Robust Uncoded Video Transmission over Wireless Fast Fading Channel

Hao Cui
EEIS, USTC
hcuipro@mail.ustc.edu.cn

Chong Luo
Microsoft Research
cluo@microsoft.com

Chang Wen Chen
CSE, SUNY Buffalo
chencw@buffalo.edu

Feng Wu
Microsoft Research
fengwu@microsoft.com

***Abstract*—This research studies robust uncoded video transmission over wireless fast fading channel, where only statistical channel state information (CSI) is available at the sender. We observe that increasing channel diversity for high priority (HP) data is essential to improving the robustness of video transmission in fading channels. By utilizing the noise and loss resilient nature of video, we find it possible to design a more robust system by re-allocating the power and channel uses among HP and LP (low priority) data. With total power and channel use constraints, we derive an optimal resource allocation scheme under the squared error distortion criterion. In particular, we first propose a new power allocation algorithm at given channel allocation. Second, based on the proposed power allocation algorithm, we design a channel allocation algorithm to strike the tradeoff between the diversity increase of HP data and the information loss of LP data. Third, under known noise power distribution, we derive the optimal resource allocation for uncoded video multicast. Simulations show that the proposed system achieves 2dB and 5dB gain in average and outage PSNR over Softcast in video Unicast, and around 1.4dB and 4dB gain in multicast.**

I. Introduction

By the end of 2012, video constitutes 51% of mobile traffic, and this percentage is estimated to increase to 67% by 2017, according to Cisco Visual Networking Index (VNI) [1]. The statistics reflect people's increasing demand to consume videos on mobile devices, instead of conventional TV, through the advanced access technologies, such as WiFi, WiMax, and LTE.

In these digital communication systems, separate source-channel design is dominant for its optimality and implementation convenience. In particular, source coding is implemented by a succession of functional modules including motion estimation, DCT (Discrete Cosine Transform), quantization and entropy coding. The generated bit sequence is then transmitted under the protection of a channel code. It is well-known that the optimality of digital communication relies on Shannon's source-channel separation principle [2]. However, a fact is often neglected that the application of this principle requires precise CSI at the transmitter, and there is a *threshold effect* in performance. One could immediately infer that this principle does not hold in broadcast channels where receivers have diverse channels and the feedback channel for CSI is absent. Besides, an optimal channel code requires high complexity and introduces infinite delay.

In view of these problems, there has been a surge of interest recently in the uncoded or near-analog video communication systems [3], [4], [5], [6] built upon the joint source-channel coding (JSCC) approach. The basic idea is to skip quantization and entropy coding in source encoder, and directly transmit DCT coefficients over the channel. In order to balance the energy consumption of coefficients with different importance, a linear power scaling operation is applied before the amplitude modulation. Such a near-analog design differs from conventional analog system in two key steps, namely de-correlating transform and power allocation, which allows the former to achieve much higher transmission efficiency. In the sense that no digital channel coding is employed, this scheme is also referred to as an uncoded system. Analyses on Softcast [7] have shown that this uncoded scheme can achieve similar end-to-end distortion as the more complex digital methods while maintaining robustness to channel variations and lowering computational complexity. Moreover, the system intrinsically has distortion performance that degrades gracefully with channel SNR, providing substantial net performance gain for broadcast channels or when the channel statistics are not well estimated.

In spite of the above merits, existing uncoded video systems are either optimized over a simple additive white Gaussian noise (AWGN) channel model or assume a slow fading channel where the transmitter CSI is available. As mentioned, mobile videos are often requested on-the-go, such as on the bus, subway, highway, or high speed rail. In these environments, the transmitted videos will experience fast fading channels due to multi-path, shadowing and the Doppler effect. If these factors are neglected in the system design, users are likely to suffer from severe quality fluctuation. From the principle of wireless communication [8], we know that increasing the channel diversity could improve the outage capacity of fading channel. In other words, the diversity gain from multiple channel use could increase the channel quality. If there is no power constraint, it is easy to derive how to assigning the channel use according to the priorities of data. However, with the limited power and channel use resource, assigning more channel uses to high priority (HP) data means that some of the lower priority (LP) data will be dropped. There exists a trade-off between the distortion decrease of HP data and the distortion increase of LP data that makes resource allocation a challenging task.

In this research, we study the resource allocation problem for uncoded video transmission and the proposed solution can strike a good balance for the above trade-off. We start from generalizing the power allocation problem for uncoded video transmission. Our proposed solution to this problem is by no means a straightforward extension of existing work. In particular, the allocation result of our algorithm would indicate

which part of the transform coefficients should be dropped under given noise power. By utilizing the unequal priority of video data, we propose a greedy algorithm solving the combination problem for channel allocation which achieves nearly the same performance as performing a full search. Moreover, the heterogeneous noise power distribution in multicast scenario is discussed. The answers to the following three questions together with the proposed algorithms constitute the main contributions of this paper. Which part of data should be dropped? How to assign the saved channel uses to the rest of the data? How the power should be allocated after the channel reallocation?

The rest of this paper is organized as follows. In Section II, we draw a whole picture of the uncoded video transmission system and set up the system model. In Section III, we present our motivation and formulate the problem. Section IV and V describe the problems and proposed algorithms for power allocation and channel allocation, respectively. The evaluation of the proposed scheme is presented in Section VI. In Section VII, we discuss the related work on joint source-channel coding and uncoded video transmission. Finally we summarize the paper in Section VIII.

II. System Model

A. Overview of Uncoded Video Transmission

Fig.1 shows the block diagram of an uncoded video transmission system. The processing and transmission unit is a group of pictures (GOP). The size of a GOP could vary from 4, 8, 16, to 30 or 32 frames. The pictures in a GOP first subtract 128 from their pixel values and then are spatially and temporally decorrelated through 3D-DCT. The DCT coefficients with similar statistics are grouped together as a chunk. Usually, equal-sized chunk [7], [4] as shown in Fig.1 is implemented. The coefficients in a chunk are considered as i.i.d. zero-mean Gaussian distributed random variables and the variance is the average energy of the chunk. Dividing coefficients into more chunks will produce better performance, but the overhead will be higher too. Usually, 64 chunks per picture is adopted because it introduces negligible overhead with little harm on performance [18].

All the coefficients in a chunk will share the same scaling factor and channel allocation, which are decided according to the variances of the chunk. The scaled coefficients are pairwisely mapped to the amplitude of in phase and quadrature phase transmission signals. In addition, the meta data, including the variance of the chunks and the channel use assignment, should be transmitted through conventional digital method at a reliable transmission rate (e.g. BPSK with 1/2 coding). With channel assignment information, maximum ratio combining (MRC) is performed if multiple channel uses are allocated to any single coefficient. Based on the estimated CSI and the refined noise power by MRC, the coefficients can be obtained by MMSE (minimum mean square error) detection. Finally, the entire GOP can be reconstructed through inverse 3D-DCT and adding 128 in pixel values.

