The error resilience options of the server’s encoder were turned off, as the link between the source and the transcoder is assumed to be error-free.

In the experiments performed, the transcoder was configured to convert the incoming variable rate bit-stream (i.e., on average at 230 kbits/s for “Students” and 445 kbits/s for “Foreman”) to a constant bit-rate output at 128 kbits/s. The adaptation to the wireless channel errors was provided using the error resilience modules of the transcoder (i.e., DP, VPR, HEC, and AIR or SC-AIR). It can be argued that if a fixed number of AIR blocks are used for every frame, a video packet size of 700 bits is a reasonable value that can be used for most video sequences given the 128 kbits/s transcoder

output and the W-CDMA channel conditions [25, 40]. Similarly, based on the figures presented in [40] and the preliminary study performed, an AIR frequency of 10 blocks was decided to be used in the first set of experiments, where the standard AIR was used for the error-resilient video adaptation tests.

The effects of W-CDMA physical link layer on the transcoder’s downlink was simulated by corrupting the transcoded video streams with appropriate error patterns. These error patterns were produced by the in-house developed W-CDMA physical link layer simulator [39]. They are used for emulating the downlink channel conditions for a specific E_b/N_o , channel coding scheme, spreading factor (SF), propagation environments, mobile speed, and power control availability. In the experiments, because of the transcoder’s target output bit-rate selection, the SF was determined as 16. In addition, the channel coding scheme was set to 1/3 convolutional coding (CC 1/3) in Vehicular-A propagation environment at 50 km/h mobile speed without power control. The video performance in this environment is similar to that in Pedestrian-B environment with power control [41]. The channel E_b/N_o was selected as 9 dB, which corresponds to a channel BER of 6×10^{-2} [39].

Five different experiments were carried out for each video sequence in the first set of tests, whose results are presented in Section 4.2. These experiments involved progressive application of different combinations of the ER tools on the 128 kbits/s rate controlled (e.g., constant bit-rate) video streams. Firstly, a rate control transcoded nonresilient video stream was transmitted over the error-prone wireless link. In the second experiment, transcoding was performed using the VPR and HEC tools. This was followed by the repetition of the same experiment with the introduction of the DP tool. Finally, the full resilience was provided by adding AIR on top of all the other resilience tools applied. While testing the proposed transcoding system with the SC-AIR method in the second set of tests, whose results are presented in Section 4.3, the SC-AIR block of Figure 1 was employed instead of the standard AIR block in addition to the DP, VPR, and HEC tools.

TABLE 1: Error-resilient transcoding test results for different resilience tool combinations.

Avg. PSNR	ER Tools				
	No ER	VPR + HEC	VPR + HEC + DP	VPR + HEC + DP + 10AIR	Error-Free
Foreman	25.67 dB	26.29 dB	28.45 dB	31.33 dB	33.98 dB
Students	26.57 dB	27.21 dB	33.24 dB	34.02 dB	36.35 dB

In the preliminary work performed, it was observed that 25 simulation runs for each test are adequate for the correct representation of the overall channel effects on the video quality. In other words, each transcoder output was corrupted with channel errors for 25 times with the same set of seeds for each different test. The corrupted video streams were then decoded, and the resulting video quality was measured in terms of average peak-to-peak signal-to-noise ratio (PSNR). The obtained results for the two sets of tests conducted (i.e., with standard AIR and SC-AIR) are discussed in the following sections.

4.2. Error-resilient video adaptation performance with standard AIR transcoding

Exhaustive simulations were run to test the performance of the transcoder with the ER tools. The average PSNR measurements of all of the five error resilience tests are presented in Table 1, which demonstrates the relative performances of the different combinations of transcoding operations. This procedure is applied both to the “Foreman” and “Students” sequences. The PSNR results reveal some interesting findings. The use of combined VPR and HEC can provide better video quality than the non-error-resilient stream. The use of these tools improves the PSNR results on average by 0.62 dB for the “Foreman” and 0.64 dB for the “Students” streams compared to the “No ER” case. In fact, for the given source rate, channel coding and propagation environment, E_b/N_o value of 9 dB can be considered to represent a moderate channel condition. Therefore, the performance gain obtained from using the combined VPR and HEC tools is expected to become more distinctive as the channel conditions worsen (i.e., as E_b/N_o decreases).

