PAPER Wireless Multi-view Video Streaming with Subcarrier Allocation

Takuya FUJIHASHI^{†a)}, Shiho KODERA^{††b)}, Nonmembers, Shunsuke SARUWATARI^{††c)}, and Takashi WATANABE^{†d)}, Members

SUMMARY When an access point transmits multi-view video over a wireless network with subcarriers, bit errors occur in the low quality subcarriers. The errors cause a significant degradation of video quality. The present paper proposes Significance based Multi-view Video Streaming with Subcarrier Allocation (SMVS/SA) for the maintenance of high video quality. SMVS/SA transmits a significant video frame over a high quality subcarrier to minimize the effect of the errors. SMVS/SA has two contributions. The first contribution is subcarrier-gain based multi-view rate distortion to predict each frame's significance based on the quality of subcarriers. The second contribution is heuristic algorithms to decide the sub-optimal allocation between video frames and subcarriers. The heuristic algorithms exploit the feature of multi-view video coding, which is a video frame is encoded using the previous time or camera video frame, and decides the sub-optimal allocation with low computation. To evaluate the performance of SMVS/SA in a real wireless network, we measure the quality of subcarriers using a software radio. Evaluations using MERL's benchmark test sequences and the measured subcarrier quality reveal that SMVS/SA achieves low traffic and communication delay with a slight degradation of video quality. For example, SMVS/SA improves video quality by up to 2.7 [dB] compared to the multi-view video transmission scheme without subcarrier allocation.

key words: Multi-view Video, Subcarrier Allocation

1. Introduction

With the progress of wireless and video coding technology for multi-view video, the demand of watching 3D video on wireless devices increases [1,2]. To watch 3D video on the wireless devices, a video encoder transmits video frames of multiple cameras to a user node over wireless networks. The user node creates 3D video using the received video frames and view synthesis techniques, such as depth image-based rendering (DIBR) [3] and 3D warping [4]. To stream 3D video over wireless networks efficiently, the wireless and multi-view video coding techniques have been studied independently. The typical studies of multi-view video coding are Multi-view Video Coding (MVC) [5], Interactive Multiview Video Streaming (IMVS) [6,7], User dependent Multiview video Streaming (UMS) [8], and UMS for Multi-user (UMSM) [9]. These studies focus on the reduction of video

Manuscript received January 1, 2011.

Manuscript revised January 1, 2011.

b) E-mail: kodera@aurum.cs.inf.shizuoka.ac.jp

DOI: 10.1587/transcom.E0.B.1

traffic by exploiting the correlation of time and inter-camera domain of video frames.

In view of wireless networks, Orthogonal Frequency Division Multiplexing (OFDM) [10] is used in modern wireless technology (802.11, WiMax, Digital TV, etc.). OFDM decomposes a wideband channel into a set of mutually orthogonal subcarriers. A sender transmits multiple signals simultaneously at different subcarriers over a single transmission path. On the other hand, the channel gains across these subcarriers are usually different, sometimes by as much as 20 [dB] [11]. The low channel gains induce high error rate at a receiver.

When a video encoder simply transmits multi-view video over a wireless network by OFDM, bit errors occur in video transmission of low channel gain subcarriers. If these errors occur randomly in all video frames, video quality at a user node suddenly degrades [12]. We define this problem as the multi-view error propagation.

The multi-view error propagation is caused by the features of the multi-view video coding techniques. The multiview video coding techniques exploit time and inter-camera domain correlation to reduce redundant information among video frames. Specifically, the multi-view video coding techniques first encode a video frame in a camera as a reference video frame. Next, the coding techniques encode the subsequent video frame in the same and neighbor cameras by calculating the difference between the subsequent and the reference video frame. After the encoding, the coding techniques select the subsequent video frame as the new reference video frame and encode the rest of subsequent video frames. If bit errors occur in a reference video frame in a camera, the user node does not decode the subsequent video frame correctly. The incorrectly decoded video frame propagates the errors to the subsequent video frames in the same and neighbor cameras.

To prevent the multi-view error propagation, typical solutions are retransmission [13–15] and Forward Error Correction (FEC) [16, 17]. The retransmission recovers from bit errors by retransmitting a whole or partial data to the user node. However, the retransmission increases communication delay and long communication delay induces low user satisfaction. The FEC codes help the user node that suffers low channel gain subcarriers. However, the FEC codes consume data rates available to video packets and degrade video quality for the user node that does not suffer low channel gain subcarriers.

[†]The authors are with the Graduate School of Information Science and Technology, Osaka University, Japan

^{††}The authors are with the Graduate School of Informatics, Shizuoka University, Japan

a) E-mail: fujihashi.takuya@ist.osaka-u.ac.jp

c) E-mail: saru@inf.shizuoka.ac.jp

d) E-mail: watanabe@ist.osaka-u.ac.jp

The present paper proposes Significance based Multi-view Video Streaming with Subcarrier Allocation (SMVS/SA) for multi-view video streaming over a wireless network with subcarriers. SMVS/SA achieves the reduction of communication delay and video traffic while the maintenance of high video quality. The key feature of SMVS/SA is to transmit significant video frames, which have a great effect on video quality when bit errors occur in the video frames, with high channel gain subcarriers. The present paper makes two contribution. The first contribution is subcarrier-gain based multi-view rate distortion to predict the effect of each video frame on video quality when the video frame is lost. The second contribution is two types of heuristic algorithms to decide the allocation between video frames and subcarriers with low computation. The allocation achieves sub-optimal multi-view rate distortion under the different subcarrier channel gains. To evaluate the performance of SMVS/SA, we use MATLAB multi-view video encoder and GNU Radio/Universal Radio Software Platform (USRP) N200 software radio. USRP N200 measures subcarrier quality of an OFDM link for the MATLAB multiview video encoder. Evaluations using the MATLAB video encoder and MERL's benchmark test sequences reveal that SMVS/SA achieves only a slight degradation of video quality. For example, SMVS/SA maintains video quality by up to 2.7 [dB] compared to existing approaches.

The remainder of the present paper is organized as follows. Section 2 presents a summary of related research. We present the details of SMVS/SA in Section 3. In Section 4, evaluations are performed to reveal the suppression of communication delay and the maintenance of video quality for the proposed SMVS/SA. Finally, conclusions are summarized in Section 5.

2. Related Research

This study is related to joint source-channel coding and multi-view rate distortion based video streaming.