B. Definitions and Assumptions

This paper focuses on the resource allocation problem, which is the dashed rectangle in Fig.1. Without loss of generality, equal-sized chunk division is assumed. Unless otherwise specified, the bold letter in this paper denotes a vector of the corresponding scalar.

Suppose that the DCT coefficients of a GOP are divided into N chunks. Let c_n be a coefficient from the n^{th} chunk. For resource allocation, the mean energy $\lambda_n = \mathbb{E}[c_n^2]$ of each chunk is estimated by averaging the energy of coefficients in the chunk. Based on the mean energy, the linear power scaling factor g_n for the n^{th} chunk can be derived. The transmitted signal for coefficient c_n is

$$s_n = g_n c_n$$

Hence, the transmission power of the n^{th} chunk will be $\mu_n = g_n^2 \lambda_n$. The transmission signal is modulated by a pair of scaled coefficients from the same chunk,

$$x_n = (s_n + i \cdot s'_n)/\sqrt{2}$$

Without loss of generality, we set the total power constraint $P = N$, so that the average transmission power for the GOP is normalized to unit power.

Suppose that M uncorrelated channels are available for the transmission. Each channel is in fast fading state and is modeled as Rayleigh fading channel. Assume that only the statistics of the channel fading is known at the sender, which is $h_m \sim \mathcal{CN}(0,1)$. In addition, the channel noise is modeled as AWGN, i.e. $e_m \sim \mathcal{CN}(0,\sigma^2)$ and is assumed to be independent with the fading gain h_m.

Let $\mathbf{k} = (k_1, k_2, \ldots, k_N)^T$ be a channel allocation, where k_n is the number of channels assigned to the n^{th} chunk. The k_n channels assigned to signal x_n form a vector channel $\mathbf{H}_n = (h'_1, h'_2, \ldots, h'_{k_n})^T$. The received signals are k_n noisy versions of x_n:

$$\mathbf{Y}_n = \mathbf{H}_n x_n + \mathbf{E}_n$$

where $\mathbf{E}_n = (e'_1, e'_2, \ldots, e'_{k_n})^T$ is the noise vector on corresponding channel. Before detection of the coefficients, maximum ratio combining is applied to improve the channel quality.

$$\hat{x}_n = \frac{\mathbf{H}_n^*}{\|\mathbf{H}_n\|^2}\mathbf{Y}_n = x_n + \hat{e}_n$$

where $\hat{e}_n \sim \mathcal{N}(0, \sigma^2/\|\mathbf{H}_n\|^2)$.

Through low complexity linear MMSE detection, the coefficients can be recovered from the real part and imaginary part of $\hat{x}_n$.

$$\hat{s}_n = \frac{g_n \lambda_n}{g_n^2 \lambda_n + \sigma^2/\|\mathbf{H}_n\|^2} Re\{\hat{x}_n\}$$

Define the distortion as the Euclidean norm $\varepsilon_n = \|\hat{s}_n - s_n\|^2$. The minimized mean distortion condition on the CSI, number of channel use and noise power will be

$$\mathbb{E}[\varepsilon_n | \mathbf{H}_n, \sigma^2] = \frac{\lambda_n \sigma^2}{\|\mathbf{H}_n\|^2 g_n^2 \lambda_n + \sigma^2} = \frac{\lambda_n}{\|\mathbf{H}_n\|^2 \rho_n + 1} \quad (1)$$

where $\rho_n = \mu_n/\sigma^2$ is the signal to noise power ratio.

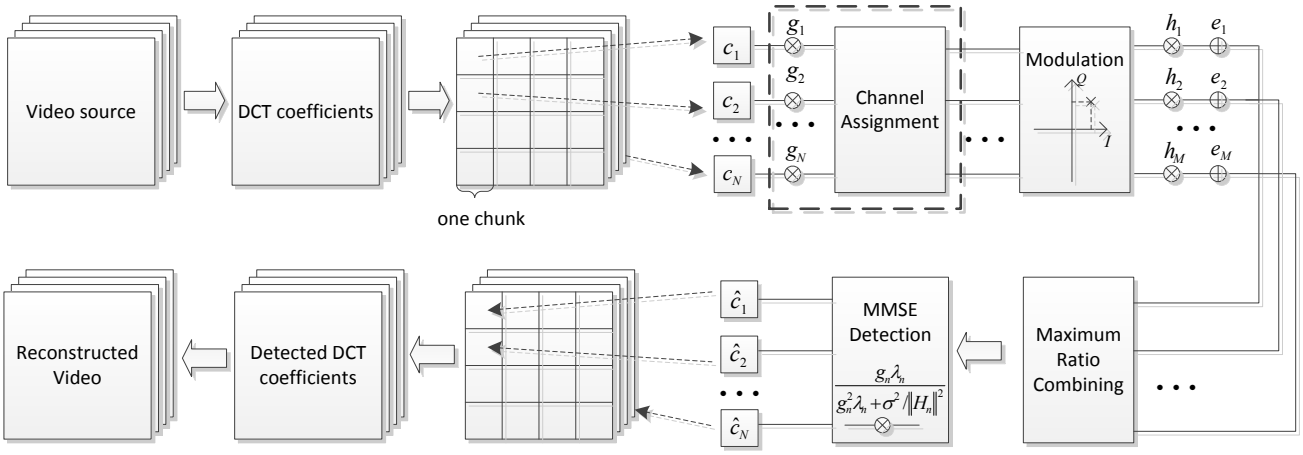


Fig. 1. Overview of an uncoded video transmission system

TABLE I. RESULTS OF HEURISTIC EXAMPLE.

	$\sigma^2 = 0.1$	$\sigma^2 = 0.01$	$\sigma^2 = 0.001$
choice 1	10.995	1.976	0.291
choice 2	6.407	0.850	0.111
choice 2+	6.407	0.826	0.101

III. PROBLEM FORMULATION

A. Motivation

The DCT coefficients of a natural image usually show heavily unequal energy distribution. Losing some low-energy coefficients may have negligible effects on the image quality. It has been shown [4] that discarding some low-priority coefficients and spending the saved power on high-priority coefficients may increase the received video quality. A question will be, is it possible to improve the video quality and stability by utilizing both the saved transmission power and channel use?

Let's consider a simple case. Three symbols $\{10,\ 1,\ 0.1\}$ are transmitted. The order of magnitude difference among different part of the DCT coefficients may be even larger. Considering two transmission choice. The first one, all symbols are transmitted, i.e. the channel assignment is $\mathbf{k} = \{1,\ 1,\ 1\}$. When the power allocation for AWGN [7] is used, the transmission power will be $\boldsymbol{\mu} = \{2.7027,\ 0.2703,\ 0.0270\}$. The second one, drop the least important symbol 0.1 and assign the saved channel to symbol 10, i.e. the channel assignment is $\mathbf{k} = \{2,\ 1,\ 0\}$ and the transmission power is $\boldsymbol{\mu} = \{1.4286, 0.1429,\ 0\}$. For different noise power level, taking expectation over fading gain $\|\mathbf{H}_n\|^2$ in (1), we compute the total mean distortion $\varepsilon_1 + \varepsilon_2 + \varepsilon_3$ as shown in table I.