On the other hand, the use of DP brings in a considerable performance gain in the decoded video quality. However, it should be noted that DP is always used with VPR, and optionally together with HEC. Here, a combined DP, VPR, and HEC resilience tool is utilised for providing robust transmission of the video streams over the error-prone W-CDMA channel. The addition of DP reduces the amount of video information discarded by the decoder in the case of errors affecting the texture information. As can be observed from the figures presented in Table 1, the DP tool together with VPR and HEC provides an improved PSNR performance on average by 2.16 dB for the “Foreman” and 6.03 dB for the “Students” streams, over the “VPR + HEC only” case. The difference in the performance gains is due to the motion activity differences between the two sequences. For the bit-rate of 128 kbits/s at the transcoder output, only 3.1% of the total compressed bits are motion information for the

“Students” stream while this figure is 11.3% for the “Foreman” stream. Thus, the “Students” stream contains much more texture information than “Foreman.” This means that the likelihood of channel errors corrupting the motion information in the “Foreman” stream is considerably higher than that in the “Students” stream, which has a limiting factor on improving the transcoding video quality using DP.

DP alone is not capable of salvaging the video packet contents when an error occurs in the motion partition of this video packet. Thus, the combined DP, VPR, and HEC tool results in an average PSNR figure of 33.24 dB for the “Students” stream (i.e., around 6.5 dB better than the “No ER” performance), and 28.45 dB for the “Foreman” stream (i.e., around 3 dB better than the “No ER” performance).

The final ER tool used in the experiments was the standard AIR. It is shown that the addition of the AIR technique to the combination of the other tools enables a considerable performance gain in the decoded video quality, as can be seen both in the PSNR measurements and subjective tests. The overall performance gains in the transcoded video qualities using all of the ER tools amount up to 5.66 dB for “Foreman” and 7.45 dB for “Students” over transcoding without ER. Furthermore, it can be observed from the objective results that while AIR addition makes a significant difference on the quality of the decoded “Foreman” stream (i.e., average PSNR result is improved by 2.88 dB) compared to the “HEC + VPR + DP” transcoding case, the “Students” stream’s average PSNR result is improved by 0.78 dB only. From this result, it can be argued that video streams with relatively higher motion activity can obtain a higher benefit from AIR than those with lower motion activity. Meanwhile, the effectiveness of the DP tool also decreases as the texture to motion information ratio in a video packet decreases, which in return limits the overall performance.

The PSNR results have also been supported with the associated subjective test results, as shown in Figures 9(a)–9(e) for the “Foreman” stream. These results demonstrate that the use of ER transcoding, particularly with the addition of AIR tool significantly improves the perceptual quality. Similar subjective test results have also been obtained for “Students.”

4.3. Error-resilient video adaptation performance with SC-AIR transcoding

The usefulness of the ER transcoding tools and in particular the effectiveness of the AIR algorithm in limiting the temporal error propagation in video communications over W-CDMA networks have been demonstrated in the previous section. Subsequently, a series of tests were performed



FIGURE 9: Subjective test results of the 90th frames of the transcoded “Foreman” stream for $E_b/N_o = 9$ dB: (a) error-free; (b) No ER; (c) HEC + VPR; (d) HEC + VPR + DP; (e) HEC + VPR + DP + 10AIR.