2.1 Joint source-channel coding

There are many studies about joint source-channel coding for single-view video. The existing studies can be classified into two types. In the first type, a video encoder calculates frame/group of picture (GOP)-level distortion based on the features of networks to predict single-view video quality at a user node before transmission. [18] defines a model for predicting the distortion due to bit errors in a video frame. [18] uses the model for adaptive video encoding and rate control under time-varying channel conditions. [19–21] propose a distortion model for single-view video and the model takes the features of subcarriers into consideration. [22] proposes a GOP-level distortion model based on the error propagation behavior of whole-frame losses. [23] takes loss burstiness into consideration for the GOP-level distortion model.

In the second type, a video encoder allocates video frames to network resource based on bit-level significance



Fig.1 System model of multi-view video streaming over wireless network

of each video frame. Typical studies are SoftCast [24], Par-Cast [25], and FlexCast [26]. SoftCast [24] exploits DCT coefficients for significance prediction of each single-view video frame. SoftCast allocates each DCT coefficient to subcarriers based on the significance and channel gains of the subcarriers. SoftCast transmits the DCT coefficients by analog modulated OFDM symbols. ParCast [25] extends the SoftCast's design to MIMO-OFDM. FlexCast [26] focuses on bit-level significance of each single-view video frame. FlexCast adds rateless codes to bits based on the significance to minimize the effect of channel gain differences among subcarriers.

SMVS/SA follows the same motivation to jointly consider sourced compression and error resilience. SMVS/SA extends their concepts to multi-view video streaming. SMVS/SA focuses on GOP-level significance and channel gain differences among subcarriers to improve 3D video delivery quality over wireless networks.

2.2 Multi-view rate distortion based video streaming

Several studies have been proposed for the maintenance of high 3D video quality. [27] introduces an end-to-end multiview rate distortion model for 3D video to achieve optimal encoder bitrate. [27] only analyzes 3D video with left and right cameras. [12] proposes the average error rate based multi-view rate distortion to analyze the distortion with multiple cameras. [28] proposes network bandwidth based multi-view rate distortion for bandwidth constrained channels. The basic concept of the proposed subcarrier-gain based multi-view rate distortion is based on these studies. SMVS/SA considers the channel gain differences among subcarriers for multi-view rate distortion to maintain high video quality in a real wireless network.

3. Significance based Multi-view Video Streaming with Subcarrier Allocation (SMVS/SA)

3.1 Overview

There are three requirements for multi-view video streaming over wireless networks: reduction of video traffic, suppression of communication delay, and the maintenance of high video quality. To satisfy all of the above requirements, we propose Significance based Multi-view Video Streaming with Subcarrier Allocation (SMVS/SA). The key idea of SMVS/SA is to transmit significant video frames, which have a great effect on video quality, by high channel gain subcarriers.

Figure 1 shows a system model of SMVS/SA. Several cameras are assumed to be connected to a video encoder by wire, and the video encoder is connected to an access point by wired networks. The access point is connected to a user node by a wireless network with subcarriers. The wireless network has different channel gains among the subcarriers. The video encoder previously transmits a encoded multi-view video sequence to the access point. The access point decodes the received multi-view video and waits for a request packet from the user node. The user node transmits a request packet to the access point by OFDM. When the access point receives the request packet, the access point encodes the multi-view video based on the received request packet. The access point transmits the encoded multi-view video to the user node by OFDM.

SMVS/SA consists of request transmission, video encoding, significance prediction, heuristic calculation, sorting and video transmission, and video decoding.

(1) **Request Transmission:** A user node periodically transmits a request packet and channel state information to an access point to play back multi-view video continuously. The details of request transmission are described in Section 3.2. (2) Video Encoding: When the access point receives the request packet, the access point encodes a multi-view video sequence in one Group of Group of Pictures (GGOP) based on the request packet. GGOP is the group of GOP, which is the set of video frames and typically consists of eight frames, for each camera. The details of video encoding are described in Section 3.3.

(3) Significance Prediction: After video encoding, the access point predicts which video frames should be transmitted in high channel gain subcarriers. To predict the significance of each video frame, SMVS/SA proposes subcarriergain based multi-view rate distortion. The details of significance prediction are described in Section 3.4.

(4) **Heuristic Calculation:** The disadvantage of the subcarrier-gain based multi-view rate distortion is high computation complexity. To reduce the computational complexity, SMVS/SA proposes two types of heuristic algorithms: First and Concentric Allocation. The details of the heuristic algorithms are described in Section 3.5.

(5) Sorting and Video Transmission: The access point allocates video frames to subcarriers based on the predicted significance. After the allocation, the access point modulates the allocated video frames by OFDM and transmits the modulated video frames to the user node. The details of sorting and video transmission are described in Section 3.6. (6) Video Decoding: When the user node receives the OFDM modulated video frames, the user node decodes the video frames by standard H.264/AVC MVC decoder. After the video decoding, the user node plays back multi-view video on display. The details of video decoding are described in Section 3.7.

3.2 Request Transmission

A user node transmits a request packet to an access point when the user begins to watch multi-view video or receives video frames in one GGOP. Each request packet consists of two fields: requested camera ID and Channel State Information (CSI). The requested camera ID field indicates the set of cameras which need to create 3D video at the user node. The requested camera ID field is an array of eight-bit fields. The CSI field is based on 802.11n Channel State Information packet [29]. The CSI describes the channel gain, which is Signal-to-Noise Ratio (SNR), of RF path between the access point and the user node for all subcarriers. The CSI is reported by the 802.11 Network Interface Card (NIC) in a format specified by the standard. When the access point receives the request packet, the access point knows the recent channel gain of each subcarrier with high accuracy.

3.3 Video Encoding

After the access point received the request packet, the access point encodes multi-view video based on the requested camera ID field in the request packet. Figure 2 shows the prediction structure of SMVS/SA where the requested camera ID field is $\{1, 2, 3\}$.

The access point encodes an anchor frame of an initial camera in requested cameras into I-frame and the subsequent video frames into P-frames. The initial camera is camera 1 in Fig. 2. I-frame is a picture that is encoded independent from other pictures. P-frame encodes only the differences from an encoded reference video frame and has lower traffic than I-frame. Specifically, the access point divides a currently coded video frame and the reference video frame into several blocks. The access point finds the best matching block between these video frames and calculates the differences [30]. After encoding the video frames of the initial camera, the access point encodes video frames of the other requested cameras.

The anchor frames of the requested cameras are encoded into P-frame using an anchor frame at the same time in the previous camera. The subsequent video frames are also encoded into P-frames. To encode a subsequent video frame, the access point selects two encoded video frames that are the previous time in the same camera and the same time in the previous camera. The access point tries to encode the subsequent video frame using each encoded video frame and calculate the distortion of video encoding. The access point decides the reference video frame of the subsequent video frame from two encoded video frames. The reference video frame achieves the lowest distortion of video encoding. After the video encoding of all video frames in one GGOP, the access point obtains bit streams of each video frame.



