ECast: An enhanced video transmission design for wireless multicast systems over fading channels

Zhilong Zhang, Danpu Liu, Xiaoli Ma, and Xin Wang

Abstract—Conventional wireless video multicast schemes face a challenge due to their stair-shaped performance curves called “cliff effect”[11] and the limited ability to accommodate multiple users with diverse channel conditions. Recently, a series of analog video coding approaches aiming to solve the problem has been proposed. In this paper, our goal is to enhance such schemes through a novel cross-layer design. A joint subcarrier matching and power allocation scheme which is proved to be optimal in terms of minimizing the mean square error is proposed. The scheme can dynamically adapt to both channel conditions and video contents, and at the same time maintain continuous quality scalability. Furthermore, the feedback overhead of the channel state information from multiple users is addressed and a novel analog feedback method is proposed. The video qualities at various receivers are controllable through assigning different weights to different users. Therefore, this scheme is suitable for wireless video multicast systems. Simulation results validate the effectiveness of our proposal.

Index Terms—Analog video coding, power allocation, subcarrier matching, wireless video multicast, analog feedback.

I. INTRODUCTION

In traditional wireless video multicast systems, the video coding layer and the wireless transmission layer are often designed separately. Source coding and channel coding are employed as individual units. Little information is exchanged between the two layers. Most of the systems, such as digital video broadcasting (DVB)[2] and some adaptive cross-layer proposals[3, 4], have stair-shaped performance curves and are fragile to transmission errors. When casting video signals, a transmitter encodes video sources at a specific coding rate to serve all the users in a multicast cell. Users under weak channel conditions cannot reconstruct videos correctly which causes the severe degradation of quality of experience (QoE); meanwhile, those under strong channel conditions cannot obtain further improved performance. Although by adopting scalable video coding (SVC) [5], weak receivers can afford low resolution video signals while strong receivers can decode high quality video signals, SVC cannot serve well all users with various channel conditions. SVC only replaces one large stair with multiple small stairs, and does not entirely mitigates the cliff effect[6].

Recently, a novel video coding technique called SoftCast was proposed in [7,9]. SoftCast adopts a special analog video coding structure. Video data are first compressed by 3D-Discrete Cosine Transform (DCT). DCT coefficients are grouped into chunks and different amounts of power are allocated to minimize the mean square error (MSE). Every two DCT coefficients from the same chunk are mapped into one dense quadrature amplitude modulation (QAM) symbol and are transmitted directly through an OFDM-based physical layer. Unlike in conventional video coding schemes, such as H.264 and MPEG, the entropy coding part is absent in the framework, because it makes the video quality sensitive to transmission errors. The key idea of SoftCast is to maintain the linear relation between pixel values and transmitted samples. The graceful design leads to a roughly linear relation between the peak signal to noise ratio (PSNR) and the signal to noise ratio (SNR), and makes the system robust and scalable.

After SoftCast, a large number of proposals have aimed to improve the original analog coding scheme in various ways. To reduce interframe redundancy, 3D-DCT is replaced with 3D wavelet in [6] and employed motion estimation is introduced in [10]. A wireless video transmission framework called DCast which was proposed in An improved scheme called SIRCast based on DCast was proposed in [11]. In [12], Xiong et al. proposed an adaptive chunk division method which adapted well to the energy distribution of DCT coefficients. WSVC which presented a hybrid digital-analog joint source-channel coding scheme that integrated the advantages of digital coding and analog coding was proposed in [13]. The analog coding part was similar to SoftCast.

On the other hand, many efforts have been devoted to improving the original wireless transmission scheme. A collision prevention mechanism was proposed to allow SoftCast to operate efficiently in IEEE 802.11 networks in [14]. ParCast [15] was proposed to enhance video transmission by introducing MIMO. ParCast applied MIMO precoding to improve the link quality and mapped important video components to more reliable channels. Cooperative communications were introduced in [16] which extended the work of [13]. Receivers could transmit messages between each other, and those with better SNR served as relays.

The aforementioned results provided various ways to improve the performance of SoftCast. Gaussian channels were assumed in most proposals. However, signals often experience fading during wireless propagation. In the OFDM-based physical layer, a large frequency band is divided into...
multiple narrowband subcarriers with different channel fading factors. Subcarriers, as significant wireless resources, should be allocated properly. Among these proposals, only ParCast mentioned mapping important video components to more reliable channels, but it focused only on video unicast. The method was not applicable for multicast systems because of the high channel state information (CSI) feedback overhead from multiple users. In addition, the optimality of its mapping scheme was not proved.

In this paper, we focus on how to optimally allocate subcarrier and power according to both video contents and channel conditions, while coping with the large feedback overhead. Through a cross-layer design of both the video layer and the OFDM-based physical layer, an enhanced video multicast scheme called ECast is proposed. The main contributions are threefold.

- First, a cross-layer resource allocation problem is formulated, and a joint carrier matching and power allocation scheme is derived which is proved to be optimal;
- Second, an analog feedback method is proposed to limit the feedback overhead for video multicast;
- Third, a quality-controlled method is introduced. By assigning different importance factors considering fairness or paid prioritization, we can either protect users with poor channel conditions or enhance the performance of high-paying users.