to demonstrate the effectiveness of the SC-AIR transcoding algorithm in comparison to the standard AIR algorithm-based video transcoding. Two versions of the same error-resilient transcoder architecture were used; one was set to execute VPR, HEC, DP, and AIR, as presented in Section 4.2, while the other operated with the VPR, HEC, DP, and SC-AIR algorithms. In this way, it is possible to relate the performances of AIR and the proposed SC-AIR algorithms for the aforementioned channel conditions. However, in order to make a fair comparison, the optimum operation point of the AIR algorithm needed to be determined. For this purpose, numerous preliminary tests were performed to find the most effective fixed AIR frequency. The AIR algorithm was tested for each sequence and under various channel conditions (i.e., between $E_b/N_o = 6$ dB and 10 dB), and a range of fixed AIR rates were listed experimentally, which produced the highest average PSNR values. Depending on the video sequence, the intra-refresh rate of the standard AIR algorithm was incremented in steps of 2 to 4 macroblocks to find the optimum operation points. Having determined four possible optimum operation points for the standard AIR algorithm, the SC-AIR performance was tested with the same video sequences and the channel errors. Tables 2 and 3 show the results of the comparison tests performed for the SC-AIR and the standard AIR algorithms using the “Foreman” and “Students” sequences, respectively.

As can be seen from these results, transcoding with the SC-AIR algorithm has resulted in better video qualities (i.e., in terms of PSNR) than those of the transcoding with the standard AIR algorithm in all of the experiments performed. However, it should be noted that with the channel condition feedback, the SC-AIR algorithm calculates the

TABLE 2: Transcoding test results for the “Foreman” stream using standard AIR and SC-AIR.

Foreman					
$E_b/N_o = 6$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	24	28	32	36	avg. 29
Avg. PSNR (dB)	21.56	21.81	22.15	22.11	22.37
$E_b/N_o = 7$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	24	28	32	36	avg. 25
Avg. PSNR (dB)	26.49	26.38	26.43	26.54	26.81
$E_b/N_o = 8$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	12	16	20	24	avg. 18
Avg. PSNR (dB)	29.93	30.21	30.33	30.06	30.41
$E_b/N_o = 9$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	8	12	16	20	avg. 17
Avg. PSNR (dB)	31.01	31.25	31.28	31.16	31.36
$E_b/N_o = 10$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	4	8	12	16	avg. 7
Avg. PSNR (dB)	33.04	33.08	32.86	32.66	33.15

optimum number of intra-refresh blocks dynamically and automatically. Conversely, the standard AIR algorithm rates are fixed and configured manually to match the performance

TABLE 3: Transcoding test results for the “Students” stream using standard AIR and SC-AIR.

Students					
$E_b/N_o = 6$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	12	15	18	21	avg. 5
Avg. PSNR (dB)	26.62	26.74	26.83	26.30	26.96
$E_b/N_o = 7$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	3	6	9	12	avg. 4
Avg. PSNR (dB)	29.80	30.54	30.47	30.45	30.73
$E_b/N_o = 8$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	3	6	9	12	avg. 3
Avg. PSNR (dB)	32.69	32.76	32.62	32.44	33.13
$E_b/N_o = 9$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	8	12	16	20	avg. 2
Avg. PSNR (dB)	34.27	34.31	34.23	34.04	34.62
$E_b/N_o = 10$ dB					
Intra-refresh algorithm	AIR	AIR	AIR	AIR	SC-AIR
Intra-blocks per frame	2	4	6	8	avg. 1
Avg. PSNR (dB)	35.71	34.47	35.39	35.34	35.78

of the SC-AIR as much as possible, as described in the previous paragraph. As can be seen from the results presented in the two tables, maximum of 1.31 dB and 0.81 dB improvements were noted for the “Students” and “Foreman” streams, respectively. These results show that even with the manual configuration of intra-refresh rates for each video sequence and the channel condition, the standard AIR-based transcoding was outperformed by the SC-AIR transcoding. Moreover, in an actual video communications scenario, such manual configuration will clearly not be feasible, and the difference in the performances of the two algorithms will hence be much higher.

