

Adaptive Multiple Description Coding and Transmission of Uncompressed Video over 60GHz Networks *

Zhifeng He

zzh0008@tigermail.auburn.edu

Department of Electrical & Computer Engineering, Auburn University, Auburn, AL, USA

Shiwen Mao

smao@ieee.org

While many cognitive radio (CR) techniques are developed to better utilize the allocated spectrum, the massive unlicensed bandwidth in the 60GHz band also provides great potential for supporting new bandwidth intensive applications. In this paper, we investigate the problem of streaming uncompressed High Definition (HD) videos over 60GHz networks. We present an adaptive multiple description (MD) coding (MDC) technique based on Priority Encoding Transmission (PET) that exploits the different significance of the pixel bits, and an interleaving based transmission scheme to combat the bursty losses due to blockage. A nonlinear integer programming problem is formulated and solved with a heuristic approach for determining the sub-optimal coding and transmission parameters. The proposed scheme is adaptive to the dynamic 60GHz link conditions for enhanced video quality. The performance of the proposed scheme is validated with simulations.

I. Introduction

With the dramatic advances in wireless networking technology, there is an exponentially increasing demand for wireless data service. In particular, mobile video is predicted to grow at a compound annual growth rate (CAGR) of 90% from 2011 to 2016 [4]. Cognitive radio (CR) has been recognized as an important technology for enhancing spectrum efficiency, while many dynamic spectrum access techniques are developed to better utilize the allocated spectrum [21]. On the other hand, millimeter wave (mm-wave) communications in the 60GHz band has gained considerable interest from academia, industry, and standards bodies [14]. There is huge unlicensed bandwidth (i.e., up to 7 GHz) in the 60GHz band that is available in most parts of the world. In addition to indoor use, FCC recently updated the rules for the unlicensed 60GHz band, which will allow higher emission limits for 60GHz devices that operate outdoors, thus enabling broader deployment of point-to-point broadband systems [5]. The massive unlicensed bandwidth provides great potential to meet the surging wireless video demand, as well as supporting new bandwidth demanding applications [6, 18].

Recently, the problem of transmitting uncompressed High Definition (HD) videos via 60GHz channels has attracted intensive interest. The immense bandwidth of the 60GHz channel enables streaming of uncompressed HD videos with high data rates [7], such as the 1080p video, which has 1920×1080 pixels (each of which has 24 bits) per video frame and the frame size is about 6 MBs. Such high speed wireless links can not only replace the cumbersome HDMI cables, but also relieve the computational burden for video decoding at the display devices (e.g., projectors or HDTVs).

Although the larger bandwidth is a great advantage of 60GHz networks, the 60GHz transmissions are highly directional to overcome the high attenuation, making it sus-

ceptible to blockage of the line-of-sight path by obstacles or pedestrians. Thus the memory of the channel poses a negative effect on the performance of the system, and channel conditions such as packet loss rate may change over time [14]. How to adjust the operation parameters of the transmission system is critical to maintain a certain level of QoS under such conditions.

In this paper, we investigate the problem of streaming uncompressed HD video over 60GHz networks. Considering the fact that bits of different positions in a pixel have different levels of significance on recovering the original video frame [8], it is desirable to improve the quality of reconstructed video frames by offering more protection for the more important bits. We adopt multiple description (MD) coding (MDC) to combat the lossy 60GHz channel. In MDC, a video is encoded into multiple descriptions, each can guarantee a low but acceptable video quality. More important, the more descriptions received, the better the video quality. MDC has been used mostly for compressed videos in lossy wireless or wireline networks in the literature [10, 11].

In this paper, we adopt MD coding for uncompressed videos, exploiting the different significance of the different bits in a pixel. In particular, in order to ensure the quality of the reconstructed video, we divide the bits in a pixel into multiple segments according to their respective significance in enhancing video quality, and provide higher protection to bits of greater importance by assigning more forward error correction (FEC) symbols to them. This way, the decoder has a higher probability of recovering the most important bits, which affect the quality of the reconstructed image to a larger extent. Such an MD coder is termed MD-FEC. How to partition the bits, i.e., how many segments the bits should be divided into, and how many bits a particular segment should have, affects the significance of a particular segment and the probability of decoding the segment.

To combat the bursty errors while transmitting over the 60GHz channel, interleaving among multiple video frames is applied in our scheme. We formulate a nonlinear integer programming problem, which can be solved to find

*This work is based on an earlier work: Multiple Description Coding for Uncompressed Video Streaming over 60GHz Networks, in ACM CRAB'13, ©ACM, 2013. <http://doi.acm.org/10.1145/2508478.2508485>.

Table 1: Notation

<i>Symbol</i>	<i>Description</i>
M	Number of rows of the image matrix
N	Number of columns of the image matrix
C	Number of rows of the sub-image matrix
D	Number of columns of the sub-image matrix
Q	Number of sub-images of an image
S	Number of segments the 8 bits are divided into
m_i	Length of the i -th segment, $1 \leq i \leq S$
k	Number of bits constituting a symbol
g_i	Number of groups the data symbols of stream i are divided
L	Code word length of the RS code
$P_e(G)$	Probability of an error occurring in the good state
$P_e(B)$	Probability of an error occurring in the bad state
$P_{g b}$	Channel state transition probability from good to bad
$P_{b g}$	Channel state transition probability from bad to good
H	Number of images interleaved together
Δ	Maximum tolerable time of recovering the image
Λ_l	Minimum packet size
Λ_u	Maximum packet size

the optimal partition of the pixel bits in the MD coder, as well as the optimal interleaving depth, for given channel conditions. To reduce computational complexity, we solve the problem with a heuristic approach for sub-optimal solutions. The proposed scheme is adaptive to the dynamic 60GHz link conditions for enhanced video quality. It is evaluated with simulations and is shown to outperform an existing scheme reported in [17] with considerable gains.

The remainder of the paper is organized as follows. Related work is reviewed in Section II. We present the proposed MD coding and transmission scheme in Section III and simulation results in Section IV. Section V concludes the paper. The notation used in this paper is summarized in Table 1.

II. Related Work

In this section, we briefly review related work on video transmission over 60GHz channels, and two MD-FEC techniques and Reed-Solomon coding, which are the basis of our proposed scheme.

Considering the similarity in the most important bits of the neighboring pixels in an image, the authors of [8] propose a technique to correct the error in the most important bits in a pixel by comparing those bits of the pixel with the bits at the same positions in its neighboring pixels. Shao et al. in [16] also develop a method to recover the bits

of a pixel using the neighboring pixels. With this scheme, the neighboring pixels are encapsulated into different packets that are transmitted separately, so that a lost pixel can still be approximated by other received packets containing the neighboring pixels. However, there is no measures on providing better protection for the most important bits during transmission, which affects the quality of reconstructed video frames.

In [17], an Uncompressed Video streaming over Wireless (UVoW) system is proposed for transmission of uncompressed video on 60GHz channels. In this system, the 8 bits representing a pixel are divided into two 4-bit parts according to their significance in recovering the video: MSB (most significant bits) and LSB (least significant bits). The MSB portions are better protected than the LSB portions by being transmitted through a channel of better condition. The encoder retransmits the MSB part if it has an error, but corrupted LSB parts will not be retransmitted. However, there is no evidence that the system achieves the best performance by cutting the 8 bits into two 4-bit segments, and there is no investigation on, when channel condition changes, how to adjust the operational parameters to make it adaptive to the changing channel conditions.