Simulation results show that the proposed scheme outperforms the performance of typical analog video coding systems significantly and is suitable for wireless video multicast.

The remainder of this paper is organized as follows. Section II describes the system design. Section III elaborates the problem setting and the optimal solution of joint carrier matching and power allocation. The system model is simplified, where the number of chunks is assumed to be equal to the number of subcarriers. Section IV extends the simplified model to a general scenario where the two numbers are unequal. Bandwidth adaption and user mobility are also considered. In Section V the LLSE decoder is introduced and a novel analog feedback scheme is proposed based on the theoretical results of Section III-C. Simulation results verify our proposal in Section VI and a summary concludes the paper in Section VII.

II. DESIGN OVERVIEW

In this section, we illustrate the system design of the considered wireless video multicast system. The system is composed of one service access point (SAP) as the transmitter and multiple users as the receivers. Time division duplexing (TDD) is assumed. Both downlink and uplink channels are based on OFDM. All users share an uplink feedback channel, through which CSI is fed back to the SAP. Meanwhile, video data are transmitted from the SAP to multiple users through a downlink multicast channel. There are two categories of video data in the downlink: content data and meta data. Only a small amount of important information is treated as meta data. Meta data are transmitted through traditional communication methods with high redundancy, such that the transmission obtains more protection and is nearly lossless. The majority of video data are content data which are transmitted in lossy mode. They are more tolerant to noise, fading and interference. The content data to be transmitted are divided into chunks with equal sizes. Symbols in the same chunk are carried by at most one subcarrier within a channel coherent time, so that they experience similar channel attenuation.

Fig. 1 depicts the framework of the proposed system. Modules that make the approach outperform others are highlighted in the figure and will be detailed in the following sections. Meanwhile, given that the system is based on the structure of SoftCast\textsuperscript{1} for integrity, modules that are similar to those adopted by SoftCast are also briefly described below. The rest of this section overviews how to transmit different categories of data based on the framework. Note that the proposed scheme can also be used in video systems that incorporate an analog coding part, such as WSVC\textsuperscript{13}.

A. Content data transmitted in lossy mode

Content data are the dominant data which are transmitted in lossy mode. The main components for transmitting content data are described as follows.

For the transmitter:

- **Linear Transform**: Linear transform is applied to a group of pictures (GOP) to compress video signals. The transform matrix could be 2D/3D DCT\textsuperscript{8}, 3D wavelet\textsuperscript{6} or any other operator matrices with good performance for compression.
- **Channel Adaption**: The role of channel adaption is to adapt the video layer to the transmission layer. Coefficients obtained after the linear transformation are grouped into chunks. Source bit rates, wireless channels and user mobility are jointly considered. The mean and the variance of each chunk and a bitmap that indicates the discarded chunks are sent as meta data. The detail of channel adaption will be elaborated in Section IV.
- **Subcarrier Matching and Power Allocation**: Subcarrier matching and power allocation (or power scaling) are performed at chunk level. The SAP collects channel conditions from each user, determines the optimal subcarrier matching order and calculates the amount of power allocated to each chunk. Matching information and power scaling factors are transmitted as meta data. The joint subcarrier matching and power allocation scheme is the key technology we focus on in this paper, which will be detailed in Section III.
- **Dense QAM Mapping**: A pair of two coefficients in the same chunk after power scaling is mapped to a 64K-QAM symbol. Specifically, each coefficient is quantized into an 8-bit integer and every two integers compose one complex number of 64K possible values. The dense QAM mapping without channel coding makes the end-to-end MSE roughly linear to the impact of wireless channels.

\textsuperscript{1}Compared with traditional wireless video transmission systems, the proposed scheme has the same limitation and benefit as SoftCast as discussed in \textsuperscript{8}. We have not stated them in the paper.
- **OFDM process**: 64K-QAM symbols are allocated to subcarriers according to the optimal matching solution. Subsequently, symbols of each subcarrier undergo the IFFT and D/A processes. Then the obtained analog signals are modulated with carrier waves to generate the final transmitted video signal.

A receiver performs the inverse processes of the transmitter:

- **Inverse OFDM**: A receiver receives the signal and reconstructs the modulated complex symbols of both the content data and the meta data. To assist the decoder in inverting the received signal, meta data are decoded firstly.

- **Inverse QAM**: Scaled coefficients are reconstructed by inverse dense QAM. Each complex value is decoupled back into two real values.

- **Demodulation and CSI feedback**: Coefficients are demodulated by LLSE (see Section V-A). The CSI feedback scheme will be briefly described in Section II-C and detailed in Section V.

- **Chunks combination**: Demodulated chunks are combined into a group of transformed coefficients. Those discarded at the transmitter are set to zero at their GOP positions.

- **Inverse Transform**: The coefficients are transformed by the inverse linear transform, and a reconstructed GOP is finally obtained.

Without any nonlinear processes in our system, the end-to-end performance has a nearly linear relation with the SNR. As will be shown, the performance is largely improved by adopting the proposed scheme.

