GRADIENT BASED IMAGE TRANSMISSION AND RECONSTRUCTION USING NON-LOCAL GRADIENT SPARSITY REGULARIZATION

Hangfan Liu[1], Ruiqin Xiong[1], Siwei Ma[1], Xiaopeng Fan[2], Wen Gao[1]

[1] Institute of Digital Media, Peking University, Beijing 100871, China
[2] Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China
Email: {liuhf, rqxiong, swma, wgao}@pku.edu.cn, fxp@hit.edu.cn

ABSTRACT

Most existing image coding and communication systems aim to minimize the mean square error (MSE) of the pixels reconstructed at receivers. However, the quality metric MSE has long been criticized for not being consistent with the perception of human vision systems. This paper considers a gradient-based image SoftCast (G-Cast) scheme, based on the recent advancements in image quality assessment which indicate that gradient similarity is highly correlated with perceptual image quality. To reconstruct the image from the received noisy gradient data, we exploit the statistical characteristics of image gradients. Instead of using the very simple Laplacian distribution for image gradient as in the total variation (TV) model, we further exploit the non-local similarity of image patches. A non-local gradient sparsity regularization (NLGSR) method is developed and solved using augmented Lagrangian method. Experimental results show that the proposed scheme provides promising perceptual image quality, and the NLGSR reconstruction scheme outperforms the existing schemes remarkably.

***Index Terms*—** SoftCast, G-Cast, gradient sparsity, total variation, non-local similarity

1. INTRODUCTION

Most existing image transmission schemes, including SoftCast [1, 2], use mean square error (MSE) as the fidelity measure. However, it has been widely recognized that MSE does not reflect the visual quality perceived by human vision system (HVS) in many cases [3]. Inspired by observations in recent literatures [4, 5] that image gradients convey important visual information, we proposed a gradient-based image SoftCast (G-Cast) scheme [6] for wireless visual communication. G-Cast transmits image gradients instead of image pixels in order to convey the visual information. This paper mainly focuses on how to reconstruct the image from the received noisy gradient data and a very small set of low-frequency data.

Similar with other restoration problems that attempt to recover the original image from its degraded observation, prior image model is very important for the reconstruction process in G-Cast. Intuitively, the total variation (TV) model [7, 8, 9] can be adopted to describe the characteristic of gradient data. Due to the local smoothness of natural images, the gradient value, representing the variation at some position of an image, will be zero or close to zero in most cases. This can also be seen as sparsity in gradient domain. This observation justifies the underlying assumption of the TV model that the gradient data of an image conform to an i.i.d. zero-mean Laplacian distribution.

The above hypothesis about the distribution of image gradients is reasonable to some extent. Nevertheless, it may not be sufficiently accurate, in that the statistics of a natural image might not be stationary. On the contrary, the distributions of gradient data often vary from one region to another. Gradient data at different locations do not have to share the same distribution, i.e. they may differs from each other in mean and variance, if Laplacian distribution is still assumed. Specifically, the zero-mean assumption does not hold, especially for regions rich in textures and edges.

In the usage of TV regularization, we allow each pixel to have a separate gradient distribution. This raised another problem: how to decide the parameters for the distribution of gradient at each pixel? In the proposed model, we estimate the statistics and derive a sparse distribution adaptively by exploiting in gradient domain a set of non-locally searched patches which are similar to the patch centered at current location. Non-local similarity of a natural image [10] is well-known by now, it makes sense that the non-local constraint also stands in gradient domain.

The remainder of this paper is organized as follows. Section 2 briefly introduces the gradient based SoftCast scheme as well as the basic solution to its restoration problem. Section 3 discusses our proposed restoration method in detail, showing how the parameters of gradient distribution are estimated and how the optimization problem is numerically tack-

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).

led. Experimental results are reported in Section 4 and Section 5 concludes the paper.