In [3], Bosco develops a cross-layer adaptive scheme for HD videos streaming over 60GHz channels, where the sender adjusts the modulation and coding schemes (MCS) in response to channel variations to maintain good video quality. A method of cooperation with relay for transmission of HD videos over 60GHz channels is developed in [7]. The source and relay use different error correcting codes so that error correcting performance can be enhanced, and spatial diversity gain is obtained by using the relay. As in [17], the message is also cut into two fixed parts and the optimal segmentation of the message is out of concern in this related work [7].

MD video coding has been shown to be highly effective for video over lossy networks [2, 9, 11, 12, 19]. Among the many MD coding techniques, MD-FEC is a popular one based on Priority Encoding Transmission (PET) [1], which assigns different levels of protection to data segments according to their respective priority defined by the user [19]. The higher the priority, the smaller number of packets the decoder needs to decode it, so that the higher priority segments have a greater probability of being decoded. The basis of this algorithm is Reed-Solomon (RS) code, which is a sub-class of the Bose-Chadhuri-Hocquenghem (BCH) codes. RS coding provides an effective way of recovering lost data symbols from a portion of data symbols received [15].

III. System Model, Problem Statement and Solution Procedure

We consider streaming uncompressed HD video over a 60GHz wireless link (e.g., from a storage device to an HDTV or a video projector). The MD-FEC coding scheme for uncompressed HD video is introduced in Section III.A. The 60GHz channel model and the interleaving based transmission strategy is presented in Section III.B. Then

R								G								B							
1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8

Figure 1: Structure of a pixel in the uncompressed video.

we present our problem formulation in Section III.C and solution procedure in Section III.D.

III.A. MD-FEC Coding of Uncompressed Video

Consider an uncompressed video frame of $M \times N$ pixels. Each pixel consists of $3 \times 8 = 24$ bits, while each 8-bit block corresponding to one of the R/G/B color components as shown in Fig. 1. The j -th bit in each of the 8-bit block represents a value of 2^{8-j} . Obviously, the j -th bit is more important than the $(j+i)$ -th bit, for all $i \geq 1$, because it has a greater influence on the color depth of the R/G/B component of the pixel.

Next we divide the frame evenly into Q sub-images, each of which has $C \times D$ pixels, i.e., $Q = \frac{MN}{CD}$, for the purpose of:

- the packet size will not become too large so that the packet transmission time can be fitted in the unit time slot in our channel model;
- the number of data symbols of a data stream will not become too large so that RS codes can be applied to the data symbols.

Now let's take the r -th sub-image as an example; each of the following procedures will be performed on other sub-images as well. First, each of the 8 bits representing the R/G/B color in a pixel is divided into S segments, $1 \leq S \leq 8$, and the length of (i.e., the number of bits in) segment i , is denoted as m_i , $1 \leq i \leq S$. Rearrange all the i -th segments of the R/G/B color of all the pixels in a sub-image into a new segment called *data stream i* , which has $\frac{3m_i MN}{Q}$ bits. We define the *significance* of data stream i of sub-image r as the summation of the values of all the data bits in segment i of sub-image r , and denote it as w_i^r .

$$w_i^r = \frac{3MN}{Q} \sum_{l=1}^{m_i-1} 2^{8-\sum_{j=1}^{i-1} m_j - l}. \quad (1)$$

Due to uncompressed video, there is no reason to prefer one pixel over another when semantic content is not considered. We thus have $w_i^r = w_i^q$ for any $r \neq q$, $1 \leq r \leq Q$, $1 \leq q \leq Q$. Fig. 2 shows an example of cutting each of the 8 bits of a pixel in a sub-image into 3 segments and combining all the first segments into data stream 1 of the sub-image.

From (1), we can see that $w_i^r \geq w_{i+1}^r$, for all $1 \leq r \leq Q$, $1 \leq i \leq S-1$, which implies that recovering data stream i is more important than recovering data stream $(i+1)$ for the purpose of enhancing the quality of reconstructed frames at the receiver. Therefore, we assign higher protection to data stream i than data stream $(i+1)$ to make the probability of recovering data stream i higher than that of recovering

data stream $(i+1)$. We illustrate how to achieve this goal below.

First, we let k bits in a data stream form a *data symbol*. For example, we can take $k = 8$ for the convenience of computer processing. Then data stream i can be represented by $\frac{3m_i MN}{Q^k}$ data symbols. These $\frac{3m_i MN}{Q^k}$ symbols of data stream i is further divided into g_i groups, each of which has $\frac{3m_i MN}{Q^k g_i}$ data symbols. We apply RS code to each group so that the total number of symbols (including both data and FEC symbols) is L , $L \leq 2^k - 1$, which is a constraint of RS coding.

This way, the number of FEC symbols assigned to each group of data stream i is $L - \frac{3m_i MN}{Q^k g_i}$. The j -th symbol of all the groups of all the data streams compose the j -th *description* of the frame, and is transmitted in one packet. This procedure is illustrated in Fig. 3. To assign higher protection to data stream i than data stream $(i+1)$, the error correcting capacity, or the amount of FEC symbols for data stream i should be greater than that of data stream $(i+1)$, i.e., $L - \frac{3m_i MN}{Q^k g_i} \geq L - \frac{3m_{i+1} MN}{Q^k g_{i+1}}$, which implies that

$$\frac{m_i}{g_i} \leq \frac{m_{i+1}}{g_{i+1}}. \quad (2)$$

As shown in Fig. 3, each description (or packet) consists of video data bits and FEC bits from all the data streams, and is equally important for reconstructing the frame. The quality of the reconstructed frame is proportional to the number of descriptions received. If $\frac{3m_1 MN}{Q^k g_1}$ or more error-free descriptions are received, then all the groups of data stream 1 can be decoded, which guarantees a basic quality for this video frame. In general, if the decoder receives at least $\frac{3m_i MN}{Q^k g_i}$ error-free descriptions, then all the data streams up to i can be decoded, resulting in an improved quality for the frame. When more than $\frac{3m_S MN}{Q^k g_S} \leq L$ error-free descriptions are received, the video frame can be reconstructed without any error.

III.B. 60GHz Channel and Transmission Schedule

After MD-FEC coding, the video packets are then transmitted over the 60GHz link from the sender (e.g., a storage device) to the receiver (e.g., an HDTV or video projector). As discussed, the 60GHz channel is highly directional and susceptible to blockage of the link-of-sight path. Therefore, we model the 60GHz link with the Gilbert-Elliot (GE) model, which is a discrete-time two-state Markov chain as shown in Fig. 4 [20]. The two states are: good state (denoted as G in the figure) and bad state (denoted as B). The probabilities that an error occurring at the good state and bad state are denoted as $P_e(G)$ and $P_e(B)$, respectively. The transition probabilities from G to B and from B to G are $P_{g|b}$ and $P_{b|g}$, respectively. Because the Gilbert-Elliot channel is a discrete-time model instead of a continuous-time model, we assume time is slotted and the transmission time of a packet over the 60GHz link is less than or equal to the duration of one time slot [2, 11].

