Multi-view Video Streaming with Mobile Cameras

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

***Abstract*—Multi-view video system includes three sections: acquisition, transmission, and display. This paper focuses on the acquisition of multi-view video. Existing multi-view video acquisition studies exploit multi-camera arrays mutually connected by wires. However, this imposes the limitations of places and objects. To overcome the limitations, we exploit multiple mobile cameras and wireless networks for multi-view video acquisition. The acquisition of the multi-view video needs to achieve a reduction in video traffic while maintaining high video quality for communication between mobile cameras and an access point. This paper proposes Multi-view Video Streaming with Mobile Cameras (MVS/MC) to satisfy these requirements. MVS/MC has two features: packet overhearing and transmission order control. First, each mobile camera overhears other cameras' video packets, and encodes its own video frames using the overheard video packets. Second, the access point controls the transmission order of the mobile cameras, thus realizing bidirectional inter-view prediction. Bidirectional inter-view prediction exploits the inter-camera domain correlation among the mobile cameras to further remove the redundant information. Evaluations using multi-view video sequences show that, compared with existing methods, MVS/MC reduces the volume of traffic with only a slight degradation in video quality. For example, MVS/MC reduces traffic by 52 % compared to existing methods when PSNR is 36 dB.**

I. INTRODUCTION

The development of 3D video technology has led to a new scene representation technique known as multi-view video. Multi-view video provides an immersive perception of a 3D scene, and has paved the way for many emerging 3D applications, such as free viewpoint video [1], [2], 3DTV, and immersive teleconferencing [3].

Figure 1 shows the structure of a multi-view video system. The acquisition section captures a scene using multiple synchronized cameras located at different spatial locations (viewpoints). The transmission section encodes the resulting video sequences, and transmits them to the display section, which displays the 3D scene.

A number of transmission technologies have been developed. For instance, Multi-view Video Coding (MVC) was issued as an amendment to H.264/MPEG-4 AVC [4]–[6], and Interactive Multi-view Video Streaming (IMVS) reduces multi-view video traffic for stored and playback streaming [7], [8]. User Dependent Multi-view Video Transmission (UDMVT) reduces multi-view video traffic for live streaming [9]–[11], and User dependent Multi-view video Streaming for Multi-user (UMSM) reduces multi-view video traffic for live streaming with multiple users [12]–[14].

Previous studies into the display of multi-view video focus on either the decoder or the display. Typical decoder-level studies of image display include depth image-based rendering [15] and 3D warping [16], while FTV [2] and integral 3D television [17] are display-level approaches.

Earlier studies of video acquisition exploit multi-camera arrays. A multi-camera array consists of multiple cameras that are mutually connected by wires. This imposes a limitation on multi-camera arrays, as it makes it inherently difficult to take the cameras outdoors to capture a scene.

To overcome this limitation, the present paper exploits wireless networks and mobile cameras for multi-view video acquisition. Multi-view video acquisition using wireless networks has two requirements: a reduction in network traffic, and high video quality. These requirements affect user satisfaction and application quality.

To this end, we propose Multi-view Video Streaming with Mobile Cameras (MVS/MC). MVS/MC has two features: packet overhearing, and transmission order control. First, each mobile camera overhears other cameras' communication, and receives the cameras' video frames. Each mobile camera encodes its own video frames with the overheard video frames, thus reducing the volume of video traffic. Second, an access point controls the order in which the mobile cameras transmit data. The transmission order enables bidirectional inter-view prediction among the mobile cameras, and this achieves further traffic reduction. Evaluations using a JMVC encoder and the MERL benchmark test sequences reveal that MVS/MC achieves low traffic volume with only a slight degradation in video quality.

The remainder of this paper is organized as follows. Section II presents a summary of current multi-view video acquisition techniques. We describe the concept of MVS/MC in Section III. In Section IV, we report the results of evaluations that reveal the reduction in traffic volume and measure the video quality of the proposed MVS/MC. Finally, our conclusions are summarized in Section V.

