POWER-DISTORTION OPTIMIZATION FOR WIRELESS IMAGE/VIDEO SOFTCAST BY TRANSFORM COEFFICIENTS ENERGY MODELING WITH ADAPTIVE CHUNK DIVISION

Ruiqin Xiong[1], Feng Wu[2], Xiaopeng Fan[3], Chong Luo[2], Siwei Ma[1], Wen Gao[1]

[1] Institute of Digital Media, Peking University, Beijing 100871, China
[2] Microsoft Research Asia, Beijing 100080, China
[3] Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China

ABSTRACT

Traditional communication systems usually suffer from the threshold effect when channel signal-to-noise ratio (CSNR) fluctuates unpredictably in wireless and mobile scenarios. The SoftCast scheme, however, provides graceful quality transition in wide CSNR range. In SoftCast, input image is decorrelated by a transform and modulated directly to a dense constellation for transmission, leaving out the conventional quantization, entropy coding and channel coding. A key point of SoftCast is that the transmission power needs to be allocated among the transform coefficients unequally, according to the energy of coefficients. Importantly, the energy diversity used to guide power allocation should be shared between the sender and the receiver for correct decoding. This paper addresses the power distortion optimization problem, introducing a new adaptive chunk division scheme to describe the energy diversity among coefficients. A concrete algorithm is developed to determine the chunk boundaries that achieve optimal transmission power usage. Experimental results show that the proposed scheme can improve the performance of the original SoftCast by 4∼8dB using a smaller number of chunks.

Index Terms**—** wireless visual communication, SoftCast, power allocation, transform coefficients modeling, optimization

1. INTRODUCTION

Communication system based on source and channel coding generally requires the channel statistics to be known at the time of encoding, in order to choose an appropriate coding rate. If the actual channel quality falls below a threshold, the decoding process tends to break down completely; if the channel quality increases beyond that threshold, such system cannot provide further improvement in the quality of received signal. This "threshold effect" brings great challenges for the design of wireless and mobile communication system.

Recently, a scheme named *SoftCast* [1, 2, 3] was proposed for wireless video. Unlike typical image or video coders that compress input signal into a binary stream, SoftCast transforms the image signal into a stream of real numbers from which exact reconstruction is possible, leaving out the conventional quantization and entropy coding. SoftCast also abandons the conventional channel coding. Instead, it modulates the number stream directly to a dense constellation for transmission. The transmission in SoftCast is lossy in nature and the noise level in the received numbers is commensurate with the channel signal-to-noise ratio (SNR). The most prominent advantage of SoftCast is that it provides graceful quality transition in very wide channel SNR range and can serve various clients of different channel conditions simultaneously, using the same transmitted signal in the air. For this reason, SoftCast has attracted much research attention in recent years [4, 5, 6, 7, 8, 9].

To achieve the best performance, SoftCast allocates transmission power among the transform coefficients unequally, by scaling each coefficient individually according to its energy. Importantly, the energy diversity used to guide power allocation should be shared between the sender and the receiver via meta data for correct decoding. To limit the overhead, SoftCast divides the coefficients into a set of chunks of the same size and perform scaling at chunk level. This turns out to be inefficient in terms of power usage. In this paper, we address the problem by introducing an adaptive chunk division scheme, exploiting the fact that the energy of transform coefficients decay rapidly from low frequency to high frequency. A concrete algorithm is developed to optimize the chunk division boundaries, based on a mathematical formulation of the overall performance of SoftCast. Experimental results indicate that the proposed approach can improve the performance of the original SoftCast scheme significantly.

The paper is organized as follows. Section 2 reviews the SoftCast scheme. Section 3 and 4 describe the proposed chunk division scheme and optimization algorithm. Section 5

Corresponding author: R. Xiong (rqxiong@pku.edu.cn). This work was supported in part by the National Natural Science Foundation of China (61073083, 61370114, 61121002), National Basic Research Program of China (973 Program, 2009CB320904), Beijing Natural Science Foundation (4112026, 4132039) and Research Fund for the Doctoral Program of Higher Education (20100001120027, 20120001110090).

shows experimental results and Section 6 concludes the paper.