The major strength of the SC-AIR algorithm comes from its ability to compute the activity levels in a video scene while also taking into account the time varying W-CDMA channel conditions, so as to adjust its refreshing strategy accordingly. In this respect, its advantages are twofold: firstly, temporal error propagation is decreased by intra-refreshing the right number of macroblocks for each frame, and secondly, the compromise in the compression efficiency is minimised by preventing over-refreshing to take place (i.e., in better channel conditions).

4.4. Complexity analysis

In our early work, steps taken to simplify the inherent complexity of the error-resilient CPDT architecture were addressed [25]. In this section, the relative run-time complexities of different error resilience tools are presented. For this purpose, the rate controlled (i.e., at 128 kbits/s) and error-

resilient transcoding operation was executed on a PC with an Intel Pentium 4 processor at 3.19 GHz, with 512 kbytes of L2 cache, 1 Gbyte of memory running the Windows XP operating system. The operation speed of the transcoder is represented with the number of frames transcoded per second (fps). With the channel condition set to $E_b/N_o = 9$ dB, the error resilience tools were progressively turned on and operation speed for each case was noted. The run-time complexity analysis of the transcoder for both the “Foreman” and “Students” sequences are given in Table 4.

The obtained results reveal that the DP tool is the most complex of all other resilience operations while VPR and HEC tools constitute negligible complexity. These figures also show that the SC-AIR algorithm does not introduce measurable run-time complexity compared to the standard AIR algorithm. In other words, the quality gain that can be obtained by employing the SC-AIR algorithm in the transcoder does not pose an additional computational cost compared to the standard MPEG-4 error resilience intra-refresh tool (i.e., AIR algorithm). As a result, the delay overhead introduced to the end-to-end communications delay is kept to minimum. Thus, it can be argued that the SC-AIR transcoding operation is ideal for use in low-latency video communications over the 3G wireless access networks.

5. CONCLUSIONS

In this paper, a comprehensive transcoding system has been presented to provide video adaptation for both error-resilient and rate-controlled video access/distribution from fixed to 3G wireless networks. Error-resilient video adaptation has been provided by a combination of a number of error resilience tools, such as HEC, VPR, DP, AIR, and SC-AIR. It has been shown that the developed error-resilient video transcoder is efficient in providing protection for the compressed video streams prior to their transmission over noisy channels in heterogeneous inter-network communication scenarios. The performance of the error-resilient video adaptation system was tested over a simulated 3G W-CDMA channel, and the results have shown that the optimum quality performance is attainable when all the tools are used together. Compared to the nonresilient error-prone case, the gain in the measured video quality performance (i.e., in terms of PSNRs) has been demonstrated to be around 5 to 8 dB depending on the scene activities of the tested video sequences. With this finding in hand, a new W-CDMA specific intra-refresh algorithm, namely SC-AIR, was also developed and incorporated into the transcoding system, which optimised the number of intra-blocks in a transcoded video frame based on the scene activity and time varying wireless channel conditions. Through exhaustive experimentations, it has been demonstrated that for a given noisy 3G wireless channel condition, SC-AIR transcoding can yield better performances in terms of received video qualities compared to those obtained using constant frequency AIR algorithm (i.e., conventional AIR tool). The strength of the SC-AIR-based transcoding comes from the fact that the SC-AIR algorithm has modelled the effects of the simulated W-CDMA channel

TABLE 4: Complexity analysis of the transcoder's ER tools.

ER Tools	No ER	VPR + HEC	VPR + HEC + DP	VPR + HEC + DP + AIR	VPR + HEC + DP + SC-AIR
Foreman@128 kbits/s	204 fps	204 fps	175 fps	168 fps	168 fps
Students@128 kbits/s	232 fps	232 fps	195 fps	195 fps	195 fps

and the video scene activity thoroughly, and thus the operation of the transcoder has been designed to be fully adaptive to the varying channel conditions as well as to the scene activity inherently embedded into the video information. In addition, it was shown that SC-AIR does not introduce measurable run-time complexity compared to the standard AIR algorithm. Thus, it is envisaged that SC-AIR-based transcoding can be employed in low-latency video adaptation applications.

