Joint Carrier Matching and Power Allocation for Wireless Video with General Distortion Measure

Zhilong Zhang, Danpu Liu, and Xin Wang

Abstract—In this paper, we present a cross-layer design for a family of OFDM-based video communications by jointly considering application layer information and the wireless channel conditions. Compared with traditional cross-layer designs, our proposed method targets to efficiently transmit video data generated with some emerging techniques for better wireless transmissions, where the video data are divided into multiple chunks and each chunk contributes independent distortion to the entire video quality. To minimize the end-to-end distortion, we formulate a generalized optimization problem and derive a joint optimal carrier matching and power allocation scheme. Rather than depending on specific video encoding method as done in the conventional work, we intend our design to be applicable to a general family of new video schemes. We apply our proposed method to two applications, the enhanced analog coding and the uncompressed video transmission over OFDM. In both applications, the performance can be improved by adopting our scheme. Simulation results validate the effectiveness of our approach in achieving significantly better PSNR and visual quality compared to reference schemes.

Index Terms—Analog video coding, Carrier matching, Cross-layer design, Power allocation, Uncompressed video transmission.

1 Introduction

Orthogonal frequency division multiplexing (OFDM) has been widely adopted in wireless communication systems due to its high frequency efficiency, robustness against wireless fading and easy implementation based on fast fourier transform algorithms [1]. On the other hand, with the dramatic increase of the demand for video services, video data have dominated the wireless traffic and will keep growing rapidly in the next few years [2]. In light of these two trends, it is of importance to optimize the performance of wireless video transmissions over OFDM.

Generally, the encoded video data are divided into multiple components before transmission, and transmission errors in different components result in unequal distortion to the overall video quality. Intuitively, for better video transmission quality, unequal protection (UEP) can be taken, so important components will receive a higher level of protection. In an OFDM system, the available bandwidth is divided into subcarriers, which generally experience different attenuation as a result of frequency selective fading. To improve the quality of experience for wireless video transmission over OFDM, a cross-layer design of the video coding layer and the OFDM-based physical layer is usually considered, where wireless resources such as power and carriers are allocated according to both video contents and channel conditions.

A number of studies have been carried out to deal with the problems of resource allocation for different types of video applications over OFDM. Since video data encoded by different coding methods have their unique structures, different schemes have to be derived for transmitting different types of video sources. For instance, there are multiple video coding layers in Scalable Video Coding (SVC) [3–5]. The data from the base layer is more critical than those from the enhancement layers and should be allocated with more wireless resources. In MPEG-4 [6], I frames (reference frames) contain necessary information for decoding B and P frames, hence are highly protected during the wireless transmission. Besides, many cross-layer studies have been made on popular video or image schemes, including D-VB [7], video streaming over LTE [8], JPEG [9], JPEG2000 [10, 11] and DWT-based video data [12, 13].

Recently, several novel cross-layer designs for wireless video transmission are proposed. To make the video quality totally scalable to channel conditions, analog video coding schemes [14–21] are derived. To reduce the transmission delay and improve the video quality, uncompressed video signals are transmitted directly over high frequency bands [22–27]. Unlike traditional cross-layer designs which usually aim at adapting existing video coding standards to characteristics of physical layer, video encoding schemes in these systems are redesigned. For example, in SoftCast [21], video data are compressed by 3D-DCT and DCT coefficients are transmitted through raw OFDM; in WirelessHD [22], the video bits are divided into most significant bits (MSBs) and less significant bits (LSBs) and transmitted directly without compression. Neither reference encoding nor complex nonlinear processes are included in the aforementioned systems, which leads to benefits such as no error propagation, low latency and high quality scalability. Although the compressing efficiency is somewhat reduced, the cost is affordable in many wireless communication systems.

In these novel systems, the coded video data can be

- *This work was supported by the National Natural Science Foundation of China under Grant 61171107 and Grant 61271257. The work of X. Wang was supported in part by the NSF under Grant 1247924.*
- *Z. Zhang and D. Liu are with Beijing Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: jerome.zl.zhang@gmail.com, dpliu@bupt.edu.cn).*
- *X. Wang is with the Department of Electrical and Computer Engineering, Stony Brook University, NY 11790, USA (e-mail: xwang@ece.sunysb.edu).*

divided into independent chunks and the end-to-end distortion can be expressed by a weighted sum of each chunk's individual distortion, which we call *linear video distortion structure* in this paper, as will be defined in Definition 1. Given the benefits of using these video schemes for wireless communications, it would be of great interest to optimally transmit video data over OFDM with the concurrent consideration of carrier matching and power allocation. Despite some existing efforts, the algorithms designed are often heuristic and far from optimal. In addition, there is a lack of a general mathematical framework to model the wireless resource allocation problem for these video applications.

In this paper, we aim to derive a generalized joint power and carrier allocation scheme to optimally transmit the family of video data with linear distortion structure. To the best of our knowledge, no existing efforts consider such a problem without relying on specific encoding model. It is usually hard to obtain an optimal solution for a traditional cross-layer design over video with highly complex encoding. In contrast, as long as a video encoder has the aforementioned linear distortion structure, our proposed scheme is applicable and can be proven to be optimal in terms of minimizing the total distortion.

Our main contributions are twofold

- We provide a generalized cross-layer design to enable efficient communications for the family of video with linear distortion structure. We propose a joint carrier matching and power allocation scheme, and prove that it can minimize the total distortion. The optimality does not depend on any specific measures, and our scheme is general and applicable to any video applications with linear distortion structure and non-increasing distortion functions.
- We introduce two real-time[1] video transmission systems as example applications to demonstrate the effectiveness and generality of our proposed schemes: an enhanced analog video coding scheme and an uncompressed video transmission method. Our proposed joint carrier matching and power allocation can be applied to improve the performance of both.

The rest of the paper is organized as follows. We introduce the related work in Section 2, and then provide the background and motivation of our work in Section 3. In Section 4, we present our system model and problem formulation. In Section 5, we first describe our proposed method with theoretical analyses, and then derive an optimal solution to minimize the total distortion and prove its optimality. In Section 6, two special examples of our proposed scheme are given. We evaluate the performance of our design in section 7, and conclude the work in Section 8.

Notations of symbols used in the paper are summarized in Table 1.

2 Related Work

Recent years have witnessed much interest in the cross-layer design of video transmission over OFDM. The transmission of video or image data with traditional coding formats has been widely studied, including but not limited to SVC [3–5], MPEG-4 [6], JPEG[9], JPEG2000 [10, 11] and generalized DWT-based video data [12, 13]. Given that different parts of video data have unequal importance, wireless resource should be allocated according to not only channel conditions but also video contents. The importance of video data is determined by the characteristic of video encoder. For example, for SVC, the base layer is more important than enhanced layers and more protection is provided to the former [4]; for MPEG-4 [6], packets containing I frames are important than those containing B and P frames, and channels are allocated according to frame importance and channel status. As different problems need to be solved when transmitting video data with different formats, usually, it is hard to obtain the unified optimal wireless resource allocation.

In light of the problem of conventional cross-layer design for transmissions of video with traditional video encoders, we consider transmitting video data encoded by a family of recently emerging video encoders which have a linear distortion structure. Two categories of video transmissions fall into this family, analog video coding and uncompressed video transmissions.

Analog video coding is a novel video transmission framework. Compared with conventional cross-layer designs, the scheme has graceful performance due to the linear structure of video coding layer and physical layer. The key idea underlying analog video coding is to ensure that distortions of transmitted symbols are linearly related to distortions of reconstructed pixel values. SoftCast as the basic analog video coding framework is proposed in [21, 28, 29], where 3D-DCT and raw OFDM are adopted, and the entropy coding part in conventional video encoders

TABLE 1
Symbol Notation

Symbols	**Descriptions**
$f(x)$	Distortion function
a_j, h_j	Coefficient and gain of channel j
λ_i	Importance factor of chunk i
p_j	Transmit power of channel j
b_{ij}	Matching indicator for chunk i and channel j
c_i	Arrangement of channel gains
N	Number of available carriers
N_c	Number of chunks to be transmitted
N_b	Number of bits per pixel
σ^2	Thermal noise power
P_{tot}	System transmit power constraint
d_i	Distortion of chunk i
D	Distortion of the entire video signal
γ	Signal to noise ratio
$e_{ij}^{LLSE}, e_{ij}^{ZF}$	Distortion of chunk i carried by carrier j under LLSE scaling, ZF scaling (application 1)
M	Modulation level (application 2)
N_s	Number of sub-streams per bit stream (application 2)

1. "Real-time" means the video data are lively generated and encoded at the transmitter, then decoded and viewed at the receiver.

is absent. SoftCast makes video quality gratefully scale with the channel quality and eliminates the "cliff effect". After that, many efforts have been devoted to improving the performance [14–20]. In [18], 3D-DCT is replaced by 3D wavelet transform to achieve better compressing performance. Motion estimation is introduced in [16] and improved by [17]. A chunk division method adapting to the energy distribution of transmitted coefficients is proposed in [19]. A hybrid scheme integrating the advantages of traditional digital coding and the novel analog coding called WSVC is proposed in [14]. The aforementioned results have provided various strategies to improve the performance. However, with OFDM sub-carriers entirely modeled as Gaussian channels, these schemes may suffer in a practical environment where signals are often subject to fading. ParCast(+) [15, 30] attempts to map the important video components to more reliable channels at the transmitter in order to achieve the graceful performance, and the optimality of the mapping scheme is guaranteed by the rearrangement inequality. We will show that the subchannel mapping and power scaling in ParCast(+) is a special case of our proposed scheme. Moreover, our proposed method can instantiate a more accurate power scaling method which can improve the performance in the low SNR range, where the optimality of the matching cannot be easily proven by using the rearrangement inequality.

