Efficient Error-Resilient Multicasting for Multi-View 3D Videos in Wireless Networks

Chi-Heng Lin*, De-Nian Yang*, Ji-Tang Lee† and Wanjiun Liao†

*Academia Sinica, Taiwan, †National Taiwan University, Taiwan
Email: *{toptoptop3344@gmail.com, dnyang@iis.sinica.edu.tw}, †{r03942049, wjliao}@ntu.edu.tw

***Abstract*—With the emergence of 3D mobile and VR devices, mobile 3D video services are becoming increasingly important for video service providers, such as Youtube and Netflix, while multi-view 3D videos have the potential to inspire a variety of innovative applications. However, enabling multi-view 3D video services may overwhelm WiFi networks when every view of a video are multicasted. In this paper, therefore, we propose to leverage depth-image-based rendering (DIBR) in multi-view 3D, which allows each mobile client to synthesize the desired view from nearby left and right views, in order to effectively reduce the bandwidth consumption. Moreover, when each client suffers from packet losses, retransmissions incur additional bandwidth consumption and excess delay, which in turn undermines the quality of experience in video applications. To address the above issue, we first discover the merit of view protection via DIBR for multi-view video multicast using a mathematical analysis and then design a new protocol, named Multi-View Group Management Protocol (MVGMP), to support the dynamic join and leave of users and the change of desired views. The simulation results demonstrate that our protocol effectively reduces bandwidth consumption and increases the probability for each client to successfully playback the desired views in a multi-view 3D video.**

***Index Terms*—3D Video, Wireless Networks, Multi-View, DIBR**

I. Introduction

The IEEE 802.11 [1] WiFi standard has achieved massive market penetration due to its low cost, easy deployment and high bandwidth. Also, with the recent emergence of 3D mobile and VR devices, such as Amazon's 3D Fire Phone, Oculus Rift, HTC Vive, LG's Optimus 3D, and Sharp's Lynx, mobile 3D video services are expected to become increasingly important for video service providers such as Youtube and Netflix. In contrast to traditional stereo single-view 3D video formats, multi-view 3D videos allow users to select the viewing angles and thus are expected to stimulate the development of innovative applications in television, movies, education, and advertising [2]. Previous research on the deployment of 3D videos in wireless networks has mostly focused on improving 3D video quality for single-view 3D videos [3], [4], but multi-view 3D videos, which typically offer 5, 16 and 32 different viewing angels [5] have attracted much less attention.

Multi-View 3D videos are expected to significantly increase the network load when all views are transmitted. One promising way to remedy the bandwidth issue is to exploit depth-image-based rendering (DIBR) in multi-view 3D, in which the idea is to synthesize the desired view from one left view and one right view [2], because adjacent left and right views with a sufficiently small angle usually share many similar scenes and objects. Several schemes for bit allocation between the texture and depth map [6] and rate control with layered encoding for a multi-view 3D video [7] have been proposed to ensure that the quality of the synthesized view is very close to the original view (i.e., by minimizing total distortion or maximizing quality). Therefore, exploiting DIBR in clients eliminates the need to deliver every view of a video in a network. For practical situations, the computation overhead and extra energy consumption incurred by DIBR is small enough to be supported by current mobile devices [7], [8].

By leveraging DIBR, only a subset of views are required to be multicasted in a network. However, multi-view 3D video multicast with DIBR brings new challenges in *view selection* for WiFi networks due to view synthesis and wireless erasure. Firstly, the number of skipped views between the left and right views in DIBR needs to be constrained to ensure high quality of the synthesized view [2]. In other words, since each transmitted view is multicasted to multiple clients, it is crucial to carefully select the transmitted views so that the desired view of each user can be synthesized with a left view and a right view close to each other. DIBR has a quality constraint [2], which specifies that the left and right views are allowed to be at most R views away (i.e., $R-1$ views skipped between them) to ensure that every view between the left and right view can be successfully synthesized with good quality. Therefore, each new user cannot arbitrarily choose a left and a right view for synthesis with DIBR. The second challenge is that WiFi networks frequently suffer from wireless erasure (i.e., packet losses), and different clients suffer from different loss probabilities due to varying channel conditions [11], [12], [13]. In 2D and single-view 3D videos, the *view loss probability* for each user can be easily derived according to the selected bit-rate, channel, and the setting of MIMO (e.g., antennas, spatial streams) in 802.11 networks. For multi-view 3D videos, however, when a video frame is lost for a user i subscribing a view k_i, we observe that the left and right views multicasted in the network to other users can natively serve to *protect* view k_i, since the user i can synthesize the desired view from the two views using DIBR. However, the view synthesis will fail if only one left view or one right view is received successfully by the client. Therefore, a new research problem is to derive the *view failure probability*, which is the probability that each user does not successfully receive and synthesize his/her desired view.

Nevertheless, to the best of our knowledge, there is no related research on multi-view 3D video multicast over wireless network. For instance, the energy efficiency of video multicast in 4G wireless network with Scalable Video Coding(SVC) and Adaptive Modulation and Coding(AMC) is

978-1-5090-1328-9/16/$31.00 ©2016 IEEE

proposed [14]. Flexible transport of 3D videos [15] for digital video broadcasting(DVB) is studied to improve the video rates and resolutions. However, these works focus on conventional single-view 3D videos, and the correlation of different views in a multi-view video has not been considered.