**B. Meta data transmitted in lossless mode**

Meta data assist receivers in decoding the received signals. They are generated in different video encoding processes, consisting of the mean and the variance of each chunk, a bitmap that indicates the discarded chunks, chunk mapping information and scaling factors. Meta data are coded using conventional digital communication schemes consisting of variable length coding (VLC), forward error correction (FEC) and binary phaseshift keying (BPSK) mapping, and decoded by receivers with the inverse processes.

**C. CSI transmitted in analog mode**

The CSI feedback from multiple users is transmitted through a shared uplink channel. As the upcoming WiFi standard and LTE both use the channel feedback, assuming the availability of channel conditions at the SAP is reasonable. However, the problem is not about the availability of CSI but how to limit the large overhead of feedback from numbers of users. In our proposal, users do not transmit channel states themselves, but transmit the strengths of channel gains through parallel tone signals. All feedback from different users is transmitted simultaneously. They are superposed at the SAP. The SAP does not need to separate each tone from others, but uses the superposed feedback directly. The detail of this limited feedback scheme and the reason why it works will be described in Section V.

### III. Joint Subcarrier Matching and Power Allocation

In this section, we discuss the joint subcarrier matching and power allocation scheme and analyze its benefits to the entire system. Firstly, a non-linear zero-one optimization problem for video unicast is formulated and the optimal solution is obtained. Then, the results for unicast are extended to multicast. We focus only on the system with equal number of chunks and subcarriers in this section. It is a simplified and basic model, and can be easily extended to a more general case (see Section V).

**A. Problem formulation for unicast**

Power scaling or power allocation is a key component in traditional analog video coding schemes[7,9]. DCT coeffi-
coefficients are scaled before transmission and the scaling factors are obtained by solving an optimization problem. In general, the transmitted chunks are often assumed to experience AWGN channels in the system design and wireless fading is handled by the OFDM-based physical layer itself.

When considering channel fading, we have to derive a new method to calculate the scaling factors and at the same time determine how to match the chunks to the subcarriers. We address the joint resource allocation problem for the system with equal number of chunks and subcarriers in this subsection. Assume $N$ chunks should be transmitted and $N$ subcarriers are available. Each subcarrier carries only one chunk, i.e. only one-to-one matching is allowed.

A coefficient in the $i^{th}$ chunk is selected as a typical value. Assume it is scaled from $x_i$ to $y_i = g_i x_i$ before transmission, where $g_i$ is the scaling factor for this chunk. Define the power of a chunk as its variance. We denote by $\lambda_i$ the variance of chunk $i$ and $\mu_i$ its power after applying the gain. Assuming the mean of each chunk has been removed to get a zero-mean distribution, we obtain that $\lambda_i = E[x_i^2]$, $\mu_i = E[y_i^2]$ and $g_i = \frac{\lambda_i}{\sqrt{\mu_i}}$. All coefficients of the chunk are allocated to the $j^{th}$ subcarrier and experience the same fading during transmission. The receiver receives $\hat{y}_{ij} = a_{ij} y_i + n$, where $a_{ij}$ is the channel fading coefficient of the $j^{th}$ subcarrier, and $n$ is a random Gaussian variable with zero mean and variance $\sigma^2$. The receiver decodes

$$\hat{x}_{ij} = \frac{\hat{y}_{ij}}{a_{ij} g_i} = x_i + \frac{n}{a_{ij} g_i},$$

and the expected MSE of the $i^{th}$ chunk carried by the $j^{th}$ subcarrier is obtained by

$$e_{ij} = E[(\hat{x}_{ij} - x_i)^2] = \frac{\lambda_i \sigma^2}{|a_{ij}|^2 \mu_i}.$$  \hfill (2)

Given that larger $\lambda_i$ contributes more to the MSE, $\lambda_i$ can be considered as the importance of the $i^{th}$ chunk. The end-to-end MSE of a GOP is expressed by

$$MSE = \sum_i \sum_j b_{ij} e_{ij} = \sum_i \sum_j b_{ij} \lambda_i \sigma^2/|a_{ij}|^2 \mu_i,$$  \hfill (3)

where $b_{ij}$ is a binary value denoting whether the $i^{th}$ chunk is allocated to the $j^{th}$ subcarrier. $b_{ij} = 1$ means that the $i^{th}$ chunk is allocated to the $j^{th}$ subcarrier and $b_{ij} = 0$ means the opposite. Let $P_{tot}$ be the total power budget. An optimization problem is formulated as

$$\begin{align*}
\text{minimize} & \quad \sum_i \sum_j b_{ij} \lambda_i \sigma^2/|a_{ij}|^2 \mu_i, \\
\text{subject to} & \quad \sum_j \mu_i \leq P_{tot}, \\
& \quad \sum_j b_{ij} = 1, \\
& \quad \sum_i b_{ij} = 1, \\
& \quad b_{ij} \in \{0, 1\}, \\
& \quad \mu_i \geq 0,
\end{align*}$$  \hfill (4)