Uncompressed video transmissions have been discussed recently[31]. The transmissions are characterized as low latency and high quality, and are usually applied in the short-range wireless communications where the available bandwidth is huge. The complexity and the delay are reduced at the cost of the large wireless bandwidth. Generally, video data are divided into multiple bit streams with different levels of significance, and unequal wireless resources are allocated to the streams. Most of the prior works achieve unequal protection by using adaptive modulation and coding at the physical layer [24–27]. More forward error correction symbols are assigned to bits of greater importance. Although such proposals provide a graceful performance, the latency is hard to be guaranteed, because time-consuming processes such as interleaving are usually included in these schemes. In this paper, to support unequal power allocation (UPA), higher power values are allocated to bits of greater importance [32, 33], thus the decoder has a high probability of recovering the more important bits. Video data without compression are usually delivered over 60GHz band [34] and ultra-wide band (UWB) [35], both of which can be based on OFDM. Considering the characteristics of OFDM systems, we propose a joint carrier matching and unequal power allocation scheme in the second application. To our best knowledge, no proposals have studied the optimal joint carrier matching and power allocation for uncompressed video transmissions over OFDM.

3 Background and Motivation

One aim of traditional video encoding design is to achieve high compressing efficiency. Usually, video encoders have a complex coding structure including a transform encoding part and an entropy encoding part, such as H.264 and MPEG-2, as shown in Fig. 1. The transform encoding part compresses video data in a transform domain. Subsequently,

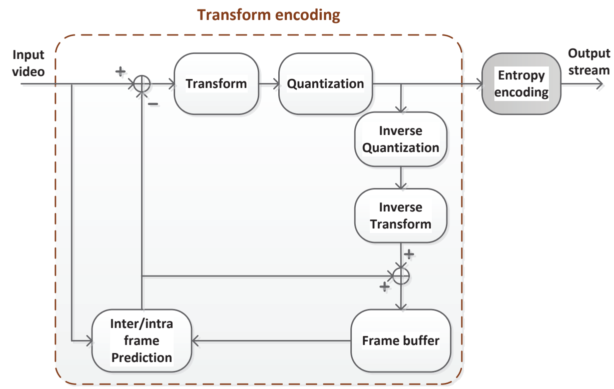


Fig. 1. Structure of traditional video encoder.

the entropy encoding part further compresses the transform coefficients using entropy coding schemes, such as Huffman coding. The entropy coding is often non-linear. If one bit is not decoded correctly, all coefficients may not be reconstructed successfully. Therefore, such kind of encoded video data are susceptible to errors in wireless transmissions. Recently, several novel video coding structures are proposed to make video signals better adaptable to different types of wireless transmissions. To make the video quality more scalable, analog coding schemes [14–20] get rid of the entropy encoding part to ensure the linear relationship between the transform coefficients and the transmitted signals. Besides, uncompressed video transmission [22–27] even eliminates the transform encoding. It only divides pixels into multiple streams and transmits the streams directly without compression.

These simplified video encoders allow more accurate calculation of the distortion of video quality, and provide a chance to derive an optimal cross-layer design, which is difficult for wireless transmissions of traditional encoded video data. Taking advantage of the features of these wireless-friendly video encoders, we can build more accurate models to efficiently transmit video data over wireless channels. We will present the applications of our design in Section 6.

To transmit such video data over an OFDM-based channel with the power and frequency resources limited, an issue to address is how to allocate power and subcarrier to video chunks. As shown in Fig. 2, in the video layer, a sequence of frames are encoded into a data stream and divided into chunks of various importance. In the physical layer, signals carried by different carriers can experience different attenuation due to the frequency selected fading and the channel gains fluctuate in different time slots. In the example, three typical chunks i, j, k are matched to carriers q, n, m, and there is also a need to determine the transmission power for each subcarrier. There are various ways to allocate subcarrier and power. For example, for a chunk assigned to a subcarrier with a good channel condition, we can allocate the chunk with relatively low transmission power or relatively high transmission power. Which strategy to use? Since wireless resources are limited, all chunks compete for wireless resources to reduce their distortion during the transmission. Intuitively, important chunks should get more resources. However, too large a distortion of less important chunks also influences the end-

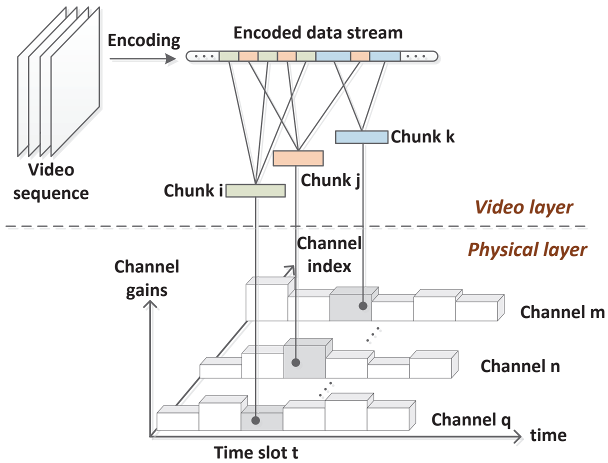


Fig. 2. Illustration of the considered video transmission scenario.

to-end video quality significantly. The questions are: Which resource matching order is better? How to make the carrier matching and the power allocation to ensure the optimal performance? Is there a simple but efficient scheme for real-time applications? Since there is a lack of such studies in previous efforts, in this paper, we intend to address these questions.

4 System Model and Problem Formulation

We consider a generalized cross-layer design for a family of OFDM-based video communications. In this section, we first present the models for the video layer and the OFDM-based transmission layer, then formulate a mixed binary optimization problem to minimize the total distortion.

4.1 Video application models

We are interested in a series of video encoding technologies. Like traditional schemes, the encoded video data are divided into multiple chunks with unequal amount of importance. Transmission errors occurring in different chunks lead to different distortion to the entire video signal. However, compared with traditional schemes, the encoded video data generated from our considered applications have a linear distortion structure as defined below.

Definition 1. *The video data are considered to have a linear distortion structure if different chunks of the data are independent and the total video distortion is separable.*

The two major characteristics in this definition are explained in the following.

First, the chunks are independent. Each chunk contributes independent distortion to the entire video quality. The distortion of a chunk caused by transmission errors does not affect any other chunks. If a chunk cannot be decoded correctly because some information is lost when transmitting another chunk, the two chunks are not considered as independent. Most traditional video codecs, such as H.264, do not have this characteristic. For example, in H.264, if an I frame in a GOP is failed to be decoded, its corresponding P and B frames cannot be reconstructed successfully. However, some recently emerging methods, such as analog coding schemes (see Section 7.1) and uncompressed video (see Section 7.2), have such a characteristic.

Second, the distortion is separable. The quality degradation of a video is caused by all chunks' distortion. The entire distortion of a video can be calculated by a weighted sum of each chunk's individual distortion. The end-to-end total distortion can be expressed by

$$D = \sum_{i=1}^{N_c} \lambda_i d_i, \tag{1}$$

where N_c denotes the number of chunks, d_i is the distortion of the ith chunk and λ_i denotes the weighting factor or significance factor of the chunk. Generally, the significance factor reflects the contribution of each chunk to the end-to-end distortion and can be estimated numerically or determined based on the coding design.

4.2 System and channel models

A generalized OFDM-based video communication system considered in this paper is depicted in Fig.3. An input video sequence is encoded and divided into multiple chunks. Assume time is slotted and channel status does not change within a time slot. Only one chunk is transmitted over a subcarrier in a time slot. The transmit power of the chunk carried by carrier j is denoted by p_j. The total power constraint is given as

$$\sum_{j=1}^{N} p_j \le P_{tot}, \tag{2}$$

where P_{tot} denotes the total transmit power for the system and N is the number of available carriers. In the OFDM modulator, chunks are modulated on orthogonal channels to form a transmit signal. The signal of each carrier experiences frequency-nonselective and slow fading. Correspondingly, the output of the demodulator for the jth carrier is expressed as

$$y_j = \sqrt{p_j} a_j x_j + n_j, \tag{3}$$

where a_j denotes the fading coefficient for the jth carrier which can be estimated. n_j represents the Gaussian additive noise with zero mean and variance σ^2. x_j is a typical symbol of the chunk carried by carrier j.

Different chunks may suffer from different distortion during wireless transmission. If the modulation levels and the channel coding schemes are the same for the parallel channels (e.g., modulation in WiFi for a burst or in LTE for a single user), the distortion of a specific chunk is determined only by its fading coefficient and transmit power. When a chunk is assigned to the jth carrier, the distortion of this chunk is expressed by

$$d_j = f(h_j p_j), \tag{4}$$

where $h_j = |a_j|^2$ denotes the channel gain and $f(x)$ is the unified distortion function. Given that more wireless resources generally lead to low distortion, $f(x)$ is generally non-increasing. Obviously, for a specific chunk, larger h_j and p_j result in smaller distortion. By allocating different carriers and different amount of power to the chunks with various importance, unequal protection is provided.