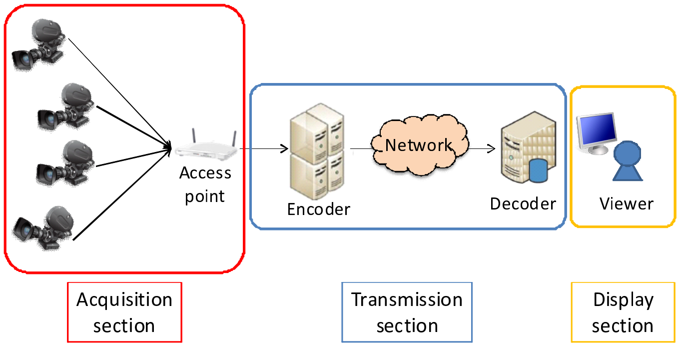


Fig. 1. Multi-view system

II. MULTI-VIEW VIDEO ACQUISITION

Multi-view video acquisition over wireless networks enables us to view indoor/outdoor scenes from every angle via freely switchable viewpoints. Users can create a 3D video using multi-view video sequences. Figure 2 shows the system model of multi-view video acquisition with wireless networks. Several mobile cameras are connected to an access point through these wireless networks, and the access point is connected to an encoder by a wired network. Each mobile camera transmits its own video frames to the access point. Once the access point has received video frames from multiple mobile cameras, it transmits the received video frames to the encoder.

To play a multi-view video smoothly, the acquisition system should satisfy two requirements. The first is that the volume of video traffic is sufficiently low to allow effective transmission over the wireless network. The amount of multi-view video traffic is greater than that of single-view video. In simple terms, the volume of N-view video traffic is N times greater than that of single-view video. However, wireless networks have a lower data rate than wired networks, because of their narrow bandwidth and interference. When mobile cameras transmit multi-view video over wireless networks, the low data rate increases the transmission delay between the mobile cameras and the access point. Long transmission delays will frustrate users.

The second requirement is the maintenance of high video quality. The video quality effectively measures the degree of video degradation that has been decoded from the raw video. Maintaining high video quality represents a trade-off with the aim of reducing video traffic. If the degradation is small, the acquisition system is applicable to numerous applications. However, high video quality necessitates a high volume of video traffic.

The simplest method for realizing multi-view video acquisition with wireless networks involves each mobile camera transmitting its own video to the access point independently. However, the data rate of wireless networks decreases when multiple mobile cameras transmit to the access point, as the bandwidth is shared among the cameras. This induces long transmission delays, leading to low user satisfaction.

Video traffic can be reduced if each mobile camera degrades the frame rate and the quantization parameter of its own video. The quantization parameter indirectly represents the relation between video traffic and quality. When the quantization parameter is high, the original values in each video frame are more likely to be quantized to zeroes. However, not surprisingly, this degradation induces low video quality.

One method of reducing video traffic and maintaining high video quality is Distributed Multi-view Video Coding (DMVC) [18]–[20]. DMVC is an encoding-level approach that exploits the inter-camera domain correlation for multi-view video streaming over wireless networks. DMVC exploits distributed source coding for encoding, and transmits video with side information. This side information includes the camera position and angle. Distributed source coding achieves the same compression ratio, as each mobile camera encodes its own video using video from other mobile cameras. Typical theories of distributed source coding are Slepian–Wolf theory [21] and Wyner–Ziv theory [22].

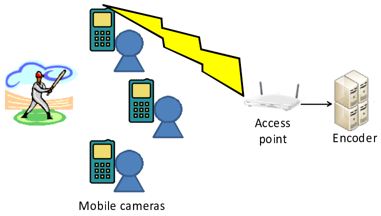


Fig. 2. Multi-view video acquisition with wireless networks

III. MULTI-VIEW VIDEO STREAMING WITH MOBILE CAMERAS (MVS/MC)

To satisfy the two requirements discussed in Section II, we propose MVS/MC. MVS/MC exploits a feature of wireless networks whereby a node can overhear packets transmitted by its neighbors. Each camera node reduces its own video traffic by calculating the differences between its own video and the overheard video. MVS/MC is a transmission-level approach, although it can be combined with an encoding-level approach such as DMVC [18]–[20].

A. Overview of MVS/MC

MVS/MC requires initialization, transmission order control, encoding, transmission, and decoding.