2. REVIEW OF G-CAST

2.1. Gradient Based Image SoftCast (G-Cast)

G-Cast [6] transmits an image via a base layer and an enhancement layer. The base layer sends the low frequency components in order to provide a coarse description of the image. For this purpose, the input image is transformed into frequency domain and then a small number of its low frequency coefficients are coded and transmitted using conventional digital communication techniques. The enhancement layer delivers the gradient information of the image so that visual details can be observed. For this purpose, the image gradient is first extracted from the input image via a gradient transform, then processed by Walsh-Hadamard transform to reduce its peak-to-mean ratio, and finally modulated to a dense constellation for raw OFDM transmission, in the same way as done in SoftCast [1, 2]. The transmission for the enhancement layer is lossy in nature, with its noise level commensurate with the channel signal-to-noise (CSNR) condition. When both the gradient information and the low frequency coefficients are available, the decoder recover the image via a gradient-based reconstruction procedure.

2.2. TV Based Reconstruction for G-Cast

The original image can be seen as a vector $\mathbf{u}$. Denote the vertical and horizontal finite difference operators by D^{v} and D^{h}. Suppose the transmitted $D^{\mathrm{v}}\mathbf{u}$ and $D^{\mathrm{h}}\mathbf{u}$ are polluted by i.i.d. Gaussian noise, which can be formulated as:

$$\mathbf{d}^{\mathrm{v}} = D^{\mathrm{v}}\mathbf{u} + \mathbf{n}^{\mathrm{v}},\ \mathbf{d}^{\mathrm{h}} = D^{\mathrm{h}}\mathbf{u} + \mathbf{n}^{\mathrm{h}}, \tag{1}$$

where $\mathbf{n}^{\mathrm{v}}$ and $\mathbf{n}^{\mathrm{h}}$ are additive Gaussian white noise. Write $D = [D^{\mathrm{v}}; D^{\mathrm{h}}]$, $\mathbf{d} = [\mathbf{d}^{\mathrm{v}}; \mathbf{d}^{\mathrm{h}}]$ and $\mathbf{n} = [\mathbf{n}^{\mathrm{v}}; \mathbf{n}^{\mathrm{h}}]$ for simplicity. Employing traditional TV as the image prior knowledge to regularize the solution leads to the minimization problem:

$$\min_{\mathbf{u}} \frac{\mu}{2}\|D\mathbf{u} - \mathbf{d}\|_2^2 + \sum_i \|D_i\mathbf{u}\| \quad \text{s.t. } E \circ \mathcal{F}(\mathbf{u}) = \mathbf{m}, \tag{2}$$

where $\|D\mathbf{u} - \mathbf{d}\|_2^2$ is l^2 data-fidelity term, $D_i\mathbf{u} \in \mathbb{R}^2$ is the gradient of $\mathbf{u}$ at pixel i, μ is a regularization parameter controlling the trade-off between two competing terms, $\mathbf{m}$ is the low-frequency coefficients of $\mathbf{u}$, $\mathcal{F}$ stands for two-dimensional discrete Fourier transform, E represents the matrix to extract the $M \times M$ block at the top left corner and "$\circ$" denotes component-wise multiplication.

Typical algorithms like SALSA [11] and Split Bregman [12] would transform the problem like (2) into

$$\min_{\mathbf{u}} \frac{\mu}{2}\|\mathbf{w} - \mathbf{d}\|_2^2 + \sum_i \|w_i\| \text{ s.t. } E \circ \mathcal{F}(\mathbf{u}) = \mathbf{u}, w_i = D_i\mathbf{u} \tag{3}$$

by introducing the auxiliary variable $\mathbf{w}$, which is the lexicographically stacked version of w_i. The similar scheme is adopted in NLGSR as will be explained in section 3, so the detailed procedure is omitted here to avoid repetition.