In this paper, we first analyze the view failure probability and compare it with the traditional view loss probability, which is the probability that a view fails to be sent to a user without DIBR. We then propose the Multi-View Group Management Protocol (MVGMP) for multi-view 3D multicast. When a user joins the video multicast group, it can exploit our analytical result to receive the most suitable right and left views as well, so that the view failure probability is guaranteed to stay below a threshold. On the other hand, when a user leaves the video multicast group, the proposed protocol carefully selects and withdraws a set of delivered views to reduce the network load, so that the video failure probability for other users will not exceed the threshold. Bandwidth consumption can be effectively reduced since it is not necessary to deliver all the views subscribed by the clients.

The rest of the paper is organized as follows. Section II describes the system model. Section III analyzes the view loss probability and view failure probability. Section IV presents the proposed protocol. Section V shows the simulation results, and Section VI concludes this paper.

II. System Model

This paper considers single-cell video multicast in IEEE 802.11 networks, where the views transmitted by different bit-rates are associated with different loss probabilities [11], [12], [13]. Currently, many video services, such as Youtube and Netflix, require reliable transmissions since Flash or MPEG DASH [9] are exploited for video streaming. Nevertheless, the current IEEE 802.2 LLC protocol for IEEE 802.11 networks does not support reliable multicast transmissions [19], and error recovery therefore is handled by Layer-3 reliable multicast standards, such as the PGM protocol[20].

A multi-view 3D video can be encoded by various encoding schemes [21], [22]. Each view in a video consists of a texture image and a depth map of the corresponding viewing angle. The idea of DIBR in multi-view 3D is to synthesize a view according to its neighbor left view and neighbor right view. Since the angle between the neighbor left and right views is relatively small, it is expected that the video objects in the synthesized view can be slightly warped (i.e., morphed) from those in the two neighbor views. Effective techniques in computer vision and image processing have been proposed to ensure high video quality and limit the processing delay [23].

For example, suppose there are three multicast views, i.e., views 1, 3, and 4 subscribed by all clients. In the original WiFi multicast without DIBR, AP separately delivers each view in a multicast group to the corresponding clients, and the three views are separately recovered or retransmitted during packet losses. In contrast, our approach enables a subscribed view to be synthesized by neighbor left and right views with DIBR, while the quality constraint in DIBR states that there are at most $R-1$ views between the neighbor left and right views, and R can be set according to [2]. When $R=3$ in the above example, the lost of view 3 can be recovered by view 1 and 4, since view 3 can be synthesized by view 1 and 4 accordingly. In other words, a user can first try to synthesize the view according to the left view and right view when a subscribed view is lost, by joining the multicast groups (originally for other users) corresponding to the left and right views.

The intuition behind our idea is *traffic protection* from neighbor views. A user can join more multicast groups to protect the desired view without extra bandwidth consumption in the network, because the nearby left view and right view may be originally multicasted to other users that subscribe the views. However, more traffic will be received, and the trade-off will be explored in the next section.

III. Analytical Solution

In this section, we present the analytical results for multi-view 3D multicast with DIBR in IEEE 802.11 networks. We first study the scenario of single-view subscription for each user and then extend it to multi-view subscription. Table I summarizes the notations in the analysis. Based on the mathematical analysis, a new protocol is proposed in the next section to dynamically subscribe different views for each user.

A. Single View Subscription

In single-view subscription, each user i specifies only one desired view k_i. Each view can be sent once or multiple times if necessary. Let $p_{i,r}$ represent the *view loss probability*, which is the probability that user i does not successfully receive a view in bit-rate r. We define a new probability $P_\varepsilon^{(i)}$ for multi-view 3D videos, called *view failure probability*, which is the probability that user i fails to receive and synthesize the desired view because the view and nearby left and right views for synthesis are all lost. In other words, the view loss probability considers only one view, while the view failure probability jointly examines the loss events of multiple views.

Theorem 1. *For single-view subscription, the view failure probability for user i is*

$$
\begin{aligned}
P_\varepsilon^{(i)} = & \prod_{r\in D_i} p_{i,r}^{n_{k_i,r}} \times \Big(\mathbf{1}\{k_i=1\}+\mathbf{1}\{k_i=M\}\Big) \\
& + \sum_{k=1}^{R-1}\Bigg((1-\prod_{r'\in D_i} p_{i,r'}^{n_{k_i-k,r'}}) \prod_{r_1,r_2\in D_i}\prod_{l=1}^{\min(R-k,M-k_i)} \\
& \prod_{q=0}^{k-1} p_{i,r_1}^{n_{k_i-q,r_1}} p_{i,r_2}^{n_{k_i+l,r_2}} \mathbf{1}\{M-1\geq k_i\geq k+1\}\Bigg) \\
& + \prod_{r_3\in D_i}\prod_{q=0}^{\min(R-1,k_i-1)} p_{i,r_3}^{n_{k_i-q,r_3}} \mathbf{1}\{M-1\geq k_i\geq 2\}
\end{aligned}
$$

where $\mathbf{1}x$ denotes the indicator function, which is 1 if statement x is true, otherwise 0.

Proof. The view failure event occurs when both of the following two conditions hold: 1) user i does not successfully receive the desired view, and 2) user i fails to receive any feasible set consisting of a left view and a right view with the view distance at most R to synthesize the desired view. The probability of the first condition is $\prod_{r\in D_i} p_{i,r}^{n_{k_i,r}}$ when the the desired view k_i of user i is transmitted by n_{k_i} times. Note

TABLE I: Notations in Analysis.