1) When a mobile camera enters the communication area of an access point, the mobile camera starts the process of initialization. The details of initialization are described in Section III-B.
2) After each mobile camera has been initialized, the access point determines their transmission order. The transmission order decision is based on positional information of each mobile camera, which is received during initialization. The access point then broadcasts the transmission order to the mobile cameras. The details of transmission order control are described in Section III-C.
3) Each mobile camera encodes its own video using video overheard from other mobile cameras. The details of encoding are described in Section III-D.
4) A mobile camera transmits the encoded video in one Group of Pictures (GOP) to the access point according to the received transmission order. Each GOP is a video frame set, typically consisting of eight frames. Other mobile cameras overhear the transmitted video. After one GOP has been transmitted by each mobile camera, the access point determines the transmission order for the next GOP. The details of video transmission are described in Section III-E.
5) The video received by the mobile cameras or the access point is decoded by a standard H.264/AVC decoder. The details of decoding are described in Section III-F.

B. Initialization

Before video transmission, an access point assigns a unique ID to each mobile camera. The access point periodically

TABLE I. NOTATION

Parameters, functions	Description
C	Mobile camera ID set in the communication area of the access point
$order[i]$	Mobile camera ID array
size (C)	Number of elements in set C

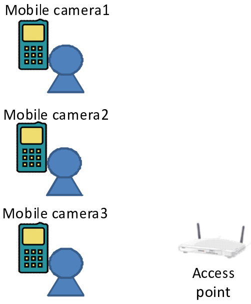


Fig. 3. Example showing three mobile cameras

transmits a beacon packet to its own communication area, informing the mobile cameras they have entered this region. When a mobile camera receives a beacon packet, it returns information on its own position and is assigned a unique ID by the access point. The position information is based on GPS data.

C. Transmission order control

To reduce the amount of redundant information passed among the mobile cameras, the access point determines their transmission order based on the positional information. The transmission order is based on bidirectional inter-view prediction in H.264/AVC. Bidirectional inter-view prediction uses the inter-camera domain correlation among the mobile cameras to further reduce the redundant information [4]–[6].

We explain the transmission order control procedure for N mobile cameras within the communication area of an access point. Table I describes the notation used in the Algorithm. The algorithm consists of two parts: a starting mobile camera decision and a transmission order decision. For the starting mobile camera decision, the access point determines which mobile camera is first to transmit its own video to the access point. The starting mobile camera x is farthest from the access point.

For the transmission order decision, the access point determines the subsequent transmission order for all mobile cameras. When mobile cameras are positioned so as to use bidirectional prediction, the access point outputs a transmission order that realizes bidirectional inter-view prediction. To determine the transmission order, the access point first selects mobile camera y that is closest to the starting mobile camera x. The access point then calculates size (C) to confirm that mobile camera y is able to encode its own video with bidirectional inter-view prediction using other mobile cameras' video.

If size (C) is greater than or equal to 1, mobile camera y is able to encode its own video using bidirectional inter-view prediction. To exploit the bidirectional inter-view prediction, mobile camera y needs to overhear two video sequences before its own transmission. The first video sequence is that from mobile camera x. The second is the video from the camera closest to mobile camera y and does not assign the transmission order. Following the above conditions, the access point selects mobile camera z, and determines the transmission order from mobile camera $x \rightarrow$ mobile camera $z \rightarrow$ mobile camera y in order. The transmission order realizes bidirectional inter-view prediction. After the order has been determined, the access point regards mobile camera z as mobile camera x, and repeats the above process to determine the transmission order among the other cameras so as to realize bidirectional inter-view prediction.

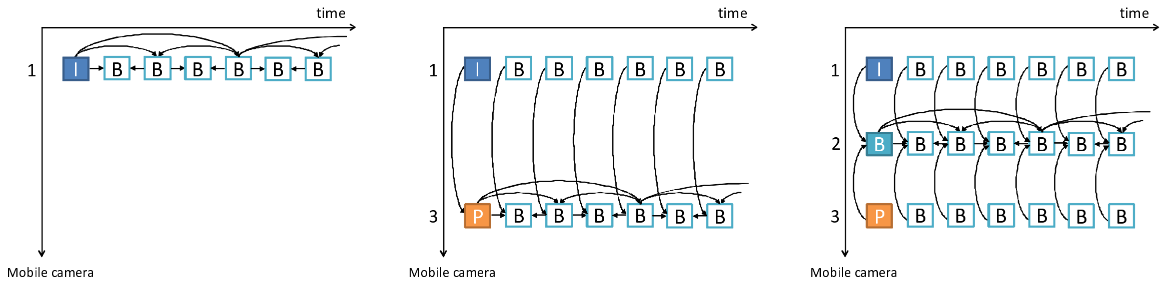
If size (C) is 0, mobile camera y is not able to encode its own video by bidirectional inter-view prediction. In this case, the access point determines that y is the last camera to transmit its video. The access point then terminates transmission order control, and broadcasts the transmission order to all mobile cameras.