Fig. 2 Prediction structure where the requested camera ID is {1, 2, 3}.

3.4 Significance Prediction

After video encoding, the access point predicts the significance of each video frame. To predict the significance, the present paper proposes subcarrier-gain based multi-view rate distortion. The subcarrier-gain based multi-view rate distortion predicts the effect of each video frame on video quality when the communication of the video frame is failed. The access point maintains high video quality under different channel gains of subcarriers by means of calculating the minimum multi-view rate distortion as

$$\arg\min_{P} D_{\text{GGOP}}(P) \tag{1}$$

where D_{GGOP} is the proposed multi-view rate distortion in one GGOP, *P* is $N_{\text{camera}} \times N_{\text{GOP}}$ matrix of success rate. The minimum multi-view rate distortion reveals which video frames should be transmitted by the high channel gain subcarriers to maintain high video quality. N_{camera} and N_{GOP} denote the number of requested cameras and the length of each GOP, respectively.

Assumption: The number of video frames in one GGOP is smaller than the number of subcarriers in OFDM. In wireless video transmission, distortion induced by the error of the frame itself occurs in video frames due to communication errors, including channel fading, interference, and noise. Specifically, even when one bit error occurs in the encoded bit stream of one video frame, a user node decodes the video frame incorrectly and experiences the distortion. Even when one bit error occurs in the encoded bit stream, SMVS/SA regards the video frame as loss. Since every bit error regards as the frame loss, our model indirectly includes the distortion in the frame loss. The reason of regarding one bit error as whole frame loss is that even one bit error induces "cliff effect" [24] in the corresponding video frame. Cliff effect is the phenomenon that one bit error causes the collapse of the whole frame decoding because current video compression includes entropy encoding. At the user node, SMVS/SA assumes that a proper error concealment operation is performed on lost video frames. The error concealment operation resorts to either temporal or inter-camera concealment. SMVS/SA performs the error concealment operation for a video frame when errors occur in bits of the video frame. Consequently, the success rate is equivalent to the video frame success rate.

Definition: Let $D_{\text{GGOP}}(P)$ be the overall subcarrier-gain based multi-view rate distortion in one GGOP at the user node. $D_{\text{GGOP}}(P)$ is defined as network-induced distortion, denoted by $D_{\text{network}}(P, s, t)$. They are expressed as:

$$D_{\text{GGOP}}(P) = \sum_{s=1}^{N_{\text{camera}}} \sum_{t=1}^{N_{\text{GOP}}} D_{\text{network}}(P, s, t)$$
(2)

$$D_{\text{network}}(P, s, t) = p(s, t) \cdot D_{\text{encoding}}(s, t) + (1 - p(s, t)) \cdot D_{\text{loss}}(s, t)$$
(3)

$$D_{\text{encoding}}(s,t) = E\{[F_{i}(s,t) - \hat{F}_{i}(s,t)]^{2}\}$$
(4)

where $D_{\text{encoding}}(s, t)$ is the encoding-induced distortion, $F_i(s, t)$ is the original value of pixel *i* in M(s, t), $\hat{F}_i(s, t)$ is the reconstructed values of pixel *i* in M(s, t) at the access point, and $p(s, t) \in P$ is the success rate for the frame at camera *s* and time *t*. The value of p(s, t) is based on the channel gain of a subcarrier. Moreover, $E\{\cdot\}$ denotes the expectation taken over all the pixels in frame M(s, t). M(s, t)denotes the frame at camera *s* and time *t*.

As can be seen from equation (4), encoding-induced distortion refers to the Mean Square Error (MSE) between the original frame and the reconstructed video frame at the access point. The network-induced distortion consists of the distortion when communication is successful and failed. $D_{\text{loss}}(s, t)$ denotes the distortion when the communication is failed. When the communication of the video frame is successful, the received bit stream is error-free because SMVS/SA regards every bit error as the frame loss. Therefore, the distortion of the received frame is only encoding.

On the other hand, $D_{loss}(s, t)$ is expressed as:

$$D_{\text{loss}}(s,t) = E\{\left[\hat{F}_{i}(s,t) - \breve{F}_{i}(s,t)\right]^{2}\} + D_{\text{previous}}$$
(5)

where $\breve{F}_{i}(s, t)$ is expressed according to the reference video frame as:

$$\breve{F}_{i}(s,t) = \begin{cases} \hat{F}_{\text{conceal}(i)}(s-1,t) & \text{if } \hat{F}_{\text{conceal}(i)} \in M(s-1,t) \\ \hat{F}_{\text{conceal}(i)}(s,t-1) & \text{else.} \end{cases}$$
(6)

where conceal(i) is the index of the matching pixel in the reference video frame for error concealment operation [31]. $D_{\text{previous}}(s, t)$ is based on a reference video frame of M(s, t) for the error concealment operation. When M(s, t) exploits a video frame at the previous time in the same camera as the reference video frame, $D_{\text{previous}}(s, t)$ is expressed as:

$$D_{\text{previous}}(s,t) = D_{\text{network}}(P,s,t-1)$$
(7)

When M(s, t) exploits a video frame at the same time in the previous camera as the reference video frame, $D_{\text{previous}}(s, t)$ is expressed as:

$$D_{\text{previous}}(s,t) = D_{\text{network}}(P, s-1, t)$$
(8)

3.5 Heuristic Calculation

The minimum subcarrier-gain based multi-view rate distortion reveals which video frames should be transmitted by the high channel gain subcarriers to achieve the highest video quality. However, the computational complexity of the multi-view rate distortion is high. Specifically, an access point needs to calculate the minimum networkinduced distortion, which is equation (2), from all combinations of the subcarriers and the video frames in one GGOP. As the result, the computational complexity of equation (2) is $O\{(N_{\text{camera}} \cdot N_{\text{GOP}})\}$. To reduce the computational complexity, SMVS/SA proposes two heuristic algorithms: 1) First Allocation and 2) Concentric Allocation. These heuristics focus on the feature of the multi-view video coding technique: the video quality of a subsequent video frame suddenly degrades when the reference video frame is lost. Therefore, the heuristics first allocate a high channel gain subcarrier for early reference video frames to maintain video quality of subsequent video frames.

