Unequal Error Protection for Backward Compatible 3-D Video Transmission over WiMAX

Chaminda T.E.R. Hewage, Z. Ahmad, Stewart. T. Worrall, S. Dogan, W.A.C. Fernando, and A. Kondoz
I-LAB, Centre for communications Systems Research, University of Surrey, Guildford, GU2 7XH, UK
(e.thushara, zaheer.ahmad, s.worrall, s.dogan, w.fernando, a.kondoz)@surrey.ac.uk

Abstract—In this paper, an Unequal Error Protection (UEP) scheme for the transmission of 3-D (Three-Dimensional) video over WiMAX communication channel is proposed. The colour plus depth map stereoscopic video is coded with backward compatibility using a Scalable Video Coding (SVC) architecture, where users with conventional video decoders/receivers can receive the conventional 2-D (Three-Dimensional) video stream whereas users with SVC decoders/receivers and necessary 3-D video displays may render 3-D video. The proposed error protection scheme is based on the perceptual importance of the coded 3-D video components. The UEP method allocates more protection to the colour video packets than the depth map packets in order to receive good quality 2-D/3-D video. The protection levels are assigned according to their perceptual importance to the decoded model quality. An UEP scheme is based on the estimated distortion of the colour image sequence. The objective and perceptual quality evaluations show that the proposed UEP scheme improves the quality of 2-D video while achieving pleasing quality for 3-D viewers.

I. INTRODUCTION

The data losses that occur during transmission affect the reconstructed quality of 2-D video. The effects of transmission errors on the perceived quality of 3-D video could not be less than for the equivalent 2-D video applications, because the errors will influence several perceptual attributes (e.g., naturalness, presence, depth perception, eye-strain, viewing experience, etc.), associated with 3-D viewing. Therefore, 3-D video content needs to be protected when transmitted over unreliable communication channels.

The initial deployment of 3-D video services should ideally be backward compatible with 2-D video applications, as most consumers only have 2-D video receivers [1]. In the case of backward compatible 3-D video applications, conventional 2-D video users may receive poor quality video if the components of 3-D video are equally protected. For example, equally protected colour and depth map video transmission may introduce more errors to the colour video, which is viewed by the larger percentage of consumers. Therefore, the protection levels can be unequally allocated, giving more priority for the important data. This paper proposes a UEP scheme for colour plus depth map 3-D video transmission over WiMAX through allocating different transmission power to individual bit-streams according to their perceptual importance.

UEP is one of the most effective techniques for addressing the quality degradation caused by channel errors. This technique focuses on providing added protection for the most important component of the coded data. The UEP schemes which exploit the characteristics of 2-D video are common in research literature. For example, Hannuksela et al. propose an UEP scheme based on Region Of Interest (ROI) coding [2]. The proposed UEP schemes for scalable video streaming over wireless communication channels [3-4] take the error sensitivity of each coded layer into consideration to provide unbalanced protection. UEP schemes using the perceptual importance of 3-D video have not been thoroughly investigated to date. A prioritized left-and-right view stereoscopic video streaming with estimated distortion and UEP is discussed in [5]. Moreover, the study described in [6] proposes an UEP scheme for progressively transmitted 3D models based on their importance to the decoded model quality. An UEP scheme using redundant motion information (i.e., more protection for the motion information of the coded data) for colour and depth map stereoscopic video streaming is described in [7]. Most of the UEP schemes are implemented with the use of channel coding. However, this introduces more overhead to the encoded bit-stream. The amount of overhead data (i.e. channel coding) added to the encoded 3-D video bit-stream will be much higher than for conventional 2-D video due to the availability of more than one video bit-stream in coded 3-D video. Moreover, this will place more demand on system resources such as bandwidth. Therefore, alternative concepts for achieving UEP have emerged, such as allocating different transmission powers to different parts of the video. For example, the study described in [8] allocates different transmission powers to individual bits according to their bit error sensitivities. Moreover, this method allows optimum use of allocated power for the video service. The UEP scheme proposed in this paper achieves different levels of protection through allocation of unequal transmission power for the 3-D video components. With this approach, the 3-D video packets with more power are more likely to survive.
than the packets with less power. The coded 3-D video data is categorized based on perceptual importance and distortion estimates before allocation of unequal error protection levels.