Is that the whole story? No. We find that the existing power allocation becomes sub-optimal because of the channel allocation. Based on choice 2, if we subtract some transmission power from symbol 10 to compensate the symbol 1, the distortion can be further decreased. With the new power allocation, say $\boldsymbol{\mu} = \{1.3,\ 0.4,\ 0\}$, the result is shown as the third row in Table I. It indicates that the diversity gain could save power to help the transmission of other symbols. Therefore, both the channel use and transmission power should be explored to enhance the video transmission in fast fading channel.

B. Problem Statement

Given a set of N chunks with average energy denoted by $\lambda_1 ... \lambda_N$. Without loss of generality, we assume $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N$. Consider the transmission of N coefficients drawn from the N chunks respectively. Let M and P be the number of available channels and the total transmission power, and σ^2 be the noise power, we shall find out the optimal channel allocation $\mathbf{k}$ and power allocation $\boldsymbol{\rho}$ that minimize the total mean distortion. Mathematically, the problem is formulated into:

$$\begin{array}{ll} \min & \sum_{n=1}^{N} \mathbb{E}[\varepsilon_n] \\ \text{s.t.} & \sum_{n=1}^{N} k_n = M \\ & k_n \geq 0,\ k_n \in \mathbb{Z} \\ & \sum_{n=1}^{N} k_n \rho_n = \frac{P}{\sigma^2} \\ & \rho_n \geq 0,\ \rho_n \in \mathbb{R} \end{array} \tag{2}$$

Here, $\varepsilon_n = f(\lambda_n, \rho_n, \mathbf{H}_n)$ is a multivariable function and H_n depends on the channel allocation k_n. The expectation $\mathbb{E}[\varepsilon_n]$ is taken over the source, the channel fading gain and the noise.

The optimization in (2) is a mixed discrete and continuous programming problem. No general method can be applied to derive the solution. Therefore, we propose a heuristic approach which divides this problem into two sub-problems: optimal power allocation under a determined channel assignment and channel allocation based on the optimal power allocation.

IV. POWER ALLOCATION

The first sub-problem is the optimal power allocation under a given channel allocation $\mathbf{k}$ and noise power σ^2. It is a relaxed problem of (2) that the objective of the optimization is replaced by the conditional expectation of distortion.

$$\begin{array}{ll} \min & \sum_{n=1}^{N} \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2] \\ \text{s.t.} & \sum_{n=1}^{N} k_n\rho_n = \frac{P}{\sigma^2} \\ & \rho_n \geq 0,\ \rho \in \mathbb{R} \end{array} \tag{3}$$

where the expectation is taken over the channel fading gains

$$\begin{aligned} \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2] &= \int \mathbb{E}[\varepsilon_n|\mathbf{H}_n,\mathbf{k},\sigma^2] d(\mathbb{P}(\mathbf{H}_n)) \\ &= \int \mathbb{E}[\varepsilon_n|\mathbf{H}_n,\sigma^2] d(\mathbb{P}(\mathbf{H}_n)) \end{aligned} \tag{4}$$

A. Convexity

First, we shall prove that (3) is a constraint convex optimization problem.

Under independent Rayleigh fading, the fading gain $\|\mathbf{H}_n\|^2$ is a sum of k_n independent Chi-square random variables. It is Gamma distributed, i.e. $\|\mathbf{H}_n\|^2 \sim \Gamma(k_n, 1)$, the probability density function (pdf) is

$$p(k_n, x) = \frac{1}{\Gamma(k_n)} x^{k_n-1} e^{-x} \tag{5}$$

Hence,

$$\mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2] = \begin{cases} \int_0^\infty \frac{\lambda_n}{\rho_n t+1} p(k_n,t)dt & k_n \geq 1 \\ \lambda_n & k_n = 0 \end{cases} \tag{6}$$

$$\mathbb{E}[\varepsilon_n^2|\mathbf{k},\sigma^2] = \begin{cases} \int_0^\infty \frac{\lambda_n^2}{(\rho_n t+1)^2} p(k_n,t)dt & k_n \geq 1 \\ \lambda_n^2 & k_n = 0 \end{cases} \tag{7}$$

From Leibniz integral rule, we can derive the first order and second order partial derivative for $k_n \geq 1$.

$$\frac{\partial \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2]}{\partial \rho_n} = -\int_0^\infty \frac{\lambda_n t}{(\rho_n t+1)^2} p(k_n,t)dt \tag{8}$$

$$\frac{\partial^2 \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2]}{\partial \rho_n^2} = \int_0^\infty \frac{2\lambda_n t^2}{(\rho_n t+1)^3} p(k_n,t)dt \tag{9}$$

When $\rho_n \geq 0$, $\forall t > 0$, $\frac{2\lambda_n t^2}{(\rho_n t+1)^3} > 0$. Therefore, $\frac{\partial^2 \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2]}{\partial \rho_n^2} > 0$, i.e. $\mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2]$ is a strict convex function of ρ_n. Similarly, we can derive that $\mathbb{E}[\varepsilon_n^2|\mathbf{k},\sigma^2]$ is also a strict convex function of ρ_n. When $\varphi(x)$ is a convex function, (3) is a convex optimization problem.

B. Solution

In general, constraint convex optimization problem can be solved by Lagrangian method. However, we observe that the derivatives are integral functions and it is difficult to get the solution to the Lagrangian constraint equations. We propose to use gradient descent method to solve the problem as shown in Algorithm 1. For the sake of presentation, we only consider $\varphi(x) = 0$ in Algorithm 1 to illustrate the main idea. The algorithm for other $\varphi(x)$ can be easily derived.

In the description of Algorithm 1, we denote $*$ as the element wise multiply operation and $\cdot$ as the inner product

Algorithm 1: Optimal power allocation

Data: $\boldsymbol{\lambda}$, $\mathbf{k}$, P, σ^2
Result: ρ_n^+, $(n = 1, 2, \ldots, N)$
1 Initialization: $\rho_n^+ = \frac{P\sqrt{\lambda_n}}{\sigma^2 \sum_{i=1}^N k_i\sqrt{\lambda_i}} I(k_n)$;
2 **repeat**
3 $\quad \boldsymbol{\kappa} = \mathbf{k} * I(\boldsymbol{\rho}^+)$;
4 $\quad \boldsymbol{\omega} = -\frac{\partial \mathbb{E}[\varepsilon_n|\boldsymbol{\kappa},\sigma^2]}{\partial \boldsymbol{\rho}}\big|_{\boldsymbol{\rho}=\boldsymbol{\rho}^+}$;
5 $\quad \boldsymbol{\theta} = \alpha_1(\boldsymbol{\omega} - \frac{\boldsymbol{\kappa}\cdot\boldsymbol{\omega}}{\|\boldsymbol{\kappa}\|^2}\boldsymbol{\kappa})$;
6 $\quad \boldsymbol{\rho}^+ = \alpha_2 \max\{\boldsymbol{\rho}^+ + \delta\boldsymbol{\theta},\ \mathbf{0}\}$;
7 $\quad D^{(t)} = \sum_{n=1}^N \mathbb{E}[\varepsilon_n|\boldsymbol{\kappa},\sigma^2]$;
8 **until** $\|D^{(t)} - D^{(t-1)}\| < \xi$;

operation. $I(x)$ is an indicator function that $I(x) = 1$ if $x \neq 0$ otherwise $I(x) = 0$. At first, ρ_n is initialized with the optimal power allocation for AWGN channel. During the iteration, the power allocation should be evolved through the direction $\boldsymbol{\theta}$ that has the largest projection [19] on gradient descent direction $\boldsymbol{\omega}$.