2. REVIEW OF SOFTCAST

In a typical conventional visual communication system, as shown in Fig. 1, input image is compressed into a stream of bits, using transform, quantization and entropy coding. The bit stream is protected by some channel code and mapped to a constellation using quadrature amplitude modulation (QAM) (e.g. BPSK, 4-QAM, 16-QAM and 64-QAM) for OFDM transmission. In the case of 16-QAM, for example, four bits are extracted from the stream each time and mapped to one of the 16 candidate points in the constellation.

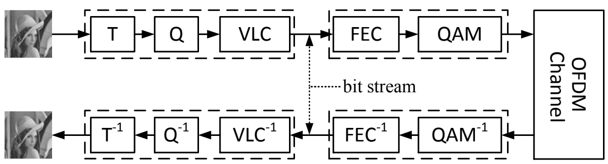


Fig. 1. Diagram of conventional visual communication.

In SoftCast, as shown in Fig. 2, the compression stage is solely a transform to decorrelate the image signal, producing a stream of transform coefficient numbers. The transmission stage scales each coefficient individually, applies a Walsh-Hardmard Transform (WHT) to whiten the whole stream, and modulates the resulted numbers directly to a dense constellation (e.g. 64k-QAM) for OFDM transmission. A pair of real numbers is extracted from the stream each time and mapped to a point in the dense QAM constellation, using the two numbers as the I- and the Q- components, respectively. The scaling operation serves the purposes of power allocation and unequal protection against channel noises. The scaling factors are determined by a power-distortion optimization (PDO) procedure, and will be shared by the SoftCast sender and receiver via a limited number of meta data.

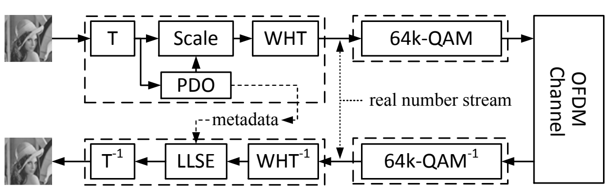


Fig. 2. Diagram of the SoftCast scheme [1, 2, 3].

3. POWER-DISTORTION OPTIMIZATION FOR SOFTCAST

Suppose $\mathbf{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$ are the coefficients to transmit. To achieve efficient power usage, the encoder scales each x_i by a factor g_i and sends out $y_i = g_i \cdot x_i$ using a dense constellation and OFDM[1]. After demodulation, the receiver gets $\hat{y}_i = y_i + n_i$, where n_i is additive white Gaussian noise (AWGN) with variance σ_n^2. The decoder gets an estimation of x_i by $\hat{x}_i = \hat{y}_i / g_i = x_i + n_i/g_i$.

In this process, the expected distortion in $\hat{x}_i$ is $D_i = \mathrm{E}[(\hat{x}_i - x_i)^2] = \sigma_n^2 / g_i^2$. The expected transmission power for sending x_i is $P_i = \mathrm{E}[y_i^2] = g_i^2 \cdot \mathrm{E}[x_i^2]$. Therefore, $D_i \cdot P_i = \sigma_n^2 \cdot E[x_i^2]$. To achieve optimal performance, the transmission power is allocated among $\{x_i\}$ by

$$\text{(P1):} \quad \text{minimize} \quad \sum_i D_i \quad \text{s. t.} \quad \sum_i P_i \leqslant P_{\text{total}} \qquad (1)$$

The problem is easily solved by setting $\partial D_i / \partial P_i$ to a constant. This eventually leads to $P_i \propto \sqrt{\mathrm{E}[x_i^2]}$ and $g_i \propto (E[x_i^2])^{-1/4}$. Using the optimal power allocation, the total distortion in the reconstructed image is

$$D_{\text{total}} = \sum_i D_i = \frac{\sigma_n^2}{P_{\text{total}}} \left(\sum_i \sqrt{\mathrm{E}[x_i^2]} \right)^2 \qquad (2)$$

Equivalently, we have

$$\text{PSNR}_{\text{dB}} = c + \text{CSNR}_{\text{dB}} - 10\log_{10} \left(\sum_i \sqrt{\mathrm{E}[x_i^2]} \right)^2 \qquad (3)$$

with $c = 10\log_{10}(255^2 N)$. For a general signal $\mathbf{x}$, We define the "activity" of $\mathbf{x}$ by $H(\mathbf{x}) = \sum_i \sqrt{E[x_i^2]}$.