Description	Notation
R	Quality constraint of DIBR
M	Total number of views
k_i	The view desired by user i
D_i	A set of the available data rates for user i
$n_{j,r}$	Number of multicast transmissions for view j transmitted by rate r
$p_{i,r}$	The view loss probability for user i under rate r
$P_\varepsilon^{(i)}$	The probability that user i cannot obtain the desired view either by direct transmission or by DIBR
$p_{j,r}^{\text{AP}}(n)$	The probability that AP multicasts a view j n times under bit-rate r
α_i	The percentage of the desired views that can be received or synthesized successfully by user i
p_{select}	The probability that a user selects each view

that if the desired view of user i is view 1 or view M, i.e., $k_i = 1$ or $k_i = M$, user i is not able to synthesize the desired view with DIBR, and thus the view failure probability can be directly specified by the first condition. For every other user i with $M-1 \geq k_i \geq 2$, we define a set of non-overlapping events $\{\mathcal{B}_k\}_{k=0}^{R-1}$, where $\mathcal{B}_k$ with $k > 0$ is the event that the nearest left view received by user i is $k_i - k$, but user i fails to receive a feasible right view to synthesize the desired view. On the other hand, $\mathcal{B}_0$ is the event that the user i fails to receive any left view. Therefore, $\bigcup_{k=0}^{R-1} \mathcal{B}_k$ jointly describes all events for the second condition.

For each event $\mathcal{B}_k$ with $k > 0$,

$$P(\mathcal{B}_k) = (1 - \prod_{r' \in D_i} p_{i,r'}^{n_{k_i-k,r'}}) \prod_{r_1,r_2 \in D_i} \prod_{l=1}^{\min(R-k,M-k_i)} \prod_{q=0}^{k-1} p_{i,r_1}^{n_{k_i-q,r_1}} p_{i,r_2}^{n_{k_i+l,r_2}} \mathbf{1}\{M-1 \geq k_i \geq k+1\}$$

The first term $1 - \prod_{r' \in D_i} p_{i,r'}^{n_{k_i-k,r'}}$ indicates that user i successfully receives view $k_i - k$, and the second term

$$\prod_{r_1,r_2 \in D_i} \prod_{l=1}^{\min(R-k,M-k_i)} \prod_{q=0}^{k-1} p_{i,r_1}^{n_{k_i-q,r_1}} p_{i,r_2}^{n_{k_i+l,r_2}}$$

means that user i does not successfully receive any left view between $k_i - k$ and k and any right view from $k_i + 1$ to $k_i + \min(R-k, M-k_i)$. It is necessary to include an indicator function in the last term since $\mathcal{B}_k$ will be a null event if $k_i \leq k$, i.e., user i successfully receives a view outside the view boundary. Finally, the event $\mathcal{B}_0$ occurs when no left view is successfully received by user i.

$$P(\mathcal{B}_0) = \prod_{q=0}^{\min(R-1,k_i-1)} \prod_{r_3 \in D_i} p_{i,r_3}^{n_{k_i-q,r_3}} \mathbf{1}\{M-1 \geq k_i \geq 2\}$$

The theorem follows after summarizing all events. □

Remark: The advantage of a multi-view 3D multicast with DIBR can be clearly seen when comparing the view loss probability and view failure probability. The latter probability attaches a new term (i.e., the probability of $\bigcup_{k=0}^{R-1} \mathcal{B}_k$) to the view loss probability, where a larger R reduces the probability of the second term. For multi-view 3D with DIBR, therefore, the view failure probability is much smaller than the view loss probability, also shown in Section V.

B. Multiple View Subscription

In the following, we explore the case of a user desiring to subscribe to multiple views. We first study the following two scenarios: 1) every view is multicasted; 2) only one view is delivered for every $\widetilde{R}$ views, $\widetilde{R} \leq R$, and thus it is necessary for a user to synthesize other views accordingly. We first define α_i, which represents the percentage of desired views that can be successfully received or synthesized by user i.

$$\alpha_i = \frac{\sum_{k_i \in \mathcal{K}_i} \mathbf{1}\{\text{user } i \text{ can obtain view } k_i\}}{|\mathcal{K}_i|}$$

where $\mathcal{K}_i$ denotes the set of desired views for user i.[1] By using Theorem 1, we can immediately arrive at the following corollary.

Corollary 1.

$$\mathbb{E}[\alpha_i] = \frac{\sum_{k_i \in \mathcal{K}_i} (1 - P_\varepsilon^{(i)}(k_i))}{|\mathcal{K}_i|} \tag{1}$$

where $P_\varepsilon^{(i)}(k_i)$ is given in Theorem 1.