ACKNOWLEDGMENTS

The work presented was developed within VISNET I and VISNET II, European Networks of Excellence (<http://www.visnet-noe.org>), funded under the European Commission IST FP6 programme. The authors would like to thank the special issue guest editors and anonymous reviewers for their constructive suggestions.

REFERENCES

- [1] S.-F. Chang and A. Vetro, "Video adaptation: concepts, technologies, and open issues," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 148–158, 2005.
- [2] "Special Issue on Universal Multimedia Access," *IEEE Signal Processing Magazine*, vol. 20, no. 2, March 2003.
- [3] "Special Issue on Multimedia Adaptation," *Signal Processing: Image Communication*, vol. 18, no. 8, September 2003.
- [4] MPEG-21—Part-7: ISO/IEC 21000-7, "Information technology—multimedia framework—part 7: Digital Item Adaptation," October 2004.
- [5] J. Bormans, J. Gelissen, and A. Perkis, "MPEG-21: the 21st century multimedia framework," *IEEE Signal Processing Magazine*, vol. 20, no. 2, pp. 53–62, 2003.
- [6] I. Burnett, R. van de Walle, K. Hill, J. Bormans, and F. Pereira, "MPEG-21: goals and achievements," *IEEE Multimedia*, vol. 10, no. 4, pp. 60–70, 2003.
- [7] A. Vetro, "MPEG-21 digital item adaptation: enabling universal multimedia access," *IEEE Multimedia*, vol. 11, no. 1, pp. 84–87, 2004.
- [8] J. Magalhães and F. Pereira, "Using MPEG standards for multimedia customization," *Signal Processing: Image Communication*, vol. 19, no. 5, pp. 437–456, 2004.
- [9] T. Warabino, S. Ota, D. Morikawa, et al., "Video transcoding proxy for 3G wireless mobile Internet access," *IEEE Communications Magazine*, vol. 38, no. 10, pp. 66–71, 2000.
- [10] S. Dogan, S. Eminsoy, A. H. Sadka, and A. M. Kondo, "Personalised multimedia services for real-time video over 3G mobile networks," in *Proceedings of the 3rd IEEE International Conference on 3G Mobile Communication Technologies (3G '02)*, no. 489, pp. 366–370, London, UK, May 2002.
- [11] Z. Lei and N. D. Georganas, "Video transcoding gateway for wireless video access," in *Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering (CCECE '03)*, vol. 3, pp. 1775–1778, Montreal, Quebec, Canada, May 2003.
- [12] E. Amir, S. McCanne, and R. Katz, "An active service framework and its application to real-time multimedia transcoding," in *Proceedings of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM '98)*, pp. 178–189, Vancouver, BC, Canada, August–September 1998.
- [13] M. Ott, G. Welling, S. Mathur, D. Reininger, and R. Izmailov, "The JOURNEY active network model," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 3, pp. 527–537, 2001.
- [14] 3GPP: TR 30.03, "Selection procedures for the choice of radio transmission technologies of the UMTS," V3.2.0 1998-04.
- [15] H. Holma and A. Toskala, *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*, John Wiley & Sons, London, UK, 3rd edition, 2004.
- [16] C. Kodikara, S. Worrall, A. H. Sadka, and A. M. Kondo, "Performance enhancement of real-time video communications over UTRAN by adaptive resource allocation: system level evaluation," in *Proceedings of the 5th IEEE International Conference on 3G Mobile Communication Technologies (3G '04)*, no. 503, pp. 427–431, London, UK, October 2004.
- [17] MPEG-4—Part-2: ISO/IEC 14496-2, "Information technology—generic coding of audio-visual objects—part 2: visual," July 2001.
- [18] R. Talluri, "Error-resilient video coding in the ISO MPEG-4 standard," *IEEE Communications Magazine*, vol. 36, no. 6, pp. 112–119, 1998.
- [19] A. Vetro, C. Christopoulos, and H. Sun, "Video transcoding architectures and techniques: an overview," *IEEE Signal Processing Magazine*, vol. 20, no. 2, pp. 18–29, 2003.
- [20] J. Xin, C.-W. Lin, and M.-T. Sun, "Digital video transcoding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 84–97, 2005.
- [21] I. Ahmad, X. Wei, Y. Sun, and Y.-Q. Zhang, "Video transcoding: an overview of various techniques and research issues," *IEEE Transactions on Multimedia*, vol. 7, no. 5, pp. 793–804, 2005.
- [22] G. de los Reyes, A. R. Reibman, S.-F. Chang, and J. C.-I. Chuang, "Error-resilient transcoding for video over wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1063–1074, 2000.
- [23] S. Dogan, A. Cellatoglu, M. Uyguroglu, A. H. Sadka, and A. M. Kondo, "Error-resilient video transcoding for robust internetwork communications using GPRS," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 6, pp. 453–464, 2002.
- [24] I. K. Kim, N. I. Cho, and J. Nam, "Error resilient video transcoding based on the optimal multiple description of DCT coefficients," in *Proceedings of International Workshop on Advanced Image Technology (IWAIT '03)*, pp. 77–80, Nagasaki, Japan, January 2003.
- [25] S. Eminsoy, S. Dogan, and A. M. Kondo, "Real-time services for video communications," in *Proceedings of IEEE International Conference on Visual Information Engineering (VIE '03)*, no. 495, pp. 262–265, Guildford, UK, July 2003.