We assume that there are three mobile cameras in the communication area of an access point. Figure 3 shows the positional relation of the mobile cameras and the access point. The set C consists of the IDs of mobile cameras 1, 2, and 3. The access point regards mobile camera 1 as the starting mobile camera, because this is farthest from the access point. The access point sets the ID of mobile camera 1 to $order[1]$, and removes this ID from C because the starting mobile camera first transmits its own video to the access point. Next, the access point selects the ID of mobile camera 2, which is closest to camera 1, from C. The access point removes mobile camera 2's ID from C and calculates size (C). The result of size (C) represents the number of mobile cameras that have not been assigned a transmission order. Because size (C) is 1, the access point selects mobile camera 3 from C. Mobile camera 3 is the closest to mobile camera 2 of the mobile cameras in C. The access point sets mobile camera 3's ID to $order[2]$ and mobile camera 2's ID to $order[3]$ to realize bidirectional inter-view prediction for mobile camera 2's encoding. The access point removes mobile camera 3's ID from C and calculates size (C) to confirm the number of mobile cameras that have not been assigned a transmission order. Because size (C) is 0, the access point terminates the transmission order control algorithm. Thus, the final transmission order is mobile camera 1 $\rightarrow$ mobile camera 3 $\rightarrow$ mobile camera 2.

D. Encoding

When each mobile camera receives the transmission order from the access point, it encodes its own video according to the transmission order. Each mobile camera encodes the 1-GOP video based on H.264/AVC.

Each mobile camera overhears the communication from all other cameras, thus enabling a reduction in video traffic. Figure 4 shows the prediction structure of MVS/MC. This prediction structure assumes that the number of mobile cameras, their position, and the transmission order are the same as in Figure 3. Figure 4(a) shows the prediction structure of mobile camera 1. The anchor frame of the structure is encoded using an I-frame, which is a picture that is encoded independently of the other pictures.



(a) Prediction structure of mobile camera 1 (b) Prediction structure of mobile camera 3 (c) Prediction structure of mobile camera 2

Fig. 4. MVS/MC's prediction structure

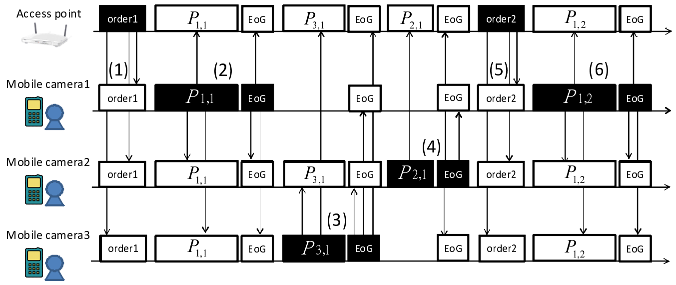


Fig. 5. Timing diagram of MVS/MC

Figure 4(b) shows the prediction structure of mobile camera 3. This camera encodes its own video using the overheard video from camera 1. The anchor frame of camera 3's video is encoded using a P-frame, which encodes only the differences from camera 1's I-frame, and thus requires less bandwidth than the I-frame.

Figure 4(c) shows the prediction structure of mobile camera 2. Mobile camera 2 encodes its own video using the overheard video from cameras 1 and 3. The anchor frame of camera 2's video is encoded as a B-frame. B-frames encode the differences based on both camera 1's I-frame and camera 3's P-frame, and thus require less bandwidth than the P-frame.

E. Video transmission

Each mobile camera transmits its own encoded video to an access point according to the transmission order determined by the access point.

Figure 5 shows the timing diagram of MVS/MC. Figure 5 assumes that the transmission order determined by the access point is mobile camera 1 → mobile camera 3 → mobile camera 2. $P_{i,j}$ represents the video packet of mobile camera i in GOP j.