which is a mixed binary programming problem. Obtaining a global optimal solution is difficult because of the computational complexity. However, when $\{b_{ij}\}$ is given, the problem is reduced to a simple convex optimization problem

$$\begin{align*}
\text{minimize} & \quad \sum_i \lambda_i \sigma^2/c_i \mu_i, \\
\text{subject to} & \quad \sum_i \mu_i \leq P_{tot}, \\
& \quad \mu_i \geq 0,
\end{align*}$$  \hfill (5)

where $\{c_i\}$ is a rearrangement of $\{|a_{ij}|^2\}$ depending on the subcarrier matching scheme. The optimal power allocation of (5) can be obtained by KKT conditions \cite{17} as

$$\mu_i = \frac{P \sqrt{\lambda_i}}{\sum_i \sqrt{\lambda_i} c_i}$$  \hfill (6)

and scaling factors are calculated accordingly:

$$g_i = \sqrt{\frac{\lambda_i}{\mu_i}} = \sqrt{\frac{P}{\sum_i \sqrt{\lambda_i} c_i}}.$$  \hfill (7)

Therefore, the optimal joint subcarrier matching and power allocation can be obtained by finding the minimal objective function among all subcarrier matching possibilities, and the corresponding subcarrier matching and power allocation are jointly optimal. Unfortunately, the complexity is $O(N!)$.

In this paper, a low-complexity matching scheme is proposed and its optimality can be proved.

**B. Problem solution for video unicast**

In order to obtain the optimal solution, a system including only two chunks and two subcarriers is considered firstly. Then, this system is extended to a system including $N$ chunks and $N$ subcarriers, where $2 \leq N < \infty$.

When the numbers of chunks and subcarriers are both two, solving the problem (4) is equivalent to obtaining the minimum of the following two optimization problems, each of which

---

2Wireless fading includes large scale fading (path loss), shadow fading and small scale fading (multipath effect). Here, the channel coefficient is the product of all fading factors.
indicates a subcarrier matching scheme:

**P1:**

\[
\begin{align*}
\text{minimize} & \quad \text{err}_1 = \frac{\lambda_1 \sigma^2}{h_1 \mu_1} + \frac{\lambda_2 \sigma^2}{h_2 \mu_2}, \\
\text{subject to} & \quad \mu_1 + \mu_2 \leq P, \\
& \quad \mu_1 \geq 0, \mu_2 \geq 0,
\end{align*}
\]

\( (\hat{\mu}_1, \hat{\mu}_2) \) is one of the feasible solutions of **P1.** The difference of \( \text{err}_1(\hat{\mu}_1, \hat{\mu}_2) \) and \( \text{err}_2(\mu_1', \mu_2') \) is calculated by

\[
	ext{err}_1(\hat{\mu}_1, \hat{\mu}_2) - \text{err}_2(\mu_1', \mu_2') = \frac{\lambda_1 \sigma^2}{h_1 \left( \mu_1' + \left( 1 - \frac{h_1}{h_2} \right) \mu_2' \right)} + \frac{\lambda_2 \sigma^2}{h_2 \mu_1} - \frac{\lambda_2 \sigma^2}{h_2 \mu_1'} - \frac{\lambda_1 \sigma^2}{h_1 \mu_1'} = \frac{\lambda_1 \sigma^2}{h_1} \frac{(h_2 - h_1)(h_2 \mu_1' - h_1 \mu_2')}{h_2 \mu_1'} = \frac{\lambda_1 \sigma^2}{h_1 b_2^2 \mu_1'} \frac{h_2 - h_1}{h_2 \mu_1'}. \tag{11}
\]

Given that \( h_1 \mu_2' > h_2 \mu_1' \) and \( h_1 < h_2 \), (11) has a non-positive value. As a result, it is finally obtained that \( \text{err}_1(\hat{\mu}_1, \hat{\mu}_2) \leq \text{err}_1(\mu_1', \mu_2') \leq \text{err}_2(\mu_1', \mu_2') \).

The above two cases all indicate that **P1** has a better solution than **P2.** Therefore, **P1** is the optimal matching scheme, and the optimal solution of **P1** is also optimal for problem [4].

Next, we extend the method to the system including \( N \) subcarriers and \( N \) chunks. Denote the subcarrier gains and the chunk variances by \( \{h_i\} \) and \( \{\lambda_i\} \), respectively. Here, without loss of generality, it is still assumed that \( h_i \leq h_j \) and \( \lambda_i \leq \lambda_j \), if \( i \leq j \). For the global optimum, the following proposition gives the joint power allocation and subcarrier matching approach.

**Theorem 1.** For the system including \( N \) subcarriers and \( N \) chunks, the optimal subcarrier matching matches the \( i \)th chunk to the \( i \)th subcarrier where both \( \{\lambda_i\} \) and \( \{h_i\} \) have been arranged in a decreasing order, respectively. Together with the optimal power allocation for this subcarrier matching, the optimal joint subcarrier matching and power allocation scheme is obtained.

