Transform Domain Energy Modeling of Natural Images for Wireless SoftCast Optimization

Zhihai Song†, Ruiqin Xiong†, Xiaopeng Fan‡, Siwei Ma† and Wen Gao†
†Institute of Digital Media, School of EE&CS, Peking University, Beijing 100871, China
‡Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China

***Abstract*—The SoftCast scheme recently proposed for wireless visual communication avoids the threshold effect that traditional communication systems usually suffer from. It provides graceful quality transition by sending images in a sequence of whitened transform coefficients using dense-constellation modulation and analog-like transmission. A key point in SoftCast is that it allocates transmission power among coefficients unequally, according to the expected energy of each coefficient. Importantly, the energy diversity utilized by power allocation should be shared with the receiver for correct decoding. Signaling the energy for each coefficient individually is prohibitive since it requires a large set of meta data. Grouping coefficients into a few chunks and signaling the energy at chunk level, on the other hand, may compromise the efficiency of SoftCast remarkably. In this paper, we investigate the energy distribution of natural images in transform domain and propose a model to approximate this distribution. We apply the model to guide the power allocation in SoftCast. Experimental results show that the proposed method outperforms the equal-chunk approach in the original SoftCast by 2∼5 dB, and reduces the number of required meta data significantly at the same time.**

I. Introduction

Traditional communication system generally requires the channel quality to be known at the time of encoding, in order to choose appropriate coding rate. Once a signal is coded and sent out, the decoding process tends to break down if the actual channel quality falls below a threshold. On the other hand, if the channel quality increases beyond that threshold, such system cannot provide further improvement in the quality of received signal. This “threshold effect” brings challenges for the design of wireless point-to-point or broadcast visual communication systems, because the quality of wireless channel may fluctuate unpredictably and the various wireless users may have very different receiving qualities.

Recently, a scheme named *SoftCast* [1]–[3] was proposed for wireless video. Unlike typical image and video coders that compress input signal into a binary stream, SoftCast transforms the image signal into a stream of coefficient 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]–[10].

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

A key point of SoftCast is that, to achieve the best performance, SoftCast allocates transmission power among the 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 the purpose of correct decoding. To limit the overhead of meta data, SoftCast divides the coefficients into a set of chunks of equal 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 energy distribution model for natural images. We first investigate the characteristics of image transform coefficients. Based on the observation that the energy of transform coefficients decay rapidly from low frequency to high frequency, we then propose a model to approximate this distribution. Finally, we apply the proposed model to guide the power allocation in SoftCast. Experimental results indicate that the proposed approach can improve the performance of the original SoftCast scheme significantly, while reducing the number of required meta data to only 4.

The rest of the paper is organized as follows. Section II briefly reviews the SoftCast scheme. Section III discusses the statistical properties of transform coefficients and describes the proposed energy distribution model. Section IV shows experimental results and Section V concludes the paper.

II. Review of SoftCast

A. SoftCast Transmission

Fig. 1 illustrates the SoftCast scheme. The input image is first decorrelated by a 2-D transform, producing a stream of transform coefficients. The transmission stage scales each coefficient individually, applies a Walsh-Hardmard transform (WHT) to whiten the stream, and modulates the resulted

978-1-4799-3432-4/14/$31.00

numbers to a dense constellation (e.g. 64k-QAM) for raw OFDM transmission. As illustrated by Fig. 2, a pair of numbers (extracted from the stream) is mapped to a point in the constellation, using the two numbers as the I- and the Q-components respectively, and transmitted by one OFDM sub-carrier. The receiver will get a noisy version of the stream due to channel noises. The scaling operation serves the purposes of power allocation and unequal protection against 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. See [1]–[3] for the details of SoftCast.

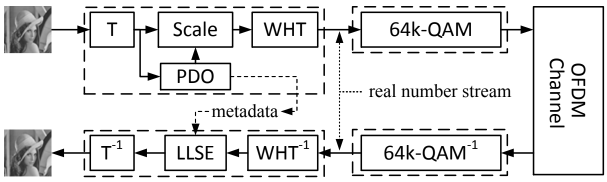

Fig. 1. Framework of the SoftCast scheme [1]–[3].