When multiple beam paths are available between the transmitter and receiver (e.g., a direct path and a second

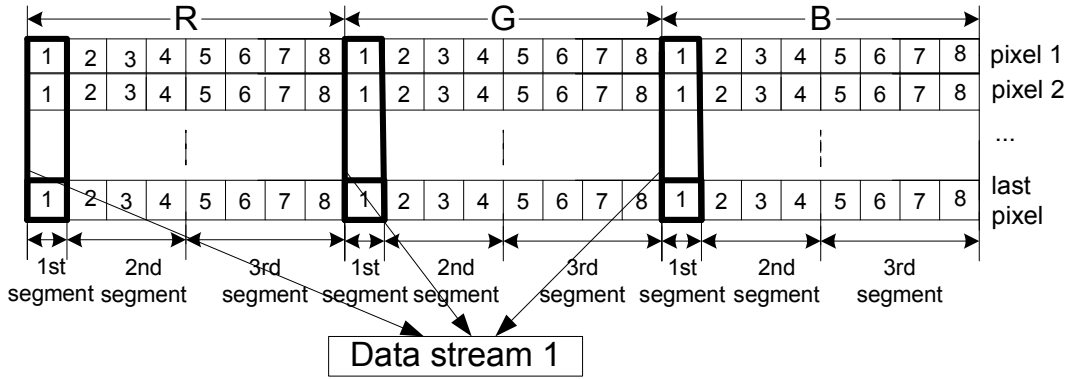


Figure 2: Structure of a data stream, assuming $S = 3$.

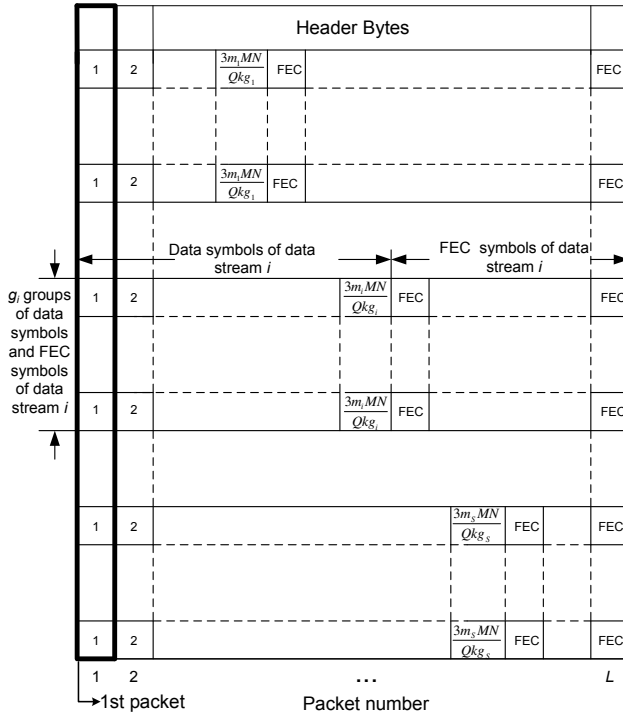


Figure 3: Structure of the packets of a sub-image.

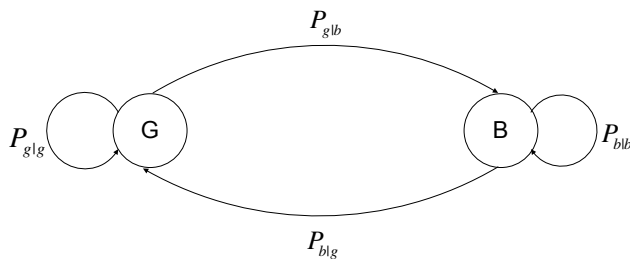


Figure 4: The Gilbert-Elliot link model.

path due to reflection on the wall), the multipath transport technique can be used for the MD video packets [9]. In this paper, we focus on the single beam path case, while the proposed technique can be extended to the case of multiple beam paths. Recall that MD-FEC coding can ensure a minimum video quality if at least $\frac{3m_1MN}{Qkg_1}$ error-free packets are received for a frame. However, if most of the packets

are transmitted within a bad period, it is possible that less than $\frac{3m_1MN}{Qkg_1}$ packets can be received and the video frame cannot be decoded. In order to reduce the bursty error rate, we adopt *interleaving* to the packets of multiple consecutive frames.

The interleaving scheme is illustrated in Fig. 5. The *interleaving depth* is the time interval between the transmissions of two consecutive packets of the same sub-image. As shown in the figure the interleaving depth is set to QH , while H is a nonnegative integer. We can adjust H to find the best interleaving depth for different channel conditions in terms of $P_e(G)$, $P_e(B)$, and the transition probabilities. when the average bad period T_{off} is large, we can increase QH by increasing H . The probability that multiple consecutive packets of a sub-image are transmitted within the bad period can be reduced. Since RS coding is applied in our scheme, the probability of failing to decode data stream i of sub-image X of image Y is the probability that the decoder receives less than $\frac{3m_iMN}{Qkg_i}$ error-free data packets, which can be expressed as [20]

$$P_i^{XY} = \sum_{d=0}^Z P_Z(d) \left\{ \sum_{e_b=0}^d \left[C_d^{e_b} P_e^{e_b}(B) (1 - P_e(B))^{d-e_b} \sum_{e_g=\max\{0, t+1-e_b\}}^{Z-d} C_{Z-d}^{e_g} P_e^{e_g}(G) (1 - P_e(G))^{Z-d-e_g} \right] \right\}, (3)$$

where $P_Z(d)$ is the probability that given the GE channel is observed at Z continuous time slots, the probability that the channel is in the bad state for d time slots; Z is the block code length of the error correcting code being used and thus $Z = L = 2^k - 1$ in our example; e_b and e_g denote the numbers of packets in error when the channel is the bad state and good state, respectively; t is the error correcting capability of the error correcting code for data stream i ; and $P_i^{XY} = P_i^{X'Y'} = P_i$, for any $X \neq X'$ and $Y \neq Y'$, due to uncompressed videos. Since the sequence numbers of all the data symbols are known, the RS code can correct up to $t = L - \frac{3m_iMN}{Qkg_i}$ known erasures as an erasure code [13].

For brevity, retransmission is not explicitly implemented in our system. Therefore, the transmission time for all the packets of a data stream i is LHQ time slots. A retransmission scheme can be translated to reduced $P_e(G)$ and $P_e(B)$, but the transmission times for the data streams

could be longer [17].

III.C. Problem Statement

Now we are ready to formulate the problem of MD video over 60GHz networks. Due to uncompressed video, the range of PSNR is from 0 (all lost) to positive infinit (error free), which makes no practical sense. On the other hand, the Mean Squared Error (MSE) of the reconstructed image reflect the distortion of the reconstructed image compared with the original image. Therefore our goal is to minimize the expected Mean Squared Error (i.e., MSE or $\sqrt{\text{MSE}}$) of the reconstructed video frames, which is a function of the various design factors as follows.

$$\text{MSE} = \frac{1}{3MN} \sum_{i=1}^M \sum_{j=1}^N \sum_{\chi=1}^3 [I(i, j, \chi) - I'(i, j, \chi)]^2. \quad (4)$$

In (4), $I(i, j, \chi)$ is a pixel in the original video frame at location (i, j) and $I'(i, j, \chi)$ is the pixel at the same location in the reconstructed video frame. The index χ indicates the color components (i.e., R, G or B). Since it is assumed that the three color components are equally important for the image quality, we don't introduce different weights for the three color components.¹ It follows that

$$\begin{aligned} & \mathbb{E}\{\sqrt{\text{MSE}}\} \quad (5) \\ &= \mathbb{E} \left\{ \sqrt{\frac{1}{3MN} \sum_{q=1}^Q \left[\sum_{i=1}^C \sum_{j=1}^D \sum_{\chi=1}^3 \left(\sum_{r=T_q}^{S_q} \frac{Q}{3MN} w_q^r \right)^2 \right]} \right\} \\ &= \mathbb{E} \left\{ \sqrt{\frac{1}{3MN} \sum_{q=1}^Q \left[3CD \left(\sum_{r=T_q}^{S_q} \frac{Q}{3MN} w_q^r \right)^2 \right]} \right\} \\ &= \mathbb{E} \left\{ \sqrt{\frac{1}{Q} \sum_{q=1}^Q \left(\sum_{r=T_q}^{S_q} \frac{Q}{3MN} w_q^r \right)^2} \right\} \\ &= \sum_{T_1=1}^{S_1} \cdots \sum_{T_Q=1}^{S_Q} \left\{ \left[\prod_{q=1}^Q P(x_{T_{q-1}} \leq X_q < x_{T_q}) \right] \times \right. \\ & \quad \left. \sqrt{\frac{1}{Q} \sum_{q=1}^Q \left(\sum_{r=T_q}^{S_q} \frac{Q}{3MN} w_q^r \right)^2} \right\} \\ &= \sum_{T_1=1}^{S_1} \cdots \sum_{T_Q=1}^{S_Q} \left\{ \left[\prod_{q=1}^Q (P_{T_q} - P_{T_{q-1}}) \right] \times \right. \\ & \quad \left. \sqrt{\frac{1}{Q} \sum_{q=1}^Q \left(\sum_{r=T_q}^{S_q} \frac{Q}{3MN} w_q^r \right)^2} \right\} \end{aligned}$$