3.5.1 First Allocation

First Allocation allocates high channel gain subcarriers for early video frames of requested cameras. An access point selects video frames of all cameras at beginning time and the same number *m* of high success rate $p_{\rm m}$ from $P_{\rm subcarriers}$. $P_{\text{subcarriers}}$ is a set of success rate in each subcarrier. The success rate is calculated by the channel gain of the subcarrier. The access point calculates the sum of proposed multi-view distortion of the video frames using each $p_{\rm m}$ from equation (3). The access point decides the best allocation between the selected video frames and $p_{\rm m}$. The best allocation is the same meaning as the achievement of minimum multi-view rate distortion. The access point sets each $p_{\rm m}$ to P, which is the same frame indexes of the allocated video frame, and removes each p_m from $P_{subcarriers}$. The access point selects video frames of all cameras at the next time and the same number m of high success rate p_m from $P_{subcarriers}$. The access point also calculates the sum of proposed multi-view distortion of each video frame using each $p_{\rm m}$ from equation (3). The access point decides the best allocation between the video frames and $p_{\rm m}$. The access point repeats the same operation over one GOP. As the results, First Allocation reduces the computation to $O(N_{\text{GOP}} \cdot N_{\text{camera}}!)$.

For example, we assume that an access point encodes multi-view video in one GGOP as shown in Fig. 2 and the number of subcarriers is the same number of the encoded video frames. The access point first selects one I-frame and two P-frames in M(1, 1), M(2, 1), and M(3, 1). The access

 N_{GOP} 1) 1 2) P 3) P 4) P 2) P 3) P 4) P 5) P 3) P 4) P 5) P 6) PRepetition

Fig. 3 One of examples in Concentric Allocation.

point also selects three high success rate p_1 , p_2 , and p_3 from $P_{\text{subcarriers}}$. The access point calculates the sum of multiview rate distortion of the selected I-frame and P-frames using p_1 to p_3 from equation (3). This example assumes that the combinations of I-frame in M(1, 1) and p_1 , P-frame in M(2, 1) and p_2 , and P-frame in M(3, 1) and p_3 achieve the lowest multi-view rate distortion. The access point sets p_1 to P(1, 1), p_2 to P(2, 1), and p_3 to P(3, 1).

Next, the access point selects P-frames in M(1, 2), M(2, 2), and M(3, 2). The access point also selects three high success rate p_4 , p_5 , and p_6 from $P_{\text{subcarriers}}$. After the selection, the access point calculates the sum of multi-view rate distortion of each P-frame using p_4 , p_5 , and p_6 from equation (3) to decide the best allocation between the video frames and the subcarriers. After the calculation, the access point sets p_4 , p_5 , and p_6 to P based on the best allocation. The access point repeats the above algorithm for all video frames in one GGOP.

3.5.2 Concentric Allocation

Concentric Allocation allocates high channel gain subcarriers for neighbor video frames of an initial camera in requested cameras. Figure 3 shows the one of examples in Concentric Allocation. We assumes that the number of cameras is smaller than the length of one GOP. The numbers located on the left side at each frame represent the operation order in the Concentric Allocation. An access point selects I-frame and the highest success rate p from $P_{\text{subcarriers}}$. The access point sets p to P(s, t), which s and t are the same frame indexes of I-frame, and removes p from $P_{\text{subcarriers}}$. Next, the access point selects *n* P-frames of the I-frame's neighborhood and the same number of high success rate $p_{\rm n}$ from $P_{\rm subcarriers}$. The access point calculates the sum of proposed multi-view distortion of each P-frame using each $p_{\rm n}$ from equation (3), and decides the best allocation between the selected P-frames and p_n . The access point sets each p_n to P, which is the same frame indexes of the allocated P-frame, and removes each p_n from $P_{subcarriers}$. The access point selects n P-frames of the previously selected P-frames' neighborhood and the same number of high success rate p_n from $P_{subcarriers}$, and repeats the above operation. When the number of selected frames is approach to the number of cameras, the access point repeatedly selects the same number of frames and subcarriers, and decides the best combination. The number of repetitions is almost close to $N_{\text{GOP}} - N_{\text{camera}}$. As the results, the computation reduces to $O\{(N_{\text{GOP}} - N_{\text{camera}}) \cdot (N_{\text{camera}}!)\}$. Even when the number of cameras is greater than the length of one GOP, the operation is just inverted and the computation becomes $O\{(N_{\text{camera}} - N_{\text{GOP}}) \cdot (N_{\text{GOP}}!)\}$. Note that when the number of cameras is the same as the length of one GOP, the computation is $O(N_{\text{GOP}}!)$ or $O(N_{\text{camera}}!)$ because the number of repetition is only one.

We assume the same prediction structure and the number of subcarriers in Sec. 3.5.1. The access point first selects the I-frame in M(1, 1), and the highest success rate p_1 from $P_{\text{subcarriers}}$. The access point sets p_1 to P(1, 1). Continuously, the access point selects P-frames in M(1, 2) and M(2, 1). These P-frames are the I-frame's neighborhood. The access point also selects three high success rate p_2 and p_3 from $P_{\text{subcarriers}}$. The access point calculates the sum of multi-view rate distortion of each P-frame using p_2 and p_3 from equation (3). This example assumes that the combinations of P-frame in M(1, 2) and p_3 , and P-frame in M(2, 1)and p_2 achieve the lowest distortion. The access point sets p_2 to P(2, 1) and p_3 to P(1, 2).

Next, the access point selects P-frames in M(1, 3), M(2, 2), and M(3, 1). These P-frames are the previously selected P-frame's neighborhood. The access point also selects three high success rate p_4 , p_5 , and p_6 in $P_{\text{subcarriers}}$. The access point decides the best allocation between the selected three P-frames and subcarriers from equation (3). The access point repeats the above algorithm for the rest of video frames in one GGOP.

3.6 Sorting and Video Transmission

After the significance prediction, an access point allocates bit streams of each video frame to subcarriers based on the prediction. Continuously, the access point transmits the bit streams to a user node over a wireless network by OFDM. The bit streams in each subcarrier are modulated equally, using BPSK, QPSK, 16 QAM, or 64 QAM, with 1, 2, 4 or 6 bits per symbol, respectively. The modulated symbols in each subcarrier are modulated by one OFDM symbol. The access point inserts up to 44 OFDM symbols into one video packet and transmits the video packets to the user node. Note that the access point allocates bit streams with different lengths to subcarriers. The bit streams with different lengths induce different transmission completion time among subcarriers and low subcarrier utilization. To improve the utilization, the access point reallocates bit streams in low channel gain subcarrier to high channel gain subcarrier when the transmission in high channel gain subcarrier is finished.

After the packet transmission, the access point transmits an EoG (End of Group of Pictures) packet to the user node. When the user node receives the EoG packet, the user node transmits the next request packet to the access point.

3.7 Video Decoding

When a user node receives an EoG packet, the user node starts demodulation and multi-view video decoding for received video packets. The demodulator converts each subcarrier's symbols into the bits of each bit stream from constellations of several different modulations (BPSK, QPSK, 16 QAM, 64 QAM). The access point assembles the demodulated bit streams in respective subcarriers. The subcarrierbased assembled bit streams are equivalent to the bit streams of each video frame. Next, the user node decodes the subcarrier-based assembled bit streams using the standard H.264/AVC MVC decoder. If bit streams in a video frame have errors, the user node exploits error concealment operation. After the decoding, the user node creates 3D video using the decoded video frames of multiple cameras. Finally, the user node plays back 3D video on display.