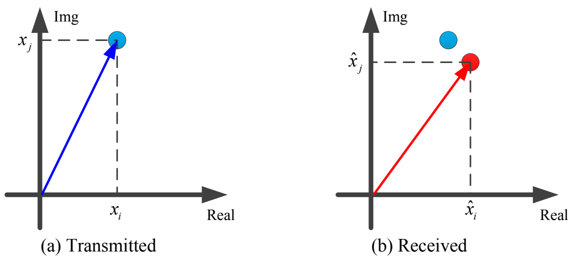

Fig. 2. Modulation for raw OFDM transmission.

B. Power-Distortion Optimization

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

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, the distortion-power relationship is $D_i \cdot P_i = \sigma_n^2 \cdot E[x_i^2]$, or $D_i(P_i) = \sigma_n^2 \cdot E[x_i^2]/P_i$. To achieve optimal performance, the transmission power is allocated among $\{x_i\}$ by

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

[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.

[2]The receiver may employ a linear least-square error (LLSE) estimator to derive $\hat{x}_i$, if σ_n is known [1]–[3]. However, this aspect is usually ignored by the power allocation procedure since σ_n is unknown by the sender.

The problem is easily solved by setting $\partial D_i / \partial P_i$ to a constant. This eventually leads to

$$P_i = C\sigma_n \sqrt{\mathrm{E}[x_i^2]} \text{ or } P_i \propto \sqrt{\mathrm{E}[x_i^2]},$$

and

$$g_i = \sqrt{C\sigma_n} (E[x_i^2])^{-1/4} \text{ or } g_i \propto (E[x_i^2])^{-1/4}.$$

Using the optimal power allocation, the total distortion in the reconstructed image is

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

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

III. Energy Modeling for SoftCast Optimization

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 meta data g_i for each coefficient individually may introduce significant overhead. Therefore, SoftCast groups coefficients into a set of chunks and perform scaling at chunk level, sending one g_i for each chunk. 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 illustrates the chunk division approach in the original SoftCast. Typically, 64 chunks are used for a whole image.

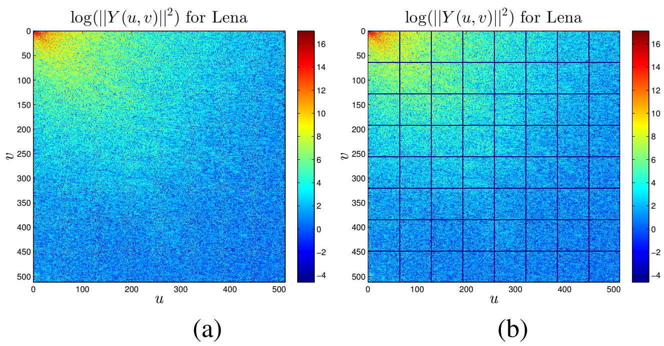

Fig. 3. Chunk division in SoftCast for power allocation. (a) The distribution of transform coefficient energy. (b) The equal-size chunk division in SoftCast. $Y(u, v)$ represents the transform coefficients of image.

In this section, we introduce a model to approximate the energy distribution in transform domain. Suppose $Y(u, v)$ is the transform domain (e.g. DCT) representation of an image. We have the following observations (as shown in Fig. 3): (1) The energy of $Y(u, v)$ decays rapidly from low frequency region to high frequency region, as the value of $\sqrt{u^2 + v^2}$ increases; (2) The energy distribution of $Y(u, v)$ varies with the angle of frequency (i.e. $\theta = \arctan(u/v)$) only slightly.