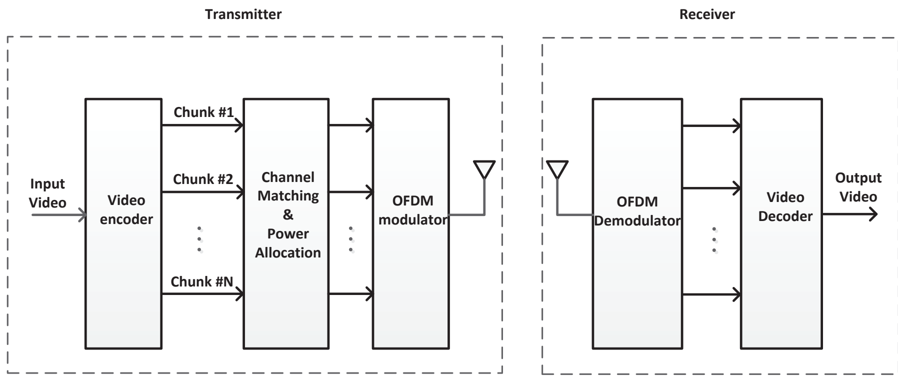


Fig. 3. A generalized framework of the considered OFDM-based video transmission system.

4.3 Problem formulation

For the convenience of presentation, we assume that all chunks have the same size, but our scheme does not depend on the chunk sizes. We consider one-to-one matching where the number of chunks transmitted in a time slot equals that of the carriers. The scheme can be easily extended to the case of one-to-many or many-to-one matching. If the number of carriers is larger, the carriers with good conditions will be selected to transmit video data. If the number of chunks is larger, assuming that m chunks are carried by one carrier in one time slot, then we duplicate each channel gain m times before matching.

In a time slot, N chunks are transmitted over N carriers, where $2 \leq N \leq \infty$. The problem is how to allocate power to a specific chunk and which carrier is used to carry it. The two types of resource should be jointly determined according to channel fading and chunk importance. According to (1) and (4), the total distortion is expressed by

$$D = \sum_{i=1}^{N}\sum_{j=1}^{N} b_{ij}\lambda_i f(h_j p_j), \tag{5}$$

where b_{ij} is a binary variable indicating whether the ith chunk is allocated to the jth carrier. $b_{ij} = 1$ means the ith chunk is allocated to the jth carrier; $b_{ij} = 0$ means the opposite. Intuitively, to minimize the total distortion, important chunks should be assigned to carriers with high channel gains and transmitted with more power. However, as discussed earlier, it may lead to the overall higher degradation. The goal of our work is to match the chunks with carriers and determine the amount of power allocated to each chunk. According to (2) and (5), our optimization problem is formulated as follows

$$\min_{\{b_{ij}\},\{p_j\}} \quad \sum_{i=1}^{N}\sum_{j=1}^{N} b_{ij}\lambda_i f(h_j p_j), \tag{6}$$

$$\text{subject to} \quad \sum_{i=1}^{N} p_i \leq P_{tot}, p_i \geq 0, \tag{7}$$

$$\sum_{j=1}^{N} b_{ij} = 1, \sum_{i=1}^{N} b_{ij} = 1, b_{ij} \in \{0,1\}, \tag{8}$$

where $\{b_{ij}\}$ determines the carrier matching order and $\{p_j\}$ determines the allocated power.

5 Optimal Joint Carrier Matching and Power Allocation

In this section, we propose a scheme for optimal joint carrier matching and power allocation. To solve the problem in (6), we first propose an optimal matching approach to reduce the original problem into a simpler form, and then solve the new problem. The optimality of our proposed matching approach is first proved for systems with two chunks and two carriers, and then is extended for systems with N chunks and N carriers.

5.1 A brute-force approach

The problem in (6) is a mixed integer programming problem and the global optimal solution is hard to be obtained directly. We first consider how to simplify the problem.

Given that each subcarrier carries only one chunk, i.e. only one-to-one matching is permitted, if $\{p_j\}$ is fixed, the problem is reduced to a typical assignment problem, which can be solved by adopting the Hungarian or Kuhn-Munkres algorithm [36, 37]. However, as the transmit power cannot be jointly optimized together with the matching, only a suboptimal solution can be obtained. Therefore, this method is not in our options in this paper.

When $\{b_{ij}\}$ is fixed, the problem is reduced into the following simplified form

$$\begin{aligned} \min_{\{p_j\}} \quad & \sum_{j=1}^{N} c_j f(h_j p_j), \\ \text{subject to} \quad & \sum_{j=1}^{N} p_j \leq P_{tot}, p_j \geq 0, \end{aligned} \tag{9}$$

where $\{c_j\}$ is a rearrangement of $\{\lambda_i\}$ determined by the carrier matching scheme. Obviously, the problem (9) is easier to solve than the problem (6). An intuitive approach to obtain an optimal solution is given in Lemma 1.

Lemma 1. *By systematically enumerating all possible candidates of carrier matching, a group of problems with the form* (9) *is obtained. Among all candidate optimal solutions, the one with the lowest objective value is the optimal solution of* (6).

The method denoted by Lemma 1 is a brute-force approach and the optimality can be guaranteed. In principle, there are $N!$ possibilities to consider. Although the original complex problem is reduced into a group of simpler problems, the number of candidates is prohibitively large. Therefore, the method is unapplicable in a real-time communication system due to the computational complexity. However, this intuitive approach sheds some lights on deriving a more applicable method.

5.2 Optimal matching between two chunks and two carriers

To find a simple optimal matching method between two chunks and two carriers, let h_1 and h_2 denote the two channel gains, and λ_1 and λ_2 denote the importance of the two chunks. The system power constraint is P_{tot}. We further assume that $0 \le h_1 \le h_2$, $0 \le \lambda_1 \le \lambda_2$ and $P_{tot} > 0$.

According to Lemma 1, we first enumerate all possible candidates of carrier matching. The number of possible candidates is two. Obtaining the optimal solution of problem (6) is equivalent to solving the following two problems separately and choosing the solution with the lowest object value from the two optimal solutions.

$$\begin{aligned} \min_{p_1,p_2} \quad & D_1(p_1,p_2) = \lambda_1 f(h_1p_1) + \lambda_2 f(h_2p_2), \\ \text{subject to} \quad & p_1 + p_2 \le P_{tot}, \\ & p_1 \ge 0, p_2 \ge 0, \end{aligned} \tag{10}$$

$$\begin{aligned} \min_{p_1,p_2} \quad & D_2(p_1,p_2) = \lambda_1 f(h_2p_1) + \lambda_2 f(h_1p_2), \\ \text{subject to} \quad & p_1 + p_2 \le P_{tot}, \\ & p_1 \ge 0, p_2 \ge 0. \end{aligned} \tag{11}$$

Note that each problem indicates a carrier matching scheme. Equation (10) denotes that the chunk with small significance λ_1 is matched to the channel with small gain h_1, and accordingly λ_2 to h_2. Problem (11) denotes another matching that λ_1 is matched to h_2 and λ_2 to h_1.

Obviously, (10) and (11) have the same feasible region. The feasible solutions of the two problems have the following characteristic denoted by Lemma 2, which is simple but useful in the proof of our proposal.

Lemma 2. *Given that* (p_1, p_2) *is a feasible solution of* (11), $(\alpha p_1 + (1-\beta)p_2, (1-\alpha)p_1 + \beta p_2)$ *is a feasible solution of* (10), *where* $0 \le \alpha, \beta \le 1$.

Proof. Since (p_1, p_2) is a feasible solution of (11), we obtain that $p_1 \ge 0, p_2 \ge 0$ and $p_1 + p_2 \le P_{tot}$. Therefore,

$$\alpha p_1 + (1-\beta)p_2 + (1-\alpha)p_1 + \beta p_2 = p_1 + p_2 \le P_{tot}.$$

In addition, given that $0 \le \alpha, \beta \le 1$, we obtain

$$0 \le \alpha p_1 + (1-\beta)p_2 \le 1,$$
$$0 \le (1-\alpha)p_1 + \beta p_2 \le 1.$$

Note that the constrains of (10) are all satisfied. Therefore, $(\alpha p_1 + (1-\beta)p_2, (1-\alpha)p_1 + \beta p_2)$ is a feasible solution of (10). □

Now, we propose the optimal matching method in Lemma 3.

Lemma 3. *For the system with only two carriers and two chunks, given that the unified distortion function* $f(x)$ *is non-increasing, the optimal carrier matching is: arrange both the* $\{\lambda_i\}$ *and* $\{h_j\}$ *in a decreasing order, respectively, and match chunks to carriers one by one according to this order.*

Proof. The matching method denoted by (10) is actually the method proposed in Lemma 3. Therefore, what we need to prove is that (10) always has a lower optimal distortion compared with (11), whatever the constants λ_1, λ_2, h_1 and h_2 are equal to. Assume the optimal solutions of (10) and (11) are (p_1^*, p_2^*) and (p_1', p_2'), respectively. We divide the whole proof into the following two cases. We will prove that $D_1(p_1^*, p_2^*) \le D_2(p_1', p_2')$ is always satisfied.