This paper is organized as follows. Section II describes the proposed UEP scheme for 3-D video in detail. The power allocation module associated with this method is also presented. The experimental work and results are discussed in Section III. Section IV concludes the paper.

II. THE UNEQUAL ERROR PROTECTION SCHEME

The colour and depth map based 3-D video representation is widely utilized in research and standardization activities due to its simplicity and adaptability [1,9,10]. The coded colour image sequence is projected to the 3-D space, based on the depth map information, using the Depth Image-Based-Rendering (DIBR) method described in [11]. The colour video is the only texture information that is directly viewed by the users. Therefore, the loss of colour video packets will be more annoying to users (i.e. both 2-D and 3-D viewers) than the loss of depth video packets. Furthermore, the colour video itself provides certain depth clues (e.g., identification of objects in the foreground and background) during 2-D viewing. Therefore, the colour video stream should be allocated more protection compared to the depth map packets in general.

Accurate depth maps are required to generate good quality stereoscopic video using the DIBR technique. However, the quality of the depth map does not need to be very high to deliver the good quality 3-D video. Coarsely quantized depth images achieve similar stereoscopic video quality compared to 3-D video rendered with good quality depth images [12]. Furthermore, the effect of depth map transmission errors on the quality of reconstructed 3-D video is insignificant compared to the quality degradation due to the loss of colour video packets [12]. The study in [12] also highlights the importance of prioritizing colour video data ahead of depth map information in order to improve the perceived 3-D video quality. Moreover, the quality evaluation carried out in [13] shows that the effect of packet losses on depth perception is lower compared to the effect of packet losses on overall image quality. The scheme proposed in this paper therefore allocates more protection for the colour image packets than for the depth map packets in order to improve the quality of received 2-D/3-D video. The protection levels are determined based on the allocated transmission power for colour and depth video data.

The proposed UEP scheme is shown in Fig. 1. The colour plus depth stereoscopic video is coded using a SVC configuration based on the scalable extension of H.264/AVC. The colour and depth video are coded at the base and enhancement layers respectively. This configuration is backward compatible for 2-D video. A performance analysis of this encoding architecture compared to other 3-D coding configurations is provided in [14]. The extractor module separates the colour and depth packets into two packet streams before feeding them into the WiMAX transmission system. The UPA module, shown in Fig. 1, allocates different power levels to the colour and depth bit-streams. The functionality of the UPA module is described in the following section.

Figure 1. The proposed UEP scheme based on UPA

A. Prioritization of 3-D video packets

The UPA module in Fig. 1 receives channel feedback information from the WiMAX transmission system. It is assumed that the scheme is to operate on pre-encoded video, and therefore the packet loss rate mainly depends upon the transmission rate, transmission power, interference from other users, and losses in the channel. However, power ‘p’ is the dominant factor, and the distortion is denoted by D(r,p), where r denotes the rate information. The problem statement to minimize this distortion due to channel characteristics can be defined as:

\[
\min D(r,p) \text{ subject to } \sum_{m=1}^{M} r_m \leq R_t \text{ and } \sum_{m=1}^{M} r_m \leq R_c
\]

where \(P_t\) and \(R_c\) are the total power and rate budget for a user, and \(M\) is the number of users in the cell.

A suboptimal solution to this problem is proposed below for fast implementation. \(N\) numbers of subcarriers need to be distributed among \(L\) video layers for \(M\) users. The transmitter has information about the channel conditions, which is updated periodically via feedback channels. First, \(N\) subchannels are distributed among \(M\) users, based upon rate, starting with distribution of equal power to all users, such that:

\[
P_{m,n} = \frac{P_{\text{final}}}{N}
\]

where \(P_{m,n}\) is the power of the \(m^{th}\) user’s \(n^{th}\) subchannel. The rate assigned to a user is given by:

\[
R_{m} = \sum_{n=1}^{N} \log_2(1 + P_{m,n} G_{m,n} R_{c,n} B)
\]

where \(G_{m,n}\) is the \(n^{th}\) channel gain of the \(m^{th}\) user with a noise density of \(N_0\). The overall available bandwidth is given by \(B\), and \(R_c\) is the transmission rate of the user \(m\). For \(M\) users, the optimization of the power can be expressed as:

\[
\max P_{m,n} \left( \sum_{m=1}^{M} \sum_{n=1}^{N} \log_2 \left(1 + P_{m,n} G_{m,n} R_{c,n} B \right) \right)
\]

The solution of (4) may be found using the methods given in [15]. The first step of the algorithm is the initialization of power levels, and then determination of the number of subchannels for each user in the cell.