Based on the above observations, we believe that the expectation of $Y^2(u, v)$ can be reliably estimated based on $\rho(u, v) = \sqrt{u^2 + v^2}$, i.e. the distance of between this coefficient and the

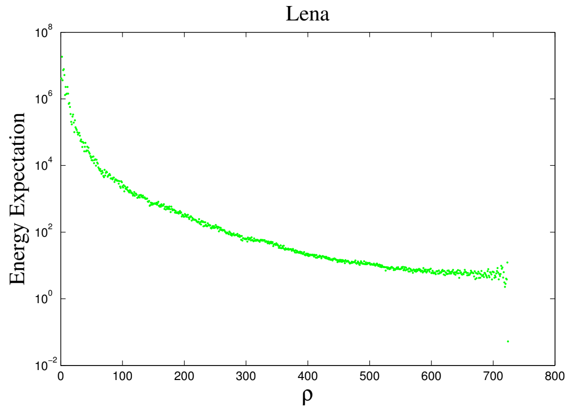

Fig. 4. Relationship between the energy expectation $E(\rho)$ and ρ.

DC coefficient in the transform coefficient plane. We establish an energy distribution model according to the following steps. Firstly, we evaluate the mean energy of transform coefficients which have the similar distance $\rho(u,v)$. To do this, we define a set of "virtual" chunks, $\mathbb{L}_k, k = 1, 2, \ldots,$ by

$$\mathbb{L}_k = \{(u,v)|k-1 < \rho(u,v) \le k\},$$

each of which includes the coefficients with similar $\rho(u,v)$ values. Then we calculate the energy expectation for each $\mathbb{L}_k$ by

$$E(k) = \text{Mean}\{Y^2(u,v)|(u,v) \in \mathbb{L}_k\}.$$

As stated before, for natural images, the value of $E(k)$ decays with k rapidly. This is illustrated in Fig. 4 more clearly. To model this distribution, we propose to employ the following function to approximate $E(k)$:

$$\tilde{E}(k) = a(k+b)^c + d.$$

The parameters a, b, c, d will be optimized so that the estimated energy expectation $\tilde{E}(k)$ becomes as close to the real energy expectation $E(k)$ as possible. Once these parameters are determined, they are transmitted to the receiver as meta data. At the same time, the energy distribution model established by a, b, c, d will be used to guide power allocation, by setting $E[||Y(u,v)||^2] = \tilde{E}(\rho(u,v))$.

IV. Experimental Results

In this section, we conduct some experiments to evaluate the proposed scheme. We first investigate the accuracy of the proposed energy distribution model. Fig. 5 illustrates the estimated transform coefficient energy based on the proposed model. To further show the accuracy of the model, we compare the actual energy $E(k)$ with the estimated $\tilde{E}(k)$. From these figures, we can see that the proposed model is accurate enough for modeling the energy diversity among transform coefficients.

Now we investigate the performance of SoftCast using the proposed energy distribution model, and compare it with the original SoftCast scheme. Fig. 7 summarizes the simulation results. We can see that the proposed method outperforms the original SoftCast scheme by $2 \sim 5$ dB. The reconstructed images are illustrated in Fig. 8 to provide a subjective performance comparison. Since the power-usage of the proposed method is very efficient, we consider a relatively low channel signal-to-noise ratio, i.e. CSNR=0dB. We can see the proposed approach can generate reconstruction images with much better perceptual qualities.

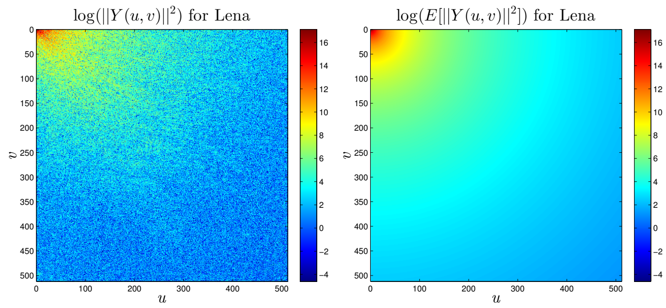

Fig. 5. The energy map of transform coefficients. Left: the actual energy. Right: the energy estimated by the model.