Case 1: $h_1p_2' \le h_2p_1'$
According to Lemma 2, by setting $\alpha = \beta = 0$, we obtain a feasible solution (p_2', p_1') of (10). The difference of $D_1(p_2', p_1')$ and $D_2(p_1', p_2')$ is calculated by

$$\begin{aligned} & D_1(p_2', p_1') - D_2(p_1', p_2') \\ & = \lambda_1 f(h_1p_2') + \lambda_2 f(h_2p_1') - \lambda_1 f(h_2p_1') - \lambda_2 f(h_1p_2') \\ & = (\lambda_1 - \lambda_2)\left[f(h_1p_2') - f(h_2p_1')\right]. \end{aligned} \tag{12}$$

Given that $\lambda_1 \le \lambda_2$, $h_1p_2' \le h_2p_1'$ and $f(x)$ is non-increasing, we obtain that $D_1(p_2', p_1') - D_2(p_1', p_2') \le 0$. Since the optimal solution (p_1^*, p_2^*) results in an even lower distortion, i.e. $D_1(p_1^*, p_2^*) \le D_1(p_2', p_1')$, we finally obtain that $D_1(p_1^*, p_2^*) \le D_2(p_1', p_2')$.

Case 2: $h_1p_2' > h_2p_1'$
Given that $0 < h_1 \le h_2$, it is obtained that $0 \le \frac{h_1}{h_2} \le 1$. According to Lemma 2, by setting $\alpha = 1$ and $\beta = \frac{h_1}{h_2}$, we obtain a feasible solution $(\hat{p_1}, \hat{p_2})$ of (10), where $\hat{p_1} = p_1' + \left(1 - \frac{h_1}{h_2}\right)p_2'$ and $\hat{p_2} = \frac{h_1}{h_2}p_2'$. The difference of $D_1(\hat{p_1}, \hat{p_2})$ and $D_2(p_1', p_2')$ is calculated by

$$\begin{aligned} & D_1(\hat{p_1}, \hat{p_2}) - D_2(p_1', p_2') \\ & = \lambda_1 f\left(h_1p_1' + \left(1 - \frac{h_1}{h_2}\right)h_1p_2'\right) + \lambda_2 f(h_2\frac{h_1}{h_2}p_2') \\ & \quad - \lambda_1 f(h_2p_1') - \lambda_2 f(h_1p_2') \\ & = \lambda_1\left[f\underbrace{\left(h_1p_1' + \left(1 - \frac{h_1}{h_2}\right)h_1p_2'\right)}_{A} - f(\underbrace{h_2p_1'}_{B})\right]. \end{aligned} \tag{13}$$

The difference of A and B in (13) is calculated by

$$\begin{aligned} A - B \quad & = h_1p_1' + \left(1 - \frac{h_1}{h_2}\right)h_1p_2' - h_2p_1' \\ & = \frac{h_2 - h_1}{h_2}(h_1p_2' - h_2p_1') \overset{(a)}{\ge} 0, \end{aligned} \tag{14}$$

where (a) is obtained due to the assumptions that $0 < h_1 \le h_2$ and $h_1p_2' > h_2p_1'$. Given that $f(x)$ is non-increasing, from (13) and (14) we obtain that $D_1(\hat{p_1}, \hat{p_2}) \le D_2(p_1', p_2')$. Since the optimal solution (p_1^*, p_2^*) results in an even lower distortion, i.e. $D_1(p_1^*, p_2^*) \le D_1(\hat{p_1}, \hat{p_2})$, we finally establish the relationship that $D_1(p_1^*, p_2^*) \le D_2(p_1', p_2')$.

In both of the above two cases, we have proved that the matching method denoted by (10) outperforms that denoted by (11). As a result, the optimal matching scheme is the method denoted by (10). □

TABLE 2
Summary of Applications

Applications	Importance factor λ_i	Distortion function $f(x)$
Enhanced analog video coding	The variance of chunk i	$f(x) = \frac{\sigma^2}{x+\sigma^2}$, for LLSE scaling; $f(x) = \frac{\sigma^2}{x}$, for ZF scaling;
Uncompressed video transmission	$4^{\lfloor \frac{i}{N_s} \rfloor}$	$f(x) = \frac{2(\sqrt{M}-1)}{\sqrt{M} log_2 \sqrt{M}} Q\left(\sqrt{\frac{3x}{(M-1)\sigma^2}}\right)$, for Q-BER approximation; $f(x) = \frac{\sqrt{M}-1}{\sqrt{M} log_2 \sqrt{M}} \exp\left(-\frac{1.5x}{(M-1)\sigma^2}\right)$, for bound-BER approximation;

5.3 Optimal matching between N carriers and N chunks

This subsection extends Lemma 3 to the system with N carriers and N chunks. The sets of channel gains and chunk importance are denoted by $\{h_i\}$ and $\{\lambda_i\}$, respectively. It is assumed that the two sets are ordered beforehand, i.e. $h_i \le h_j$ and $\lambda_i \le \lambda_j$, if $i \le j$. The following proposition gives the optimal carrier matching.

Proposition 1. *For the system with N carriers and N chunks, given that the unified distortion function $f(x)$ is non-increasing, the optimal matching is to match the ith chunk to the ith carrier ($1 \le i \le N$), where $\{\lambda_i\}$ and $\{h_i\}$ are decreasingly arranged, respectively.*

Proof. This proposition is proven in a contrapositive form. For simplicity of notation, let λ_i denote the ith chunk, h_j denote the jth channel and $\lambda_i \sim h_j$ denote the ith chunk is matched to the jth channel. Suppose that there is an optimal carrier matching result which includes the two matched pairs that $\lambda_i \sim h_j$ and $\lambda_k \sim h_l$ where $\lambda_i < \lambda_k$ and $h_j > h_l$, and the distortion of (6) by adopting this method is lower than that by adopting Proposition 1.

We first divide the matching and power allocation result in the hypothesis into two groups. The first group includes the above two matched pairs of chunks and carriers; the second group includes the rest $N-2$ pairs of chunks and carriers. Note that the total distortion of the first group can be further reduced by adopting Lemma 3 which implies a better matching method by rematching the pairs as $\lambda_i \sim h_l$ and $\lambda_k \sim h_j$. Keeping the matching and power allocation result of the second group fixed, the total distortion of the two groups can be reduced by rematching the first group. Obviously, the method in the hypothesis is not optimal, which is contrary to the assumption. Therefore, the optimal matching result should not include such matching pairs. Since only the method denoted by Proposition 1 satisfies the requirement, this carrier matching is optimal for the system with N carriers and N chunks. □

In conclusion, for our considered system, the optimal matching is to match the carriers according to the orders of channel gains and chunk powers, i.e., $\lambda_i \sim h_i$.

5.4 Joint carrier matching and power allocation

With the optimal carrier matching method proposed in Proposition 1, we can simplify the problem (6) into the form of (9), where $\{c_j\}$ is the descending arrangement of $\{h_j\}$ and $\{\lambda_i\}$ has been also arranged in a decreasing order. As the optimal carrier matching is ensured to achieve the global optimum performance with the transmission power optimally allocated, they are jointly optimal.

According to the information theory [38], the rate-distortion functions for a variety of sources are convex, which has been demonstrated in some existing work [39]. In addition, the distortion functions in the two examples in Section 6 are all convex. Therefore, it is rational to assume $f(x)$ is convex in this paper. The optimal power allocation of (9) can be obtained by traditional convex optimization algorithms, typically in an iterative manner such as the interior-point method [40].

Moreover, if $f(x)$ is differentiable and the derivative is invertible, a close-form solution can be obtained by using KKT conditions [40]. The Lagrangian of (9) is given by

$$L = \sum_{j=1}^{N} \lambda_j f(c_j p_j) - \mu \left(\sum_{j=1}^{N} p_j - P_{tot} \right), \tag{15}$$

where μ is the lagrangian variable. Differentiating separately by p_i and μ and setting to 0, yields:

$$p_j = \frac{1}{c_j} \left[f'^{-1} \left(\frac{\mu}{c_j \lambda_j} \right) \right]^{+}, \tag{16}$$

and μ can be solved by

$$\sum_{j=1}^{N} \frac{1}{c_j} \left[f'^{-1} \left(\frac{\mu}{c_j \lambda_j} \right) \right]^{+} = P_{tot}, \tag{17}$$

where $x^+ = \max\{x, 0\}$.

To show how our proposed algorithm works, an example is given in Fig. 4 for the case of many-to-one matching and power allocation. A total of 10 carriers with various gains are available in the physical layer, while 20 chunks with different importance factors need to transmit over these carriers within one time slot. To schedule the transmission, first the channel gains and importance factors are arranged in a decreasing ordered, respectively. Then every two chunks are mapped to one channel in this order. Give that two chunks are allocated to a same channel, we can duplicate the channel gains twice to make sure the number of channel gains equals that of important factors. After that, the transmit power is allocated optimally according to the solution of problem (9).

6 Applications

We have proposed mechanism for optimal allocation of carriers and transmission power for a family of video transmission systems. In this section, we will show how

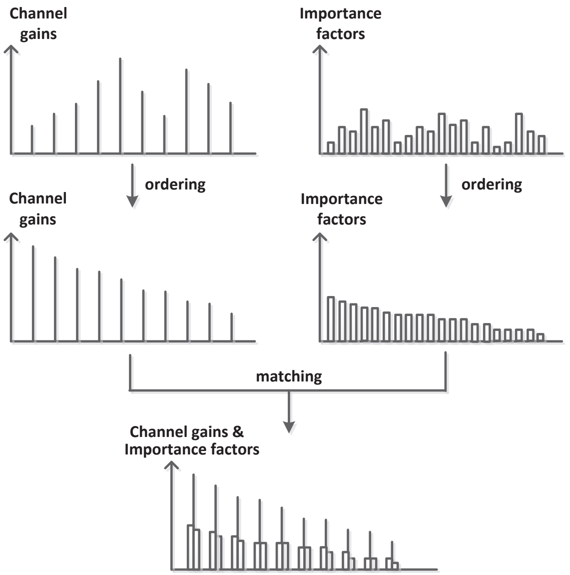


Fig. 4. Illustration of our proposed scheme. Every two chunks are carried by one channel.

some cross-layer designs can be recast as the form of (6). We use two types of applications as examples. In both applications, the mean square error (MSE) is adopted as the distortion measurement[2]. We first briefly introduce the scenarios and formulate the optimization problems with specific importance factors and distortion functions. Then, we point out that both the two problems have the same form of (6) and can be solved by adopting our proposed scheme. The applications are detailed in the following two subsections and summarized in Table 2.