Proof.

$$\mathbb{E}[\alpha_i] = \frac{\sum_{k_i \in \mathcal{K}_i} \mathbb{E}\mathbf{1}\{\text{user } i \text{ can obtain view } k_i\}}{|\mathcal{K}_i|} = \frac{\sum_{k_i \in \mathcal{K}_i} (1 - P_\varepsilon^{(i)}(k_i))}{|\mathcal{K}_i|}$$

□

Eq. (1) becomes more complicated as $|\mathcal{K}_i|$ increases. In the following, therefore, we investigate the asymptotic behavior of α_i for a large $|\mathcal{K}_i|$ and a large M (i.e., $|\mathcal{K}_i| \leq M$). To find the closed-form solution, we first consider a uniform view subscription and assume that user i subscribes to each view j with probability $p_{\text{select}} = \frac{|\mathcal{K}_i|}{M}$ independently across all views so that the average number of selected views is $|\mathcal{K}_i|$. Assume that the AP multicasts view j with rate r by n times with probability $p_{j,r}^{\text{AP}}(n)$ independently across all views and rates. In the following, we first perform an asymptotic analysis in order to acquire a theoretical closed-form solution, and we then present the insights from the theorem by comparing the results of single-view subscription and multi-view subscription. Due to the space constraint, a more general analysis that also allows each user to subscribe to a sequence of consecutive views is presented in [24].

[1] Since a retransmission is necessarily to be involved when a desired view cannot be received or synthesized, we derive the minimal number of views required to be retransmitted to obtain all desired views for each user in [24].

Theorem 2. *In multi-view 3D multicast,*

$$\alpha_i(\mathcal{K}_i) \overset{a.s.}{\to} (1-p_i)\left\{\sum_{k=1}^{R} k(1-p_i)p_i^{k-1} + p_i^R\right\} \quad (2)$$

$$\mathbb{E}[\alpha_i(\mathcal{K}_i)] \overset{a.s.}{\to} (1-p_i)\left\{\sum_{k=1}^{R} k(1-p_i)p_i^{k-1} + p_i^R\right\} \quad (3)$$

as $|\mathcal{K}_i| \to \infty$, *where* $p_i = \prod_{r\in D_i}\sum_n p_{j,r}^{AP}(n)p_{i,r}^n$

Proof. We first derive the view loss probability for user i. Suppose that the AP multicasts a view j n times via rate r. The probability that user i cannot successfully receive the view is $p_{i,r}^n$. Because the AP will multicast a view n times via rate r with probability $p_{j,r}^{\text{AP}}(n)$, the probability that user i cannot receive the view via rate r is $\sum_n p_{j,r}^{\text{AP}}(n)p_{i,r}^n$. Therefore, the view loss probability for user i is the multiplication of the view loss probabilities in all rates, i.e., $\prod_{r\in D_i}\sum_n p_{j,r}^{\text{AP}}(n)p_{i,r}^n$. For simplification, we denote p_i as the view loss probability for user i in the remainder of the proof.

Since the multicast order of views is not correlated to α_i, we assume that the AP sequentially multicasts the views from view 1 to view M. Now the scenario is similar to a tossing game, where we toss M coins, and a face-up coin represents a view successfully received from the AP. Therefore, the face-up probability of at least one coin is $1-p_i^M$. Now we mark a coin with probability p_{select} if it is face-up or if there is one former tossed face-up coin and one latter tossed face-up coin with the view distance at most R. Since the above analogy captures the mechanism of direct reception and DIBR of views, the marked coins then indicate that the views selected by user i can be successfully acquired.

To derive the closed-form asymptotic result, we exploited the *delayed renewal reward process*, in which a cycle begins when a face-up coin appears, and the cycle ends when the next face-up coin occurs. The reward is defined as the total number of marked coins. Specifically, let $\{N(t) := \sup\{n : \sum_{i=0}^{n} X_i \le t\}, t \ge 0\}$ denote the delayed renewal reward process with inter-arrival time X_n, where X_n with $n \ge 1$ is the time difference between two consecutive face-up coins, and X_0 is the time when the first face-up coin appears.

Let $R(M)$ and R_n denote the total reward earned at the time M, which corresponds to the view numbers in a multi-view 3D video. At cycle n,

$$\frac{R(M)}{M} = \frac{\sum_{n=1}^{N(M)} R_n}{M} + o(1) \quad a.s.$$

where the $o(1)$ term comes from the fact that the difference between the total reward and $\sum_{n=1}^{N(M)} R_n$ will have a finite mean. Recall that the reward earned at each cycle is the number of marked coins,

$$\mathbb{E}[R_n|X_n] = \begin{cases} p_{\text{select}}, & \text{for } X_n > R \\ X_n p_{\text{select}}, & \text{for } X_n \le R \end{cases} \quad (4)$$

since when $X_n \le R$, X_n coins can be marked (each with probability p_{select}) between two consecutive face-up coins, and thus the expected reward given X_n is $X_n p_{\text{select}}$. By contrast, only one coin can be marked with probability p_{select} when $X_n > R$, and the expectation of reward given X_n is only p_{select}.

