Wireless Multi-view Video Streaming with Subcarrier Allocation by Frame Significance

Takuya Fujihashi*, Shiho Kodera†, Shunsuke Saruwatari†, Takashi Watanabe*
*Graduate School of Information Science and Technology, Osaka University, Japan
†Faculty of Informatics, Shizuoka University, Japan

Abstract—When an access point transmits multi-view video over wireless networks with multiple subcarriers, errors occur in 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 with a high quality subcarrier to minimize the effect of the errors. Evaluations using MERL's benchmark test sequences reveal that SMVS/SA achieves a slight degradation of video quality. For example, SMVS/SA improves video quality by 8.0 [dB] compared to standard H.264/AVC MVC when the maximum packet loss ratio of each subcarrier is 10 %.

I. INTRODUCTION

With the progress of wireless technology and video coding technology for multi-view video, the demand of watching 3D video on wireless devices increases [1]. To satisfy the demand, the wireless technique and multi-view video coding technique have been studied independently. The typical studies of multi-view video coding are Multi-view Video Coding (MVC), Interactive Multi-view Video Streaming (IMVS) [2], User dependent Multi-view video Streaming (UMS) [3], and UMS for Multi-user (UMSM) [4]. These studies focus on the reduction of video traffic by exploiting the correlation of time domain and inter-camera domain of video frames.

In view of wireless networks, Orghogonal Frequency Division Multiplexing (OFDM) [5] is used in modern wireless technology. 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] [6]. The low channel gains induce low packet reception rate at a receiver.

When a video encoder simply transmits multi-view video over 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 due to 2D error propagation [7].

To minimize the effect of 2D error propagation, the present paper proposes Significance based Multi-view Video Streaming with Subcarrier Allocation (SMVS/SA) for multi-view video streaming over wireless networks. SMVS/SA achieves the reduction of communication delay and video traffic while maintaining high video quality. The key feature of SMVS/SA is to transmit significant video frames with high channel gain subcarriers. The significant video frames have a great effect on video quality when errors occur in the video frames. The present paper makes two contribution. First, we propose subcarrier-gain based 2D rate distortion to predict the effect of each video frame on video quality when the video frame is lost. Second, we propose a heuristic algorithm to decide the allocation between video frames and subcarriers with low computation. The allocation achieves sub-optimal 2D rate distortion under the different subcarrier channel gains. Evaluations using the MATLAB video encoder and MERL's benchmark test sequences reveal that SMVS/SA achieves only a slight degradation of video quality.

The remainder of the present paper is organized as follows. Section II presents a summary of related research. We present the details of SMVS/SA in Section III. In Section IV, evaluations are performed to reveal the maintenance of video quality for proposed SMVS/SA. Finally, conclusions are summarized in Section V.

II. RELATED RESEARCH

This study is related to joint source-channel coding and 2D rate distortion based video streaming.

Joint source-channel coding: There are some studies about joint source-channel coding: SoftCast [8], FlexCast [9] and ParCast [10]. SoftCast 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 extends the SoftCast's design to MIMO-OFDM. FlexCast focuses on bit-level significance of each single-view video frame. FlexCast adds rateless codes to the bit based on the significance to prevent 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 and frame-level significance to improve 3D video delivery quality over wireless networks.

2D rate distortion based video streaming: Several studies have been proposed for the maintenance of high video quality. [11] introduces an end-to-end 2D rate distortion model for 3D video to achieve optimal encoder bitrate. [11] only analyzes for 3D video with the left and right camera. [7] proposes the average packet loss based 2D rate distortion to analyze the distortion with multiple cameras. [12] proposes network bandwidth based 2D rate distortion for bandwidth constrained channels. The basic concept of the proposed subcarrier-gain based 2D rate distortion is based on these studies. SMVS/SA
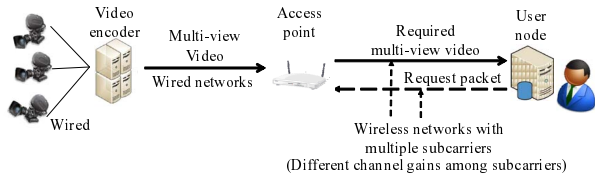
Fig. 1. System model of multi-view video streaming over wireless network

considers the channel gain differences of subcarriers for 2D rate distortion to maintain high video quality in practical wireless networks.