1-Initialize the set of users, \(A = \{1,2,3,\ldots,M\}\), and \(R_m = 0\)
2-For \(m=1\) to \(M\) \{
   Find \(n\) satisfying \(G_{m,n} \geq G_{m,j}\) for all \(j \in A\)
   Calculate \(R_m\) using (3)\}
3-While \(A \neq \emptyset\)
   \(\text{Find} m \text{ so that } R_m \leq R, \text{ for all } i \leq M\)
4-For above \(m\), find \(n\) satisfying \(G_{m,n} \geq G_{m,j}\) for all \(j \in A\)
The above algorithm will assign subcarriers to each user in the cell, in proportion with the rate requirement.

**B. Distribution of Subcarriers of a User among 3-D Video Layers**

After all users have been assigned subcarriers, each user has a power budget and a rate budget. Assume $N_{m,s}$ subchannels are available to the $m$th user for transmission of data. An extractor in the base station analyzes the importance of the data packet. Colour packets in the base layer are given priority over depth information in the enhancement layer. Based upon feedback about the packet loss ratio, the UPA algorithm calculates power for the subchannels with the help of a lookup table as shown in Table II. The table translates the packet loss rate information into distortion levels, which in turn is used to distribute the power among L video layers for that particular user. In this paper we have used only 2 layers (i.e., colour and depth-bit streams). The objective of the UPA algorithm at this stage is to minimize the received distortion:

$$\min \text{Distortion} \quad (p_{m,s}) \quad \text{subject to} \quad \sum p_{m,s} \leq P_{m,s}$$

(5)

where $P_{m,s}$ is the total power budget for user $m$. We assume all subcarriers assigned to a user have equal power. This means allocation of two subcarriers to a particular layer doubles the power, and so forth. Hence, the power increment is available in fixed step increments of $\Delta P_i$. The UPA algorithm is provided with information about the number of layers being transmitted to the user, the condition of the subcarrier channels $N_{m,s}$, and the number of subcarriers available. The algorithm for distribution of power to video layers is as follows:

1. Initialize power for all subchannels $N_{m,s}$ to $P_{m,s}/N_{m,s}$
2. Distribute video layers equally among subchannels $N_{m,s}$
3. At each feedback interval, distribute the power among layers based upon feedback information of estimated distortion. If $D_{\text{Estimated}} < D_{\text{Required}}$:
   - Increment $N_{m,s,b}$ according to Table II
   - Decrement $N_{m,s,e}$ according to Table II

where $N_{m,s,b}$ is the number of subchannels assigned to base layer, and $N_{m,s,e}$ is the number of subchannels assigned to enhancement layer.

### III. EXPERIMENTAL SETUP AND DISCUSSION

This section describes the experimental procedure and the simulation environment for evaluation of the proposed UPA scheme for backward compatible 3-D video.

**A. Parameter Settings**

The Orbi and Interview sequences (Resolution: 720x576 and frame rate: 25 frames/s) are coded offline using the JSVM reference software codec version 8.9. The basic encoding parameters are: IPPP... sequence format, a single reference frame, Content Adaptive Binary Arithmetic Coding (CABAC), and 200 frames. The average encoded bitrate is approximately 1Mbps for both sequences. The bit-stream extractor module of JSVM is utilized to extract base and enhancement layers and to analyze the transmission bandwidth requirement. The lost frames of each layer are concealed using frame copy method at the JSVM decoder.