6.1 Enhanced analog video coding

Analog video coding is a novel OFDM-based video transmission framework. Power scaling plays an important role, where scaling factors are obtained by solving an optimization problem [21]. In this paper, we reformulate the power scaling problem and provide a flexible resource allocation structure where the optimality of the joint channel matching and power allocation are guaranteed theoretically.

Our work is based on the framework of SoftCast and ParCast(+) which are typical analog video coding schemes [15, 21, 28–30]. As shown in Fig. 5, a group of pictures (GOP) is first compressed by 3D-DCT. The obtained DCT coefficients are divided into equal-sized rectangular-shaped chunks, which is shown in Fig. 6. The mean is removed from each chunk to get a zero-mean distribution. Different carriers and amount of transmit power are allocated to the chunks. Then, every two coefficients in the same chunk are directly mapped into a dense QAM symbol without channel coding, which was called raw OFDM in SoftCast. Some important data, such as scaling factors, matching information and the mean and variance of each chunk, are transmitted as meta data through conventional communication technologies, such that they can be recovered losslessly. At the receiver, meta data are decoded first to assist the decoder in inverting the received signal. A LLSE (Linear Least Squares Estimation) decoder is adopted to transform the received dense QAM symbols into DCT coefficients. After 3D-IDCT, a reconstructed GOP is obtained.

The framework in Fig. 5 is a specific example of the generalized framework in Fig. 3. The modules of 3D-DCT in Fig. 5 serves as the video encoder in Fig. 3. If the transmitter and the receiver are equipped with multiple antennas, a subcarrier in our method corresponds to the subchannel in the MIMO environment [15, 30].

In this application, we aim to determine how to allocate channel and power to the chunks at the transmitter.

Proposition 2. *For our considered analog video coding system, given that the ith chunk is carried by the jth channel with gain h_j and transmit power p_j, if LLSE scaling is adopted, the expected MSE of the ith chunk is expressed by*

$$e_{ij}^{LLSE} = \frac{\lambda_i \sigma^2}{h_j p_j + \sigma^2}; \tag{18}$$

if ZF (Zero Forcing) is assumed, the expected MSE is expressed by

$$e_{ij}^{ZF} = \frac{\lambda_i \sigma^2}{h_j p_j}, \tag{19}$$

where λ_i denotes the variance of the ith chunk and σ^2 is the noise power.

Proof. See Appendix A. □

To obtain more accurate scaling factors, the LLSE scaling is employed in our work. The number of chunks and the number of available carriers are both N. According to (5) and (18), the total MSE of a GOP is expressed by

$$\text{MSE}^{LLSE} = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{b_{ij} \lambda_i \sigma^2}{h_j p_j + \sigma^2}. \tag{20}$$

To minimize the MSE denoted by (20), we formulate the following optimization problem

$$\begin{aligned} \min_{\{b_{ij}\},\{p_j\}} \quad & \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{b_{ij} \lambda_i \sigma^2}{h_j p_j + \sigma^2}, \\ \text{subject to} \quad & (7) \text{ and } (8). \end{aligned} \tag{21}$$

Obviously, (21) has a similar form to (6), with the importance factor denoted by the variance $\{\lambda_i\}$ and the distortion function $f(x)$ expressed by

$$f(x) = \frac{\sigma^2}{x + \sigma^2}, \tag{22}$$

which is convex and non-increasing. Hence (21) can be solved by using our method. The optimal scheme matches the carriers by the order of the channel gains and chunk variances, and allocates power according to (16) and (17).

In SoftCast and ParCast(+), ZF is assumed when the scaling factors are calculated at the transmitter and LLSE is adopted to decode the received signal at the receiver. In

2. Common distortion measurements include MSE, sum of absolute difference, dSSIM [41], etc. For lack of space, we cannot elaborate all of them. Since MSE is the most popular distortion measurement, we only give applications with MSE in this paper.

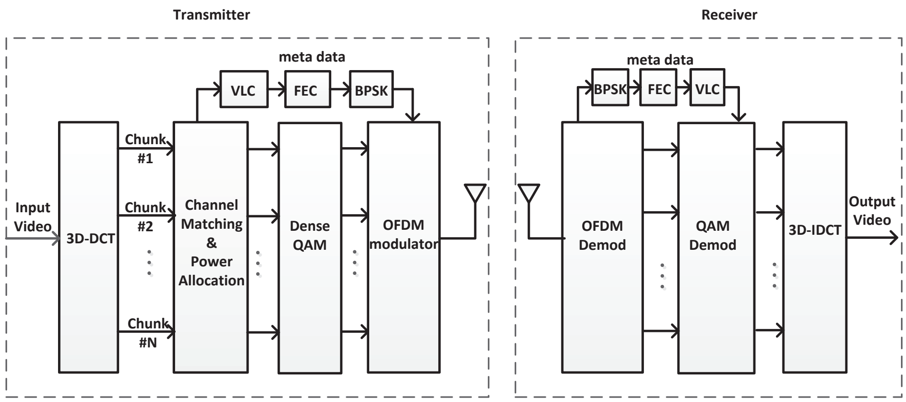


Fig. 5. Framework of the considered enhanced analog video coding system.

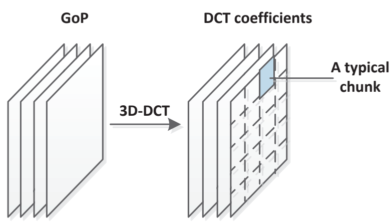


Fig. 6. Illustration of chunks in analog video coding systems.

this case, according to (5) and (19), the total MSE of a GOP is approximated by

$$\text{MSE}^{ZF} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{b_{ij}\lambda_i\sigma^2}{h_j p_j}.$$

A similar problem can be formulated only with the difference that the distortion function becomes $f(x) = \frac{\sigma^2}{x}$ which is also non-increasing and convex. Therefore, the problem under ZF assumption can be solved by adopting our proposed method. In the case that the ZF scaling is taken, the matching optimality can also be proven by using the rearrangement inequality [15, 30], while rearrangement inequality cannot be applied in LLSE scaling which works better in the low SNR range.

6.2 Uncompressed video transmission

In uncompressed video transmission systems, video data are transmitted directly over wireless channels without compression at the application layer. Usually, video data are divided into multiple bit streams with different importance, and unequal wireless resources are allocated to these streams in order to achieve high end-to-end video quality.

A novel framework is proposed as shown in Fig. 7. Video data are input into the system pixel by pixel. With the luminance of each pixel represented by N_b bits, the video data are first divided into N_b bit streams. The nth bit stream is composed of all the nth left most bits of the input pixels. Then, QAM mapping is applied independently on each bit steam and every $\log_2 M$ bits are mapped into an M-QAM symbol. The obtained QAM symbols are assigned to different subcarriers and transmitted with unequal power. At the receiver, the QAM symbols are decoded into parallel bit streams. Finally, the bit combination module combines the decoded bit streams into a series of reconstructed video pixels.

The framework in Fig. 7 is a specific example of the generalized framework in Fig. 3. The modules of bit division and QAM mapping in Fig. 7 serve as the video encoder in Fig. 3.

Different from analog video coding, uncompressed video transmission essentially assumes digital coding. As a result, the order of the QAM mapping and the power allocation in Fig. 7 are different from those in Fig. 5, and the scaling factors are not transmitted as meta data. In uncompressed video transmissions, with the help of AGC (Automatic Gain Control), the receiver can normalize the scaled QAM symbols and decode the bits, while without which the receiver cannot decode the scaled coefficients in the analog video coding schemes. In addition, in Fig. 7, all bits should be decoded to reconstruct the original pixels. Given the sufficient bandwidth in uncompressed video transmissions, to guarantee the performance of detection, a low modulation level is often taken to provide a high BER.

We aim to minimize the end-to-end MSE by allocating different subcarriers and unequal power to different bit streams. According to [33], the relationship between the MSE and each bit stream's BER can be expressed by:

$$\text{MSE} = \sum_{n=0}^{N_b-1} 4^n e_n, \tag{23}$$

where e_n is the BER of the nth bit stream. Notice that the MSE is a weighted summation of each bit stream's BERs. The weight for the nth bit stream is 4^n which also represents the significance of the nth bit stream.

Given that N_b is usually smaller than the number of available subcarriers N, we assume that N_s subcarriers are needed to transmit one bit stream. Therefore, a bit stream should be further divided into N_s substreams, where $N = N_b N_s$. In each time slot, a certain number of bits from each substream are grouped into a chunk and transmitted over OFDM. Fig. 8 shows the relationship among pixels, bit streams, sub-frames and chunks in this application.