**Proof.** Theorem will be proved in a contrapositive form. Suppose that there exists a subcarrier matching method whose matching result includes the following two pairs that \( \lambda_i \) is matched to \( h_i \) and \( \lambda_k \) is matched to \( h_l \), where \( \lambda_i < \lambda_k \) and \( h_j > h_l \). Under this supposition, the minimal MSE obtained by adopting this matching method is lower than that by adopting Theorem [4].

Fix the matching result and the amounts of power allocated to other subcarriers. The total MSE of the above two pairs can be further reduced according to Lemmas [1]. Thus, the total MSE of all the chunks can be reduced by rematching subcarriers and the matching method is obviously not optimal, which is contrary to the supposition. Therefore, the optimal matching result should include no such matching pairs. Since only the method denoted by Theorem [4] satisfies the requirement, this subcarrier matching and corresponding optimal power allocation are the optimal joint subcarrier matching and power allocation.

So far, the optimal joint subcarrier matching and power allocation for unicast is provided. The optimal scheme matches the subcarriers by the order of the channel gains and chunk
variances, and allocates optimal power under this matching accordingly. The scaling factors can be calculated by (7). Thus the mixed integer optimization problem (4) is reduced to a simplified form (5) by adopting Theorem 1.

C. Extension to video multicast

In this subsection, the results of unicast are extended into multicast. To make the video qualities at various receivers controllable, different weights or importance factors are assigned to multiple users. Define the new optimization objective as the weighted sum of all users’ MSEs. By minimizing the objective, users with high weights will get large performance improvement. We can enlarge the weights for users at cell edge or some important users to enhance their performance.

Assuming $K$ users locate randomly in a wireless video multicast cell and the $k^{th}$ user’s weight is $\beta_k$, a new optimization problem for video multicast is formulated as

$$
\text{minimize} \quad \sum_k \beta_k \sum_i \sum_j \frac{b_{ij} \lambda_i \sigma^2}{|a_{ijk}|^2 \mu_i}.
$$

subject to

$$
\sum_i \mu_i \leq P,
$$
$$
\sum_i b_{ij} = N,
$$
$$
\sum_j b_{ij} = N,
$$
$$
\mu_i \geq 0, b_{ij} \in \{0, 1\},
$$

where $|a_{ijk}|^2$ denotes the $i^{th}$ channel gain of the $k^{th}$ user, which can be obtained from the feedback channel.

We define the virtual gain of the $j^{th}$ subcarrier as the weighted harmonic mean of all users’ $j^{th}$ channel gains as follows:

$$
|\tilde{a}_j|^2 = \frac{1}{\sum_k \frac{\sigma^2}{|a_{jk}|^2}}.
$$

According to the above definition, the objective of (12) is rewritten as

$$
\sum_i \sum_j b_{ij} \lambda_i \sigma^2 |\tilde{a}_j|^2 \mu_i.
$$

Note that by replacing the objective with (14), (12) has the same form as (4) and thus can be solved by the proposed method.

In conclusion, the SAP collects all users’ CSI and calculates the virtual channel gains denoted by (13). Then, the joint carrier matching and power allocation scheme is applied in the same way as video unicast: the variances of chunks and the virtual channel gains are arranged in a decreasing order, respectively, and then matched by this order; subsequently, the optimal power is calculated for this subcarrier matching accordingly. This scheme is the optimal carrier matching and power allocation scheme for video multicast in terms of minimizing the weighted sum of the end-to-end MSEs.

IV. CHANNEL ADAPTION

In the proposed real-time wireless video multicast system, the role of channel adaption is to adapt the video layer to the transmission layer. The bit rate of the video source, the available bandwidth and user mobility are jointly considered, which are all in the scope of channel adaption. Another goal of this module is to extend the results in Section III into a general case where the number of chunks is unequal to the number of subcarriers.

Assume that the time is slotted and the length of a time slot is less than the channel coherence time. As a result, channel states stay nearly unchangeable within one time slot. In order to simplify the resource allocation process, a proper chunk size is determined such that each subcarrier carries an integral number of chunks in one time slot. As shown in Fig. 2 for real-time transmission, chunks in a GOP should be transmitted within a GOP period, which is determined by the number of frames per GOP and the frame rate. Assume that $N_s$ time slots are needed to transmit a GOP. Given that future channel states are unknown in advance at the SAP and only instantaneous values are fed back from receivers, the joint carrier matching and power allocation scheme has to be applied independently in each time slot.

A. Bandwidth adaption

The amount of information that can be transmitted is limited by the capacity of the PHY layer. When the bandwidth is larger than needed, the transmitter skips a certain number of weak subcarriers and allocate chunks to strong subcarriers in each time slot; otherwise, the least important chunks are discarded at the transmitter and reconstructed with zero values at the receiver. The chunks with least variance are considered as least important chunks, because they contribute the least to the video quality. After channel adaption, we ensure that the number of reserved chunks is an integer multiple of the number of selected subcarriers.