4. Evaluation

4.1 Evaluation Settings

To evaluate the performance of SMVS/SA, we implemented the SMVS/SA encoder/decoder on a multi-view video encoder based on MATLAB video encoder [32]. The evaluation uses multi-view video test sequences with different characteristics: "Ballroom" (faster motion), "Exit" (little motion), and "Vassar" (very little motion). The size of the video frames was 144×176 pixels for all evaluations. The test sequence was provided by Mitsubishi Electric Research Laboratories (MERL) [33], which are recommended by the Joint Video Team (JVT) as standard test sequences to evaluate the performance of multi-view video. The number of cameras was eight. The video frames of each camera were encoded at a frame rate of 15 [fps]. The GOP length of video sequence was set to eight frames. We used 250 frames per sequence for all of the evaluations. Quantization parameter value for "Ballroom" used in our experiments was 25.

The evaluation assumes that one access point and one user node were connected by a wireless network with subcarriers. The user node transmitted a request packet to the access point. The request packet includes requested camera IDs. The access point sent back the requested multi-view video in one GGOP to the user node by OFDM. The number of subcarriers was the same as the number of video frames in one GGOP. The evaluation assumed that request packet and bit streams of encoded I-frame are received error-free because these data were transmitted in the highest channel gain subcarrier.

We used the standard peak signal-to-noise ratio (PSNR) metric to evaluate multi-view video quality in one GGOP. PSNR_{GGOP} represents the average video quality of multi-view video in one GGOP as follows:

$$PSNR_{GGOP} = 10\log_{10} \frac{(2^{L} - 1)HN_{camera}N_{GOP}W}{D_{GGOP}}$$
(9)

where D_{GGOP} is the predicted or measured multi-view rate distortion in one GGOP, *H* and *W* are the height and width of a video frame, respectively. Moreover, *L* is the number of bits used to encode pixel luminance, typically eight bits. Measured D_{GGOP} means the observed distortion at the user node. The measured D_{GGOP} is used to evaluate video quality in each reference scheme. Predicted D_{GGOP} means the estimated distortion at the access point using equation (2). Figure 9 shows the differences between the predicted D_{GGOP} and the measured D_{GGOP} .

4.2 Baseline Performance

To evaluate the baseline performance of the proposed SMVS/SA, we compared the video quality and communication delay of three encoding/decoding schemes: ALL for EACH, Retransmission, and SMVS/SA.

1) ALL for EACH: ALL for EACH encodes multi-view video exploiting the time and inter-view domain correlation of video frames. The access point uses ALL subcarriers to transmit EACH encoded video frame. ALL for EACH is a baseline for performance with the simplest scheme of multi-view video streaming over a wireless network with subcarriers.

2) Retransmission: Retransmission also transmits each encoded video frame using all subcarriers. When errors occurred in a video frame, an access point retransmits the video frame using all subcarriers. Retransmission is a baseline for performance with the scheme of preventing multiview error propagation.

3) SMVS/SA: As shown in Section 3, SMVS/SA is the proposed approach. SMVS/SA allocates each encoded video frame to subcarriers using the proposed First Allocation. After the allocation, an access point transmits the video frames over a wireless network based on the allocation.

Maintenance of High Video Quality: We compared video quality to evaluate the maintenance of high video quality for the three encoding/decoding schemes described in Section 4.2.

We implemented the three encoding/decoding schemes on a multi-view video encoder and decoder. The multiview video decoder first transmits a request packet to the multi-view video encoder. The multi-view video encoder encoded the requested multi-view video sequence and allocated the encoded bit streams to subcarriers based on each encoding/decoding scheme. The error rate of each subcarrier was a random rate between 0 and p_{max} [%], which is the maximum error rate. After the allocation, the multiview video encoder transmitted the bit streams by OFDM. When an error occurred in subcarrier communication, the multi-view video decoder exploited error concealment operation to compensate the error. When the multi-view video decoder received all video frames in one GGOP, the multiview video decoder measured the video quality. We performed one thousand evaluations and obtained the average video quality.

Figure 4 shows the video quality as a function of max-

imum error rate, where the GOP length is eight [frames], the number of cameras is six, and video sequence is "Ball-room". Figure 4 shows the following:

- SMVS/SA achieves higher video quality than ALL for EACH when the maximum error rate increases. For example, SMVS/SA improves video quality by 6.7 [dB] compared to ALL for EACH, when the maximum error rate is 10 [%]. SMVS/SA transmits significant video frames in high channel gain subcarriers to minimize the effect of multi-view error propagation.
- 2) ALL for EACH has the lowest video quality of three encoding/decoding schemes. This is because ALL for EACH transmits a video frame over wireless networks using all subcarriers. If an error occurs in subcarrier communication, the video frame is lost even when the other subcarrier communication is successful. The frame loss induces multi-view error propagation among cameras and low video quality.
- 3) Retransmission achieves the highest video quality in other encoding/decoding schemes. Even when errors occur in transmitted video frames, a video encoder retransmits the video frames until a user node receives the video frames successfully. Therefore, the user node decodes the video frames without errors.

Suppression of Communication Delay: We compared communication delay between an access point and a user node to evaluate the suppression of communication delay for the three encoding/decoding schemes described in Section 4.2.

A user node transmitted a request packet for one GGOP to an access point. The access point sent back video frames in one GGOP based on the request packet by OFDM. When the user node successfully received the video frames, the user node calculated communication delay of the received video frames. If the user node detected errors in the video frames, the user node did not transmit the next request packet. In this case, the access point retransmitted the video frames of all GGOP, the user node calculated communication delay. We performed one million evaluations and obtained the average communication delay.

We assumed that the bandwidth of wireless networks was 20 [MHz] and the access point modulated bit streams in each subcarrier by 16 QAM. The duration of one OFDM symbol was 4 [μ s] and the guard interval was 800 [ns]. The number of subcarriers is 48. These settings were based on IEEE 802.11a.

Figure 5 shows the communication delay as a function of video quality, where the GOP size is eight [frames], the number of cameras is six, and video sequence is "Ballroom". Figure 5 shows the following:

 As the maximum error rate increases, SMVS/SA achieves lower communication delay than Retransmission. For example, SMVS/SA reduces communication delay by 41.3 [%] compared to Retransmission,



Fig. 4 PSNR vs. maximum error rate.

when the maximum error rate is 5 [%]. This is because SMVS/SA maintains high video quality without retransmission by transmitting significance video frames over high channel gain subcarriers.