$$\begin{array}{ll} \max & \sum_{i=1}^{N} \omega_i\theta_i \\ \text{s.t.} & \sum_{i=1}^{N} \kappa_i\theta_i = 0 \\ & \sum_{i=1}^{N} \theta_i^2 = 1 \end{array} \tag{10}$$

Through Lagrangian method, we can derive a closed form of the direction.

$$\theta_n = \alpha_1(\omega_n - \kappa_n \frac{\sum_{i=1}^N \kappa_i\omega_i}{\sum_{i=1}^N \kappa_i^2}) \tag{11}$$

where $\alpha_1 > 0$ is the normalization factor that $\sum_{i=1}^N \theta_i^2 = 1$.

It should be noted that the power is a non-negative scalar and should be lower bound by zero. Due to this bounding, normalization will be performed if the power of any chunk is evolved below zero. α_2 is the normalization factor that $\alpha_2 \sum_{i=1}^N \kappa_i\rho_i = \frac{P}{\sigma^2}$. In addition, we virtually set the corresponding channel allocation to zero by line 3 so as to stop evolving for the zero transmission power chunk.

Although the Algorithm 1 could converge in limited number of iterations, the integral computation for the expectation and its partial derivative becomes the bottleneck. Through analysis, we find an efficient way to compute the integral and eliminate the concerns on computation cost.

Let's define

$$\psi(k,x) = \int_0^\infty \frac{1}{t+x} t^k e^{-t} dt \tag{12}$$

and

$$\eta(k,x) = \int_0^\infty \frac{1}{(t+x)^2} t^k e^{-t} dt \tag{13}$$

Based on these definition, (6) can be written as

$$\begin{aligned} \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2] &= \begin{cases} \int_0^\infty \frac{\lambda_n}{\rho_n t+1} \frac{1}{\Gamma(k_n)} t^{k_n-1} e^{-t} dt & k_n \geq 1 \\ \lambda_n & k_n = 0 \end{cases} \\ &= \begin{cases} \frac{\lambda_n}{\Gamma(k_n)} \frac{1}{\rho_n} \psi(k_n - 1, \frac{1}{\rho_n}) & k_n \geq 1 \\ \lambda_n & k_n = 0 \end{cases} \end{aligned} \tag{14}$$

The partial derivative is

$$\frac{\partial \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma]}{\partial \rho_n} = \begin{cases} \frac{\lambda_n}{\Gamma(k_n)}(-\frac{1}{\rho_n^2})\eta(k_n,\frac{1}{\rho_n}) & k_i \geq 1 \\ 0 & k_i = 0 \end{cases} \tag{15}$$

Functions $\psi(k,x)$ and $\eta(k,x)$ can be implemented by lookup tables to save the computational cost. For each integer number k, the key (quantized x) of the lookup table is selected so that the discrete function values are with an interval 0.001. When looking up the table, the input x is rounded to the nearest key and the corresponding function value is obtained. The lookup tables for $k = 1, \ldots, 100$ can be stored in 1 MB memory.

V. Channel Allocation

The second sub-problem is to find the optimal channel allocation based on the solution of power allocation. In solving the problem, we consider both unicast or multicast scenarios.

A. Unicast

For unicast, the noise power can be estimated and fed back to the sender. Based on the noise power, the sender looks for the optimal channel assignment to combat the variation caused by fast fading channel.

$$\begin{array}{ll} \min & \sum_{n=1}^{N} \mathbb{E}[\varepsilon_n|\sigma^2] \\ \text{s.t.} & \sum_{n=1}^{N} k_n = M \\ & k_n \geq 0,\ k_n \in \mathbb{Z} \\ & \sum_{n=1}^{N} k_n \rho_n = \frac{P}{\sigma^2} \\ & \rho_n \geq 0,\ \rho_n \in \mathbb{R} \end{array} \tag{16}$$

As mentioned before, in the power allocation procedure, the power for some chunks may be evolved to zero. This means that the channels for these chunks can be assigned to other chunks to increase their diversity gain. Although the power allocation indicates whether or not there exists available channels, where to assign these channels still needs an exhaustive search. To avoid the complexity of testing all the combinations, we propose a greedy algorithm based on the asymptotic analysis of Equation (4).

According to the definition in (12), we find that $\psi(k,x)$ is an recursive function regarding k. When $k \geq 1$, we can derive

$$\begin{aligned} \psi(k,x) &= \int_0^\infty \frac{t}{t+x} t^{k-1} e^{-t} dt \\ &= \int_0^\infty t^{k-1} e^{-t} dt - x \int_0^\infty \frac{1}{t+x} t^{k-1} e^{-t} dt \\ &= \Gamma(k) - x\psi(k-1,x) \end{aligned} \tag{17}$$

where $\Gamma(\cdot)$ is the Gamma function. When $k = 0$, we have

$$\begin{aligned} \psi(0,x) &= \int_0^\infty \frac{1}{t+x} e^{-t} dt = e^x \int_x^\infty \frac{1}{t} e^{-t} dt \\ &= -e^x Ei(-x) \end{aligned} \tag{18}$$

where $Ei(-x)$ is the exponential integral function. It can be represented by two series according to the value of x. For small x, by integrating the Taylor series,

$$-e^x Ei(-x) = -e^x \left[\gamma + \log x + \sum_{m=1}^{+\infty} \frac{(-x)^m}{m\Gamma(m+1)} \right] \tag{19}$$