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

A. 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 encoder node is connected to an access point by wired networks. The access point is connected to a user node by wireless networks with multiple subcarriers. The wireless networks have 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 decoded 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 III-B.

(2) Video Encoding: When the access point receives the request packet, the access point encodes a multi-view video sequence in 1 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 8 frames, for each camera. The details of video encoding are described in Section III-C.

(3) Significance Prediction: After the encoding, the encoder predicts which video frames should be transmitted in high channel gain subcarriers. To predict the significance of each

video frame, SMVS/SA proposes subcarrier-gain based 2D rate distortion. The details of significance prediction are described in Section III-D.

(4) Heuristic Calculation: The disadvantage of the subcarrier-gain based 2D rate distortion is high computational complexity. To reduce the computational complexity, SMVS/SA proposes a heuristic algorithm. The details of the heuristic algorithm are described in Section III-E.

(5) Sorting and Video Transmission: The access point allocates video frames to the 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 III-F.

(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 decoding, the user node plays back multi-view video on display. The details of video decoding are described in Section III-G.

B. 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 1 GGOP. Each request packet consists of three fields: watched camera ID, required camera ID, and Channel State Information (CSI). The watched camera ID is an 8-bit field that indicates the camera ID being watched by the user. The required camera ID is arrays of an 8-bit field that indicates the camera IDs required by the user. The CSI field is based on 802.11n Channel State Information packet [13]. 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.

C. Video Encoding

After the access point received the request packet, the encoder encodes multi-view video based on the watched and required camera ID fields in the request packet.

The access point encodes an anchor frame of the watched camera into I-frame and the subsequent video frames into P-frames. 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. After encoding the video frames of watched camera, the access point encodes video frames of the required cameras.

The anchor frames of the required cameras are encoded into P-frame using the same time anchor frame 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 previous time frame of the same camera and the same time frame 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 video frames. The reference video frame achieves the lowest distortion of video encoding. After the video encoding of all video frames in 1 GGOP, the access point gets bit streams of each video frame.

D. Significance Prediction

After encoding, the access point predicts the significance of each video frame. To predict the significance, the present paper proposes subcarrier-gain based 2D rate distortion. The subcarrier-gain based 2D 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 2D rate distortion as

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

where P is $N_{\text{camera}} \times N_{\text{GOP}}$ matrix of packet reception ratio. The minimum 2D rate distortion reveals which video frames should be transmitted by the high channel gain subcarriers to maintain high video quality. Denote by N_{camera} and N_{GOP} the number of required cameras and the length of each GOP, respectively.

Assumption: $N_{\text{camera}} N_{\text{GOP}}$ is the number of video frames in 1 GGOP and is smaller than the number of subcarriers in OFDM. At the user node, SMVS/SA assumes that a proper error concealment operation is performed on lost video frame. Generally, the error concealment operation resorts to either temporal or inter-camera concealment. For simplicity, SMVS/SA performs the error concealment operation for a video frame when errors occur in bits of the video frame. Consequently, the packet reception ratio is equivalent to the frame reception ratio.

Definition: Let $D_{\text{GGOP}}(P)$ be the overall subcarrier-gain based 2D rate distortion in 1 GGOP at the user node. $D_{\text{GGOP}}(P)$ is defined as encoding-induced distortion and network-induced distortion, denoted by $D_{\text{encoding}}(s,t)$ and $D_{\text{network}}(P,s,t)$. They are expressed as:

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

$$D_{\text{encoding}}(s,t) = E\{[F_{\text{i}}(s,t) - \hat{F}_{\text{i}}(s,t)]^2\} \qquad (3)$$

$$D_{\text{network}}(P,s,t) = p(s,t)D_{\text{success}}(s,t) + (1-p(s,t))D_{\text{loss}}(s,t)$$
$$(4)$$

where $F_{\text{i}}(s,t)$ is the original value of pixel i in $M(s,t)$, $\hat{F}_{\text{i}}(s,t)$ is the reconstructed values of pixel i in $M(s,t)$ at the encoder, and $p(s,t) \in P$ is the packet reception ratio for the frame at camera s and time t. The value of $p(s,t)$ is based on the channel gain of a subcarrier. Denote by $M(s,t)$ the frame at camera s and time t. Moreover, $E\{\cdot\}$ denotes the expectation taken over all the pixels in frame $M(s,t)$.