1) The access point broadcasts the transmission order for GOP 1 to all mobile cameras.
2) When the mobile cameras receive the transmission order, mobile camera 1 starts video transmission. Mobile camera 1 transmits $P_{1,1}$ to the access point. $P_{1,1}$ includes position information about camera 1 in the position field and the encoded video in the video field. Mobile cameras 2 and 3 overhear $P_{1,1}$ and decode the video. After $P_{1,1}$ has been transmitted, mobile camera 1 broadcasts an End-of-GOP (EoG) packet. The EoG packet informs other mobile cameras about the end of 1-GOP video transmission. The format of the EoG packet is the same as that of an ACK frame in IEEE 802.11 [23]. When mobile camera 3 overhears the EoG packet, it encodes its own video with the video overheard from camera 1, and commences video transmission. Mobile camera 2 stores camera 1's video, and waits for an EoG packet from mobile camera 3.
3) Mobile camera 3 transmits $P_{3,1}$ to the access point. $P_{3,1}$ includes camera 3's position information in the position field and the encoded video in the video field. Mobile camera 2 overhears $P_{3,1}$ and decodes the video. After $P_{3,1}$ has been transmitted, mobile camera 3 broadcasts an EoG packet. When mobile camera 2 overhears the EoG packet, mobile camera 2 encodes its own video using the videos overheard from cameras 1 and 3. Mobile camera 2 then commences video transmission.
4) Mobile camera 2 transmits $P_{2,1}$ to the access point. $P_{2,1}$ includes camera 2's position information in the position field and the encoded video in the video field. After $P_{2,1}$ has been transmitted, mobile camera 2 broadcasts an EoG packet.
5) When the access point receives the EoG packet from mobile camera 2, it transmits $P_{1,1}$, $P_{2,1}$, and $P_{3,1}$ to an encoder. After this transmission, the access point determines the transmission order for the second GOP based on the position information included in $P_{1,1}$, $P_{2,1}$, and $P_{3,1}$. The transmission order for the second GOP is then broadcast to all mobile cameras.
6) Mobile camera 1 transmits $P_{1,2}$ to the access point. Mobile cameras 2 and 3 overhear $P_{1,2}$ and decode the video. After $P_{1,2}$ has been transmitted, mobile camera 1 broadcasts an EoG packet.

MVS/MC repeats (2) to (6) until the end of video transmission for all GOPs.

F. Decoding

MVS/MC does not require a special decoder. Each mobile camera and the encoder exploit a standard H.264/AVC video decoder. The mobile cameras and the encoder first receive and decode the I-frame. The video frames received by the mobile

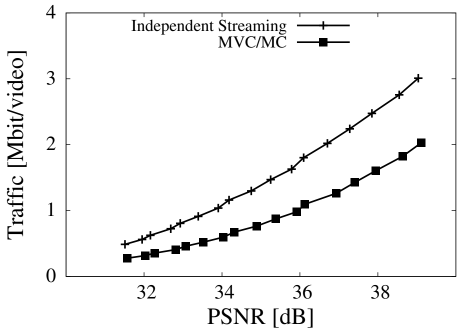


Fig. 6. PSNR vs. Traffic

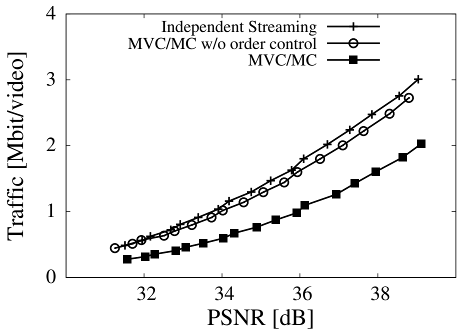


Fig. 7. PSNR vs. Traffic with different camera positions

cameras and the encoder are then encoded using the previously received video frames. The mobile cameras and the encoder decode the newly received video after decoding the previously received video. When the encoder decodes video frames from all mobile cameras, the video frames are encoded based on a multi-view video coding technique, such as MVC, IMVS, or UDMVT. Finally, the encoder transmits the encoded video frames to a user's device. The user can then play back the multi-view video.

IV. EVALUATION

A. Evaluation settings

To evaluate the traffic and video quality of MVS/MC, we implemented a MVS/MC encoder/decoder with JMVC, which is an open source project [24]. The distance between the mobile cameras in these video sequences is 19.5 [cm]. These test video sequences are provided by MERL [25]. Table II shows the encoding parameters of the evaluation. Evaluation settings are as follows: the resolution is 176×144, the frame rate is 15 fps, number of frames is 250, GOP size is 8, number of mobile cameras is 8 and quantization parameter is 24-40.