Algorithm 2: Greedy Algorithm for Channel Allocation

Data: $\boldsymbol{\lambda}$, P, σ^2
Result: k_n^+, ρ_n^+ $(n = 1, 2, \ldots, N)$
Initialization: $\mathcal{S}_t = \{1\}$, $n_d = N$, $k_n^+ = 1$;
compute ρ_n^+ and $D^{(0)}$ by Algorithm 1 with k_n^+;
repeat
  for $n_t \in \mathcal{S}_t$ **do**
    set $\boldsymbol{\kappa} = \mathbf{k}^+$, $\kappa_{n_t} = \kappa_{n_t} + 1$ and $\kappa_{n_d} = \kappa_{n_d} - 1$;
    compute $\boldsymbol{\rho}$ and D by Algorithm 1 with $\boldsymbol{\kappa}$;
  end
  $D^{(t)} = \min\{D\}$;
  if $D^{(t)} < D^{(t-1)}$ **then**
    $\mathbf{k}^+ = \boldsymbol{\kappa}$ with $\min\{D\}$;
    $\boldsymbol{\rho}^+ = \boldsymbol{\rho}$ with $\min\{D\}$;
    if $k_{n_d}^+ == 0$ **then**
      $n_d = n_d - 1$
    end
    $\mathcal{S}_t = \{1\} \bigcup \{n | k_n^+ < k_{n-1}^+, 2 \leq n < n_d\}$;
  end
until $D^{(t)} > D^{(t-1)}$;

where γ is the Euler-Mascheroni constant. For large x, there is a divergent approximation,

$$-e^x Ei(-x) \approx -\sum_{m=1}^{+\infty} \Gamma(m)(-x)^{-m} \tag{20}$$

Based on these two series, we can analyze the asymptotic property of (4) for small and large noise power, respectively. From the recursion in (17), $\psi(k,x)$ can be written as

$$\begin{aligned} &x\psi(k-1,x) = \Gamma(k) - \psi(k,x) \\ &= \Gamma(k) - (-x)^k(-e^x Ei(-x)) - \sum_{m=0}^{k-1} (-x)^m \Gamma(k-m) \end{aligned} \tag{21}$$

Let's define

$$\begin{aligned} \Psi(k,x) &\triangleq \frac{1}{\Gamma(k)} x\psi(k-1,x) \\ &= 1 - \frac{-e^x Ei(-x)}{\Gamma(k)}(-x)^k - \sum_{m=0}^{k-1} \frac{\Gamma(k-m)}{\Gamma(k)}(-x)^m \end{aligned} \tag{22}$$

When $x \to +\infty$,

$$\begin{aligned} \Psi(k,x) &\approx \sum_{m=1}^{+\infty} \frac{\Gamma(m)}{\Gamma(k)}(-x)^{k-m} - \sum_{m=1}^{k-1} \frac{\Gamma(k-m)}{\Gamma(k)}(-x)^m \\ &= 1 + \sum_{m=1}^{+\infty} \frac{\Gamma(k+m)}{\Gamma(k)}(-x)^m \\ &= 1 - \frac{k}{x} + o(\frac{1}{x^2}) \end{aligned} \tag{23}$$

When $x \to 0$, $e^x \approx 1$.

$$\Psi(k,x) \approx \frac{(-x)^k}{\Gamma(k)}\left[\gamma + \log x + o(x)\right] - \sum_{m=1}^{k-1}\frac{\Gamma(k-m)}{\Gamma(k)}(-x)^m = \frac{x}{\xi(k)} + o(x^2) \tag{24}$$

where

$$\xi(k) = \begin{cases} \frac{1}{-\log x - \gamma} & k = 1 \\ k-1 & k \geq 2 \end{cases}$$

is an increasing function of k.

Since (6) can be represented by

$$\mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2] = \lambda_n\left(1 - \frac{\psi(k_n, 1/\rho_n)}{\Gamma(k_n)}\right) = \lambda_n \Psi(k_n, 1/\rho_n) \tag{25}$$

Therefore, $\forall k_n \geq 1$, we have

$$\mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2] \approx \begin{cases} \frac{\lambda_n}{\xi(k_n)\rho_n} & \sigma^2 \to 0 \\ \lambda_n(1-k_n\rho_n) & \sigma^2 \to \infty \end{cases} \tag{26}$$

We have assumed that the mean energy of the chunks are in descending order, i.e. $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. For both cases, to achieve the minimum sum of mean distortion with the constraint $\sum_{n=1}^N k_n\rho_n = \frac{P}{\sigma^2}$, the product of $k_n\rho_n$ should be in descending order too, i.e. $k_1\rho_1 \geq k_2\rho_2 \geq \ldots \geq k_N\rho_N$. This condition could be easily achieved if the channel uses are in descending order, i.e. $k_1 \geq k_2 \geq \ldots \geq k_N$. Therefore, we develop a greedy algorithm for problem (16) which satisfies this strict constraint.

The algorithm is described in Algorithm 2. The channel allocation is evolved from the initial value of one channel use per coefficient. In each iteration, we take one channel use from the least important chunk being transmitted and try to assign it to more important chunks. In this re-allocation procedure, we keep k_n's in descending order. With this constraint, the number of possible positions to allocate the channel use is dramatically decreased. The distortions of all possible choices are evaluated, and the one with the minimum distortion will be selected as the result of current iteration. We continue this process until the minimum distortion does not decrease any more.

B. Multicast

For multicast, the noise power is heterogeneous among different receivers. The channel allocation should consider the overall performance rather than the performance of a specific receiver. The heterogeneous condition is modelled as a discrete set of noise power, the distribution of which is known to the sender. The problem can be formulated as:

$$\begin{array}{ll} \min & \sum_{n=1}^N \mathbb{E}[\varepsilon_n] \\ \text{s.t.} & \sum_{n=1}^N k_n = M \\ & k_n \geq 0,\ k_n \in \mathbb{Z} \\ & \sum_{n=1}^N k_n g_n^2 \lambda_n = P \\ & g_n \geq 0,\ g_n \in \mathbb{R} \end{array} \tag{27}$$

Algorithm 3: Channel Allocation for Multicast

Data: $\boldsymbol{\lambda}$, P, $\mathcal{S} = \{\sigma^2|\mathbb{P}(\sigma^2)\}$
Result: k_n^+, ρ_n^+ $(n = 1, 2, \ldots, N)$
for $\sigma^2 \in \mathcal{S}$ **do**
  find $\mathbf{k}$ and $\boldsymbol{\rho}$ by Algorithm 2;
  compute $\mathbb{E}[D] = \sum_{\sigma^2}\mathbb{P}(\sigma^2)\sum_{n=1}^N \mathbb{E}[\varepsilon_n|\mathbf{k},\sigma^2]$;
end
$\mathbf{k}^+ = \mathbf{k}$ with min $\mathbb{E}[D]$;
$\boldsymbol{\rho}^+ = \boldsymbol{\rho}$ with min $\mathbb{E}[D]$;

Based on Algorithm 2, the channel allocation problem for multicast session can be easily solved by Algorithm 3. Please note that, the solution can be extended to other applications, such as video pricing, by simply replacing the optimization objective.

VI. Evaluation

We carry out the evaluation based on monochrome CIF video sequences, with a resolution of 352×288, and the frame rate 30 fps (frame per second). A simple calculation would tell that the source bandwidth is 1.52 MHz (in complex symbols). Similar to the test sequence used in SoftCast [7], our test sequence is composed of multiple standard video test sequences including 'akiyo', 'bus', 'coastguard', 'crew', 'football', 'foreman', 'news', 'harbour','husky','ice'. They are indexed from 1 to 10.