Where S_q , $1 \leq q \leq Q$ is the total number of segments that each pixel of the sub-image q is divided into; T_q is

¹The same scheme can be applied to videos stored in the YUV format, where the weights for the Y, U and V components may vary and three optimization problems need to be solved.

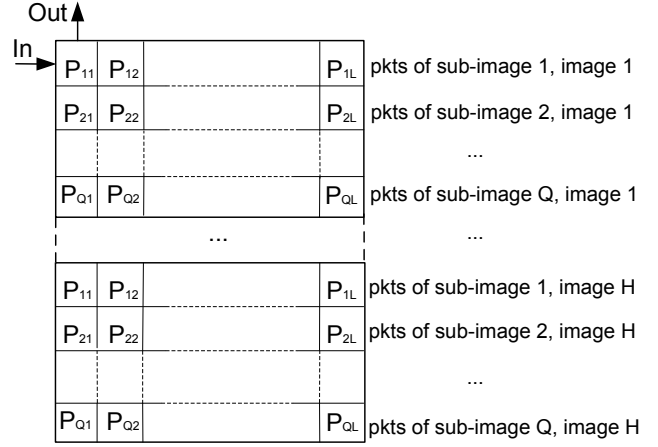


Figure 5: Structure of the interleaver.

the first segment of sub-image q that cannot be decoded; X_q is the number of packets of sub-image q that are received and not corrupted by error; x_{T_q} is the number of data symbols required to decode T_q (and denote $x_0 = 0$); $P(X_q < x_{T_q})$ is the probability that X_q is less than x_{T_q} , such that $P(X_q < x_{T_q}) = P_{T_q}$, where P_{T_q} is the probability that data T_q cannot be decoded, as given in (3). Note that we assume if a bit cannot be decoded then it is discarded at the decoder, meaning that (5) is indeed the worst-case expected $\sqrt{\text{MSE}}$.

We then formulate the following nonlinear integer programming (NLIP) problem.

$$\text{minimize } \mathbb{E}\{\sqrt{\text{MSE}}\} \quad (6)$$

$$\text{subject to: } \frac{m_i}{g_i} \leq \frac{m_{i+1}}{g_{i+1}}, \text{ for all } i \quad (7)$$

$$\frac{3m_{S_q}MN}{Qkg_{S_q}} \leq L, \text{ for all } q \quad (8)$$

$$L \leq 2^k - 1 \quad (9)$$

$$\sum_{i=1}^{S_q} m_i = 8, \text{ for all } q \quad (10)$$

$$LHQ \leq \Delta \quad (11)$$

$$\Lambda_l \leq \sum_{i=1}^{S_q} \frac{g_i k}{8} \leq \Lambda_u, \text{ for all } q. \quad (12)$$

Constraint (11) is due to the fact that the transmission time of all the packets of an image must not be larger than a tolerable transmission delay, i.e., the delay bound Δ . Constraint (12) is for the purpose of making the size of a packet adequate for transmission after adding header bytes to the packets, i.e., within the range of the minimum packet size Λ_l and the maximum packet size Λ_u .

III.D. Solution Strategy

The formulated problem aims to minimize the expected $\sqrt{\text{MSE}}$ given in (5). This goal is achieved by tuning the control parameters S_q , $\{m_1, m_2, \dots, m_{S_q}\}$, $\{g_1, g_2, \dots, g_{S_q}\}$, for all q , and H to find the optimal par-

tion of the pixel bits for MD coding and the optimal interleaving depth for transmission. And the problem is solved once when the channel condition is changed.

Checking out the search space for the optimal solution, it can be seen that: (i) the feasible values for S_q is between 1 and 8; (ii) due to constraint (10), the m_i 's are all small integers; and (iii) H is also limited by the maximum delay constraint (11). The parameters that dominate the search space of the problem are $\{g_1, g_2, \dots, g_{S_q}\}$, for all q , since the packet size can be a value between hundreds to thousands of bytes. This motivates us to reduce the search space by "taking out" $\{g_1, g_2, \dots, g_{S_q}\}$, for all q , such that the search space of the problem will be small enough to apply an exhaustive search for the remaining parameters. This way, the problem can be reduced to a sub-optimal problem having S_q , $\{m_1, m_2, \dots, m_{S_q}\}$, and H as its optimization parameters.

In our solution strategy, we reduce constraint (12) into $\sum_{i=1}^{S_q} \frac{g_i k}{8} = \Lambda$, where Λ is a constant, for all q . For a particular S_q , for all q , we have $g_i = \frac{8\Lambda}{kS_q}$, for all i . Furthermore, constraint (7) can be transformed to $m_i \leq m_j$, for all $i < j$. Then the problem is reduced to a sub-optimal problem of finding the optimal $\{m_1, m_2, \dots, m_{S_q}\}$ and H under a specific S_q value. The search space of the sub-optimal problem is now much smaller than that of the original problem, so that exhaustive search can be applied. The ideas of finding the best bits-partition scheme and interleaving depth are still maintained in the sub-optimal problem, thus the proposed scheme still being adaptive to channel dynamics. Note that for the case where $\frac{8\Lambda}{kS_q}$ is not an integer, we take $g_i = \lfloor \frac{8\Lambda}{kS_q} \rfloor$, for all $1 \leq i \leq S_q$ and the remaining space in the packet left by the data symbols are padded by redundancy.

In addition, we also aim to reduce the number of loops caused by Q , since a large number of loops will seriously drag down the speed of the search algorithm, as can be seen from (5). We cut the HD video image into many small images (different from the sub-images mentioned above) and take each of the small images as an intact image, and then interleaving is applied to these small images. Thus $Q = 1$ in the exhaustive search algorithm and our following simulations. And we omit the subscript q for brevity in the following. This way, the number of loops is reduced and simulation speed is greatly improved.

After these simplifications the computation overhead is dominated by the computation of (3), which is the probability of failing to decode a data stream, and by searching for the optimal pixel bit allocation and interleaving depth within a limited searching space. The computation can be further sped up by pre-computing the probabilities (3) under various cases and store them in a table for future use. The performance of this solution strategy will be demonstrated in the performance evaluation section.