As can be seen from equation (3), encoding-induced distortion refers to the Mean Square Error (MSE) between the original frame and the reconstructed video frame at the encoder. The network-induced distortion consists of the distortion when communication is successful and failed, denoted

by $D_{\text{success}}(s,t)$ and $D_{\text{loss}}(s,t)$, respectively. $D_{\text{success}}(s,t)$ is expressed as:

$$D_{\text{success}}(s,t) = E\{[\hat{F}_{\text{i}}(s,t) - \tilde{F}_{\text{i}}(s,t)]^2\} \qquad (5)$$

where $\tilde{F}_{\text{i}}(s,t)$ is expressed according to the type of video frame and the reference video frame as:

$$\tilde{F}_i(s,t) = \begin{cases} \hat{F}_{\text{i(s,t)}} & \text{if M(s,t) = I-frame} \\ \hat{F}_{\text{p(i)}}(s-1,t) & \text{else if } F_{\text{p(i)}} \in M(s-1,t). \\ \hat{F}_{\text{p(i)}}(s,t-1) & \text{else.} \end{cases}$$
$$(6)$$

where p(i) is the index of the matching pixel in the reference video frame.

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

$$D_{\text{loss}}(s,t) = E\{[\hat{F}_{\text{i}}(s,t) - \breve{F}_{\text{i}}(s,t)]^2\} + D_{\text{previous}} \qquad (7)$$

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

$$\breve{F}_{\text{i}}(s,t) = \begin{cases} \hat{F}_{\text{conceal(i)}}(s-1,t) & \text{if } \hat{F}_{\text{conceal(i)}} \in M(s-1,t). \\ \hat{F}_{\text{conceal(i)}}(s,t-1) & \text{else.} \end{cases}$$
$$(8)$$

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

$$D_{\text{previous}}(s,t) = D_{\text{network}}(P,s,t-1) \qquad (9)$$

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

$$D_{\text{previous}}(s,t) = D_{\text{network}}(P,s-1,t) \qquad (10)$$

E. Heuristic Calculation

The minimum subcarrier-gain based 2D rate distortion reveals which video frames should be transmitted by the high channel gain subcarriers. However, the computational complexity of network-induced distortion is high. Specifically, an access point calculates the minimum network-induced distortion, which is equation (4), from all combinations of the subcarriers and the video frames in 1 GGOP. As the result, the computational complexity of equation (4) is $O([N_{\text{camera}} N_{\text{GOP}}]!)$. To calculate sub-optimal network-induced distortion with low computation, SMVS/SA proposes a heuristic algorithm. The heuristic focuses 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 heuristic first allocates a high channel gain subcarrier for early reference video frames to prevent the degradation of subsequent video frames.

We explain the details of the proposed heuristic. An access point selects I-frame and the highest packet reception ratio p in $P_{\text{subcarriers}}$. $P_{\text{subcarriers}}$ is a set of packet reception ratio in each subcarrier. The packet reception ratio is calculated by

the channel gain of the subcarrier. 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 packet reception ratio p_{n} in $P_{\text{subcarriers}}$. The access point calculates the sum of proposed 2D distortion of each P-frame using each p_{n} from equation (4), and decides the best allocation of p_{n} which achieves the lowest distortion. 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 $P_{\text{subcarriers}}$. The access point selects m P-frames of the previously selected P-frame's neighborhood and the same number of high packet reception ratio p_{m} in $P_{\text{subcarriers}}$. The access point also calculates the sum of proposed 2D distortion of each P-frame using each p_m from equation (4), and decides the best allocation of p_m which achieves the lowest distortion. The access point repeats the heuristic for all video frames in 1 GGOP. The heuristic reduces the computation to $O([N_{\text{GOP}} - \frac{N_{\text{camera}}}{2}] \cdot N_{\text{camera}}!)$ when $N_{\text{GOP}} > \frac{N_{\text{camera}}}{2}$. Otherwise, the computation saturates at approximately $O([2N_{\text{GOP}}]!)$.