Since this paper focuses on the transmission for uncoded video, SoftCast is considered as a reference system. With sufficient bandwidth, Softcast does not drop any chunks, regardless of the receiver's noise power. When such information is available at the sender, we extend Softcast such that the sender drops less important chunks according to the noise power using the formulas given in [12]. We name this extension as Softcast+. For a fair comparison, all systems use GOP size 4 and equal chunk division with 8×8 chunks per frame. We use Rayleigh fading channel model with unit mean power for the evaluation, i.e. $h_m \sim \mathcal{CN}(0,1)$ and adopt AWGN model to generate random noise. The number of available channel uses is assumed to be equal to source bandwidth.

We use the objective video quality assessment, peak signal-to-noise ratio (PSNR) in dB, as the evaluation metric. The PSNR of a frame is $\text{PSNR} = 10\log_{10}\frac{255^2}{\text{MSE}}$, where MSE is the mean squared error over all pixels in a frame. The PSNR of a sequence is the average PSNR over all frames.

A. Micro-Benchmark

This subsection is devoted to explain the observations and conclusions drawn from analysis in previous sections. Due to space limit, we only present results of sequence 'foreman' on channel SNR 10dB. Similar results are obtained for other video sequences and on other channel SNR values.

1) Power allocation: This experiment shows our resource allocation results and demonstrates the difference between the proposed algorithm and Softcast in power allocation. Fig. 2 contains four sub-figures. The first sub-figure depicts the channel assignment for each chunk. To ensure stability, the

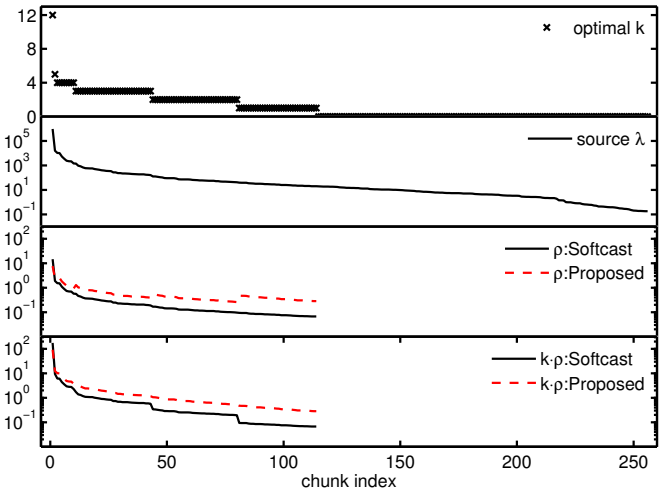


Fig. 2. Resource allocation result of the proposed scheme and the comparison with Softcast

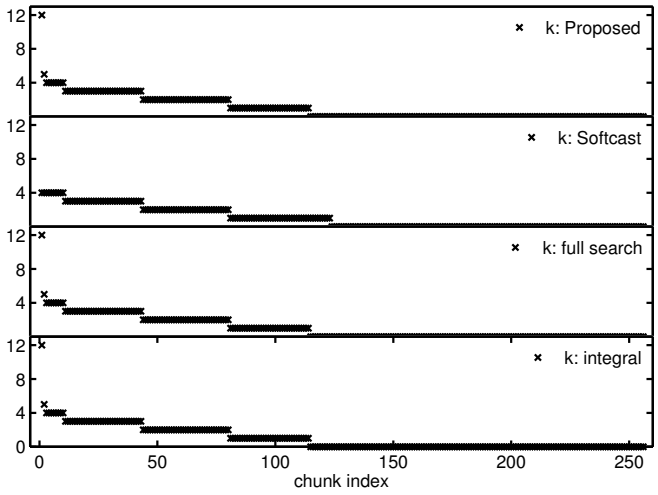


Fig. 3. Simplifications in channel allocation algorithm and implementation do not incur loss.

first and second chunk of a GOP are allocated with 12 and 5 transmission opportunities, respectively. It is clear that the channel allocation possesses a stair-step shape. Around half of the chunks are not assigned transmission opportunity in order to save the channel uses for HP chunks. Although not shown in the figure, experiments over other channel SNRs show that more chunks are discarded under poorer channel conditions and less chunks are discarded under better channel conditions.

The second sub-figure plots the energy (or variance) of each chunk in descending order. The third and fourth sub-figures compare the per-slot energy allocation (ρ) and per-coefficient energy allocation ($k\rho$) between our scheme and Softcast. An obvious difference between the two schemes is that Softcast keeps ρ proportional to λ while our scheme keeps $k\rho$ proportional to λ. Intuitively, when an HP chunk has diversity gain, its overall energy expenditure could be reduced, and the saved power could be allocated to LP chunks to improve the overall performance. This experiment shows that our power allocation result is in good agreement with the intuition.

2) Channel allocation: The channel allocation results shown in Fig. 3 note two observations. First, power allocation algorithm has a perceptible impact on the channel allocation algorithm. This supports our claim that channel allocation should be jointly designed with power allocation. The second sub-figure is obtained by replacing our power allocation with Softcast power allocation in Algorithm 2. Softcast power allocation neglects the power saving benefits of diversity gain, and continues to allocate a large amount of power to HP chunks even if they have high diversity gain. As such, the power expenditure on LP chunks will be dramatically decreased. As a result, the channel allocation algorithm becomes more conservative in increasing the diversity gain of HP chunks. Comparing the first and the second sub-figure, we can find that by using Softcast power allocation, less LP chunks are discarded, and the first a few HP chunks are only allocated with 4 channel use.