IV. Performance Evaluation

In this section, we evaluate the performance of the proposed scheme. The coding and transmission schemes are

Table 2: Simulation Parameters

<i>Parameters</i>	<i>Value</i>
M	1080
N	1920
C	60
D	80
k	6
Δ	10^6 time slots
L	51
H	$1 \sim 5$
Λ_l	800 Bytes
Λ_u	2000 Bytes
$\sum_{i=1}^S \frac{g_i k}{8}$	900 Bytes

implemented and simulated using Matlab. In the simulations, we use 1920×1080 HD uncompressed images with 24 bits per pixel (i.e., 8 bits for each R/G/B component), instead of a real HD video stream. The simulation parameters are shown in Table 2. We compare the proposed scheme with UVoW presented in [17], as discussed in Section II. Note that ARQ is not employed in both schemes for a fair comparison. The channel conditions are assumed to be known in advance, e.g., through a proper channel measurement/feedback scheme. In the simulations, we assume a time slot duration of $1 \mu\text{s}$.

IV.A. Operation of the Proposed Scheme

We first examine the operation of the proposed scheme to reveal its properties. Figure 6 visually shows the different significance of bits at different positions for reducing $\sqrt{\text{MSE}}$. In this simulation, the segmentation of the eight bits of a color component in a pixel is 1-1-1-2-3, which means the eight bits are divided unevenly into five segments, with 1, 1, 1, 2, and 3 bits, respectively. The significance of the segments are 128, 64, 32, 24, and 7, respectively. When less than 10 packets are received, the video cannot be decoded, resulting in a $\sqrt{\text{MSE}}$ of 255. When more than 10 but less than 20 packets are received, the most important bits in the pixels can be decoded, resulting in greatly reduced $\sqrt{\text{MSE}}$. The $\sqrt{\text{MSE}}$ curve has a staircase shape and the gap is getting smaller as more and more packets are received.

We learn from the figure that recovering the bits of larger values has a greater effect on lowering the $\sqrt{\text{MSE}}$ than recovering the bits of smaller values, which is the guideline for the proposed algorithm. Recall that the total number of FEC symbols is $L \sum_{i=1}^S g_i - \frac{3 \cdot 8 \cdot M \cdot N}{k}$, where the first term is the total number of symbols and the second term is the total number of data symbols in a sub-image. Therefore, the amount of FEC symbols in a sub-image is fixed. If we want to provide stronger protection to the bits of greater values by assigning more FEC symbols to them, then the bits of smaller values will be less protected, and vice versa. If the channel is bad and the packet loss rate is high, then it is more urgent to offer higher level of protection to the

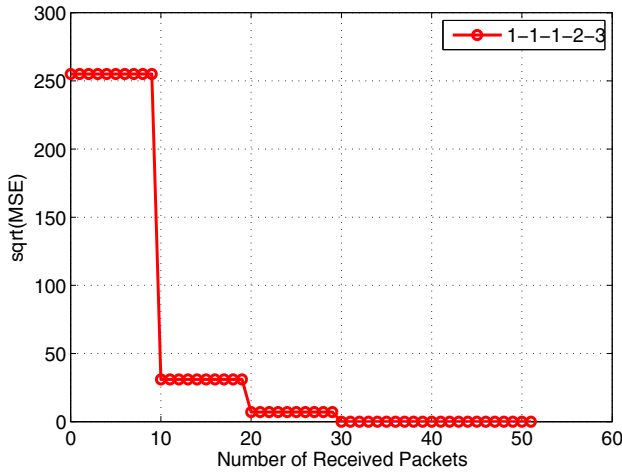


Figure 6: Significance of different bits at different positions to $\sqrt{\text{MSE}}$.

bits of greater values than to the bits of smaller values considering their respective level of significance for lowering the $\sqrt{\text{MSE}}$. When channel is good and the packet loss rate is low, then the probability of recovering the bits of greater values will be high enough even if not many FEC symbols are assigned to them, thus enough FEC symbols may be assigned to the bits of smaller values to enhance the probability of decoding them to further lower the $\sqrt{\text{MSE}}$.

This reasoning has been validated when we checked with the segmentation of the 8 bits under different channel conditions in simulations. Another question is how effective is the MDC component of the proposed scheme in improving video quality. To answer this question, we compare the performance between our proposed scheme and another scheme with interleaving only but without MDC. The results is presented in Fig. 7, under a channel condition of $P_{b|g}$ changes from 0.06 to 0.3, $P_{g|b} = 0.1$, and $P_e(G) = 0.01$. Note that according to the channel model, the average off duration is $T_{off} = 1/P_{b|g}$. It can be seen that when the channel condition is good (with short off periods), the gain achieved by MDC is not obvious. However, when the channel gets worse with increased off periods, the gap between the two curves are quickly increasing. Considerable reduction in $\sqrt{\text{MSE}}$ by MDC is observed under bad channel conditions.

The relationship between the packet loss trace and $\sqrt{\text{MSE}}$ trace is plotted in Fig. 8 under a channel condition of $P_{b|g} = 0.1$, $P_{g|b} = 0.1$, and $P_e(G) = 0.01$. The packet loss trace is plotted in dashed line using the left-hand-side y-axis, and the corresponding $\sqrt{\text{MSE}}$ is plotted in solid line using the right-hand-side y-axis. The corresponding optimal segmentation of the 8 bits is 2-3-3, and the optimal interleaving depth is 5. We can see from the above figure that when packet loss exceeds a certain threshold and thus the last two segments cannot be decoded, the $\sqrt{\text{MSE}}$ of the sub-image is about 60; otherwise, the $\sqrt{\text{MSE}}$ is 0, which means all the three segments are decoded. The statistical average $\sqrt{\text{MSE}}$ of all the tested sub-images is 3.0652 while the expected $\sqrt{\text{MSE}}$ calculated from the minimization problem is 4.1472. The difference may be explained

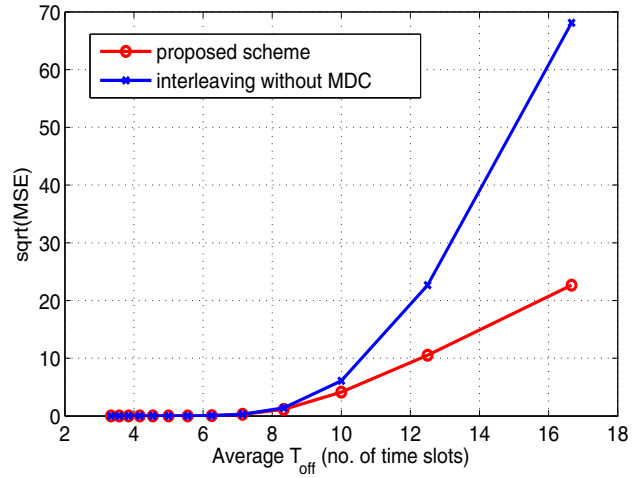


Figure 7: Impact of the MDC component in the proposed scheme under various average bad period of the 60 GHz channel.

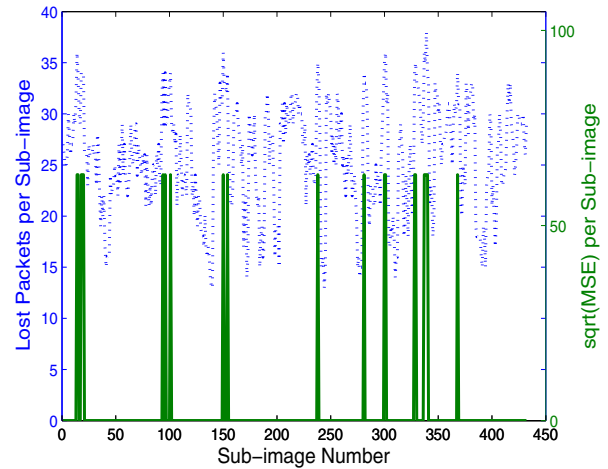


Figure 8: $\sqrt{\text{MSE}}$ vs. packet loss trace.

by a limited number of tested sub-images here, and the dependence of the channel condition for the consecutive sub-images which is not considered on the minimization problem which tends to minimize the expected $\sqrt{\text{MSE}}$ of a sub-image and doesn't consider the channel condition correlation problem of consecutive sub-images.