4. ADAPTIVE CHUNK DIVISION

4.1. Why Chunk Division in SoftCast?

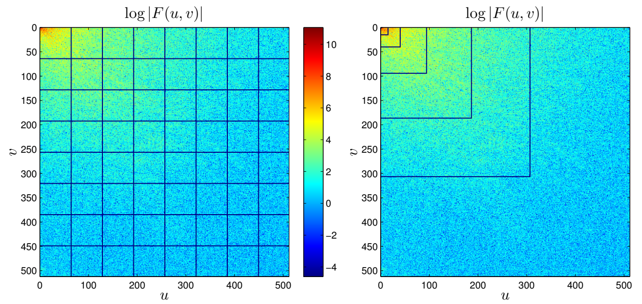
Ideally, to achieve optimal power usage, the scaling factors g_i should be selected individually according to $E[x_i^2]$. However, the receiver needs to know the scaling factors employed by the sender, for the purpose of correct decoding. Of course, sending one g_i for each coefficient may introduce significant communication overhead. Therefore, SoftCast groups the coefficients into a set of chunks and perform scaling at chunk level. In other words, all the coefficients in a chunk choose the same g_i value, based on the mean coefficient energy (i.e. $E[x_i^2]$) of that chunk. Fig. 3(a) illustrates the equal-size chunk division approach in the original SoftCast scheme. Typically, 64 chunks are used so that a total of 64 meta data is sent to the receiver via a reliable channel, to signal the value $E[x_i^2]$ of each chunk.

4.2. The Proposed Chunk Division Scheme

In this section, we consider the statistical characteristics of DCT coefficients and propose a new chunk division approach, for implementing efficient power allocation using a limited number of metadata.

Suppose $F(u, v)$ is the DCT domain representation of two-dimensional image $f(i, j)$. For most natural images, we

[1]The Walsh-Hardmard transform can be ignored during power-distortion analysis, because the WHT transform of a white noise is still a white noise.



(a) (b)

Fig. 3. Chunk division for SoftCast power allocation. (a) The scheme in original SoftCast. (b) The proposed scheme.

have the following observations (as shown in Fig. 3): (1) The amplitude of $F(u,v)$ decays rapidly from low frequency region to high frequency region, as the value of $||(u,v)||_2 = \sqrt{u^2+v^2}$ increases; (2) The amplitude of $F(u,v)$ varies with the angle of frequency (i.e. $\theta = \arctan(u/v)$) only slightly.

Based on these observations, we propose a chunk division scheme as illustrated by Fig. 3(b). To be precise, for identifying the chunk boundaries, we introduce a vector $\mathbf{r} = (r_1, r_2, \ldots, r_M)$, satisfying $0 = r_0 < r_1 < \cdots < r_{M-1} < r_M = 1$. For an $H \times W$ image, the vector $\mathbf{r}$ defines a set of similar rectangles $\mathbb{R}(r_i)$ in the (u,v)-plane:

$$\mathbb{R}(r) = \{(u,v) \mid 0 \leqslant u \leqslant r \cdot H,\ 0 \leqslant v \leqslant r \cdot W\} \tag{4}$$

Based on these rectangles, we construct a set of chunks $\mathbb{C}_i$, letting chunk $\mathbb{C}_i$ contain the coefficients in $\mathbb{R}(r_i)$ but not in $\mathbb{R}(r_{i-1})$. The proposed strategy allows chunks with nonequal sizes. For typical natural images, we prefer dense division at the up-left corner and sparse division at the bottom-right corner of the transform coefficient plane.

4.3. Chunk Division Optimization

Given the total number (denoted by M) of chunks to use, the parameters $r_i, i = 1, 2, \ldots, M-1$ should be optimized in order to achieve the best possible transmission performance. In this section, we propose an optimization algorithm for solving the parameters $\{r_i\}$.