- [26] H.-J. Chiou, Y.-R. Lee, and C.-W. Lin, "Error-resilient transcoding using adaptive intra refresh for video streaming," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '04)*, vol. 3, pp. 777–780, Vancouver, BC, Canada, May 2004.
- [27] M. Xia, A. Vetro, H. Sun, and B. Liu, "Rate-distortion optimized bit allocation for error resilient video transcoding," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '04)*, vol. 3, pp. 945–948, Vancouver, BC, Canada, May 2004.
- [28] A. Vetro, J. Xin, and H. Sun, "Error resilience video transcoding for wireless communications," *IEEE Wireless Communications*, vol. 12, no. 4, pp. 14–21, 2005.
- [29] H.-J. Chiou, Y.-R. Lee, and C.-W. Lin, "Content-aware error-resilient transcoding using prioritized intra-refresh for video streaming," *Journal of Visual Communication and Image Representation*, vol. 16, no. 4-5, pp. 563–588, 2005.
- [30] S. Dogan and A. H. Sadka, "Video transcoding for inter-network communications," in *Compressed Video Communications by A. H. Sadka*, chapter 6, pp. 215–256, John Wiley & Sons, London, UK, 2002.
- [31] TM5: ISO/IEC N0400, "Test model 5," April 1993.
- [32] G. Côté and F. Kossentini, "Optimal intra coding of macroblocks for robust H.263 video communication over the Internet," *Signal Processing: Image Communication*, vol. 15, no. 1-2, pp. 25–34, 1999.
- [33] J. Y. Liao and J. Villasenor, "Adaptive intra block update for robust transmission of H.263," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 1, pp. 30–35, 2000.
- [34] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, 2000.
- [35] C.-H. Chen, Y.-R. Lee, C.-W. Lin, and Y.-C. Chen, "Error resilience transcoding using prioritized intra-refresh for video multicast over wireless networks," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '05)*, pp. 1310–1313, Amsterdam, The Netherlands, July 2005.
- [36] C.-M. Chen, Y.-C. Chen, and C.-W. Lin, "Error-resilience transcoding using content-aware intra-refresh based on profit tracing," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '06)*, pp. 5283–5286, Kos, Greece, May 2006.
- [37] S. T. Worrall, A. H. Sadka, A. M. Kondo, and P. Sweeney, "Motion adaptive intra refresh for MPEG-4," *Electronics Letters*, vol. 36, no. 23, pp. 1924–1925, 2000.
- [38] 3GPP: TS 25.853, "Delay budget within the access stratum," release 1999, March 2001.
- [39] C. Kodikara, S. Worrall, S. N. Fabri, and A. M. Kondo, "Performance evaluation of MPEG-4 video telephony over UMTS," in *Proceedings of the 4th IEE International Conference on 3G Mobile Communication Technologies (3G '03)*, no. 494, pp. 73–77, London, UK, June 2003.
- [40] L. D. Soares, S. Adachi, and F. Pereira, "Influence of encoder parameters on the decoded video quality for MPEG-4 over W-CDMA mobile networks," in *Proceedings of IEEE International Conference on Image Processing (ICIP '00)*, vol. 2, pp. 148–151, Vancouver, BC, Canada, September 2000.
- [41] C. Kodikara, *Multimedia communications over 3G wireless communication systems*, Ph.D. thesis, Electronic Engineering Department, University of Surrey, Surrey, UK, January 2004.