IV.B. Impact of Channel Dynamics

We next focus on the impact of channel dynamics to show how the proposed scheme adapt to 60GHz channels. We compare the expected $\sqrt{\text{MSE}}$ achieved by the proposed scheme and UVoW, by varying the average duration of the channel bad period T_{off} . In the simulations, $P_{g|b}$ is fixed at 0.1 and $P_e(G)$ is fixed at 0.01 for both schemes. The results are plotted in Fig. 9. Note that using our proposed scheme, the best combination of $\{S, m_1, m_2, \dots, m_S, H\}$, which leads to the lowest expected $\sqrt{\text{MSE}}$ under the given channel condition, can be found for each different average T_{off} value. We can see from the figure that the expected $\sqrt{\text{MSE}}$ of the proposed scheme is lower than that of UVoW for the

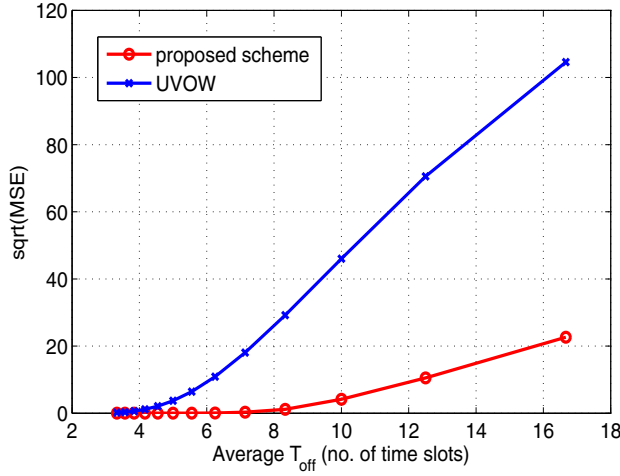


Figure 9: $\sqrt{\text{MSE}}$ vs. average bad period of the 60 GHz channel T_{off} .

entire range. When channel condition is good (i.e., the average T_{off} is less than 4), the performance gap between the two schemes is not remarkable. However, as the channel condition degrades (i.e., average T_{off} is increased beyond 4), the performance gap quickly grows, for the reason that the packet loss rate grows thus it is of greater importance to provide stronger protection to the most important bits. We observed from the result that when T_{off} is 4, the optimal segmentation is 4-4 (as in UVoW) and the optimal interleaving depth is 5; when average T_{off} is 10, the optimal segmentation is 2-3-3 and optimal interleaving depth is 5, which means that more FEC symbols are assigned to the first 2 bits, and interleaving is applied to overcome the channel burstiness. Thus the proposed scheme can adapt to changing channel conditions, to strengthen the protection for the most important bits and increase the interleaving depth to combat the temporarily blocked channel. On average, the proposed scheme achieves 19.6055 reduction in $\sqrt{\text{MSE}}$ than UVoW.

Fig. 10 demonstrates the relationship between the expected $\sqrt{\text{MSE}}$ and $P_e(G)$, the packet loss probability when the channel is in the good state. In this simulation, $P_{b|g}$ and $P_{g|b}$ are fixed at 0.1 for both schemes. As $P_e(G)$ is increased, the probability of a packet being lost in the good state is increased, which means the error correcting capability of the bits of the highest priorities should be enhanced by re-segmentation of the bits. As Fig. 10 shows, the characteristic of adaptivity enables the proposed scheme outperform UVoW in terms of $\sqrt{\text{MSE}}$ for the entire range of $P_e(G)$ examined in the study. The average $\sqrt{\text{MSE}}$ of the proposed scheme is 44.8486 lower than that of UVoW. This result also indicates that our proposed scheme is extremely suitable for the case that $P_e(G)$ is comparatively high during communication.

We look into the optimal segmentation result obtained by solving the minimization problem (6). We find that, in general, when channel is good, more descriptions are created as the bits are divided into more segments, and the length of the segments of the more significant bits are shorter. This

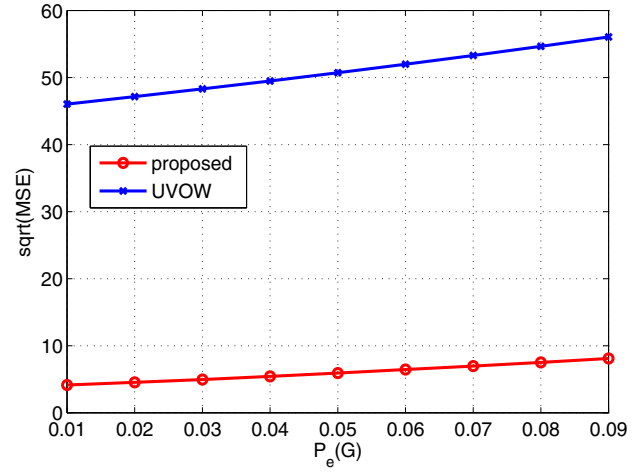


Figure 10: $\sqrt{\text{MSE}}$ vs. packet loss rate at the good state $P_e(G)$.

implies that the number of data symbols constituted by the bits of greater values are smaller compared with the number of FEC symbols assigned to protect them, so that these bits have a higher level of protection.

Fig. 11 and 12 plot the relationship between $\sqrt{\text{MSE}}$, $P_{b|g}$ and transmission delay tolerance Δ . Recall that the relationship between transmission delay tolerance (short for *delay* hereafter) and interleaving depth H is given in (11). Under a specific delay, the expected $\sqrt{\text{MSE}}$ will decrease as $P_{b|g}$ increases; this pattern also applies to the relationship between delay and expected $\sqrt{\text{MSE}}$ under a specific $P_{b|g}$. Besides, in the entire ranges of delay and $P_{b|g}$ as shown in this figure the effect of increased delay and that of increased $P_{b|g}$ on lowering the expected $\sqrt{\text{MSE}}$ do not conflict with each other, which means choosing a larger interleaving depth under a specific channel condition will be always good for improved video performance, especially when channel condition is bad (i.e., $P_{b|g}$ ranges from 0.06 to 0.1). This is because the reduction of $\sqrt{\text{MSE}}$ is much more significant than that achieved when channel condition is good (i.e., $P_{b|g}$ ranges from 0.1 to 0.3), as plotted in Fig. 12. This phenomenon motivates us to carefully design the system parameters under certain channel conditions in order to achieve a balance between delay and desired video performance.

IV.C. Perceived Video Quality

Finally, we examine the visual quality of reconstructed HD video. In the next simulation, example video images are used in our performance test. About 10 distinct images are evaluated using both the proposed scheme and UVoW. The image which reflects the simulation results of most tested images are selected to undergo about 100 distinct evaluations using both schemes in each test. We choose the results that are representative of the 100 tests to present in the following. The optimal solutions of $\{S, m_1, \dots, m_S, H\}$ corresponding to the lowest expected $\sqrt{\text{MSE}}$ are recalculated when channel condition changes, and then applied to the transmission of each sub-image. Since we take $Q = 1$

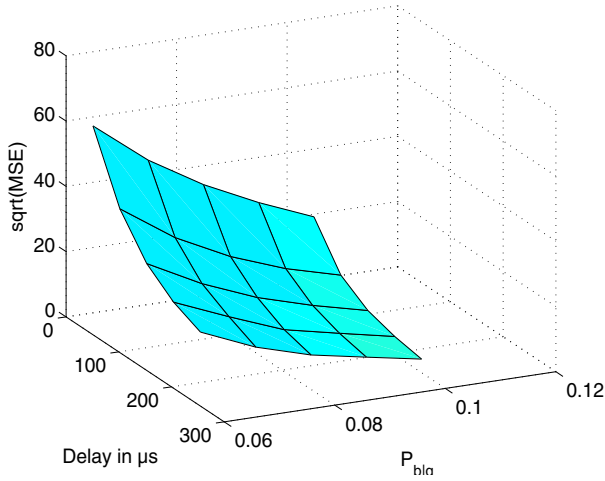


Figure 11: $\sqrt{\text{MSE}}$ vs. transmission delay tolerance and channel state transition probability $P_{b|g}$: good channel.