According to Section 3, the problem is equivalent to finding the optimal values of $\{r_i\}$ that minimize the data "activity" $H(F)$ associated with the chunk division result constructed from $\{r_i\}$. To facilitate the optimization, we define $T_E(r)$ as the total energy of the coefficients in $\mathbb{R}(r)$:

$$T_E(r) = \sum_{(u,v)\in\mathbb{R}(r)} F(u,v)^2 \tag{5}$$

and $T_N(r)$ as the total number of the coefficients in $\mathbb{R}(r)$:

$$T_N(r) = \#\left\{(u,v) | (u,v) \in \mathbb{R}(r)\right\} \tag{6}$$

With the aid of $T_E(r)$ and $T_N(r)$, the mean energy of coefficients in chunk $\mathbb{C}_i$ can be formulated by

$$E[F(u,v)|(u,v)\in\mathbb{C}_i] = \sqrt{\frac{T_E(r_i) - T_E(r_{i-1})}{T_N(r_i) - T_N(r_{i-1})}}. \tag{7}$$

Therefore, the data activity associated with the chunk division becomes:

$$\begin{aligned} &H(F, \{r_i\}) \\ &= \sum_{i=1}^{M} \{T_N(r_i) - T_N(r_{i-1})\} \cdot E[F(u,v)|(u,v)\in\mathbb{C}_i] \\ &= \sum_{i=1}^{M} \sqrt{(T_N(r_i) - T_N(r_{i-1}))(T_E(r_i) - T_E(r_{i-1}))} \end{aligned} \tag{8}$$

Here we explicitly include $\{r_i\}$ as arguments of $H(\cdot)$ in (8) to reveal its dependence on the chunk division result. Note that $T_E(r)$ and $T_N(r)$ can be calculated beforehand. For the special case $r_0 = 0$, we have $T_E(r_0) = 0$ and $T_N(r_0) = 0$.

Before presenting our algorithm, we note that the coefficients $F(u,v)$ locate only at integer positions in the (u,v)-plane. Therefore, we only need to consider a finite set of values $\mathbb{V}$ for parameter r_i. Since an exhaustive full search for joint optimization of all the $M-1$ parameters can be computationally prohibitive, we employ a more practical iterative algorithm. In this algorithm, each parameter is firstly initialized by $r_i = i/M$ (and rounded to the nearest value in $\mathbb{V}$), which corresponds to equal division in each direction. Then, the algorithm iteratively updates the parameters $r_1, r_2, \ldots, r_{M-1}$ one by one, with only one of them being optimized while the others being fixed each time. This iteration process is terminated when a maximum number of iterations is reached or when the changes in $H(F, \{r_i\})$ is small enough.

Now we consider the sub-problem of choosing the optimal r_k for a particular k, with all the other parameters $\{r_i\}_{i\neq k}$ being fixed. This is to select a rectangle boundary that divides the region $\mathbb{R}(r_{k+1}) \setminus \mathbb{R}(r_{k-1})$ into two chunks (i.e. $\mathbb{C}_k$ and $\mathbb{C}_{k+1}$) so that $H(F, \{r_i\})$ is minimized. Since all the other chunks are unchanged, the optimization problem is reduced to choosing a value $r \in \mathbb{V}$, subject to $r_{k-1} < r < r_{k+1}$, that minimizes

$$\begin{aligned} H_{\text{sub}}^{(k)}(r) = &\sqrt{(T_N(r_{k+1}) - T_N(r))(T_E(r_{k+1}) - T_E(r))} \\ &+ \sqrt{(T_N(r) - T_N(r_{k-1}))(T_E(r) - T_E(r_{k-1}))}. \end{aligned} \tag{9}$$

This problem can be easily solved.

5. EXPERIMENTAL RESULTS

In this section, we conduct some experiments to evaluate the performance of the proposed chunk division scheme and compare it with the equal-size chunk division scheme in the original SoftCast. *Lena*, *Peppers*, *Elaine*, *Barbara*, *Baboon* and *Fishingboat* (512×512, gray) are used as test images.