A WiMAX baseband error trace simulator is utilized to generate error traces for a range of SNR levels, with the parameters given in Table I [16]. System level simulations are performed to analyze the effects of unequal power allocation on prioritized colour plus depth 3-D video. This models a time varying channel including the effect of multipath for the ITU Vehicular A scenario. The total received power can be mapped to an SNR value, and the pre-simulated error trace that has the closest SNR is utilized to corrupt the video stream transmitted through the simulator. The power distribution among layers is adjusted based on the feedback about the received video distortion according to lookup tables similar to Table II. These tables are obtained experimentally, and describe the best distribution of power amongst the base and enhancement layers. The simulations are run ten times to obtain stable results.

### TABLE I. SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of Trace (s)</td>
<td>10</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>10MHz</td>
</tr>
<tr>
<td>Permutation</td>
<td>PUSC</td>
</tr>
<tr>
<td>MCS5</td>
<td>16 QAM 1/2</td>
</tr>
<tr>
<td>Channel Coding</td>
<td>CTC</td>
</tr>
</tbody>
</table>

### TABLE II. EXAMPLE LOOKUP TABLE FOR MCS5 WITH $\Delta P_i=0.55$

<table>
<thead>
<tr>
<th>SNR Level (dB)</th>
<th>PLR</th>
<th>BER</th>
<th>BL:EL (a)</th>
<th>$D_{\text{Estimated}}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.30</td>
<td>0.0453449</td>
<td>0.2806538</td>
<td>2.13</td>
<td>21.44</td>
</tr>
<tr>
<td>8.85</td>
<td>0.0273837</td>
<td>0.176291</td>
<td>2.02</td>
<td>22.00</td>
</tr>
<tr>
<td>9.40</td>
<td>0.0153232</td>
<td>0.103876</td>
<td>1.97</td>
<td>23.01</td>
</tr>
<tr>
<td>9.95</td>
<td>0.0077446</td>
<td>0.0553054</td>
<td>1.76</td>
<td>24.32</td>
</tr>
<tr>
<td>10.50</td>
<td>0.0034710</td>
<td>0.0206157</td>
<td>1.59</td>
<td>26.26</td>
</tr>
<tr>
<td>11.05</td>
<td>0.0014022</td>
<td>0.0109829</td>
<td>1.51</td>
<td>28.30</td>
</tr>
<tr>
<td>11.60</td>
<td>0.0005203</td>
<td>0.0042650</td>
<td>1.47</td>
<td>33.30</td>
</tr>
<tr>
<td>12.15</td>
<td>0.0001747</td>
<td>0.0014701</td>
<td>1.25</td>
<td>36.40</td>
</tr>
<tr>
<td>12.70</td>
<td>0.0000615</td>
<td>0.0005000</td>
<td>1.07</td>
<td>36.80</td>
</tr>
<tr>
<td>13.25</td>
<td>0.0000243</td>
<td>0.0002094</td>
<td>1.00</td>
<td>36.80</td>
</tr>
</tbody>
</table>

\(a\) BL:EL = Base Layer : Enhancement Layer Power Ratio

**B. Results and Discussion**

Table III presents the performance with and without the proposed UEP scheme for the Orbi and Interview sequences respectively. The colour image sequence achieves improved image quality with the proposed UEP scheme for all the error conditions. This is mainly due to the allocation of more power (more subcarriers) to the colour video packets than depth video packets. Consequently, the colour bit-stream packets have survived under error conditions than depth bit-stream packets. For the proposed UEP scheme, the depth quality degrades rapidly as channel errors increase. In order to evaluate the effect of impaired depth images to the perceived quality subjective evaluation tests are conducted. Two 3-D perceptual attributes namely, overall image quality and depth perception are measured. Figures 2 and 3 show improved overall image quality with the proposed UEP method. However, the subjective results show no significant effect to depth perception from the resultant depth images with the proposed UEP scheme. Moreover, the depth perception has improved at high error conditions due to the availability of improved texture information.
good quality stereoscopic video even with impaired depth map sequences, less protection has been allocated for associated depth maps. The proposed UEP scheme is implemented using a UPA algorithm which allocates transmission power for colour and depth video sequences based on their distortion estimates. The results show that, the colour image quality achieved superior image quality under a range of channel error conditions. Even though the depth quality is reduced with the proposed scheme, the effect on the perceived quality of reconstructed stereoscopic video is insignificant. Although quality has been optimized in this paper, it should also be possible to optimize the power to achieve energy efficient video transmission.

REFERENCES