Since X_n is a geometric random variable with parameter $1-p_i$, we have

$$\mathbb{E}[X_n] = 1 - p_i + 2p_i(1-p_i) + 3p_i^2(1-p_i) + \cdots = \frac{1}{1-p_i}$$

and

$$\begin{aligned}\mathbb{E}[R_n] =& p_{\text{select}}(1-p_i) + 2p_{\text{select}}p_i(1-p_i) + \cdots \\ &+ Rp_{\text{select}}p_i^{R-1}(1-p_i) + p_{\text{select}}p_i^R \end{aligned} \quad (5)$$

By theorem 3.6.1 of renewal process in [25],

$$\begin{aligned}\frac{\sum_{n=1}^{N(M)} R_n}{M} &\overset{a.s.}{\to} \frac{\mathbb{E}R_n}{\mathbb{E}X_n} \\ &= p_{\text{select}}(1-p_i)\left\{\sum_{k=1}^{R} k(1-p_i)p_i^{k-1} + p_i^R\right\}\end{aligned} \quad (6)$$

Let U_M denote the number of views selected by user i. Therefore,

$$\alpha_i = \frac{R(M)}{U_M} = \frac{R(M)}{M}\frac{M}{U_M}$$

For $\frac{U_M}{M} \overset{a.s.}{\to} p_{\text{select}}$, by the strong law of large numbers, after combining with Eq. (4), (5), (6),

$$\alpha_i \overset{a.s.}{\to} (1-p_i)\left\{\sum_{k=1}^{R} k(1-p_i)p_i^{k-1} + p_i^R\right\}$$

The proof for convergence in mean is similar. It is only necessary to replace the convergence in Eq. (6) by the convergence in mean, which can be proven by the same theorem. □

Remark: Under the above uniform view subscription, it can be observed that α_i is irrelevant to p_{select}, implying that different users with different numbers of subscription will acquire the same percentage of views. Most importantly, $\alpha_i = 1 - p_i$ for multi-view 3D multicasts without DIBR. In contrast, multi-view 3D multicasting with DIBR effectively improves α_i by $\sum_{k=1}^{R} k(1-p_i)p_i^{k-1} + p_i^R$. Since this term is strictly monotonically increasing with R, we have $\sum_{k=1}^{R} k(1-p_i)p_i^{k-1} + p_i^R > \sum_{k=1}^{1} k(1-p_i)p_i^{k-1} + p_i = 1$, which implies that the percentage of obtained views is strictly larger in statistic term by utilizing the DIBR technique.

Due to the space constraint, the second case with only one view delivered for every R view is presented in [24] since it can be extended by Theorem 2.

IV. Protocol Design

For a multi-view 3D multicast, each view is associated with a multicast group. Based on the analytical results in Section III, each client subscribes to a set of views by joining a set of multicast groups, in order to satisfy the view failure probability. To support the dynamic join and leave of users and the change of the subscribed views, we present a new protocol, named Multi-View Group Management Protocol (MVGMP), which exploits the theoretical results in Section III. The MVGMP protocol extends the current IETF Internet standard for multicast group management, the IGMP [26], by adding the view selection feature to the protocol. In MVGMP, the AP maintains a table, named *ViewTable*, for each video. The table

specifies the current multicast views and the corresponding bit-rates for each view, and each multicast view is associated with a multicast address and a set of users that choose to receive the view. ViewTable is periodically broadcasted to all users in the WiFi cell. The MVGMP includes two control messages by extending IGMP. The first message is *Join*, which contains the address of a new user and the corresponding requested view(s). The views can be the subscribed views, or the left and right views to synthesize the subscribed view. An existing user can also exploit this message to update its requested views. The second message is *Leave*, which includes the address of a leaving user and the views that no longer need to be received. An existing user can also exploit this message to stop receiving a view. Following the design rationale of the IGMP, the MVGMP is also a soft-state protocol, which implies that each user is required to periodically send the *Join* message to refresh its chosen views, so that unexpected connection drops will not create dangling states in ViewTable.

Join. When a new user decides to join a 3D video multicast transmission, it first acquires the current ViewTable from the AP. After this, the user selects the views according to Theorem 1. Specifically, the client first examines whether ViewTable has included the subscribed view. If ViewTable does not include the subscribed view, or if the view loss probability for the subscribed view in the corresponding bit-rate exceeds the threshold, the user adds a left view and a right view accordingly. To properly select a pair of nearby views, the user carefully examines the view combination to achieve the theoretical bounds in Section III, and the closet pair of view will be extracted to satisfy the constraint on the view failure probability. The computation overhead is small because R is tiny, and thus only a small number of view combinations is necessary to be examined.

When a multi-view 3D video starts, usually the current multicast views in ViewTable are not sufficient for a new user. In other words, when the view failure probability still exceeds the threshold after the user selects some existing views in ViewTable, the user may need to add the desired view to ViewTable to reduce the view failure probability. Also, the left and right views are required to be chosen again according to the analytic results in Section III to avoid receiving too many views. After choosing the views to be received, a Join message is sent to the AP. The message contains the views that the user chooses to receive, and the AP adds the user to ViewTable accordingly. To avoid receiving too many views, note that a user can restrict the maximum number of left and right views that are allowed to be received. Also, the AP can deny a Join request if the residual bandwidth is not sufficient to support a new view.

Leave and View Re-organization. On the other hand, when a user decides to stop subscribing to a multi-view 3D video, it multicasts a Leave message to the AP and to any other user that receives at least one identical view k_i. Different from the Join message, the Leave message is also delivered to other remaining users in order to minimize the bandwidth consumption, since each remaining user that receives k_i will examine if there is a chance to switch k_i to another view $\overline{k}_i$ that is still transmitted in the network. In this case, the remaining user also sends a Leave message that includes view k_i, together with a Join message that contains view $\overline{k}_i$. If a view is no longer required by any remaining users, the AP stops delivering the view. Therefore, the MVGMP can effectively reduce the number of multicast views.