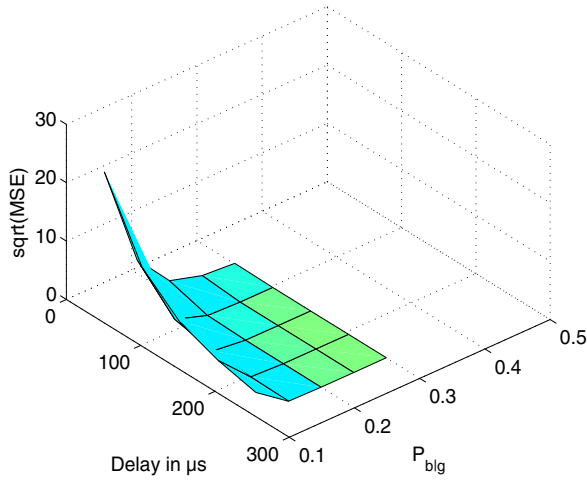


Figure 12: $\sqrt{\text{MSE}}$ vs. transmission delay tolerance and channel state transition probability $P_{b|g}$: bad channel.

to reduce the complexity of solving the optimization problem, we tend to obtain the optimal solution on sub-image basis, and the optimal solution is the same for each sub-image in our simulations under the same channel.

In the simulations, the entire image is divided into smaller sub-images and we show only 1/16 of the sub-images (termed partial image in the following). This is because the size of an entire HD image will be too large to fit into the paper, and the quality of the partial images is representative since all the partial images are treated equally in uncompressed video streaming. The partial images as shown in Fig. 13. The partial image in Fig. 13(a), (b), and (c) are produced by the proposed scheme and the partial images in Figs. 13(d), (e), and (f) are produced by UVoW. Fig. 13(a) and (d) are obtained under a good channel condition; Fig. 13(b) and (e) are obtained under a bad channel condition; Fig. 13(c) and (f) are obtained under a severe channel condition.

It can be seen that when the channel condition is good, both schemes work fine by producing high quality received

videos. When the channel is in a bad condition (i.e., $P_{b|g} = 0.1$), the partial images produced by our scheme are still good with a comparatively much high quality, while a portion of the partial images using UVoW are black (meaning that those pixels are completely lost). Under such bad channels, the most important bits can still be received and decoded in our scheme, but are completely lost in UVoW. This is because the protection to the most important bits is adaptively enhanced after re-segmentation of the pixel bits to adapt to the degraded channel condition.

We also observed that the optimal segmentation is 2-3-3 and the optimal interleaving depth is 5 under the bad channel condition. That is, the two most significant bits are better protected. The change in the color depth in Fig. 13(b) (indicated by the circle) is caused by the loss of the 6 least significant bits during transmission over the 60 GHz channel. The data stream reconstituted by the last 6 bits cannot be decoded due to the transmission loss, but the first 2 most significant bits are correctly decoded since more FEC symbols are assigned to them to provide a higher level of protection, and consequently a basic visual quality is still guaranteed. Under the same channel condition, the UVoW partial image in Fig. 13(e) has a few black blocks, which are caused by the loss of all the 8 bits in those pixels.

When the channel condition is even worse (i.e., $P_{b|g} = 0.06$), the visual quality of the reconstructed image using our proposed scheme (in Fig. 13(c)) is still much better than that using the UVoW (in Fig. 13(f)). The optimal segmentation is now 1-1-1-1-4 and the optimal interleaving depth is 5 under such channel condition. As can be seen from Figs. 13(c), although the color depth has a little change since the last four bits are lost (e.g., see the up right corner), and there are 3 black blocks, the visual quality is still acceptable compared with the UVoW image in Fig. 13(f), which has much more black blocks.

These results clearly demonstrate the advantages of being adaptive both at the MD video coder and at the transmission scheduler to the varying channel conditions.

V. Conclusion

In this paper, we investigated the problem of streaming uncompressed HD video over 60GHz wireless networks. We developed an MD-FEC coding scheme that partitions the pixel bits and an interleaving based transmission strategy to minimize the expected $\sqrt{\text{MSE}}$ of the reconstructed video quality. The performance of the proposed scheme was evaluated with simulations and shown to outperform an existing scheme with considerable gains.

Acknowledgments

This work is supported in part by the US National Science Foundation under Grant CNS-1320664, and through the NSF Broadband Wireless Access & Applications Center (BWAC) site at Auburn University. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the foundation.