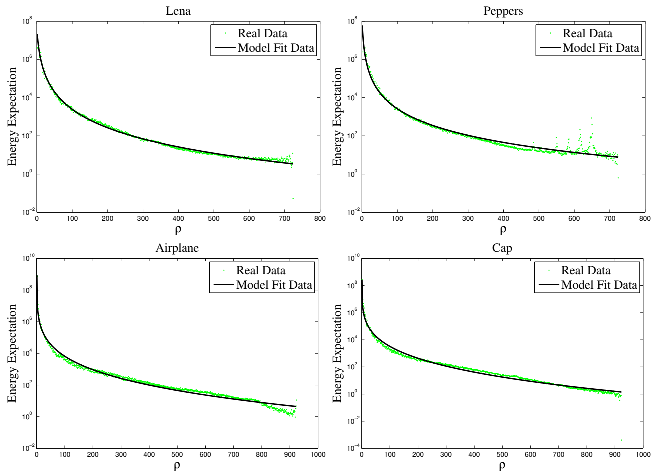

Fig. 6. Comparison between the actual energy $E(k)$ and the estimated energy $\tilde{E}(k)$ by the model.

V. Conclusions and Discussions

SoftCast provides graceful quality transition for wide channel SNR range. However, both the SoftCast sender and receiver needs to know the energy diversity among the coefficients, in order to achieve optimal power allocation and correct decoding. Signaling the energy of each coefficient (approximately) using a limited number of meta data is very important for SoftCast. This paper investigates the transform coefficient energy distribution of natural images and propose a model to approximate this distribution. Experimental results verified the efficiency of the model.

References

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

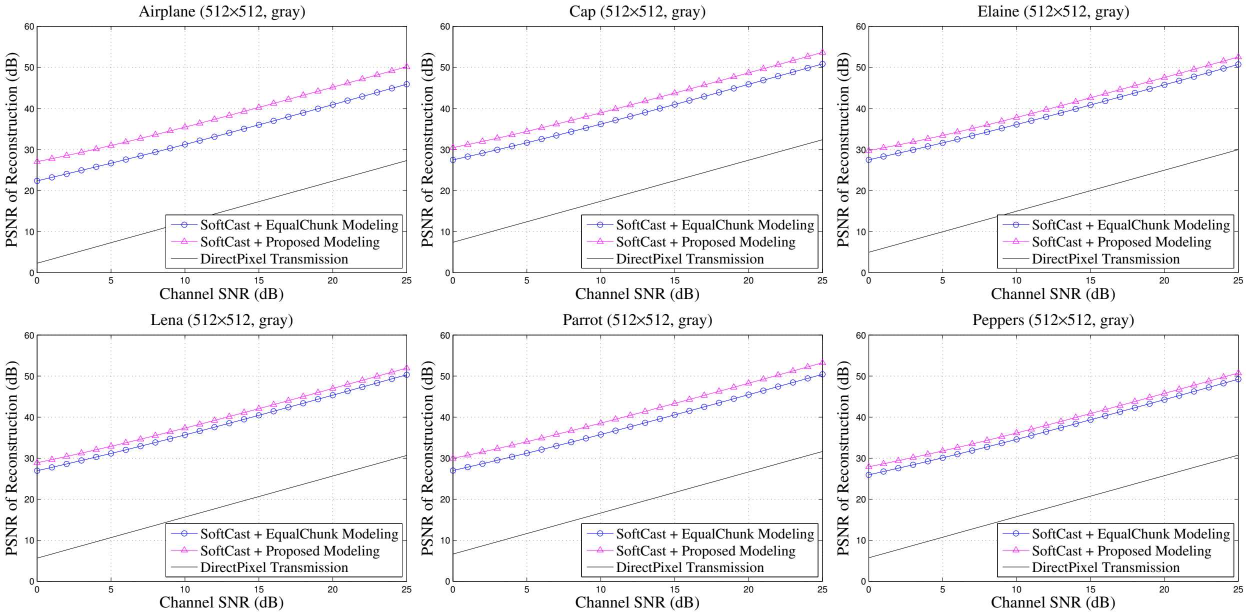

Fig. 7. PSNR results of SoftCast transmission using the proposed method.