- 2) As the maximum error rate increases, the communication delay of Retransmission increases rapidly. To receive video frames without errors at a user node, a video encoder retransmits the video frames repeatedly. The retransmission increases communication delay because the retransmitted video frames have high traffic.
- **3)** ALL for EACH achieves the lowest communication delay even when the maximum error rate increases. This is because ALL for EACH transmits each video frame by all subcarriers.
- 4.3 Effect of Different Subcarrier Allocation

Section 4.2 revealed the baseline performance of SMVS/SA using First Allocation. To evaluate the performance of SMVS/SA in more details, we compared video quality and computational complexity for four subcarrier allocation: Brute Force, Random, First Allocation, and Concentric Allocation.

1) Brute Force: Brute force is the upper bound of video quality for multi-view video streaming over a wireless network with subcarriers. Brute force calculates video quality from all combinations of the subcarriers and the video frames in one GGOP and selects the best combination.

2) Random: Random is the simplest method of subcarrier allocation. A video encoder allocates each encoded video frame to subcarriers randomly.

3) First Allocation: First Allocation is our proposed heuristic allocation described in Section 3.5.1.

4) Concentric Allocation: Concentric Allocation is also our proposed heuristic allocation described in Section 3.5.2. **Video Quality:** We first compared the video quality of the proposed SMVS/SA for four subcarrier allocation described in Section 4.3.

As in the evaluation in Section 4.2, we implemented



Fig. 5 Communication delay vs. maximum error rate.

SMVS/SA with different subcarrier allocation on MATLAB video encoder and decoder. We performed one thousand evaluations and obtained the average video quality.

Figure 6 shows the video quality as a function of maximum error rate, where the GOP length is eight [frames], the number of cameras is six, and video sequence is "Ballroom". Figure 6 shows the following:

- Even when the maximum error rate increases, the video quality of First Allocation approaches that of brute force. For example, the difference of video quality between First Allocation and brute force is up to 0.57 [dB] when the maximum error rate is 10 [%]. First Allocation achieves high video quality without the calculation of all combinations of subcarriers and video frames.
- 2) The video quality of Concentric Allocation is lower than that of First Allocation. Concentric Allocation concentrically allocates high channel gain subcarriers to the neighbor video frames of an initial camera. When a video encoder allocates subcarriers to anchor frames of other cameras, Concentric Allocation allocates low channel gain subcarriers to the anchor frames as the distance between the initial camera and a camera increases. The high error rate of anchor frames induces lower video quality than First Allocation.

Computational Complexity: To evaluate the overhead of each subcarrier allocation, we compared the computational complexity of the proposed subcarrier-gain based multiview rate distortion for four subcarrier allocation.

An access point encoded a multi-view video sequence and calculated the proposed multi-view rate distortion by the four subcarrier allocation. We measured the number of the calculations of the network-induced distortion, which is equation (2), per one GGOP as computational complexity.

Figure 7 shows the computational complexity per one GGOP as a function of the number of requested cameras, where the GOP length is eight [frames]. Figure 7 shows the following:



Fig. 6 PSNR vs. maximum error rate for different subcarrier allocation.



Fig. 8 Computational complexity vs. number of cameras.

- As the number of requested cameras increases, First and Concentric Allocation reduce the computation of significance prediction. The proposed heuristic calculation decides sub-optimal allocation between video frames and subcarriers for maintenance of high video quality with low overheads.
- 2) As the number of requested cameras increases, the computation of brute force calculation increases exponentially. The brute force calculation decides the best allocation between video frames and subcarriers to achieve the highest video quality. However, the enormous computation induces high overheads for significance prediction.

Next, we compared the computational complexity of Random, First Allocation, and Concentric Allocation in more details. We assumed that the number of requested cameras in this evaluation was up to 16. Note that the original test sequence consisted of eight cameras. To evaluate computational complexity in more than eight cameras, we copied the original test sequence in order.

Figure 8 shows the computational complexity per one



Fig. 7 Computational complexity vs. number of cameras.



Fig. 9 Predicted and measured PSNR vs. maximum error rate

GGOP for the three subcarrier allocation as a function of the number of requested cameras, where the GOP length is eight [frames]. Figure 8 shows the following:

- The computational complexity of Concentric Allocation is lower than First Allocation. Concentric Allocation handles a small number of combinations between video frames and subcarriers at each calculation. Concentric Allocation reduces overheads for decision of suboptimal allocation between video frames and subcarriers compared to First Allocation even when the number of requested cameras increases.
- 2) When the number of requested cameras is more than the GOP length, the computational complexity of Concentric Allocation approaches $O(N_{\text{GOP}}!)$. Concentric Allocation concentrically calculates multi-view rate distortion from the first video frame of an initial camera. After Concentric Allocation calculated the rate distortion for the last video frame in the initial camera, Concentric Allocation selects N_{GOP} video frames of other cameras and the same number of high channel gain subcarriers. Concentric Allocation repeatedly calculates

the rate distortion for N_{GOP} video frames until Concentric Allocation calculates the rate distortion for the first video frame of the edge camera.

- 3) As the number of requested cameras increases, the computation of First Allocation increases exponentially. First Allocation calculates network-induced distortion for video frames of all requested cameras at each time. When the number of cameras is large, First Allocation needs to handle a large number of combinations between video frames and subcarriers at each time.
- 4) Random achieves the lowest computational complexity in other subcarrier allocation. Random allocates subcarriers to video frames regardless of the channel gain of subcarriers and the significance of video frames.

4.4 Significance Prediction Accuracy

We evaluated the accuracy of the proposed significance prediction. If the accuracy is low, an access point incorrectly allocates video frames to high-low channel gain subcarriers. As the result, the video quality of a multi-view video sequence will degrade.

An access point predicted the quality of video frames in one GGOP by the proposed subcarrier-gain based multiview rate distortion. The access point calculated the proposed multi-view rate distortion based on the error rate of each subcarrier. To decide the error rate, the access point generated a random rate between 0 and p_{max} [%] for each subcarrier. After the calculation, the access point allocated video frames to subcarriers based on the prediction and transmitted the video frames to a user node. Errors occurred in the video frames during the communication based on the error rate of each subcarrier. When the user node received the video frames, the user node measured actual video quality in one GGOP. We performed one thousand evaluations and obtained the average video quality.