Discussion. Note that the MVGMP can support the scenario of a user changing the desired view, by first sending a Leave message and then a Join message. Similarly, when a user moves, and thus the channel condition changes, it will send a Join message to receive additional views if the channel condition deteriorates, or a Leave message to stop receiving some views if the channel condition improves. Moreover, when a user handovers to a new WiFi cell, it first sends a Leave message to the original AP and then a Join message to the new AP. If the network connection to a user drops suddenly, the AP removes the information corresponding to the user in ViewTable when it does not receive the Join message (see soft-state updates as explained earlier in this section) for a period of time. Therefore, the MVGMP also supports the silent leave of a user from a WiFi cell. Moreover, our protocol can be extended to the multi-view subscription for each client by replacing Theorem 1 with Theorem 2. The fundamental operations of Join/Leave/Reorganize remain the same since each view is maintained by a separate multicast group.

TABLE II: Simulation Settings.

Parameter	Default Value
Carrier Frequency	5.0 GHz
The unit of Channel Time	$1ms$
Channel Bandwidth	20MHz
AP Tx Power	20.0 dBm
OFDM Data Symbols	7
Video Bit-Rate (Per View)	800kbps
Transmission Data Rates	{6.5, 13, 19.5, 26, 39, 52, 58.5, 65} Mbps defined in 802.11n spec. [1]

V. Simulation Results

In the following, we first describe the simulation setting and then compare the MVGMP with the current multicast scheme.

A. Simulation Setup

We evaluate the channel time of the MVGMP in a series of scenarios with NS3[27] 802.11n package. The channel time of a multicast scheme is the average time consumption of a frame in WiFi networks. To the best knowledge of the authors, there has been no related work on channel time minimization for multi-view 3D video multicast in WiFi networks. For this reason, we compare the MVGMP with the original WiFi multicast scheme, in which all desired views are multicasted to the users.

We adopt the setting of a real multi-view 3D dataset Book Arrival [5] and the existing multi-view 3D videos [28] with 16 views, where the texture video quantization step is 6.5, and depth map quantization step is 13, and the PSNR of the synthesized views in DIBR is around 37dB [29]. The video rates for reference texture image and its associated depth map are assigned as $r_t = 600$kbps and $r_d = 200$kbps, respectively, and thus $r = 800$ kbps. The DIBR quality constraint is 3, $R = 3$. The threshold of each user is uniformly distributed in (0, 0.1]. Each user randomly chooses one preferred view from three preference distributions: Uniform, Zipf, and Normal