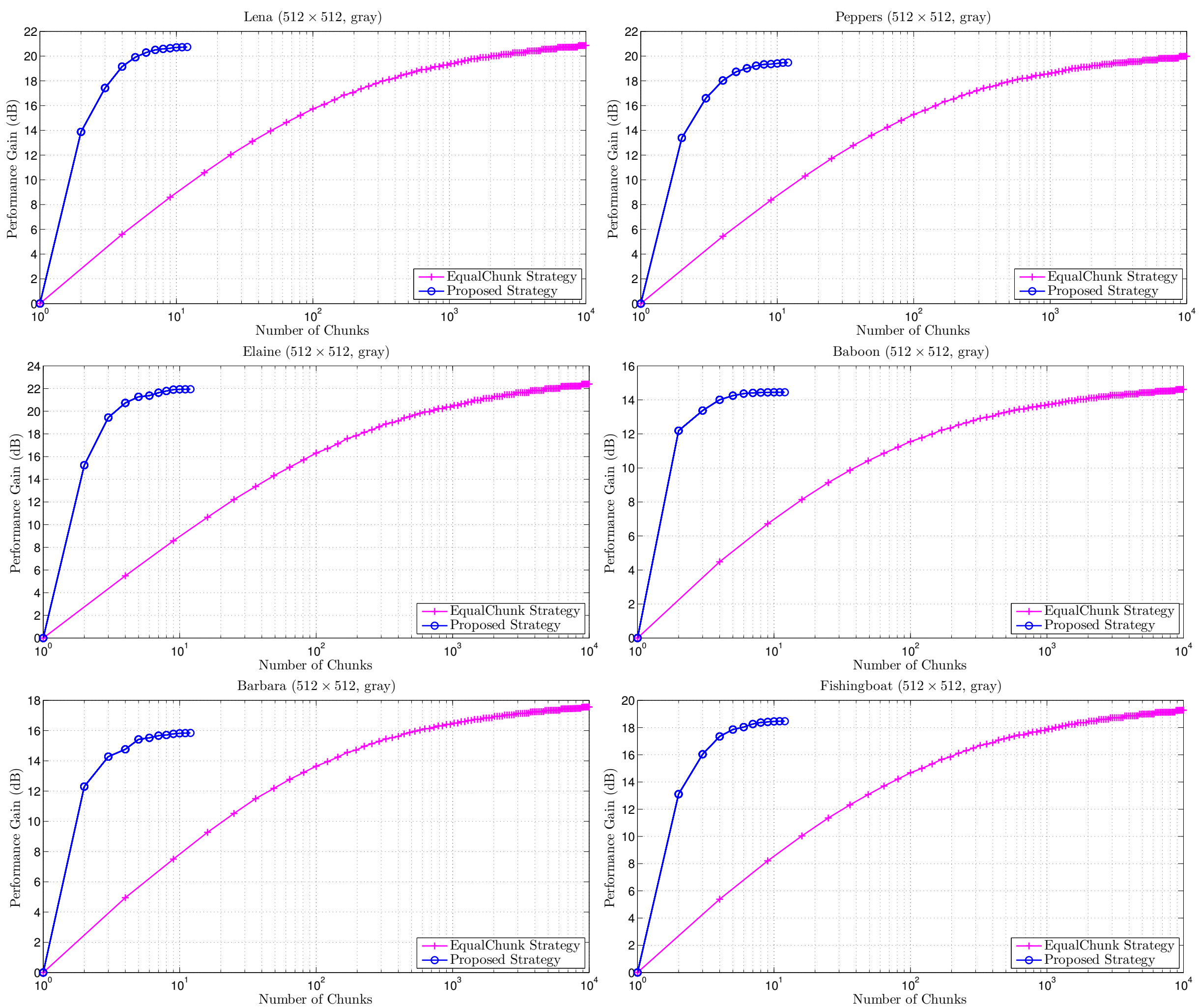


Fig. 4. Performance comparison between the proposed approach and the equal-chunk approach.

We first evaluate the benefit of employing chunk division schemes, based on the formulation (3). The benefit is measured by the performance gain in the reconstruction quality, compared with the case that all coefficients are transmitted without chunk division (i.e. using the same scaling factor), subject to the same channel SNR condition and the same channel usage. The results are shown in Fig. 4.