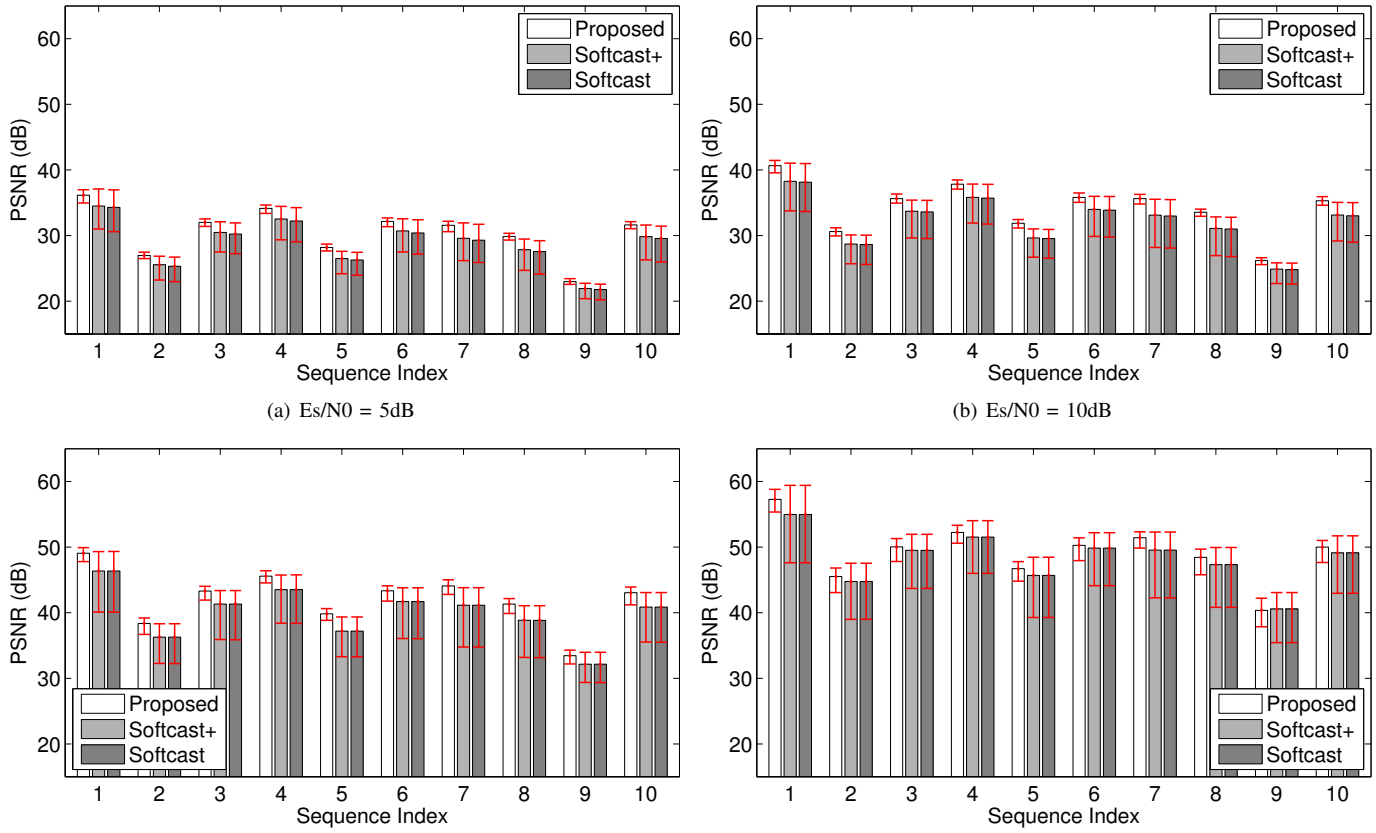
The second observation is that the simplifications we made in channel allocation algorithm and implementation have a negligible impact on the results. In the proposed channel allocation algorithm (Algorithm 2), we progressively increase the diversity gain for HP chunks, and the number of candidate chunks in each iteration is limited by the proposed constraint $k_1 \geq k_2 \geq ... \geq k_N$. In this experiment, we test the full search option and present the result in the third sub-figure. In addition, we have implemented the integral computation of expected distortion and the first derivative of expected distortion through lookup tables. We evaluate the effect of such approximation through comparing the result with using precise computation. Almost identical channel allocation result in the first and fourth sub-figures shows that using lookup tables do not incur any loss.

B. System Comparison

We implement a video transmission system based on the proposed resource allocation algorithms, and compare its performance against Softcast and Softcast+. Both Unicast and Multicast scenarios are considered.

1) Unicast: For Unicast transmission, it is assumed that the sender has the information of receiver noise power. At typical signal-to-noise ratios, including 5dB, 10dB, 20dB and 30dB, we carry out 1000 test runs for each GOP of the test videos. Then the average PSNR and the PSNR fluctuation range are computed and compared among our system, Softcast and Softcast+. The PSNR fluctuation range is defined by the 5^{th} and 95^{th} percentiles of the achieved PSNR in 1000 test runs. The 5^{th} percentile PSNR is also called the *outage PSNR* as an indication to the robustness of a video transmission system.

From the results shown in Fig. 4, we find that our system not only achieves better average performance but is more stable than Softcast and Softcast+. Softcast+ only has marginal performance improvement over Softcast. At the four typical SNRs, our system achieves 1.9 dB, 2.2 dB, 2.2 dB and 0.9 dB gain over Softcast in average PSNR, respectively. In addition, our system achieves striking 4.2 dB, 5.2 dB, 5.9 dB and 5.0 dB gain over Softcast in outage PSNR.



(a) Es/N0 = 5dB

(b) Es/N0 = 10dB

(c) Es/N0 = 20dB

(d) Es/N0 = 30dB

Fig. 4. Comparing the video quality and stablity with Softcast and Softcast+.

2) Multicast: Multicast is an important application scenario of wireless video transmission. In a multicast session, receivers have different receive noise power. Therefore, it is impossible to optimize the performance for all receivers. This experiment evaluates the performance tradeoffs achieved among different receivers. To run our system and Softcast+, we would set a target SNR and allocate resources accordingly. Then, the performance at receives whose actual SNRs range from 5dB to 30dB are evaluated.

Table II provides the average PSNR and the variance of PSNR (over 10 test sequences) achieved by the three comparing systems (index 1, 2 and 3 denotes the proposed system, Softcast and Softcast+ respectively). As Softcast is a one-size-fits-all design, its result only occupies one row in the Table. As a matter of fact, Softcast assumes zero noise power at the receiver, so its performance is almost identical to Softcast+ when the latter's target SNR is 30dB. It can be seen from the Table that our system achieves a much smaller variance than the other two systems while the average PSNR performance is superior. When the target SNR is 30dB, our system achieves 1.35dB, 1.64dB, 1.70dB and 0.94dB gain in average PSNR over Softcast for receivers with actual SNR of 5dB, 10dB, 20dB and 30dB. The gain in outage PSNR is as large as 5.3 dB on average. Overall, our system achieves the design goal of robust and efficient video transmission.

TABLE II. THE MEAN AND VARIANCE PAIR OF VIDEO PSNR ACHIEVED BY THREE COMPARING SYSTEMS (1: PROPOSED, 2: SOFTCAST, 3: SOFTCAST+)

	Target	Actual Es/N0			
		5dB	10dB	20dB	30dB
1	5dB	30.56, 0.32	33.81, 0.26	37.55, 0.10	38.35, 0.03
	10dB	30.48, 0.39	34.30, 0.35	40.18, 0.21	42.11, 0.07
	20dB	29.96, 0.62	33.95, 0.62	42.14, 0.59	48.11, 0.41
	30dB	30.04, 0.76	33.76, 0.79	41.63, 0.83	49.23, 0.96
2	any	28.69, 1.69	32.12, 1.79	39.93, 2.00	48.29, 2.33
3	5dB	28.94, 1.64	32.22, 1.69	39.12, 1.62	44.38, 1.14
	10dB	28.82, 1.68	32.23, 1.76	39.86, 1.91	47.50, 1.97
	20dB	28.70, 1.69	32.14, 1.79	39.94, 1.99	48.28, 2.32
	30dB	28.69, 1.69	32.13, 1.79	39.93, 2.00	48.29, 2.33

VII. RELATED WORK

A. Uncoded video transmission systems

The pioneer work Softcast arouses a surge of interest in uncoded video transmission systems. Follow-up works include ParCast [4], Dcast [9], [5], and Cactus [6]. These systems share the same core modules with Softcast, but differ from each other in the way to remove or utilize the source redundancy. Among these works, Parcast is the most related one to our research. It concerns the uncoded video transmission across MIMO-OFDM channel. The basic idea is to separate the MIMO-OFDM channel into a set of orthogonal subchannels and then match the more important source (belonging to a Gaussian distribution with larger variance) to higher gain subchannel. Let s_i be the i^{th} largest channel gain, then the optimal power allocation is to scale each coefficient with a factor proportional to $(\lambda_i s_i^2)^{-\frac{1}{4}}$. Although Parcast considers fading channels, it

assumes availability of transmit CSI and therefore can only be applied to Unicast communications.