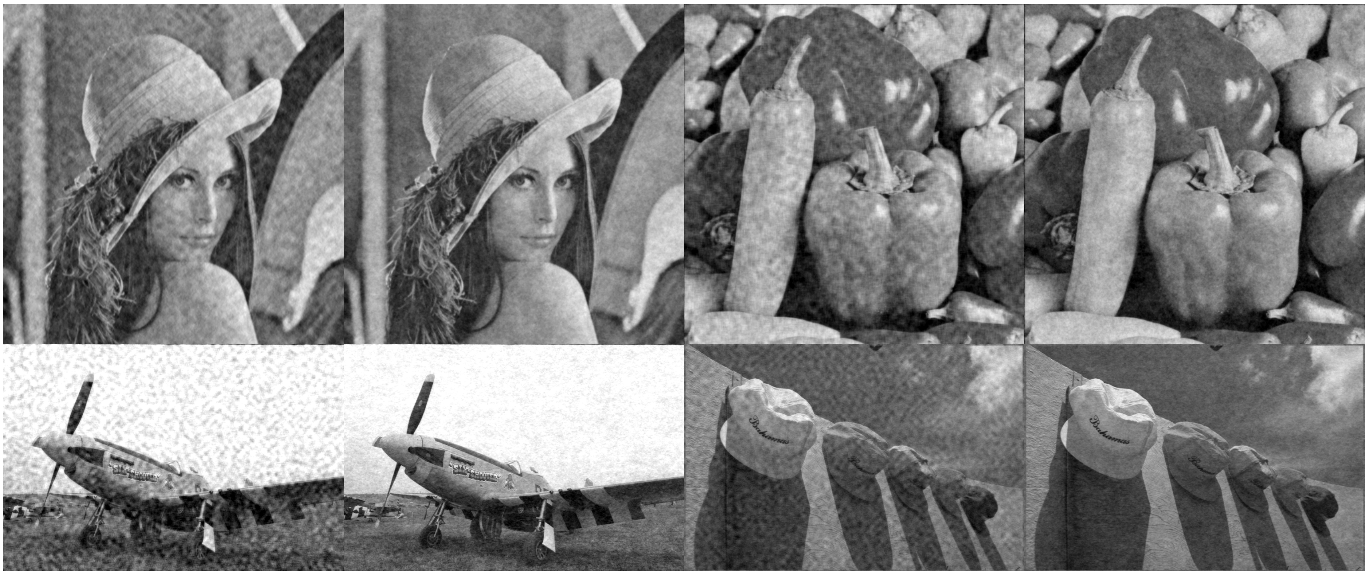

Fig. 8. Reconstruction images of SoftCast (CSNR=0dB). Left: the original SoftCast using equal-chunk division (64 chunks). Right: the proposed method.

[2] S. Jakubczak and D. 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.

[3] S. Jakubczak, H. Rahul, and D. Katabi, "One-size-fits-all wireless video," in *The eigth ACM SIGCOMM HotNets Workshop*, New York, NY, USA, 2009.

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

[5] X. Liu, W. Hu, Q. Pu, F. Wu, and Y. Zhang, "Parcast: soft video delivery in mimo-ofdm wlans," in *Internation Conference on Mobile Computing and Networking (Mobicom)*, 2012, pp. 233–244.

[6] X. Fan, R. Xiong, F. Wu, and D. Zhao, "Wavecast: Wavelet based wireless video broadcast using lossy transmission," *IEEE Visual Communications and Image Processing (VCIP)*, pp. 1–6, 2012.

[7] R. Xiong, X. Fan, F. Wu, and W. Gao, "Analysis of uncoded image communication over noisy channels with unequal protection," *SPIE Electronic Imaging 2013, Visual Information Processing and Communication IV*, p. 866603, Feb. 2013.

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

[9] 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," *IEEE Visual Communications and Image Processing (VCIP), 2013*, pp. 1–6, Nov 2013.

[10] X. Fan, F. Wu, D. Zhao, and C. O. 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.