The average BER of the nth bit stream is expressed by

$$e_n = \frac{1}{N_s}\sum_{j=1}^{N} b_{nj} f_e\left(\frac{h_j p_j}{\sigma^2}\right), \tag{24}$$

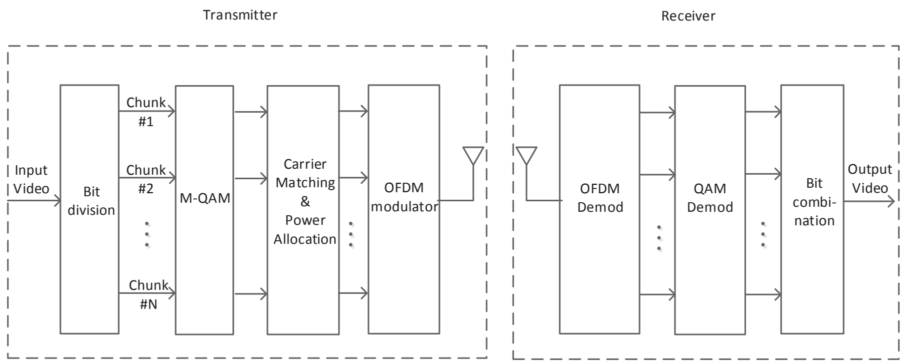


Fig. 7. Framework of our proposed uncompressed video transmission system.

where b_{nj} is a binary value denoting whether the nth stream is transmitted over the jth subcarrier. b_{nj} satisfies the constraints that $\sum_{j=1}^{N} b_{nj} = N_s$ which denotes N_s subcarriers are used to transmit a bit stream and $\sum_{n=0}^{N_b-1} b_{nj} = 1$ which denotes a subcarrier can only carry one bit stream. The function $f_e(\gamma)$ is the BER function of a specific modulation and coding combination in AWGN channels, where γ denotes the SNR. h_j denotes the gain of the jth subcarrier and p_j is the transmit power of the jth subcarrier.

Taking (24) into (23), we obtain

$$\begin{aligned} \text{MSE} &= \sum_{n=0}^{N_b-1}\sum_{j=1}^{N} b_{nj} 4^n f_e\left(\frac{h_j p_j}{\sigma^2}\right) \\ &\overset{(b)}{=} \sum_{i=1}^{N}\sum_{j=1}^{N} b_{ij} 4^{\lfloor \frac{i}{N_s} \rfloor} f_e\left(\frac{h_j p_j}{\sigma^2}\right), \end{aligned} \tag{25}$$

where $\lfloor \cdot \rfloor$ denotes an integer rounding-down operator. (b) is obtained by replacing the group of binary valuables $\{b_{nj}\}$ with a group of new binary valuables $\{b_{ij}\}$ which satisfies that $\sum_{j=1}^{N} b_{ij} = 1$ and $\sum_{i=1}^{N} b_{ij} = 1$.

To obtain the minimal MSE denoted by (25), an optimization problem is formulated as

$$\begin{aligned} \min_{\{b_{ij}\},\{p_j\}} \quad & \sum_{i=1}^{N}\sum_{j=1}^{N} b_{ij} 4^{\lfloor \frac{i}{N_s} \rfloor} f_e\left(\frac{h_j p_j}{\sigma^2}\right), \\ \text{subject to} \quad & \text{(7) and (8).} \end{aligned} \tag{26}$$

Note that (26) has the same form as (6). Specifically, the importance factor is $4^{\lfloor \frac{i}{N_s} \rfloor}$ and the distortion function is $f_e\left(\frac{x}{\sigma^2}\right)$. Given that the BER function $f_e(\gamma)$ is generally convex and non-increasing, the assumptions of (6) are satisfied and hence (26) can be solved by our proposal.

For M-QAM, the BER function can be approximated by [42]

$$f_e(\gamma) \cong \frac{2(\sqrt{M}-1)}{\sqrt{M}\log_2\sqrt{M}} Q\left(\sqrt{\frac{3\gamma}{M-1}}\right) \tag{27}$$

$$\leq \frac{\sqrt{M}-1}{\sqrt{M}log_2\sqrt{M}} \exp\left(-\frac{1.5\gamma}{M-1}\right), \tag{28}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{-\frac{t^2}{2}} dt$ denotes the Q-function. Both (27) and (28) are non-increasing and convex. (27) is an accurate approximation of the BER function which is called Q-BER approximation in this paper. Given that there does not exist a closed-form invertible function of $Q(x)$, therefore, the problem (26) can only be solved by an iterative term if (27) is adopted. (28) is a tight upper bound of the BER function which is called bound-BER approximation in this paper. If (28) is adopted, a closed-form solution of (26) can be obtained according to (16) and (17).

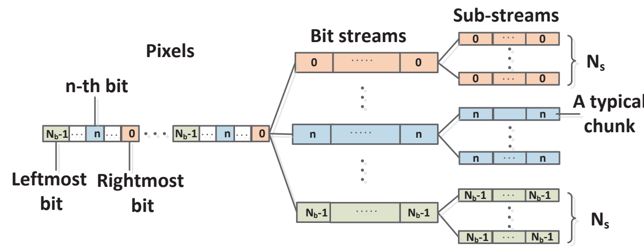


Fig. 8. Illustration of pixels, bit streams, sub-streams and chunks in our proposed uncompressed video transmission system.

7 Simulation Results

In this section, we evaluate the performance of our proposed joint carrier matching and power allocation scheme. We simulate the two applications described in section 6 respectively. In both applications, a general OFDM system with 64 subcarriers is adopted. The fading coefficients are assumed to be independent and identically distributed circularly symmetric complex Gaussian random variables with zero mean and unit variance. Assume the CSI is available at the transmitter in both cases. PSNR is employed as the performance metric:

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

where N_b is the number of bits per pixel and is set to 8. Given that the PSNR is entirely determined by the MSE, maximizing the PSNR is equivalent to minimizing the MSE. Therefore, our proposed schemes also provide the optimal performance in terms of PSNR. Video contents may influence the performance of video encoders, and video algorithms often achieve different PSNR improvement when operated on different video samples. Therefore, to evaluate the performance of our proposed algorithms, 16 common CIF video samples are combined to form an integrated video sequence. The integrated sequence includes akiyo, news, stefan, tempete, bus, mobile, soccer, city, coastguard, crew, foreman, waterfall, harbour, ice, flower and football.

7.1 Performance enhancement in analog video coding

We simulate the enhanced analog video coding scheme according to the structure in Fig. 5. SoftCast and ParCast are selected as the baselines. Without considering various channel gains of subcarriers, the carrier matching scheme in SoftCast can be viewed as a random carrier matching algorithm. Moveover, the most recent version of SoftCast [21] adopts Hadamard transform after power allocation to improve the performance. Although we focus on subcarrier matching in OFDM-based video transmission systems, our proposed method can be applied to the subchannel matching in MIMO systems, such as ParCast(+). To compare the performance with ParCast(+), we assume that both the transmitter and the receiver are equipped with 2 antennas. Therefore, only a half of bandwidth is needed to transmit the same number of chunks. Since we focus on the channel matching and power allocation scheme, we only employ modules related to our proposal in the simulation. Some enhanced schemes in ParCast+, such as temporal transform with motion alignment, are not implemented for simplicity. For all schemes in this simulation, every 4 video input frames are grouped into a GOP, with 3D-DCT applied over each GOP. The 3D-DCT coefficients in a GOP are divided into 256 chunks. To reduce the bandwidth consumption, chunks with variances nearly zero are discarded and reconstructed with zero values. The ratio of chunks discarded is determined based on the available bandwidth and video data rate. In our simulation, we discard 25% of total chunks.

As shown in Fig. 9, compared to SoftCast without Hadamard transform, there is a 6 - 8 dB gain of our proposed scheme in a large SNR range. Note that in both the two schemes, the same power allocation method is adopted and the Hadamard transform is absent. SoftCast can be viewed as a special case of our proposal without channel matching. Therefore, it is reasonable to conclude that the performance gain mainly comes from channel matching. Our proposed scheme outperforms ParCast by 1-2 dB in the low SNR range. Hence, if the channel state information and noise level are available at the transmitter, the LLSE scaling method is suggested. Note that, the Hadamard transform enhances the performance of original SoftCast by 1-2 dB in low SNR ranges, and is an effective option to improve the performance when the channel state information is not available at the transmitter.

Fig. 10 shows the frame quality of different schemes. The SNR is set to 10 dB. The test frame is a snapshot from the transmitted video sequence under evaluation. Obviously, the two enhanced analog coding schemes shown in Fig. 10(a) and Fig. 10(b) offer much better visual quality than that of conventional SoftCast schemes shown in Fig. 10(c) and Fig. 10(d). There is no big visual quality difference between our proposed method and ParCast, because the performance of the two schemes tends to be similar when S-NR is high. Comparing Fig. 10(c) and Fig. 10(d), we observe that the Hadamard transform improves the visual quality in fading channel. However, the improvement is much lower compared with the schemes adopting channel matching.