3. NON-LOCAL GRADIENT SPARSITY REGULARIZATION

This section will introduce a more advanced image prior model based on traditional TV to solve the reconstruction problem in G-Cast receiver. Considering the fact that the gradient values are mostly close to zero in flat area while can be quite large in regions rich in textures, the characteristics of image content may vary from location to location, so the underlying assumption of TV that all gradient data in an image conform to the same i.i.d. zero-mean Laplacian distribution is not accurate. Therefore, we replace the traditional TV expression $\mathrm{TV}(\mathbf{u}) = \sum_i \left(|D^{\mathrm{v}}\mathbf{u}| + |D^{\mathrm{h}}\mathbf{u}|\right)$ with an extended term

$$\mathcal{J}(\mathbf{u}) = \sum_i \left(\frac{\sqrt{2}}{\sigma^{\mathrm{h}}_{\Delta_i}} \left|D^{\mathrm{h}}_i\mathbf{u} - m^{\mathrm{h}}_{\Delta_i}\right| + \frac{\sqrt{2}}{\sigma^{\mathrm{v}}_{\Delta_i}} \left|D^{\mathrm{v}}_i\mathbf{u} - m^{\mathrm{v}}_{\Delta_i}\right| \right), \tag{4}$$

implying that each gradient has a separate Laplacian distribution. $\sigma^{\mathrm{h}}_{\Delta_i}$ and $\sigma^{\mathrm{v}}_{\Delta_i}$ are the standard deviations of the horizontal and the vertical gradients at pixel i, while $m^{\mathrm{h}}_{\Delta_i}$ and $m^{\mathrm{v}}_{\Delta_i}$ denotes corresponding expectations. Since basic TV can be seen as a kind of sparsity regularization in gradient domain and the proposed extension exploits the non-local similarity for parameter estimation, the extended scheme is named non-local gradient sparsity regularization (NLGSR).

3.1. Adaptive Parameter Estimation

One of the key issues in NLGSR is to decide the distribution parameters σ_{Δ_i} and m_{Δ_i} for each gradient. This paper follows the wisdom of non-local estimation, which has been widely recognized since the application of non-local means [10]. It is reasonable to infer that such non-local property should also exist in gradient domain. Specifically speaking, in order to find out the gradient distribution at pixel i, a set of blocks most similar to the block centered at location i are searched out within the gradient picture, then the center data of these blocks can be regarded as samples of the distribution we struggle to learn. The similarity of two blocks are measured by the l_2-distance $d(i,j) = \|\mathbf{b}_i - \mathbf{b}_j\|/L$, where $\mathbf{b}_i$ is the gradient block centered at i and L is the block size. NLGSR retrieves K blocks that have smallest distances with $\mathbf{b}_i$ and record their center locations in set $\mathcal{S}_i$. Then the estimation of parameters $m^{\mathrm{h}}_{\Delta_i}$ and $\sigma^{\mathrm{h}}_{\Delta_i}$ are calculated as

$$m^{\mathrm{h}}_{\Delta_i} = \frac{1}{|\mathcal{S}_i|} \sum_{j \in \mathcal{S}_i} D^{\mathrm{h}}_j\mathbf{u}, \tag{5}$$

$$\sigma^{\mathrm{h}}_{\Delta_i} = \sqrt{\frac{1}{|\mathcal{S}_i|} \sum_{j \in \mathcal{S}_i} \left(D^{\mathrm{h}}_j\mathbf{u} - m^{\mathrm{h}}_{\Delta_i}\right)^2}, \tag{6}$$

$m^{\text{v}}_{\Delta_i}$ and $\sigma^{\text{v}}_{\Delta_i}$ are calculated in the same way.

Theoretically, both block matching and calculation of (5) and (6) call for the noise-free gradient data. Unfortunately, such clean data is not available at the G-Cast receiver, so we use an alternative $\mathbf{u}^{\text{basic}}$, generated by existing methods like traditional TV, to get the relatively clean gradient data.