We evaluate the traffic and video quality of three encoding/decoding schemes: Independent Streaming, MVS/MC w/o transmission order control, and MVS/MC.

1) Independent Streaming
Independent Streaming encodes the video of each mobile camera independently, and transmits the video to the access point. Independent Streaming gives the baseline performance, using the simplest method for multi-view video acquisition with wireless networks.

TABLE II. EVALUATION SETTINGS

Resolution	176×144
Frame rate	15 fps
Number of frames	250
GOP size	8 frames
Number of mobile cameras	8
Quantization parameter (QP)	24-40

2) MVS/MC w/o order control
MVS/MC w/o transmission order control supports only the packet overhearing technique of MVS/MC. Even when the position of a mobile camera changes, each mobile camera transmits its own encoded video in the previously assigned order.
3) MVS/MC
As described in Section III, MVS/MC is the proposed approach. MVS/MC supports both overhearing and transmission order control.

B. PSNR vs. Traffic

We compared the volume of traffic required for different levels of video quality. This evaluates the baseline performance of traffic reduction while maintaining high video quality for the three encoding/decoding schemes described in Section IV-A. We use the standard peak signal-to-noise ratio (PSNR) metric to evaluate the video quality. The PSNR represents the video quality of multi-view video as follows:

$$PSNR = 20\log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right)$$

where MAX is the largest pixel value and MSE is the mean squared error between all pixels of the decoded and original videos.

We implemented Independent Streaming and MVS/MC on the JMVC encoder. This evaluation used the Ballroom video sequence, and we assumed that the position of each mobile camera was fixed. The JMVC encoder encodes the video frames of a mobile camera with or without the video frames of other mobile cameras, depending on the encoding/decoding scheme. The video frames were encoded using different QP values to evaluate the effect on traffic volume and video quality. When QP is large, the traffic volume and quality of the video frames are low. Finally, the JMVC encoder calculates the average traffic volume for each PSNR.

Figure 6 shows the traffic produced by each encoding/decoding scheme as a function of PSNR. Figure 6 shows the following:

1) MVS/MC reduces traffic compared to Independent Streaming for the same video quality. For example, when the PSNR is 36 [dB], MVS/MC reduces traffic by 700 [Kbits/video] compared with Independent Streaming. This is because MVS/MC removes redundant information using the overheard video from other mobile cameras.
2) As PSNR increases, the difference between the traffic volume produced by MVS/MC and Independent

Streaming becomes larger. For example, when the PSNR is 32 [dB], MVS/MC reduces traffic by 240 [Kbits/video] compared with Independent Streaming, but when the PSNR is 39 [dB], MVS/MC reduces traffic by 980 [Kbits/video]. When the PSNR is high, the video traffic increases because the video from each mobile camera is similar to the original video. As a result, the traffic produced by Independent Streaming increases greatly with an increase in PSNR. The volume of redundant video information increases when the video from each mobile camera is almost the same as the original. MVS/MC exploits this redundant information during the encoding process, thus achieving a considerable reduction in traffic volume.

C. *Effect of transmission order control*

Sections IV-Bdiscussed the performance of MVS/MC with packet overhearing and transmission order control. This section evaluates the contribution of packet overhearing and transmission order control in more detail by comparing the traffic volume using MVS/MC to that of MVS/MC w/o transmission order control.

As in the evaluation in Section IV-B, we implemented the three encoding/decoding schemes on the JMVC encoder. To evaluate the effect of transmission order control, we randomly exchanged the positions of the mobile cameras, and evaluated the average traffic of the three encoding/decoding schemes over 100 evaluations.

Figure 7 shows the traffic produced by each encoding/decoding scheme as a function of PSNR. From this, we can conclude the following:

1) MVS/MC achieves the lowest traffic of the three encoding/decoding schemes, even when the position of each mobile camera is changed. For example, MVS/MC reduces traffic by 700 [Kbits/video] compared to MVS/MC w/o transmission order control.
2) The difference in traffic volume between MVS/MC w/o order control and Independent Streaming is relatively small. This is because there is little redundant information among the mobile cameras when their positions are changed and the transmission order is fixed. This shows that the transmission order control gives a strong advantage in achieving low video traffic and high video quality.