F. Sorting and Video Transmission

After the significance prediction, the access point allocates bit streams of each video frame to subcarriers based on the prediction and transmits the bit streams to a user node over wireless networks by OFDM. The bit streams in each subcarrier are modulated equally, using BPSK, QPSK, QAM-16, or QAM-64, with 1, 2, 4 or 6 bits per symbol, respectively. The modulated symbols in each subcarrier are modulated by 1 OFDM symbol. The access point inserts up to 44 OFDM symbols into 1 video packet and transmits the video packets to the user node.

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

G. 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 sub-carrier's symbols into the bits of each bit stream from constellations of several different modulations (BPSK, QPSK, QAM-16, QAM-64). The access point assembles the demodulated bit streams in respective subcarriers. The subcarrier-based 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. Finally, the user node plays back multi-view video on display.

IV. EVALUATION

A. Evaluation Settings

To evaluate the video quality of SMVS/SA, we implemented the SMVS/SA encoder/decoder with MATLAB. The evaluation used a multi-view video test sequence: "Ballroom"

(faster 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) [14]. 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 each sequence was set to eight frames. We used 250 frames per sequence for all of the evaluations. The Quantization Parameter (QP) value used in our experiment was 25.

The evaluation assumes that one access point and one user node were connected by wireless network with multiple subcarriers. The number of subcarriers was the same as the number of video frames in 1 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 evaluated the video quality of three encoding/decoding schemes: H.264/AVC MVC, SMVS/SA w/o Significance Prediction, SMVS/SA.

1) H.264/AVC MVC: H.264/AVC MVC encodes multi-view video exploiting the time domain and inter-view domain correlation of video frames. The access point transmits each encoded video frame using all subcarriers. H.264/AVC MVC is a baseline for performance with the simplest scheme.

2) SMVS/SA w/o Significance Prediction: SMVS/SA w/o Significance Prediction transmits each encoded video frame by randomly allocated subcarriers. SMVS/SA w/o Significance Prediction is a baseline for performance with subcarrier allocation of the proposed approach.

3) SMVS/SA: As shown in Section III, SMVS/SA is the proposed approach. SMVS/SA predicts the significance of each video frame by the proposed subcarrier-gain based 2D rate distortion. After the prediction, SMVS/SA allocates each encoded video frame to subcarriers based on the prediction and transmits the video frames over wireless networks.

We used the standard peak signal-to-noise ratio (PSNR) metric to evaluate multi-view video quality in 1 GGOP. $\text{PSNR}_{\text{GGOP}}$ represents the average video quality of multi-view video in 1 GGOP as follows:

$$\text{PSNR}_{\text{GGOP}} = 10\log_{10}\frac{(2^{\text{L}} - 1)HN_{\text{camera}}N_{\text{GOP}}W}{D_{\text{GGOP}}} \quad (11)$$

where D_{GGOP} is the measured 2D rate distortion in 1 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.

B. Baseline Performance

We compared the computational complexity reduction of the proposed 2D rate distortion for greedy and proposed heuristic calculation. We measured the computation of the network-induced distortion, which is equation (4), for each calculation and plotted the logarithm of the computation.

Figure 2 shows the logarithm of the computational in 1 GGOP as a function of the number of cameras. Figure 2 shows that as the number of cameras increases, the computation of greedy calculation increases exponentially. The greedy
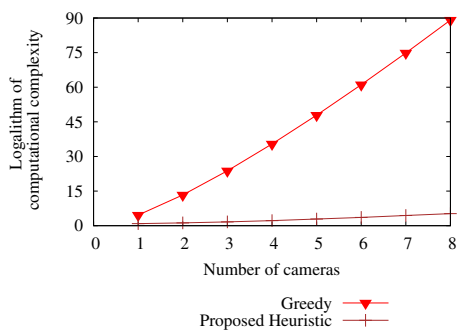
Fig. 2. Logarithm of computational complexity vs number of cameras
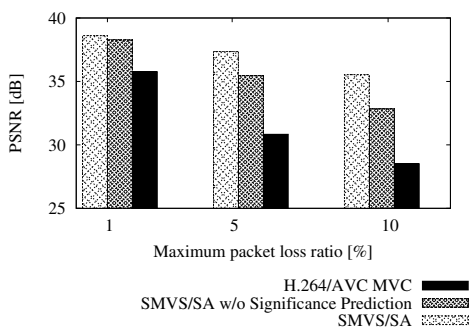

Fig. 3. PSNR vs. Packet loss ratio

calculation calculates the best combinations between video frames and subcarriers for the high video quality maintenance. However, the enormous computation induces high overheads for significance estimation.