3.2. Algorithm

Suppose the variation of Gaussian white noise in (1) is σ_n^2. Based on NLGSR, the MAP estimate of $\mathbf{u}$ can be deduced as:

$$\begin{aligned}\min_{\mathbf{u}} \ & \frac{1}{2\sigma_n^2}\left(\|D^{\text{h}}\mathbf{u}-\mathbf{d}^{\text{h}}\|_2^2+\|D^{\text{v}}\mathbf{u}-\mathbf{d}^{\text{v}}\|_2^2\right)\\ & +\sum_i\left(\frac{\sqrt{2}}{\sigma^{\text{h}}_{\Delta_i}}\left|D^{\text{h}}_i\mathbf{u}-m^{\text{h}}_{\Delta_i}\right|+\frac{\sqrt{2}}{\sigma^{\text{v}}_{\Delta_i}}\left|D^{\text{v}}_i\mathbf{u}-m^{\text{v}}_{\Delta_i}\right|\right)\\ \text{s.t. } & E\circ\mathcal{F}(\mathbf{u})=\mathbf{m}. \end{aligned} \tag{7}$$

This model is hard to solve directly because of the non-linearity and non-differentiability of the NLGSR term. Making use of variable splitting technique [11, 13], the problem becomes a constrained optimization:

$$\begin{aligned}\min_{\mathbf{u}} \ & \frac{1}{2\sigma_n^2}\left(\|\mathbf{w}^{\text{h}}-\mathbf{d}^{\text{h}}\|_2^2+\|\mathbf{w}^{\text{v}}-\mathbf{d}^{\text{v}}\|_2^2\right)\\ & +\sum_i\left(\frac{\sqrt{2}}{\sigma^{\text{h}}_{\Delta_i}}\left|w^{\text{h}}_i-m^{\text{h}}_{\Delta_i}\right|+\frac{\sqrt{2}}{\sigma^{\text{v}}_{\Delta_i}}\left|w^{\text{v}}_i-m^{\text{v}}_{\Delta_i}\right|\right)\\ \text{s.t. } & E\circ\mathcal{F}(\mathbf{u})=\mathbf{m}, w^{\text{h}}_i=D^{\text{h}}_i\mathbf{u}, w^{\text{v}}_i=D^{\text{v}}_i\mathbf{u}. \end{aligned} \tag{8}$$

The corresponding augmented Lagrange function is

$$\begin{aligned}\mathcal{L}_A(\mathbf{u},\mathbf{w}^{\text{h}},\mathbf{w}^{\text{v}}) =\ & \frac{1}{2\sigma_n^2}\left(\|\mathbf{w}^{\text{h}}-\mathbf{d}^{\text{h}}\|_2^2+\|\mathbf{w}^{\text{v}}-\mathbf{d}^{\text{v}}\|_2^2\right)\\ & +\sum_i\left(\frac{\sqrt{2}}{\sigma^{\text{h}}_{\Delta_i}}\left|w^{\text{h}}_i-m^{\text{h}}_{\Delta_i}\right|+\frac{\sqrt{2}}{\sigma^{\text{v}}_{\Delta_i}}\left|w^{\text{v}}_i-m^{\text{v}}_{\Delta_i}\right|\right)\\ & +\frac{\beta}{2}\left(\|\mathbf{w}^{\text{h}}-D^{\text{h}}\mathbf{u}\|_2^2+\|\mathbf{w}^{\text{v}}-D^{\text{v}}\mathbf{u}\|_2^2\right)\\ & -(\mathbf{w}^{\text{h}}-D^{\text{h}}\mathbf{u})^{\text{T}}\lambda^{\text{h}}-(\mathbf{w}^{\text{v}}-D^{\text{v}}\mathbf{u})^{\text{T}}\lambda^{\text{v}}\\ & +\frac{\gamma}{2}\|E\circ\mathcal{F}(\mathbf{u})-\mathbf{m}\|_2^2-(E\circ\mathcal{F}(\mathbf{u})-\mathbf{m})^{\text{T}}\rho. \end{aligned} \tag{9}$$