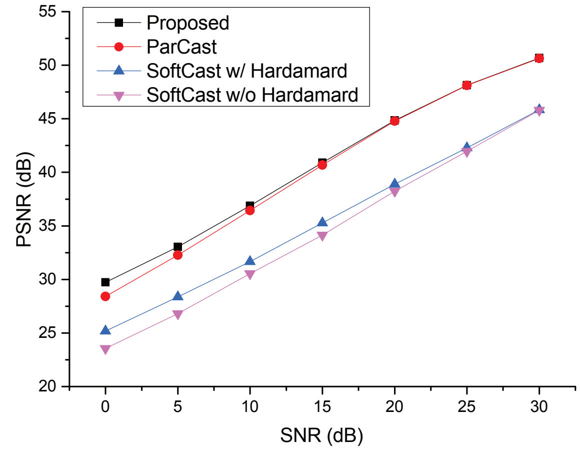
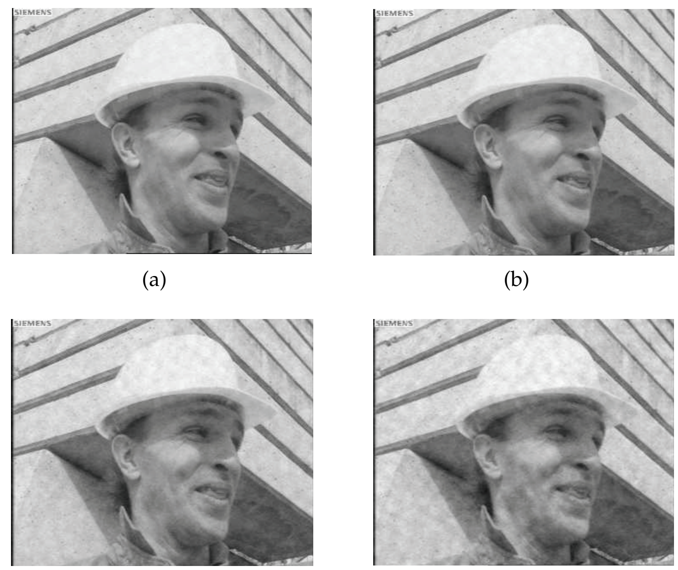


Fig. 9. Performance comparison of analog video coding schemes.



(a) (b) (c) (d)

Fig. 10. Frame quality comparison for analog video coding schemes at SNR=10 dB. (a) Proposed. (b) ParCast. (c) SoftCast with Hadamard transform. (d) SoftCast without Hadamard transform.

7.2 Performance enhancement in uncompressed video transmission

We simulate our proposed uncompressed video transmission system according to Fig.7. Our proposed method in this application can be viewed as a joint bit and power allocation method. If we do not consider the importance of bits, i.e. bits are randomly allocated to subcarriers, the best power allocation scheme is waterfilling under shannon capacity theory. However, what we focus on is a specific modulation and coding scheme for video transmission. Given that PSNR is determined by MSE and MSE is further associated with mean BER which has been discussed in Section 6.2, we adopt a minimized average BER scheme as one of our baselines [42]. To see the channel matching gain and power allocation gain separately, we also employ the methods of optimal channel matching with min-BER power allocation and random channel matching with optimal power allocation. To simplify the simulation, we adopt uncoded QPSK modulation and the BER function (28). The number of bits

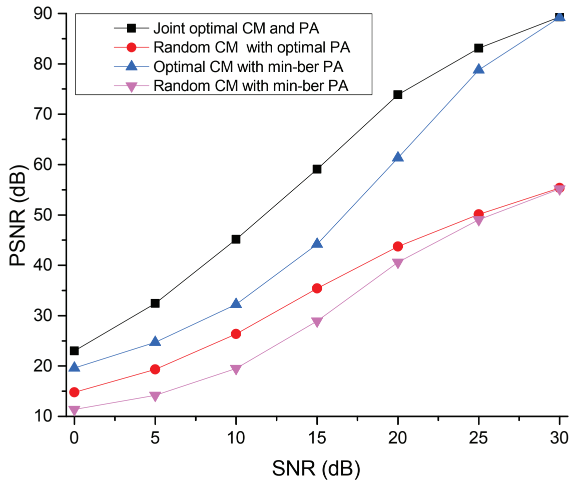


Fig. 11. Performance comparison of uncompressed video transmission schemes.

per frame is $352 \times 288 \times 8$. Then, the size of each chunk is 10032 bits.

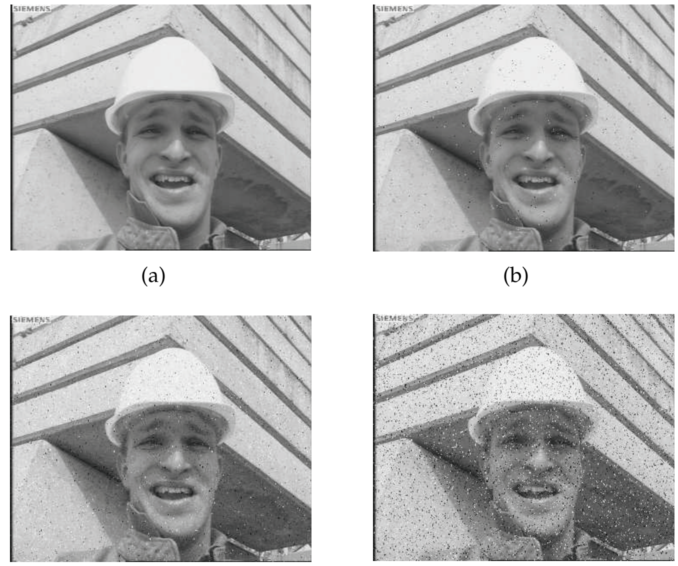
With the simplicity of implementation, uncompressed video transmission has the feature of low delay. For a real-time point-to-point application, the delay is usually resulted from the large buffering time in the processes of video encoding, wireless transmission and video decoding. Since there are no large buffers in the structure, the low end-to-end delay is guaranteed inherently. Therefore, we focus on the frame quality comparison in our performance studies.

Fig.11 shows the performance of different uncompressed video transmission schemes. Denote channel matching by CM and power allocation by PA for simplicity of notation. As shown in the figure, with the increase of SNR, the performance improvement by adopting the optimal power allocation reduces. However, the gain brought by channel matching is still significant. These results indicate that channel matching plays the most important role in our proposed scheme. The joint power and carrier allocation scheme improves the PSNR performance by 13 dB - 35dB compared with the minimized BER approach. These curves have verified the efficiency of our proposal.

Fig.12 shows the frame quality of different schemes. The SNR is set to 10 dB and a frame "Foreman" is selected from the transmitted video sequence for visual quality comparison. As shown in the four sub-figures, our proposed joint channel matching and power allocation method provides the highest visual quality. It is also observed that the video quality of the output frames is degraded by gaussian noises, especially Fig.12(c) and Fig.12(d). This is because, when the important bits are not well protected, the transmission errors in such bits will cause big difference in pixel values.

8 Conclusion

In this paper, we formulate a generalized optimization problem and propose a joint channel matching and power allocation scheme for a family of emerging video schemes that are designed to be better fit for wireless transmissions. We further apply our scheme to two types of applications, analog video coding and uncompressed video transmissions. Numerical results demonstrate that our proposed method provides graceful PNSR performance in a large SNR range compared with the baselines. Therefore, the joint channel matching and power allocation scheme is very useful for these applications.



(a) (b) (c) (d)

Fig. 12. Frame quality comparison for uncompressed video transmission systems at SNR=10 dB. (a) Joint optimal channel matching and power allocation. (b) Optimal channel matching with min-BER power allocation. (c) Random channel matching with optimal power allocation. (d) Random channel matching with min-BER power allocation.

References

[1] N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, "Generalized frequency division multiplexing for 5th generation cellular networks," *Communications, IEEE Transactions on*, vol. 62, no. 9, pp. 3045–3061, 2014.

[2] T. Cisco, "Cisco visual networking index: global mobile data traffic forecast update, 2014–2019," *Cisco Public Information*, 2015.

[3] S. Cicalo and V. Tralli, "Distortion-fair cross-layer resource allocation for scalable video transmission in ofdma wireless networks," *Multimedia, IEEE Transactions on*, vol. 16, no. 3, pp. 848–863, 2014.

[4] M. Li, Z. Chen, and Y.-P. Tan, "Joint power allocation and bit loading for enhanced svc video downlink transmissions over sdma/ofdma networks," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–6.

[5] X. Ji, J. Huang, M. Chiang, G. Lafruit, and F. Catthoor, "Scheduling and resource allocation for svc streaming over ofdm downlink systems," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 19, no. 10, pp. 1549–1555, 2009.

[6] J. Gross, J. Klaue, H. Karl, and A. Wolisz, "Cross-layer optimization of ofdm transmission systems for mpeg-4 video streaming," *Computer Communications*, vol. 27, no. 11, pp. 1044–1055, 2004.

[7] E. ETSI, "300 744 digital video broadcasting (dvb); framing structure, channel coding and modulation for digital terrestrial television," 2001.

[8] P. Zhao, Y. Liu, J. Liu, R. Yao, and S. Ci, "Ssim-based cross-layer optimized video streaming over lte downlink," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 1394–1399.

[9] M. F. Sabir, A. C. Bovik, and R. W. Heath, "Unequal power allocation for jpeg transmission over mimo systems,"

Image Processing, IEEE Transactions on, vol. 19, no. 2, pp. 410–421, 2010.

[10] M. Shayegannia, A. Mohammadi, S. Muhaidat, and M. Torki, "An ofdm based system for transmission of jpeg2000 images using unequal power allocation," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. IEEE, 2012, pp. 2064–2069.