C. Comparison

We compared the video quality to evaluate the maintenance of high video quality for the three encoding/decoding schemes described in Section IV-A.

We implemented the three encoding/decoding schemes on MATLAB video encoder. The MATLAB video encoder allocated encoded bit streams to subcarriers based on each encoding/decoding scheme. The packet loss ratio of each subcarrier was a random rate r between 0 and p_{max} [%]. p_{max} was the maximum packet loss ratio. After the allocation, the MATLAB video encoder transmitted the bit streams by OFDM. When an error occurred in subcarrier communication, MATLAB video decoder exploited error concealment operation to compensate the error. We performed 1,000 evaluations and obtained the average video quality.

Figure 3 shows the video quality as a function of packet loss ratio. Figure 3 shows the following:

First, SMVS/SA achieves the highest video quality of the three encoding/decoding schemes even when packet loss ratio increases. For example, SMVS/SA maintains video quality by 8.0 [dB] compared to H.264/AVC MVC and 3.4 [dB] compared to SMVS/SA w/o Significance Prediction when the packet loss ratio is 10 %. SMVS/SA transmits significant video frames in high channel gain subcarriers to minimize the effect of 2D error propagation.

Second, SMVS/SA w/o Significance Prediction achieves higher video quality compared to H.264/AVC MVC. The scheme transmits a video frame to the decoder by a subcarrier. If channel gains of the subcarrier is high, the communication is successful, and vice-versa. As the result, SMVS/SA w/o Significance Prediction decreases the effect of low channel gain subcarriers.

Third, H.264/AVC MVC has the lowest video quality of three encoding/decoding schemes. This is because H.264/AVC MVC 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 2D error propagation among cameras and low video quality.

V. CONCLUSION

The present paper proposes SMVS/SA for multi-view video streaming over wireless networks with multiple subcarriers. SMVS/SA maintains high video quality by transmitting significant video frames in high channel gain subcarriers. Evaluations reveal that SMVS/SA enables a small degradation in video quality.

REFERENCES

[1] M. Tanimoto and K. Suzuki, "Global view and depth (GVD) format for FTV/3DTV," in *Three-Dimensional Imaging Visualization And Display*, 2013, pp. 1–10.

[2] Z. Liu, G. Cheung, and Y. Ji, "Unified distributed source coding frames for interactive multiview video streaming," in *IEEE ICC*, 2012, pp. 2048–2053.

[3] 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.

[4] T. Fujihashi, Z. Pan, and T. Watanabe, "UMSM: a traffic reduction method on multi-view video streaming for multiple users," *IEEE Transactons on Multimedia*, vol. 16, no. 2, pp. 228–241, 2014.

[5] 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.

[6] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," in *ACM SIGCOMM*, 2010, pp. 1–12.

[7] 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.

[8] S. Jakubczak, H. Rahui, and D. Katabi, "One-size-fits-all wireless video," in *ACM HotNets*, 2009, pp. 1–6.

[9] S. T. Aditya and S. Katti, "FlexCast: Graceful wireless video streaming," in *ACM MOBICOM*, 2011, pp. 277–288.

[10] L. X. Lin, H. Wenjun, P. Qifan, W. Feng, and Z. Yongguang, "Parcast: Soft video delivery in MIMO-OFDM WLANs," in *ACM MOBICOM*, 2012, pp. 233–244.

[11] 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.

[12] 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.

[13] IEEE Standard 802.11n, *Enhancements For Higher Throughput*, 2009.

[14] ISO/IEC JTC1/SC29/WG11, *Multiview Video Test Sequences from MERL*, 2005.