B. Mobility adaption

When a receiver is moving, the channel status changes fast and the SNR also fluctuates. To adapt user mobility, shot time slots are needed. Assume a GOP is transmitted in $N_s$ time slots. Then, the reserved chunks should be divided into $N_s$ groups with one group transmitted in one time slot. To disperse the importance of video data across all time slots, let the $N_s$ groups have a nearly equal average variance. Therefore,
A chunk interleaving method is proposed in this paper. It can be viewed as a chunk interleaver which has the same effect as that of the Hadamard transform in SoftCast. Given that the chunks are sorted according to their variances, the \((2N_s i + n)\)th and \((2N_s i + N_s - n)\)th chunks are assigned to the \(n\)th time slot, where \(i = 0, 1, 2, \ldots\). Fig. 3 illustrates the method where \(N_s\) is set to 3. The \((6i + 1)\)th and \((6i + 6)\)th chunks are allocated to the first time slot, \((6i + 2)\)th and \((6i + 5)\)th chunks to the second time slot and \((6i + 3)\)th and \((6i + 4)\)th chunk to the third time slot, respectively.

\[ 
\lambda_i = \frac{\sum a_i^2}{\sigma^2 + \sum a_i^2}, 
\]

where \(y_i[k]\) is the received sample, \(\sigma^2\) denotes the noise power and \(a_i^*\) is the complex conjugate of \(a_i\).

\[ 
F_j = \sum_k F_{jk}|a_{jk}|^2 + \sigma^2 = \sum_k \frac{\beta_k}{|a_{jk}|^2} + \sigma^2, 
\]

where \(\sigma^2\) is the noise power. Although the SAP cannot distinguish each of the channel gains, it directly obtains an inverse of (13). As a result, the SAP only needs to listen to the

\[ 
F_{jk} = \frac{\beta_k}{|a_{jk}|^2}. 
\]
feedback time slot and obtains the detected power of signals over each subcarrier.

When only one user exists, the feedback that the SAP received in unicast and multicast is unified. The effect of noise power will be analysed in Section VI-C. As will be shown, the system is not sensitive to noise.

VI. SIMULATION RESULTS

In this section, the performance of our method is evaluated. Video algorithms often perform differently when tested with resolutions of 228 pixels × 352 pixels as shown in Table I. Coefficients of each frame are divided into 64 chunks. An OFDM-based PHY layer with 64 subcarriers is adopted. We assume that the video signal experiences Rayleigh fading and path loss during transmission. For Rayleigh fading, ZMCSCG random variables with unit variance are adopted as fading coefficients. For path loss, when a user is located at a distance $r_k$ from the SAP, the received signal power is attenuated by $r_k^{-\alpha}$, where $\alpha$ is the path loss factor and is set to 4 in our simulations.

Different schemes are compared using a standard metric PSNR, which is defined as

$$PSNR = 10 \log_{10} \left( \frac{2L - 1}{MSE} \right)^2,$$

where $L$ is the number of bits used to encode each pixel and is typically 8 bits.

A. Unicast Performance Comparison

Fig. 4 shows the performance comparison of video unicast. ECast has a 4 dB - 6 dB gain compared to SoftCast with Hadamard transform in a large SNR range from 0 dB to 30 dB. Hadamard transform provides a 2 dB - 3 dB precoding gain in low SNR ranges in fading channels. However, our proposal outperforms Hadamard transform. When SNR tends to be higher, the distortion caused by wireless transmission errors is smaller and chunk discarding in bandwidth adaption becomes the dominant reason for performance loss. As a result, the gain tends to be small in high SNR ranges. ECast without subcarrier

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>TEST SEQUENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>akiyo</td>
<td>bus</td>
</tr>
<tr>
<td>crew</td>
<td>flower</td>
</tr>
<tr>
<td>harbour</td>
<td>ice</td>
</tr>
<tr>
<td>soccer</td>
<td>stefan</td>
</tr>
<tr>
<td></td>
<td>tempete</td>
</tr>
<tr>
<td></td>
<td>waterfall</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>SIMULATION PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of chunks per frame</td>
<td>64</td>
</tr>
<tr>
<td>number of data subcarriers</td>
<td>64</td>
</tr>
<tr>
<td>wireless channel</td>
<td>Rayleigh fading channel</td>
</tr>
<tr>
<td>video resolution</td>
<td>352 × 288</td>
</tr>
<tr>
<td>video frame rate</td>
<td>30 frames/s</td>
</tr>
<tr>
<td>number of frames per GoP</td>
<td>32</td>
</tr>
<tr>
<td>ratio of reserved chunks</td>
<td>75%</td>
</tr>
<tr>
<td>pathloss factor $\alpha$</td>
<td>4</td>
</tr>
<tr>
<td>Gaussian noise</td>
<td>$-90$ dBm</td>
</tr>
<tr>
<td>transmission power</td>
<td>5 dBm</td>
</tr>
</tbody>
</table>

A total of 32 video frames are grouped into a GOP. DCT (3D-DCT for SoftCast and temporal DCT after 2D-DWT for WSVC) is applied on each GOP. Coefficients of each frame are divided into 64 chunks. An OFDM-based PHY layer with 64 subcarriers is adopted. We assume that the video signal experiences Rayleigh fading and path loss during transmission. For Rayleigh fading, ZMCSCG random variables with unit variance are adopted as fading coefficients. For path loss, when a user is located at a distance $r_k$ from the SAP, the received signal power is attenuated by $r_k^{-\alpha}$, where $\alpha$ is the path loss factor and is set to 4 in our simulations.