V. CONCLUSION

This paper proposed Multi-view Video Streaming with Mobile Cameras (MVS/MC) for multi-view video acquisition over wireless networks. MVS/MC achieves a reduction in traffic volume while maintaining high video quality by means of packet overhearing and transmission order control. Through a series of evaluations, it was revealed that MVS/MC enables low traffic volumes with only a small degradation in video quality.

REFERENCES

[1] K. Müller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, H. Rhee *et al.*, “3d high efficiency video coding for multi-view video and depth data,” *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3366–3378, 2013.

[2] M. Tanimoto and S. Kazuyoshi, “Global view and depth (gvd) format for ftv/3dtv,” in *Three-Dimensional Imaging Visualization And Display*, 2013, pp. 1–10.

[3] S. Beck, A. Kunert, A. Kulik, and B. Froehlich, “Immersive group-to-group telepresence,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 4, pp. 616–625, 2013.

[4] A. Vetro, P. Pandit, H. Kimata, A. Smolic, and Y.-K. Wang, *Joint Draft 8.0 on Multi-view Video Coding*, 2008.

[5] K. Muller, P. Merkle, H. Schwarz, T. Hinz, A. Smolic, and T. Wiegand, “Multi-view video coding based on H. 264/AVC using hierarchical B-frames,” in *IEEE PCS*, 2006.

[6] Text Of ISO/IEC 14496-10:2008/FDAM 1 ISO/IEC JTC1/SC29/WG11, “Multiview Video Coding,” 2008.

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

[8] H. Huang, B. Zhang, S.-H. Chan, G. Cheung, and P. Frossard, “Coding and replication co-design for interactive multiview video streaming,” in *IEEE INFOCOM*, 2012, pp. 2791–2795.

[9] Z. Pan, Y. Ikuta, M. Bandai, and T. Watanabe, “User dependent scheme for multi-view video transmission,” in *IEEE ICC*, 2011.

[10] ——, “A user dependent system for multi-view video transmission,” in *IEEE AINA*, 2011, pp. 732–739.

[11] Z. Pan, M. Bandai, and T. Watanabe, “A user dependent scheme for multi-view video live streaming,” *International Journal of Computational Information Systems*, vol. 9, no. 4, pp. 1439– 1448, 2013.

[12] T. Fujihashi, Z. Pan, and T. Watanabe, “Traffic reduction for multiple users in multi-view video streaming,” in *IEEE ICME*, 2012.

[13] ——, “UMSM: A traffic reduction method on multi-view video streaming for multiple users,” *IEEE Transactions on Multimedia*, vol. 16, no. 2, pp. 1–14, 2014.

[14] ——, “Traffic reduction on multi-view video live streaming for multiple users,” *IEICE Transactions on Communications*, vol. 96, no. 7, pp. 2034–2045, 2013.

[15] C. Fehn, “A 3D-TV approach using depth-image-based rendering (DIBR),” in *VIIP*, 2003.

[16] W. R. Mark, L. McMillan, and G. Bishop, “Post-rendering 3D warping,” in *ACM Interactive 3D graphics*, 1997, pp. 7–16.

[17] K. Hisatomi, K. Ikeya, M. Katayama, Y. Iwadate, and K. Aizawa, “Depth estimation based on stereo camera pairs of color and infrared using cross-based local multipoint filter,” *3DSA2013*, vol. 1, p. 3, 2013.

[18] G. Xun, L. Yan, W. Feng, G. Wen, and L. Shipeng, “Distributed multi-view video coding,” in *VCIP*, vol. 38, no. 11, 2006, pp. 1917–1921.

[19] X. Artigas, E. Angeli, and L. Torres, “Side information generation for multiview distributed video coding using a fusion approach,” in *IEEE NORSIG*, 2006, pp. 250–253.

[20] D. Frederic, O. Mourad, and E. Touradj, “Recent advances in multiview distributed video coding,” in *DSS*, 2007, pp. 1–11.

[21] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, 1973.

[22] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Transaction on Information Theory*, vol. 3, no. 4, pp. 45–49, 1976.

[23] IEEE Computer Society, *IEEE Standard for Information technology-Telecommunications and information exchange between systems Local and metropolitan area networks-Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2012.

[24] Joint Video Team Of ITU-T VCEG And ISO/IEC MPEG, *JMVC (Joint Multiview Video Coding) Software*, 2008.

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