Sertac Eminsoy received his B.S. degree in electrical and electronics engineering from the Eastern Mediterranean University in North Cyprus, in 1998. In 1999, he completed his M.S. degree in mobile and satellite communications at the University of Surrey, Guildford, UK. He received his Ph.D. degree from the University of Surrey in 2005 where he conducted his research on QoS support for multimedia communications. His main research interests are video transcoding, multimedia transmission over packet networks, multimedia content adaptation, and active networks. Since 2006, he has been working for NEC Electronics (Europe) GmbH, Milton Keynes, UK, where he is a Member of the Multi Media Software Solutions Team.



Safak Dogan received his B.S. degree with Honors in electronics and communications engineering from the Technical University of Istanbul, Turkey, in 1995, the M.S. degree in satellite communications engineering, and the Ph.D. degree in multimedia communications engineering, both from the University of Surrey, Guildford, UK, in 1996 and in 2001, respectively. He has been with the I-Lab Multimedia Technologies and DSP Research Group of the Centre for Communication Systems Research (CCSR) at the University of Surrey since 2001. His primary research interests are video adaptation, video transcoding, context-based media content adaptation, multimedia communication networks, low bit-rate video communications, error resilience for next generation mobile-wireless video transmissions, and multiparty audiovisual collaboration/conferencing technologies. Dr. Dogan has authored several research papers published in various peer-refereed international conferences and prestigious journals. He has acted as a referee for the Research Council of Canada, and has also been serving as a reviewer to a number of international journals and conferences in his research areas. Moreover, he has contributed to a technical book with a highly specialised video transcoding chapter. He is a Member of the IEEE.



Ahmet M. Kondo was born in Cyprus. He received the B.S. degree with Honors in engineering, the M.S. degree in telematics, and the Ph.D. degree in communication in 1983, 1984, and 1986, respectively. He became a lecturer in 1988, a Reader in 1995, and then in 1996, a Professor in Multimedia Communication Systems and Deputy Director of Centre for Communication Systems Research (CCSR), University of Surrey, Guildford, UK. He has over 250 publications, including a book on low bit-rate speech coding and several book chapters. He has graduated more than 40 Ph.D. students in the areas of speech/image and signal processing and wireless multimedia communications, and has been a Consultant for major wireless media terminal developers and manufacturers. Professor Kondo has been awarded several prizes, the most significant of which are The Royal Television Societies' Communications Innovation Award and The IEE Benefactors Premium Award. He has been on the Refereeing College for EPSRC and on the Canadian Research Councils. He is a Member of the IEEE and the IET.