Figure 9 shows the predicted and measured PSNR of First and Concentric Allocation as a function of the maximum error rate where the GOP length is eight [frames] and the number of cameras is six, and video sequence is "Ballroom". Figure 9 shows the following:

- When error rate is low, the differences between predicted and measured PSNR of heuristics are small. When the maximum packet loss ratio is 1 [%], the differences between the predicted and measured PSNR in First Allocation are up to 0.39 [dB]. An access point is able to predict each video frame quality accurately based on the channel gains of each subcarrier.
- 2) As the maximum error rate increases, the differences between predicted and measured PSNR become larger. When the the maximum packet loss ratio is 10 [%], the differences between the predicted and measured PSNR in First Allocation are 1.67 [dB] (95 [%] confidence interval, 1.47-1.87 [dB]). As the maximum error rate increases, a video decoder exploits an early video frame for error concealment operation when a video frame is



Fig. 10 Improvement of PSNR from ALL for EACH vs. maximum error rate for different video sequences.

lost. On the other hand, an access point predicts the significance of a video frame using error rate and the previous time/camera video frame. The distortion between the lost video frame and the early video frame is significantly larger than the distortion between the lost video frame and the previous time video frame. The large distortion induces the large differences between predicted and measured PSNR.

4.5 Effect of Different Video Sequences

Section 4.2 and 4.3 revealed the performance of SMVS/SA using the Ballroom video sequence. However, the performance may change when user requests different scenes. To evaluate the effect of multi-view video contents on video quality, we compared video quality for different video sequences.

As in the evaluation in Section 4.2, we implemented ALL for EACH and First Allocation on MATLAB video encoder and decoder. The only difference from the evaluation in Section 4.2 is that the encoder encodes the video frames of "Exit" and "Vassar". After the evaluation, we compared the video quality of First Allocation to that of ALL for EACH.

Figure 10 shows the improvement of video quality from ALL for EACH as a function of maximum error rate for different video sequences, where the GOP length is eight [frames] and the number of cameras is six. Figure 10 shows the following:

 SMVS/SA maintains high video quality independent of the video sequence. Note that the degree of improvement varies with the motion of the video sequence. For example, First Allocation improves video quality by 6.3 [dB] compared to ALL for EACH for "Exit" when the maximum error rate is 10 [%] and video sequence is "Exit". If a video frame is loss, a user node exploits previously received video frames for error concealment. When a video sequence is fast motion, dis-





Fig. 11 Experimental equipment. Fig. 12 C

Fig. 12 Channel trace environment. na

rigite video quality in a real whereas herwork.

tortion between the lost and the received video frames is large. The large distortion induces low video quality.

2) First Allocation improves video quality by 2.8 [dB] compared to ALL for EACH for "Vassar" when the maximum error rate is 10 [%]. Vassar is less motion and the distortion between the lost and the received video frames is small. Therefore, the improvement of video quality becomes lower even when the maximum error rate increases.

4.6 Trace-driven Simulation

Section 4.2 and 4.3 discussed the performance of SMVS/SA with the random error rate of each subcarrier. This section evaluates the performance of SMVS/SA in a real wireless network. We compared video quality for five schemes using a trace-driven simulator based on MATLAB video encoder: ALL for EACH, Random, Brute force, First Allocation, and Concentric Allocation.

We traced the channel quality of IEEE 802.11a OFDM link for the trace-driven simulator. To trace the OFDM link, we used two GNU Radio/USRP N200 transceivers [34] with XCVR 2450 RF frond-ends [35] and control PCs as shown in Fig. 11. The USRP N200 is a software radio that allows the channel trace of each subcarrier. When coupled with XCVR 2450 radio front-ends, the USRP allows channel trace at 5.11 [GHz]. To trace the channel quality of each subcarrier by USRP N200, we run a program based on RawOFDM [36].

We built our channel trace environment at our laboratory in Shizuoka University, Japan. The two USRP N200 transceivers and PCs are placed at one room as shown in Fig. 12. Each USRP N200 connected GPSDO Kit [37] to synchronize between the two USRPs. All channel traces were conducted in the 5.11 [GHz] test experiment band with 2 [MHz] bandwidth, which is licensed by Ministry of Internal Affairs and Communications, Japan. The transmission power of the USRP N200 with XCVR 2450 is about -8.36 [dBm]. Each USRP N200 and PC are connected by wire as an access point and a user node, respectively. The access point transmits modulated symbols to the user node over subcarriers every 4 $[\mu s]$. The access point exploited 16 QAM for modulation and 48 subcarriers. The user node recorded the bit errors of each subcarrier's symbols for one minute

An access point allocates encoded video frames to subcarriers based on the recorded bit errors in each subcarrier. After the allocation, the access point modulated the video frames in each subcarrier using 16 QAM with 4 bits per symbol. The access point transmitted the modulated symbols by OFDM to a user node. The transmitted symbols in each subcarrier are lost based on the recorded bit errors in each subcarrier. Specifically, the maximum error rate of the subcarriers is approximately 10 [%]. When the user node received the symbols of a video frame and bit errors occurred in the symbols, the user node regarded the video frame as the lost video frame. The user node exploited error concealment operation for the video frame. When the user node received all video frames in one GGOP, the user node measured the video quality of the received video frames.

Figure 13 shows the video quality of each scheme, where the GOP length is eight [frames] and the number of cameras is six, and video sequence is "Ballroom". Figure 13 shows the following:

- First Allocation achieves higher video quality than other encoding/decoding schemes in a real wireless network. For example, First Allocation improves video quality by 2.7 [dB] compared to ALL for EACH and 2.2 [dB] compared to Random. First Allocation minimizes the effect of the high error rate subcarriers by allocating significant video frames to other low error rate subcarriers.
- 2) Each encoding/decoding scheme achieves higher video quality compared to the results in Section 4.2 and 4.3. This is because errors do not occur in the half of subcarriers and a user node receives more video frames compared to the above evaluations.

5. Conclusion

The present paper proposes SMVS/SA for multi-view video streaming over a wireless network with subcarriers. SMVS/SA maintains high video quality by transmitting significant video frames in high channel gain subcarriers.

SMVS/SA has two contributions. The first one is subcarriergain based multi-view rate distortion to predict each frame's significance. The second one is heuristic algorithms to decide sub-optimal allocation between video frames and subcarriers with low overheads. Evaluations using MATLAB video encoder and USRP N200 software radio reveal that SMVS/SA enables a small degradation in video quality in a real wireless network.

In the future work, we will examine the effect of the number of subcarriers on the performance of SMVS/SA. Evaluations assume the number of subcarriers is the same number of video frames in 1 GGOP. We will also extend the proposal to multiple user environments to provide multiview video services to mobile devices efficiently.