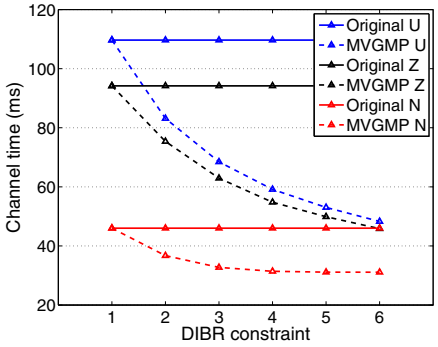

Fig. 1: Synthesis Range

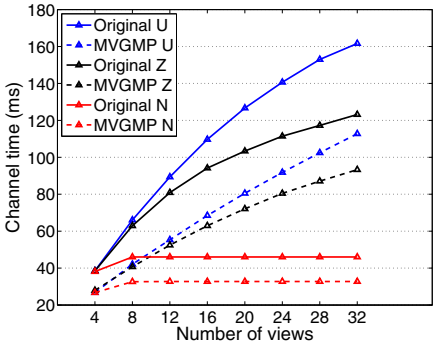

Fig. 2: Number of Views in a Video

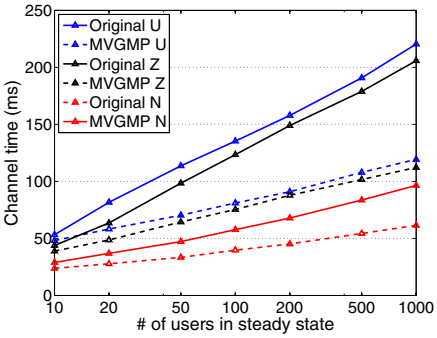

Fig. 3: Number of Users

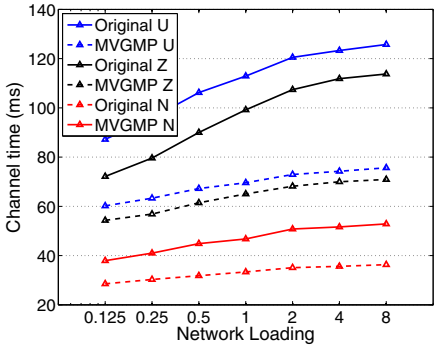

Fig. 4: Network Load

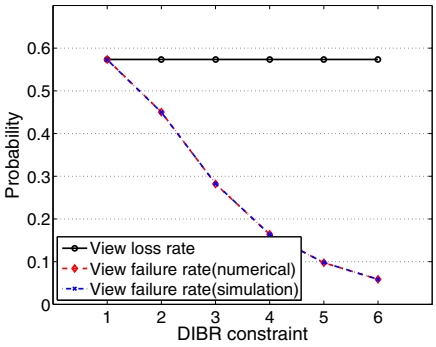

Fig. 5: View Failure Probability

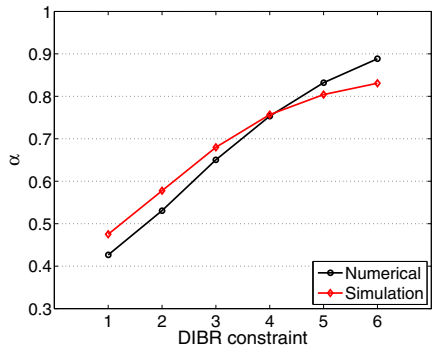

Fig. 6: Ratio of Successfully Received Views

distributions. There is no specifically hot view in the Uniform distribution. In contrast, the Zipf distribution, $f(k;s;N) = (\frac{1}{k^s})/\sum_{n=1} N(\frac{1}{n^s})$, differentiates the desired views, where k is the preference rank of a view, s is the the exponent characterizing the distribution, and N is the number of views. The views with smaller ranks are major views and thus more inclined to be requested. We set $s = 1$ and $N = |V|$ in this paper. In the Normal distribution, central views are accessed with higher probabilities. The mean is set as $|V|/2$, and the variance is set as 1 throughout this study.

We simulate a dynamic environment with 50 client users located randomly in the range of an AP (30 meters for 20.0dBm Tx power). After each frame, there is an arrival and departure of a user with probabilities λ and μ, respectively. In addition, a user changes the desired view with probability η. The default probabilities are $\lambda = 0.2$, $\mu = 0.3$, $\eta = 0.4$. TABLE II summarizes the simulation setting consisting of an 802.11n WiFi network with a 20MHz channel bandwidth. In the following, we first compare the performance of the MVGMP with the current WiFi multicast scheme in different scenarios and then compare the analytical and simulation results.

The relationship between the setting of R and the video quality has been studied in [2], [23]. Therefore, due to the space constraint, we focus on the channel time and view failure probability with different R here, and the corresponding video quality can be simply acquired according to [2], [23].

B. Scenario: Synthesized Range

Fig. 1 evaluates the MVGMP with different settings of R. Compared with the current WiFi multicast, the channel time is effectively reduced in the MVGMP as R increases. Nevertheless, it is not necessary to set a large R because the improvement becomes marginal as R exceeds 3. Therefore, this finding indicates that a small R (i.e., limited quality degradation) is sufficient to effectively reduce the channel time in WiFi.

C. Scenario: Number of Views

Fig. 2 explores the impact of the numbers of views in a video. The channel time in both schemes increases when the video includes more views, because more views need to be transmitted. This result shows that MVGMP consistently outperforms the original WiFi multicast scheme with different numbers of views in a video.

D. Scenario: Number of Users in Steady State

Fig. 3 evaluates the channel time with different numbers of users in the steady state. We set $\lambda = \mu = 0.25$, so that the expected number of users in the network remains the same. The channel time was found to grow as the number of users increases. Nevertheless, the increment becomes marginal since most views will appear in *ViewTable*, and thus more users will subscribe to the same views in the video.

E. Scenario: Utilization Factor

Fig. 4 explores the impact of the network load. Here, we change the *loading ratio* $\rho := \frac{\lambda}{\mu}$, i.e., the ratio between the arrival probability λ and departure probability μ. Initially, new multicast users continuously join the 3D video stream until the network contains 50 users. The results indicate that the channel time increased for both multicast schemes. Nevertheless, the MVGMP effectively reduces at least 40% of channel time for all three distributions.

F. Impact of User Preferences

From Fig. 1 to Fig. 4. the results clearly show that Uniform distribution requires the most channel time compared with Zipf and Normal distributions. This is because in Zipf and Normal distributions, users prefer a sequence of hot views, and those views thus have a greater chance to be synthesized by nearby views with DIBR.

G. Analytical Result

Fig. 5 and Fig. 6 compare the simulation results from NS3 and the analytical results of Theorem 1 and Theorem 2 for the Uniform distribution, where each user subscribes to each view with a probability of 0.8. The results reveal that the discrepancy among the simulation and analysis is very small. Most importantly, α increases for a larger R since each user can synthesize and acquire a desired view from more candidate right and left views when the desired view is lost during the transmissions.

VI. Conclusion

With the emergence of mobile and VR devices, this paper proposes to leverage DIBR in multi-view 3D video to faciliate view protection in WiFi multicast networks. We first investigated the merits of view protection via DIBR and showed that the view failure probability is much smaller than the view loss probability, while the multi-view subscription for each client was also studied. Thereafter, we proposed the Multi-View Group Management Protocol (MVGMP) to handle the dynamic joining and leaving for a 3D video stream and the change of the desired view for a client. The simulation results demonstrated that with DIBR our protocol effectively reduces 50% of the bandwidth consumption and increases the probability for each client for 50% to successfully playback the desired view in a multi-view 3D video.

Acknowledgment

This work was supported in part by MOST 105-2622-8-009-008.

References

[1] IEEE, "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," *IEEE Standard*, Mar. 2012.

[2] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View Generation with 3D Warping Using Depth Information for FTV," *Signal Processing: Image Communication*, vol. 24, no. 1-2, pp. 65–72, Jan. 2009.

[3] M. Nasralla, O. Ognenoski, and M. Martini, "Bandwidth Scalability and Efficient 2D and 3D Video Transmission over LTE Networks," in *IEEE International Conference on Communications*, Jun. 2013, pp. 617–621.