B. Theoretical work on analog JSCC

There has been a large body of theoretical work on analog JSCC, among which linear coding and single-letter coding have attracted the most attention for their simplicity and low delay. At the intersection of these two coding schemes, there is the uncoded transmission strategy in which no compression or channel coding is used and the source samples are transmitted by appropriately scaling according to the transmitter power constraint P. Surprisingly, such a simple strategy has been shown to achieve optimality in certain practical cases [10], [11]. A famous example is to transmit a uniform-distributed binary source with Hamming distance distortion metric over a binary symmetric channel. Another one is to transmit a Gaussian source with squared-error distortion metric over an AWGN channel.

The uncoded transmission strategy concerned in this research is slightly different from the above two examples, because the source under consideration is modeled by mixed-Gaussian. Source samples belonging to different Gaussian distributions should be scaled by different factors to achieve optimality. Lee and Petersen [12] have carried out comprehensive investigation on this optimal linear coding problem, and Softcast and Parcast are straightforward implementations based on their obtained results. Unfortunately, Lee and Petersen do not consider the case when the fading channel parameter is not available at the transmitter.

Considering the fading channels, and under a more realistic assumption that the receiver has perfect CSI while the transmitter has only statistical information about the channel state, Kashyap et al. [13] have derived the performance of linear coding scheme. They find that, for the Rayleigh fading channel, while linear coding is suboptimal in general, it is close to optimal in the low SNR regime. Xiao et al. [14] have also considered the linear coding of a discrete memoryless Gaussian source transmitted through a discrete memoryless fading channel with AWGN. They show that among all single-letter codes, linear coding achieves the smallest MSE. However, these works do not consider the mixed-Gaussian source.

Although there are other non-linear JSCC strategies [15] and *semianalog* strategies [16], [17] that have shown near-optimal performance, we focus this research on linear analog (near-uncoded) coding strategies for their simplicity. In particular, we are interested in transmitting mixed-Gaussian source over fast fading channels, a problem rooted from wireless video communications.

VIII. Conclusion

This paper presents the design of robust uncoded video transmission over fast fading channels. Based on the observation that increasing the channel diversity of HP data is essential to increasing the transmission robustness, we have designed a channel allocation algorithm under total resource constraint. Further, we have found that channel allocation and power allocation should be jointly performed to achieve the optimum performance. In future, we plan to study the resource allocation problem in MIMO fast fading channel.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2011-2016," tech. rep., http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ ns827/white_paper_c11-520862.pdf, Feb 2012.

[2] C. E. Shannon, "A Mathematical Theory of Communication," *Bell system technical journal*, vol. 27, 1948.

[3] S. Jakubczak and D. Katabi, "A Cross-layer Design for Scalable Mobile Video," in *Proceedings of ACM Mobicom'11*, pp. 289–300, ACM, 2011.

[4] X. L. Liu, W. Hu, Q. Pu, F. Wu, and Y. Zhang, "ParCast: Soft Video Delivery in MIMO-OFDM WLANs," in *Proceedings of ACM Mobicom'12*, pp. 233–244, ACM, 2012.

[5] X. Fan, F. Wu, D. Zhao, and O. Au, "Distributed wireless visual communication with power distortion optimization," *IEEE Transactions on CSVT*, vol. 23, no. 6, pp. 1040–1053, 2013.

[6] H. Cui, Z. Song, Z. Yang, C. Luo, R. Xiong, and F. Wu, "Cactus: A Hybrid Digital-Analog Wireless Video Communication System," in *Proceedings of ACM MSWiM'13*, ACM, 2013.

[7] S. Jakubczak, J. Sun, D. Katabi, and V. Goyal, "Performance Regimes of Uncoded Linear Communications Over AWGN Channels," in *Proceedings of the 45th CISS*, pp. 1–6, 2011.

[8] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, June 2005.

[9] X. Fan, F. Wu, D. Zhao, O. Au, and W. Gao, "Distributed soft video broadcast (dcast) with explicit motion," in *Proceedings of IEEE Data Compression Conference*, pp. 199–208, 2012.

[10] T. Goblick Jr, "Theoretical limitations on the transmission of data from analog sources," *Information Theory, IEEE Transactions on*, vol. 11, no. 4, pp. 558–567, 1965.

[11] M. Gastpar, B. Rimoldi, and M. Vetterli, "To Code, or not to Code: Lossy Source-Channel Communication Revisited," *Information Theory, IEEE Transactions on*, vol. 49, no. 5, pp. 1147–1158, 2003.

[12] K.-H. Lee and D. Petersen, "Optimal linear coding for vector channels," *IEEE Transactions on Comm.*, vol. 24, no. 12, pp. 1283–1290, 1976.

[13] A. Kashyap, T. Basar, and R. Srikant, "Minimum distortion transmission of gaussian sources over fading channels," in *Proceedings of the 42nd IEEE Conf. on Decision and Control*, vol. 1, pp. 80–85, IEEE, 2003.

[14] J.-J. Xiao, Z.-Q. Luo, and N. Jindal, "Linear Joint Source-Channel Coding for Gaussian Sources through Fading Channels," in *Proceedings of IEEE Globecom'06*, pp. 1–5, 2006.

[15] G. de Oliveira Brante, R. Demo Souza, and J. Garcia-Frias, "Analog Joint Source-Channel Coding in Rayleigh Fading Channels," in *Proceedings of IEEE ICASSP'11*, pp. 3148–3151, 2011.

[16] Y. Kochman and R. Zamir, "Analog Matching of Colored Sources to Colored Channels," *Information Theory, IEEE Transactions on*, vol. 57, no. 6, pp. 3180–3195, 2011.

[17] U. Mittal and N. Phamdo, "Hybrid digital-analog (hda) joint source-channel codes for broadcasting and robust communications," *Information Theory, IEEE Transactions on*, vol. 48, no. 5, pp. 1082–1102, 2002.

[18] R. Xiong, F. Wu, J. Xu, and W. Gao, "Performance Analysis of Transform in Uncoded Wireless Visual Communication," in *Proceedings of IEEE ISCAS'13*, 2013.

[19] J. B. Rosen, "The Gradient Projection Method for Nonlinear Programming. Part I. Linear Constraints," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 1, pp. pp. 181–217, 1960.