Different schemes are compared using a standard metric PSNR, which is defined as

$$PSNR = 10 \log_{10} \left( \frac{2L - 1}{MSE} \right)^2,$$

where $L$ is the number of bits used to encode each pixel and is typically 8 bits.

A. Unicast Performance Comparison

Fig. 4 shows the performance comparison of video unicast. ECast has a 4 dB - 6 dB gain compared to SoftCast with Hadamard transform in a large SNR range from 0 dB to 30 dB. Hadamard transform provides a 2 dB - 3 dB precoding gain in low SNR ranges in fading channels. However, our proposal outperforms Hadamard transform. When SNR tends to be higher, the distortion caused by wireless transmission errors is smaller and chunk discarding in bandwidth adaption becomes the dominant reason for performance loss. As a result, the gain tends to be small in high SNR ranges. ECast without subcarrier
matching assigns chunks to subcarriers randomly but calculate the scaling factors according to the optimal power allocation scheme considering fading. The performance gain is about 1 dB in low SNR ranges and 2 dB - 3 dB in high SNR ranges. From these curves, we obtain that subcarrier matching is the main reason for the performance enhancement. It dominates the proposed joint subcarrier matching and power allocation scheme.

By adopting the proposed method in the analog part of WSVC, the performance is improved by 2 dB in average. Given that the performance of WSVC is determined by both the analog coding and digital coding parts, the improvement is a slightly smaller. The enhanced WSVC outperforms pure analog schemes in low SNR ranges, because some important components are protected well by the digital coding part. However, with the increasing SNR, the performance increases slowly in PSNR. This is because the data generated by the digital part are over-protected in high SNR ranges. Non-scalability is inevitable because the digital part of WSVC adopts traditional video coding and wireless transmission schemes which are not scalable.

B. Mobility test

The performance is evaluated under user movement. We run a simple unicast experiment with one SAP and one moving user to compare our proposal against SoftCast. With the help of Hadamard transform which redistributes the importance of video coefficients, SoftCast performs well in user movement scenarios. By contrast, the mobility adaption in our scheme has a similar effect to that of Hadamard transform. We assume that each subcarrier carries only one chunk within a time slot, and the SNR varies uniformly from 12 dB to 6 dB and changes at every time slot.

As depicted in Fig. 5 the PSNR fluctuation of SoftCast without Hadamard is more violent than ECast and SoftCast with Hadamard. Both ECast and SoftCast with Hadamard have a similar fluctuation, but ECast outperforms SoftCast in PSNR.

In addition, the performance of different schemes is also compared numerically. As introduced in [20], temporal video quality fluctuation is perceivable only if the change exceeds a certain threshold. In our simulation, if the PSNR change between two consecutive frames exceeds 0.5 dB, it is considered as a visible change[21]. A total of 720 frames are simulated.

The results are shown in Table III. The number of perceivable changes for SoftCast with Hadamard, SoftCast without Hadamard and ECast are 17, 69 and 18, respectively. ECast and SoftCast with Hadamard have a similar PSNR fluctuation. Both of them outperform SoftCast without Hadamard in terms of quality fluctuation.

C. Inaccurate feedback

Fig. 6 shows the performance degradation caused by inaccurate CSI. The inaccuracy is involved in both the channel estimation and the channel feeding processes. We assume that the minimum mean square error estimator is adopted to estimate channel states. The estimation errors can be modeled as i.i.d. ZMCSGC variables[22]. The noises in feedback tones are also i.i.d. ZMCSGC variables. They are independent from each other. As a result, a combined Gaussian noise can be used to model the inaccuracy of each feedback tone. The AP receiving inaccurate CSI is equivalent to receiving accurate CSI plus the Gaussian noise. Fig. 6 shows the performance of inaccurate CSI feedback under different noise powers. We plot three curves with SNR 5 dB, 8 dB and 11 dB, respectively. The estimation error ratio is the variance ratio of the estimation error and the Rayleigh fading coefficient. In all cases, with the ratio increasing, the PSNR of the system degrades smoothly. No sharp drops of these performance curves have been found. Therefore, the analog feedback scheme is not sensitive to noise.

D. Multicast Performance Comparison

The performance of receivers at various distances from the SAP is compared. We assume that the SAP serves a group of three receivers. The distance between the $k^{th}$ receiver and the SAP is $r_k$, where $r_1 = 100$ m, $r_2 = 134$ m and $r_3 = 167$ m. Due to large scale fading, the three users receive different average signal powers. The transmitting power target is set to 5 dB. The received power for the $k^{th}$ user is approximately $5 - 10 \log(r_k^2)$ dBm. We further set the noise and interference power to $-90$ dBm. Then, the corresponding SNRs for the users are 15 dB, 10 dB and 5 dB, respectively.