References

- M. Tanimoto and K. Suzuki, "Global view and depth (GVD) format for FTV/3DTV," Three-Dimensional Imaging Visualization And Display, pp.1–10, 2013.
- [2] L. Do, S. Zinger, and D.W. Peter, "Conversion of free-viewpoint 3D multi-view video for stereoscopic displays," IEEE ICME, pp.1730– 1734, 2010.
- [3] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," Stereoscopic Displays and Virtual Reality Systems, pp.93–104, 2004.
- [4] W.R. Mark, L. McMillan, and G. Bishop, "Post-rendering 3D warping," Symposium On Interactive 3D Graphics, pp.7–16, 4 1997.
- [5] Text of ISO/IEC 14496-10:2008/FDAM 1 ISO/IEC JTC1/SC29/WG11, Multiview Video Coding, 2008.
- [6] Z. Liu, G. Cheung, and Y. Ji, "Unified distributed source coding frames for interactive multiview video streaming," IEEE ICC, pp.2048–2053, 2012.
- [7] H. Huang, B. Zhang, S.H.G. Chan, G. Cheung, and P. Frossard, "Coding and replication co-design for interactive multiview video streaming," IEEE INFOCOM, pp.2791–2795, 2012.
- [8] Z. Pan, M. Bandai, and T. Watanabe, "A user dependent scheme for multi-view video live streaming," Journal of Computational Information Systems, vol.9, no.4, pp.1439–1448, 2013.
- [9] T. Fujihashi, Z. Pan, and T. Watanabe, "UMSM: a traffic reduction method on multi-view video streaming for multiple users," IEEE Transactions on Multimedia, vol.16, no.2, pp.228–241, 2014.
- [10] O. Edfors, M. Sandell, and J.J.V.D. Beek, "OFDM channel estimation by singular value decomposition," IEEE Transactions on Communications, vol.46, no.1, pp.931–939, 1996.
- [11] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," ACM SIG-COMM, pp.1–12, 2010.
- [12] Y. Zhou, C. Hou, W. Xiang, and F. Wu, "Channel distortion modeling for multi-view video transmission over packet-switched networks," IEEE Transactions on Circuits and Systems for Video Technology, vol.21, no.11, pp.1679–1692, 2011.
- [13] P. Iannucci, J. Perry, H. Balakrishnan, and D. Shah, "No symbol left behind: A link-layer protocol for rateless codes," ACM MOBICOM, pp.17–27, 2012.
- [14] J. Zhang, H. Shen, K. Tan, R. Chandra, Y. Zhang, and Q. Zhang, "Frame retransmissions considered harmful:improving spectrum efficiency using micro-acks," ACM MOBICOM, pp.89–100, 2012.
- [15] K. Jamieson and H. Balakrishnan, "Ppr: Partial packet recovery for wireless networks," ACM SIGCOMM, pp.409–420, 2007.
- [16] A. Majumdar, D.G. Sachs, I.V. Kozintsev, K. Ramchandran, and M.M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," IEEE Transactions on Circuits and Systems for Video Technology, vol.12, no.6, pp.524–534, 2002.

- [17] M. Choi, W. Sun, J. Koo, S. Choi, and K.G. Shin, "Reliable video multicast over Wi-Fi networks with coordinated multiple APs," IEEE INFOCOM, pp.1–9, 2014.
- [18] Z. He, J. Cai, and C.W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol.12, no.6, pp.511–523, 2002.
- [19] J.C. Kwon and J.K. Kim, "Adaptive code rate dicision of joint source-channel coding for wireless video," IEEE Electronics Letters, vol.38, no.25, pp.1752–1754, 2002.
- [20] J. Xu, Q. Zhang, W. Zhu, X.G. Xia, and Y.Q. Zhang, "Optimal joint source-channel bit allocation for MPEG-4 fine granularity scalable video over OFDM system," IEEE International Symposium on Circuits and Systems, pp.360–363, 2003.
- [21] O. Fresned, V. Araujo, L. Casted, and G. Frias, "Analog joint sourcechannel coding for OFDM systems," IEEE Workshop on Signal Processing Advances in Wireless Communications, pp.704–708, 2013.
- [22] C. Zhang, H. Yang, S. Yu, and X. Yang, "Gop-level transmission distortion modeling for mobile streaming video," Signal Processing: Image Communication, vol.23, no.2, pp.116–126, 2008.
- [23] Z. Li, J. Chakareski, X. Niu, Y. Zhang, and W. Gu, "Modeling and analysis of distortion caused by markov-model burst packet losses in video transmission," IEEE Transactions on Circuits and Systems for Video Technology, vol.19, no.7, pp.917–931, 2009.
- [24] S. Jakubczak, H. Rahui, and D. Katabi, "One-size-fits-all wireless video," ACM HotNets, pp.1–6, 2009.
- [25] L.X. Lin, H. Wenjun, P. Qifan, W. Feng, and Z. Yongguang, "Parcast: Soft video delivery in MIMO-OFDM WLANs," ACM MOBI-COM, pp.233–244, 2012.
- [26] S.T. Aditya and S. Katti, "FlexCast: Graceful wireless video streaming," ACM MOBICOM, pp.277–288, 2011.
- [27] A.S. Tan, A. Aksay, G.B. Akar, and E. Arikan, "Rate-distortion optimization for stereoscopic video streaming with unequal error protection," EURASIP Journal on Advances in Signal Processing, vol.2009, no.7, pp.1–14, 2009.
- [28] J. Chakareski, "Transmission policy selection for multi-view content delivery over bandwidth constrained channels," IEEE Transactions on Image Processing, vol.23, no.2, pp.931–942, 2014.
- [29] IEEE Standard 802.11n, Enhancements For Higher Throughput, 2009.
- [30] S. Heiko, M. Detlev, and W. Thomas, "Overview of the scalable video coding extension of the h. 264/avc standard," IEEE Transactions of Circuits And Systems for Video Techonology, vol.17, no.9, pp.1103–1120, 2007.
- [31] Y. Liu, J. Wang, and H. Zhang, "Depth image-based temporal error concealment for 3-D video transmission," IEEE Transactions on Circuits and Systems for Video Technology, vol.20, no.4, pp.600–604, 2010.
- [32] A.A. Muhit, "H.264 Baseline Codec v2." http://www.mathworks.com /matlabcentral/fileexchange/40359-h-264-baseline-codec-v2, 2013.
- [33] ISO/IEC JTC1/SC29/WG11, Multiview Video Test Sequences from MERL, 2005.
- [34] Ettus Research, "USRP N200/N210 Networked Series," 2012.
- [35] Ettus Research, "XCVR2450 2.4-2.5 GHz, 4.9-5.9 GHz Tx/Rx." https://www.ettus.com/product/details/XCVR2450.
- [36] S. Jakubczak, "How to build and run RawOFDM." http://people.csail .mit.edu/szym/rawofdm/README.html/, 2014.
- [37] Ettus Research, "Ettus GPSDO." https://www.ettus.com/content/files/gpsdokit_4.pdf.