[4] B. Feitor, P. Assuncao, J. Soares, L. Cruz, and R. Marinheiro, "Objective Quality Prediction Model for Lost Frames in 3D Video over TS," in *IEEE International Conference on Communications*, Jun. 2013, pp. 622–625.

[5] I. Feldmann, M. Mueller, F. Zilly, R. Tanger, K. Mueller, A. Smolic, P. Kauff, and T. Wiegand, "HHI Test Material for 3-D Video," *Proc. 84th Meet. ISO/IEC JTC1/SC29/WG11, document M15413*, Apr. 2008.

[6] F. Shao, G. Jiang, M. Yu, K. Chen, and Y.-S. Ho, "Asymmetric Coding of Multi-View Video plus Depth based 3D Video for View Rendering," *IEEE Trans. on Multimedia*, vol. 14, no. 1, pp. 157–167, Feb. 2012.

[7] A. Hamza and M. Hefeeda, "Energy-Efficient Multicasting of Multiview 3D Videos to Mobile Devices," *ACM Trans. on Multimedia Computing, Communications, and Applications*, vol. 8, no. 3s, pp. 45:1–45:25, Sep. 2012.

[8] Y. Aksoy, O. Sener, A. Alatan, and K. Ugur, "Interactive 2D-3D Image Conversion for Mobile Devices," in *IEEE International Conference on Image Processing*, Sep. 2012, pp. 2729–2732.

[9] "Information Technology–Dynamic Adaptive Streaming over HTTP (DASH)," *ISO/IEC 23009-1*, Dec. 2014.

[10] S. Alcock and R. Nelson, "Application Flow Control in YouTube Video Streams," in *SIGCOMM Computer Communication Review*, Apr. 2011.

[11] A. Sheth, S. Nedevschi, R. Patra, S. Surana, E. Brewer, and L. Subramanian, "Packet Loss Characterization in WiFi-Based Long Distance Networks," in *IEEE International Conference on Computer Communications*, May 2007, pp. 312–320.

[12] J. Feng, Z. Liu, and Y. Ji, "Wireless Channel Loss Analysis - a Case Study using WiFi-Direct," in *International Wireless Communications and Mobile Computing Conference*, Aug. 2014, pp. 244–249.

[13] A. Bose and H. F. Chuan, "A Practical Path Loss Model for Indoor WiFi Positioning Enhancement," in *International Conference on Information, Communications and Signal Processing*, Dec. 2007, pp. 1–5.

[14] Ya-Ju Yu, Pi-Cheng Hsiu, and Ai-Chun Pang, "Energy-Efficient Video Multicast in 4G Wireless Systems," *Mobile Computing, IEEE Transactions on*, vol. 11, no. 10, pp. 1508–1522, Oct. 2012.

[15] C.G. Gurler, B. Gorkemli, G. Saygili, and A.M. Tekalp, "Flexible Transport of 3-D Video Over Networks," *Proceedings of the IEEE*, vol. 99, no. 4, pp. 694–707, April 2011.

[16] Jongryool Kim, Kiho Choi, Hyunyong Lee, and JongWon Kim, "Multi-View 3D Video Transport using Application Layer Multicast with View Switching Delay Constraints," in *3DTV Conference, 2007*, May 2007, pp. 1–4.

[17] E. Kurutepe and T. Sikora, "Feasibility of Multi-View Video Streaming Over P2P Networks," in *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, 2008*, May 2008, pp. 157–160.

[18] Li Zuo, Jian Guang Lou, Hua Cai, and Jiang Li, "Multicast of real-time multi-view video," in *Multimedia and Expo, IEEE International Conference on*, July 2006, pp. 1225–1228.

[19] IEEE, "Information Technology–Telecommunications and Information Exchange between Systems–Local and Metropolitan Area Networks–Specific Requirements–Part 2: Logical Link Control," *IEEE Standard*, 1998.

[20] T. Speakman *et al.*, "PGM Reliable Transport Protocol Specification," *IETF RFC 3208*, Dec. 2001.

[21] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[22] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard," in *Proceedings of the IEEE*, vol. 99, no. 4, Apr. 2011, pp. 626–642.

[23] G. Cheung, V. Velisavljevic, and A. Ortega, "On Dependent Bit Allocation for Multiview Image Coding with Depth-Image-Based Rendering," *IEEE Trans. on Image Processing*, vol. 20, no. 11, pp. 3179–3194, Dec. 2011.

[24] C.-H. Lin, D.-N. Yang, J.-T. Lee, and W. Liao, "Error-Resilient Multicasting for Multi-View 3D Videos in IEEE Wireless Networks," *CoRR*, vol. abs/1503.08726, 2015.

[25] S. M. Ross, *Stochastic Processes*, 2nd ed. Wiley, 1983.

[26] B. Cain *et al.*, "Internet Group Management Protocol, ver. 3," *IETF RFC 3376*, Oct. 2002.

[27] The ns-3 network simulator. [Online]. Available: http://www.nsnam.org/

[28] P. Merkle, A. Smolic, K. Mller, and T. Wiegand, "Multi-view video plus depth representation and coding," in *IEEE International Conference on Image Processing (IEEE ICIP)*, vol. 1, Sept 2007, pp. I – 201–I – 204.

[29] Z. Gaofeng, J. Gangyi, Y. Mei, L. Fucui, S. Feng, and P. Zongju, "Joint video/depth bit allocation for 3d video coding based on distortion of synthesized view," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Jun. 2012, p. 16.