Simulation results are given in Fig. 7. The ECast (unicast) means that only one user exists in the multicast cell, so the SAP transmits video signals only according to this user’s channel conditions. Obviously, the user obtains best performance improvement. It provides an upper bound of performance improvement for users in a multicast cell. In addition, we assign the weights of the three users to 0.01, 0.1 and 0.89 from near to far, respectively. Users that are remote from the

![Fig. 5. Performance variation under large mobility. The SNR varies uniformly from 6 dB to 12 dB.](image-url)
SAP obtain larger performance improvement, whereas those near the SAP obtain less or even negative gains. As a result, the scheme provides an effective way to guarantee fairness by assigning large weight to users at the cell edge.

E. Effect of importance factors

The effect of importance factors is evaluated. Two users with SNRs 11 dB and 17 dB are simulated. Let $0 < \omega < 1$ denote the weight of user 1 and $1 - \omega$ the weight of user 2. As shown in Fig. 8 with $\omega$ increasing, the PSNR of the first user increases rapidly between 0 and 0.1 and gradually between 0.1 and 1. At the same time, the performance curve of user 2 declines gently between 0 and 0.9 and steeply between 0.9 and 1. As a result, assigning an importance factor larger than 0.1 is suggested; otherwise, the performance may decrease largely. A crossover point exists when $\omega$ approaches 1. This finding implies that the video quality is affected by the weight, and a user with high weight but low SNR can have better performance than that with high SNR but low weight.

F. Performance changing with an increasing number of users

We assign the same weight to each user in this simulation and evaluate how a user’s performance changes with the number of users. The SNR is set to 11 dB. As shown in Fig. 9 with the increasing number of users, the performance of ECast decreases. If the number of users is large enough and the same weight is assigned to each user, then ECast will perform similarly as SoftCast and all virtual channel gains obtained by the SAP will be nearly equal to 1.

VII. Conclusion

In this paper, a cross-layer design for wireless video multicast systems is proposed. It is an enhanced analog video coding scheme. We take advantage of the SoftCast framework and focus on the joint subcarrier matching and power allocation method. With analytical argument, a low-complexity and optimal scheme is provided. Channel adaption is considered to make the scheme applicable and robust in real-time wireless video systems. Finally, a novel analog method is proposed to limit the overhead of channel feedback. Simulation results show that the proposed approach outperforms SoftCast and WSVC, and provides flexible enhancement for different users. Therefore, the proposed approach is suitable for wireless video multicast applications.

REFERENCES


Zhilong Zhang received the B.E. degree in communication engineering from the University of Science and Technology, Beijing, China in 2007, and the M.S. degree in communication and information systems from Beijing University of Posts and Telecommunications, Beijing, China in 2010. He is currently pursuing his Ph.D. degree at Beijing University of Posts and Telecommunications. During 2010 to 2012, he was a software engineer in TD Tech Ltd., Beijing, China. During 2014 to 2015, he was a visiting Ph.D student at Stony Brook University, NY, USA. His research interests include optimization theory and its applications in wireless video transmission, cross-layer design and wireless networks.

Danpu Liu received the Ph.D. degree in communication and electrical systems from Beijing University of Posts and Telecommunications, Beijing, China in 1998. She was a visiting scholar at City University of Hong Kong in 2002, University of Manchester in 2005, and Georgia Institute of Technology in 2014. She is currently working at the Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing, China. Her research involved MIMO, OFDM as well as broadband wireless access systems. She has published over 100 papers and 3 teaching books, and submitted 26 patent applications. Her recent research interests include 60GHz mmWave communication, wireless high definition video transmission and wireless sensor network.

Xiaoli Ma received the B.S. degree in automatic control from Tsinghua University, Beijing, China in 1998, the M.S. degree in electrical engineering from the University of Virginia, Charlottesville, in 2000, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, in 2003. From 2003 to 2005, she was an Assistant Professor of Electrical and Computer Engineering at Auburn University. Since 2006, she has been with the School of Electrical and Computer Engineering at Georgia Institute of Technology, Atlanta, where she is currently an Associate Professor. Her research interests include transceiver designs and diversity techniques for wireless fading channels, cooperative communications, UWB communications, synchronization for multi-carrier systems, and channel modeling, estimation, and equalization for wireless systems.
Xin Wang received the B.S. and M.S. degrees in telecommunications engineering and wireless communications engineering respectively from Beijing University of Posts and Telecommunications, Beijing, China, and the Ph.D. degree in electrical and computer engineering from Columbia University, New York, NY. She is currently an Associate Professor in the Department of Electrical and Computer Engineering of the State University of New York at Stony Brook, Stony Brook, NY. Before joining Stony Brook, she was a Member of Technical Staff in the area of mobile and wireless networking at Bell Labs Research, Lucent Technologies, New Jersey, and an Assistant Professor in the Department of Computer Science and Engineering of the State University of New York at Buffalo, Buffalo, NY. Her research interests include algorithm and protocol design in wireless networks and communications, mobile and distributed computing, as well as networked sensing and detection. She has served in executive committee and technical committee of numerous conferences and funding review panels, and is the referee for many technical journals. Dr. Wang achieved the NSF career award in 2005, and ONR challenge award in 2010.