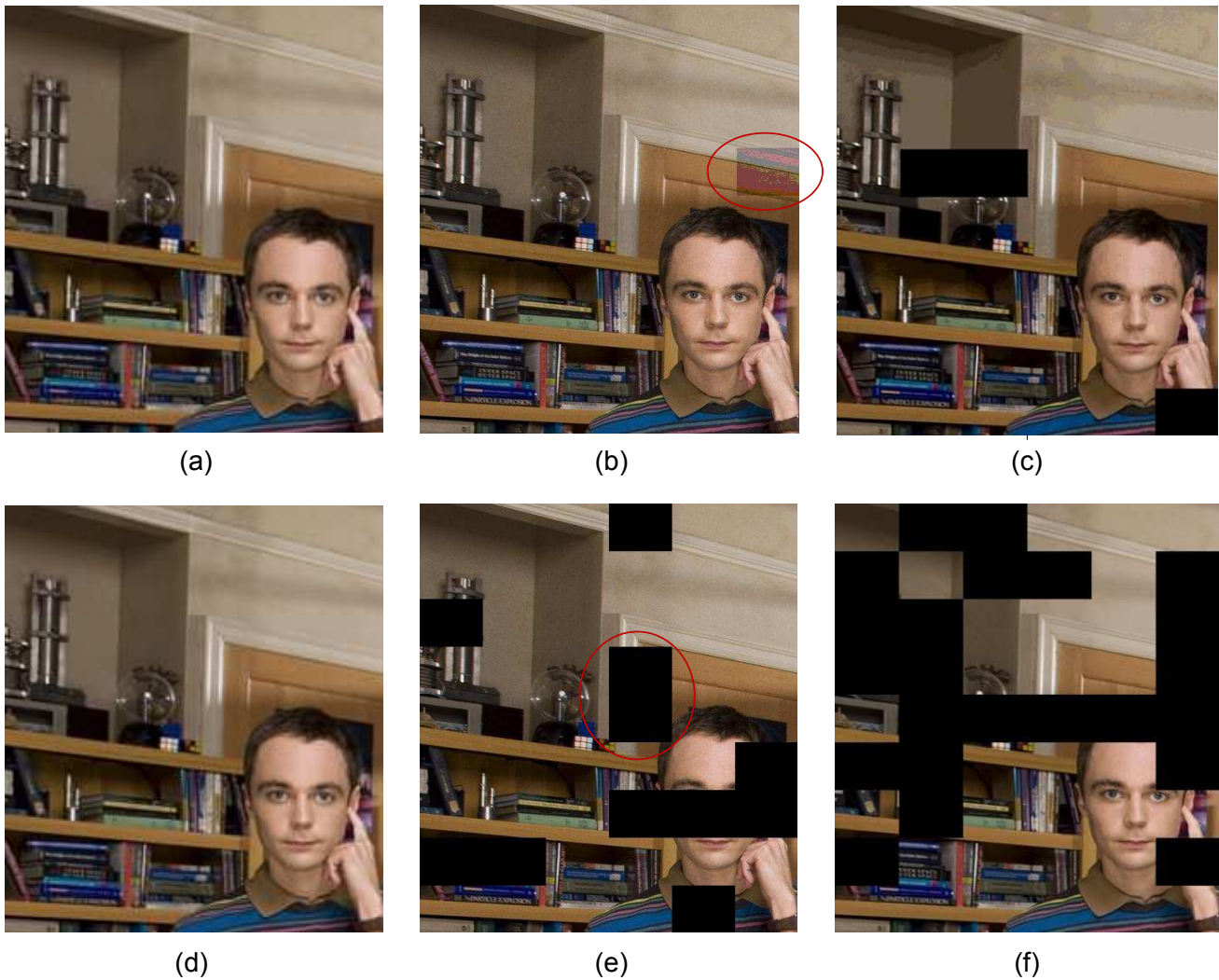


Figure 13: Visual quality comparison of the two schemes under different channel conditions. (a)(b)(c): the proposed scheme. (d)(e)(f): UVoW. (a)(d): $P_{b|g} = 0.3$; (b)(e): $P_{b|g} = 0.1$; (c)(f): $P_{b|g} = 0.06$. $P_{g|b} = 0.1$ and $P_e(G) = 0.01$ are fixed in all the simulations.

References

- [1] A. Albanese, J. Blömer, J. Edmonds, M. Luby, and M. Sudan. Priority encoding transmission. *IEEE Trans. Inform. Theory*, 42(6):1737–1744, Nov. 1996.
- [2] A. C. Begen, Y. Altunbasak, O. Ergun and M. H. Ammar. Multi-path selection for multiple description video streaming over overlay networks. *EURASIP Signal Processing: Image Commun.*, 20(1):39–60, Jan. 2005.
- [3] B. Bosco. Adaptive cross layer design and implementation for Gigabit multimedia applications using 60 GHz wireless links. Master’s thesis, Arizona State Univ., Phoenix, AZ, 2011.
- [4] Cisco. Cisco visual networking index: Global mobile data traffic forecast update, 2011-2016. [online] Available: <http://www.cisco.com>. May 2012.
- [5] Federal Communications Commission. FCC 13-112: In the Matter of Revision of Part 15 of the Commission’s Rules Regarding Operation in the 57-64 GHz Band. *ET Docket No. 07-113*, Aug. 9, 2013.
- [6] M. X. Gong, D. Akhmetov, R. Want, and S. Mao. Directional CSMA/CA protocol with spatial reuse for mmWave wireless networks. In *Proc. IEEE GLOBECOM 2010*, pages 1–5, Miami, FL, Dec. 2010.
- [7] W. Lee, K. Noh, S. Kim, and J. Heo. Efficient cooperative transmission for wireless 3D HD video transmission in 60GHz channel. *IEEE Trans. Consum. Electron.*, 56(4):2481–2488, Nov. 2010.
- [8] M. Manohara, R. Mudumbai, J. Gibson, and U. Madhoo. Error correction scheme for uncompressed HD video over wireless. In *Proc. IEEE ICME 2009*, pages 802–805, New York, NY, June/July 2009.
- [9] S. Mao, D. Bushmitch, S. Narayanan, and S. S. Panwar. MRTP: A multi-flow real-time transport proto-

- col for ad hoc networks. *IEEE Trans. Multimedia*, 8(2):356–369, Apr. 2006.
- [10] S. Mao, X. Cheng, Y. T. Hou, and H. D. Sherali. Multiple description video multicast in wireless ad hoc networks. *ACM/Springer Mobile Netw. Appl.*, 11(1):63–73, Jan. 2006.
- [11] S. Mao, S. Lin, S. Panwar, Y. Wang, and E. Celebi. Video transport over ad hoc networks: Multistream coding with multipath transport. *IEEE J. Sel. Areas Commun.*, 21(10):1721–1737, Dec. 2003.
- [12] S. Mao, S. Lin, Y. Wang, S. S. Panwar, and Y. Li. Multipath video transport over wireless ad hoc networks. *IEEE Wireless Commun.*, 12(4):42–49, Aug. 2005.
- [13] A. E. Mohr, E. A. Riskin, and R. E. Lader. Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction. *IEEE J. Sel. Areas Commun.*, 18(6):819–828, Jun. 2000.
- [14] T. S. Rappaport, J. N. Murdock, and F. Gutierrez. State of the art in 60-GHz integrated circuits and systems for wireless communications. *Proc. IEEE*, 99(8):1390–1436, Aug. 2011.
- [15] L. Rizzo. Effective erasure codes for reliable computer communication protocols. *ACM Computer Commun. Rev.*, 27(2):24–36, Apr. 1997.
- [16] H.-R. Shao and et al. Adaptive multi-beam transmission of uncompressed video over 60GHz wireless systems. In *Proc. Future Generation Commun. Netw. 2007*, pages 430–435, Jeju-Island, South Korea, Dec. 2007.
- [17] H. Singh, X. Qin, H. Shao, C. Ngo, C. Kwon, and S. S. Kim. Support of uncompressed video streaming over 60GHz wireless networks. In *Proc. IEEE CCNC 2008*, pages 243–248, Las Vegas, NV, Jan. 2008.
- [18] I.-K. Son, S. Mao, M. Gong, and Y. Li. On frame-based scheduling for directional mmWave WPANs. In *Proc. IEEE INFOCOM 2012*, pages 2149–2157, Orlando, FL, Mar. 2012.
- [19] Y. Wang, A. Reibman, and S. Lin. Multiple description coding for video delivery. *Proc. IEEE*, 93(1):57–70, Jan. 2005.
- [20] L. R. Wilhelmsson and L. B. Milstein. On the effect of imperfect interleaving for the Gilbert Elliott channel. *IEEE Trans. Commun.*, 47(5):681–688, May 1999.
- [21] Y. Zhao, S. Mao, J. Neel, and J. H. Reed. Performance evaluation of cognitive radios: metrics, utility functions, and methodologies. *Proc. IEEE*, 97(4):642–659, Apr. 2009.

Zhifeng He received the M.S. degree in Micro Electronics and Solid State Electronics from Beijing University of Posts and Telecommunications, Beijing, China, and the B.S. degree in Electronics Information Science and Technology from Shandong University of Technology, Zibo, China, in 2012 and 2009, respectively. Since 2012, he has been pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His current research interests include cognitive radio, multimedia communications and optimization.

Shiwen Mao received Ph.D. in electrical and computer engineering from Polytechnic University, Brooklyn, NY. Currently, he is the McWane Associate Professor in the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His research interests include cross-layer optimization of wireless networks and multimedia communications, with current focus on cognitive radio, small cells, 60 GHz mmWave networks, free space optical networks, and smart grid. He is on the Editorial Board of IEEE Transactions on Wireless Communications, IEEE Internet of Things Journal, IEEE Communications Surveys and Tutorials, among others. He received the 2013 IEEE ComSoc MMTTC Outstanding Leadership Award and the NSF CAREER Award in 2010. He is a co-recipient of The IEEE ICC 2013 Best Paper Award and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems. He is a senior member of the IEEE and a member of ASEE.