We first check the results of the equal-size chunk division scheme in original SoftCast. We can see that, as the number of chunks increases, a maximum chunk division gain of $15 \sim 25$dB can be achieved. The actual number of maximum chunk division gain is image dependent and reflects the strength of correlation in the image signal. We also note that, with equal-size chunk division, only about 40% and 70% of the whole gain can be achieved by using 10 and 10^2 chunks, respectively. To realize 95% of the whole gain, it requires more than 10^4 equal chunks. Such huge number of chunks (and meta data) is of course undesired. With the proposed adaptive chunk division scheme, however, employing only 2, 4 or 12 chunks can achieve a performance comparable to the original SoftCast using 10^2, 10^3 or 10^4 equal-size chunks. Therefore, the proposed chunk division approach is very efficient in differentiating the transform coefficients with different energies. Fig. 7 illustrates the chunk division results, produced by the proposed chunk division optimization algorithm.

Fig. 6 summarizes the simulated performance of SoftCast transmission using the proposed approach and the equal-size chunk approach. These simulation results confirm the results in Fig. 4 obtained from theoretical analysis in (3). Using as few as 12 chunks, the proposed strategy can achieve much better performance than the equal-size chunk approach using 256 chunks.

The reconstructed images are shown in Fig. 5 to provide a subjective performance comparison between the proposed approach and the equal-size chunk approach. Since the pro-

Fig. 5. Reconstructed images by the proposed chunk division and the equal-size chunk division at CSNR= 0dB. In this experiment, whole-frame DCT is used for decorrelation. The columns from left to right: equal-size chunk division with 16 chunks, equal-size chunk division with 64 chunks, equal-size chunk division with 256 chunks, and the proposed chunk division approach with 12 chunks.

posed approach works much better than the equal-size chunk approach, we consider a relatively low channel quality, i.e. CSNR= 0dB. In such case, the reconstruction images produced by the equal-size chunk approach are far from satisfactory. They contains very annoying artifacts and noises, especially for the case of using 64 chunks or less. The reconstruction images by the proposed approach using 12 chunks only, on the other hand, are much better. The noises remained in the decoded images are almost invisible.

6. CONCLUSIONS AND DISCUSSIONS

SoftCast is a flexible scheme for visual communication in wireless and mobile environment. It provides graceful quality degradation for very wide channel SNR range. A key issue in SoftCast is that the transmission power should be allocated among the coefficients unequally, according to the diversity in the energy of coefficients. This paper proposed a new chunk division scheme for SoftCast. Compared with the previous approach in the original SoftCast, the proposed scheme can describe the diversity in the energy of transform coefficients more accurately, using very limited number of meta data. Therefore, it facilitates efficient power allocation in SoftCast transmission. Experimental results show that the proposed method can improve both the objective and subjective quality significantly.

7. REFERENCES

[1] Szymon Jakubczak, Hariharan Rahul, and Dina Katabi, "Softcast: One video to serve all wireless receivers," in *MIT Technical Report, MIT-CSAIL-TR-2009-005*, 2009.