[11] M. Shayegannia, A. Hajshirmohammadi, S. Muhaidat, and M. Torki, "Transmission of jpeg2000 images over frequency selective channels with unequal power allocation," *Image Processing, IET*, vol. 7, no. 1, pp. 33–41, 2013.

[12] S. De, R. Saini, and A. Sharma, "Source-aware adaptive power allocation in ofdm systems for rate constrained applications," *Communications Letters, IEEE*, vol. 18, no. 10, pp. 1855–1858, 2014.

[13] S. Mahajan and A. K. Jagannatham, "Hierarchical dwt based optimal diversity power allocation for video transmission in 4g ofdma wireless systems," in *Imaging Systems and Techniques (IST), 2011 IEEE International Conference on*. IEEE, 2011, pp. 279–283.

[14] L. Yu, H. Li, and W. Li, "Wireless scalable video coding using a hybrid digital-analog scheme," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 24, no. 2, pp. 331–345, 2014.

[15] X. L. Liu, W. Hu, C. Luo, Q. Pu, F. Wu, and Y. Zhang, "Parcast+: Parallel video unicast in mimo-ofdm wlans," *Multimedia, IEEE Transactions on*, vol. 16, no. 7, pp. 2038–2051, 2014.

[16] A. Zhang, X. Fan, R. Xiong, and D. Zhao, "Distributed soft video broadcast with variable block size motion estimation," in *Visual Communications and Image Processing (VCIP), 2013*. IEEE, 2013, pp. 1–5.

[17] W. Huang, X. Fan, and D. Zhao, "Soft mobile video broadcast based on side information refining," in *Visual Communications and Image Processing (VCIP), 2013*. IEEE, 2013, pp. 1–6.

[18] X. Fan, R. Xiong, F. Wu, and D. Zhao, "Wavecast: Wavelet based wireless video broadcast using lossy transmission." in *VCIP*, 2012, pp. 1–6.

[19] R. Xiong, F. Wu, X. Fan, C. Luo, S. Ma, and W. Gao, "Power-distortion optimization for wireless image/video softcast by transform coefficients energy modeling with adaptive chunk division," in *Visual Communications and Image Processing (VCIP), 2013*. IEEE, 2013, pp. 1–6.

[20] L. Yu, H. Li, and W. Li, "Hybrid digital-analog scheme for video transmission over wireless," in *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 1163–1166.

[21] S. Jakubczak and D. Katabi, "Softcast: one-size-fits-all wireless video," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 449–450, 2011.

[22] http://www.wirelesshd.org/pdfs/WirelessHD-Specification-Overview-v1.1May2010.pdf.

[23] H. Yoon, Y. Jung, and S. Lee, "An image sequence transmission method in wireless video surveillance systems," *Wireless Personal Communications*, pp. 1–14, 2015.

[24] Z. He and S. Mao, "Adaptive multiple description coding and transmission of uncompressed video over 60ghz networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 18, no. 1, pp. 14–24, 2014.

[25] X. Mo and D. Liu, "An adaptive unequal error protection strategy for uncompressed video transmission over 60ghz wpan systems," in *Signal Processing, Communication and Computing (ICSPCC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–5.

[26] S.-E. Hong and W. Y. Lee, "Flexible unequal error protection scheme for uncompressed video transmission over 60ghz multi-gigabit wireless system," in *Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on*. IEEE, 2011, pp. 1–6.

[27] N. A. M. Nguyen and W. C. Lee, "Optimal solution of uncompressed high-definition video transmission using selection modulation method," in *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*. IEEE, 2011, pp. 10–13.

[28] S. Jakubczak, H. Rabul, and D. Katabi, "Softcast: One video to serve all wireless receivers," *Computer Science and Artificial Intelligence Laboratory Technical Report, MIT-CSAIL-TR-2009-005, MIT*, 2009.

[29] S. Jakubczak and D. Katabi, "A cross-layer design for scalable mobile video," in *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM, 2011, pp. 289–300.

[30] X. L. Liu, W. Hu, Q. Pu, F. Wu, and Y. Zhang, "Parcast: soft video delivery in mimo-ofdm wlans," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 233–244.

[31] H. Singh, J. Oh, C. Kweon, X. Qin, H.-R. Shao, and C. Ngo, "A 60 ghz wireless network for enabling uncompressed video communication," *Communications Magazine, IEEE*, vol. 46, no. 12, pp. 71–78, 2008.

[32] T. Bruggen and P. Vary, "Unequal error protection by modulation with unequal power allocation," *Communications Letters, IEEE*, vol. 9, no. 6, pp. 484–486, 2005.

[33] Z. Zhang and D. Liu, "Unequal power allocation for real-time uncompressed video transmission over wireless channels," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2015 IEEE 26th Annual International Symposium on*. IEEE, 2015, pp. 1142–1147.

[34] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-ghz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, 2011.

[35] W. Cui, P. Ranta, T. A. Brown, and C. Reed, "Wireless video streaming over uwb," in *Ultra-Wideband, 2007. ICUWB 2007. IEEE International Conference on*. IEEE, 2007, pp. 933–936.

[36] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[37] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial & Applied Mathematics*, vol. 5, no. 1, pp. 32–38, 1957.

[38] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[39] J. Huang, Z. Li, M. Chiang, and A. K. Katsaggelos, "Joint source adaptation and resource allocation for multi-user wireless video streaming," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 5, pp. 582–595, 2008.

[40] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[41] C. Yeo, H. L. Tan, and Y. H. Tan, "On rate distortion optimization using ssim," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 23, no. 7, pp. 1170–1181, 2013.

[42] C. S. Park and K. B. Lee, "Transmit power allocation for ber performance improvement in multicarrier systems," *Communications, IEEE Transactions on*, vol. 52, no. 10, pp. 1658–1663, 2004.

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. and PH.D degrees in communication and information systems from Beijing University of Posts and Telecommunications, Beijing, China in 2010 and 2016, respectively. He is currently a lecturer 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 scholar 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.

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 in 1990 and 1993, 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 serves as the associate editor of IEEE Transactions on Mobile Computing. Dr.Wang achieved the NSF career award in 2005, and ONR challenge award in 2010.

Appendix A
Proof of Proposition 2

Choose chunk i as a typical chunk and a coefficient in chunk i as a typical DCT coefficient x_i. The variance of the ith chunk is calculated by $\lambda_i = E[x_i^2]$. Given that chunk i is transmitted over the jth channel with transmit power p_j, the scaling factor for the ith chunk is obtained by $g_i = \sqrt{\frac{p_j}{\lambda_i}}$. According to (3), the coefficient x_i is scaled to $g_i x_i$ before transmission and the receiver receives

$$y_{ij} = \sqrt{\frac{p_j}{\lambda_i}} a_j x_i + n_i,$$

where a_i denotes the channel coefficient and n_i is the Gaussian noise with variance σ^2. At the receiver, a_j and σ^2 are obtained from channel estimation; λ_i and g_i are obtained from the decoded meta data. At the transmitter, a_j and σ^2 are assumed to be available from feedback.

If a LLSE decoder is adopted, the receiver decodes

$$\hat{x}_{ij} = \frac{\lambda_i g_i a_j^*}{\lambda_i |a_j|^2 g_i^2 + \sigma^2} y_{ij}.$$

The MSE of the ith chunk is calculated by

$$\begin{aligned} e_{ij}^{LLSE} &= E\left[|\hat{x}_{ij} - x_i|^2 \right] = E\left[\left| \frac{\lambda_i g_i a_j^*}{\lambda_i |a_j|^2 g_i^2 + \sigma^2} y_i - x_i \right|^2 \right] \\ &= E\left[\left| \frac{\lambda_i g_i a_j^*}{\lambda_i |a_j|^2 g_i^2 + \sigma^2} (a_j g_i x_i + n_i) - x_i \right|^2 \right] \\ &= E\left[\left| \frac{\lambda_i a_j^* g_i n_i - \sigma^2 x_i}{\lambda_i |a_j|^2 g_i^2 + \sigma^2} \right|^2 \right]. \end{aligned} \tag{29}$$

Note that the Gaussian noise n_i is independent of x_i, thus $E[x_i n_i] = E[x_i]E[n_i] = 0$ and (29) becomes

$$\begin{aligned} e_{ij}^{LLSE} &= \frac{\lambda_i^2 |a_j|^2 g_i^2 E[|n_i|^2] + \sigma^4 E[|x_i|^2]}{\left(\lambda_i |a_j|^2 g_i^2 + \sigma^2\right)^2} \\ &= \frac{\lambda_i^2 |a_j|^2 g_i^2 \sigma^2 + \sigma^4 \lambda_i}{\left(\lambda_i |a_j|^2 g_i^2 + \sigma^2\right)^2} \\ &= \frac{\lambda_i \sigma^2}{\lambda_i |a_j|^2 g_i^2 + \sigma^2} = \frac{\lambda_i \sigma^2}{|a_j|^2 p_i + \sigma^2}. \end{aligned}$$

If a ZF decoder is assumed, the receiver decodes

$$x'_{ij} = \frac{y_i}{a_j g_i} = x_i + \frac{n_i}{a_j g_i}.$$

The MSE of the ith chunk is calculated by:

$$e_{ij}^{ZF} = E\left[\sum_i |x'_{ij} - x_i|^2 \right] = \sum_i \frac{E[|n_i|^2]}{|a_j|^2 |g_i|^2} = \sum_i \frac{\lambda_i \sigma^2}{|a_j|^2 p_j}.$$