[2] Szymon Jakubczak, Hariharan Rahul, and Dina Katabi, "One-

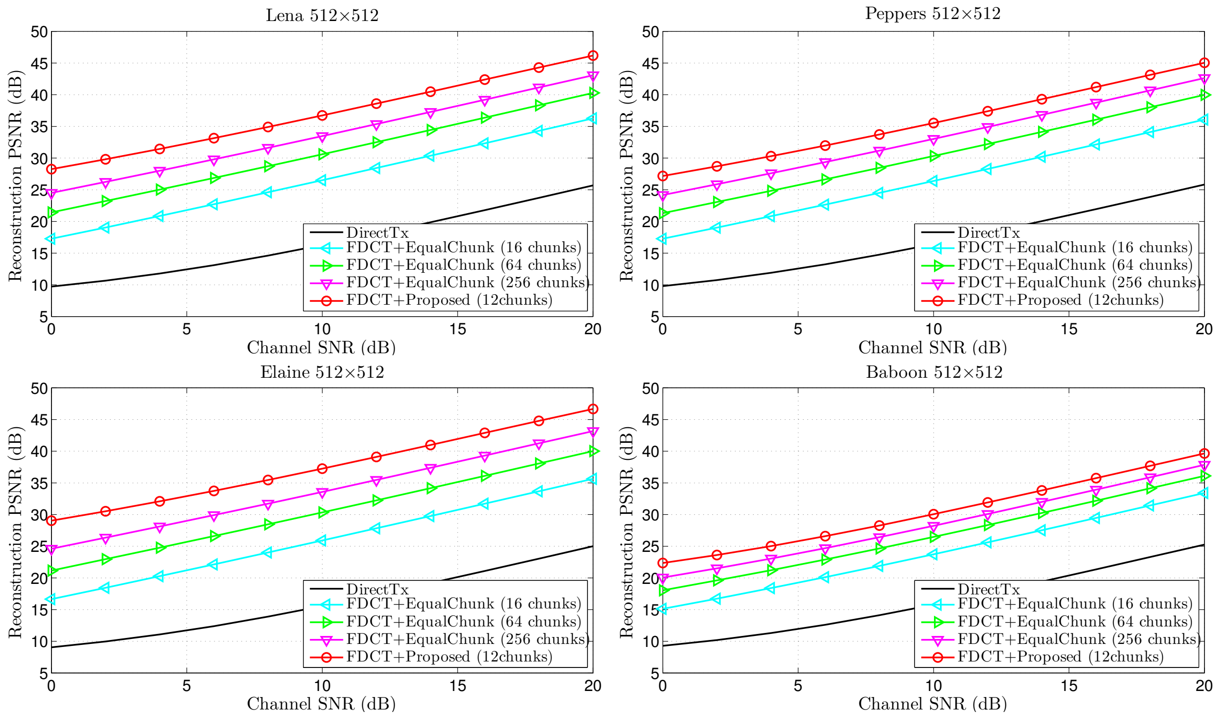


Fig. 6. Performance comparison between the proposed approach and the equal-chunk approach. Frame-DCT is considered here for decorrelation.

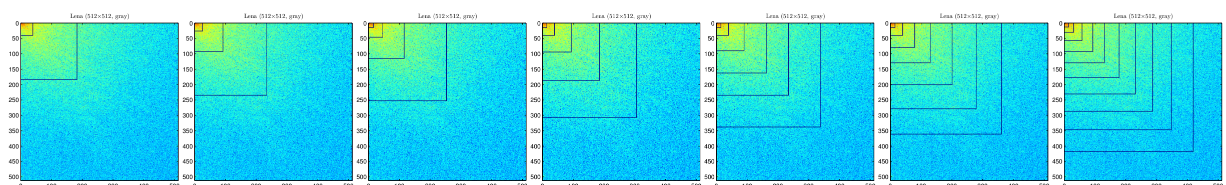


Fig. 7. Examples of chunk division results by the proposed scheme.

size-fits-all wireless video," in *The eigth ACM SIGCOMM HotNets Workshop*, New York, NY, USA, 2009.

[3] Szymon Jakubczak and Dina Katabi, "A cross-layer design for scalable mobile video," in *International conference on Mobile computing and networking (MobiCom '11)*, New York, NY, USA, 2011, pp. 289–300.

[4] Xiaopeng Fan, Feng Wu, Debin Zhao, C. Oscar Au, and Wen Gao, "Disributed soft video broadcast (dcast) with explicit motion," in *IEEE Data Compression Conference*, 2012, pp. 199–208.

[5] Xiaolin Liu, Wenjun Hu, Qifan Pu, Feng Wu, and Yongguang Zhang, "Parcast: soft video delivery in mimo-ofdm wlans," in *Internation Conference on Mobile Computing and Networking (Mobicom)*, 2012, pp. 233–244.

[6] Xiaopeng Fan, Ruiqin Xiong, Feng Wu, and Debin Zhao, "Wavecast: Wavelet based wireless video broadcast using lossy transmission," in *IEEE Visual Communications and Image Processing (VCIP)*, 2012, pp. 1–6.

[7] Ruiqin Xiong, Xiaopeng Fan, Feng Wu, and Wen Gao, "Analysis of uncoded image communication over noisy channels with unequal protection," in *SPIE Electronic Imaging 2013, Visual Information Processing and Communication IV*, Feb. 2013, p. 866603.

[8] Ruiqin Xiong, Feng Wu, Jizheng Xu, and Wen Gao, "Performance analysis of transform in uncoded wireless visual communication," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2013, pp. 1159–1162.

[9] Xiaopeng Fan, Feng Wu, Debin Zhao, and C. Oscar Au, "Distributed wireless visual communication with power distortion optimization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 6, pp. 1040–1053, 2013.