where β and γ are regularization parameters, λ^{h}, λ^{v} and ρ are Lagrange multipliers. The problem can be solved by solving (10) and (11) iteratively:

$$\left(\mathbf{u}_{(k+1)},\mathbf{w}^{\text{h}}_{(k+1)},\mathbf{w}^{\text{v}}_{(k+1)}\right)=\min_{\mathbf{u},\mathbf{w}^{\text{h}},\mathbf{w}^{\text{v}}}\mathcal{L}_A\left(\mathbf{u}_{(k)},\mathbf{w}^{\text{h}}_{(k)},\mathbf{w}^{\text{v}}_{(k)}\right) \tag{10}$$

$$\begin{aligned}\lambda^{\text{h}}_{(k+1)} &= \lambda^{\text{h}}_{(k)}-\beta_{(k)}(\mathbf{w}^{\text{h}}-D^{\text{h}}\mathbf{u}),\\ \lambda^{\text{v}}_{(k+1)} &= \lambda^{\text{v}}_{(k)}-\beta_{(k)}(\mathbf{w}^{\text{v}}-D^{\text{v}}\mathbf{u}),\\ \rho_{(k+1)} &= \rho_{(k)}-\gamma_{(k)}(E\circ\mathcal{F}(\mathbf{u})-\mathbf{m}), \end{aligned} \tag{11}$$

where k is the iteration number.

We can use alternating direction techniques [14, 15] to decompose (9) into three sub-problems, each of which can be solved efficiently. When $\mathbf{u}$ and $\mathbf{w}^{\text{v}}$ are fixed, the optimization problem (9) is reduced to

$$\begin{aligned}\mathcal{L}_A(\mathbf{w}^{\text{h}}) =\ & \frac{1}{2\sigma_n^2}\|\mathbf{w}^{\text{h}}-\mathbf{d}^{\text{h}}\|_2^2+\sum_i\frac{\sqrt{2}}{\sigma^{\text{h}}_{\Delta_i}}\left|w^{\text{h}}_i-m^{\text{h}}_{\Delta_i}\right|\\ & +\frac{\beta}{2}\|\mathbf{w}^{\text{h}}-D^{\text{h}}\mathbf{u}-\frac{\lambda^{\text{h}}}{\beta}\|_2^2. \end{aligned} \tag{12}$$

Let $\theta=1/\sigma_n^2$, and set

$$\tilde{\mathbf{w}}^{\text{h}}=\frac{\beta(D^{\text{h}}\mathbf{u}+\frac{\lambda^{\text{h}}}{\beta})+\theta\mathbf{d}^{\text{h}}}{\beta+\theta}, \tag{13}$$

$$\tilde{\mathbf{w}}^{\text{v}}=\frac{\beta(D^{\text{v}}\mathbf{u}+\frac{\lambda^{\text{v}}}{\beta})+\theta\mathbf{d}^{\text{v}}}{\beta+\theta}, \tag{14}$$

then (12) can be written as

$$\mathcal{L}_A(\mathbf{w}^{\text{h}})=\sum_i\frac{\sqrt{2}}{\sigma^{\text{h}}_{\Delta_i}}\left|w^{\text{h}}_i-m^{\text{h}}_{\Delta_i}\right|+\frac{\beta+\theta}{2}\|\mathbf{w}^{\text{h}}-\tilde{\mathbf{w}}^{\text{h}}\|_2^2. \tag{15}$$

The solution is a simple component-wise shrinkage operation:

$$\mathbf{w}^{\text{h}}=\mathbf{m}^{\text{h}}_{\Delta}+\text{shrink}(\tilde{\mathbf{w}}^{\text{h}}-\mathbf{m}^{\text{h}}_{\Delta},\frac{\sqrt{2}}{(\beta+\theta)\sigma^{\text{h}}_{\Delta}}). \tag{16}$$

Here $\text{shrink}(x,b)=\max(|x|-b,0)\cdot\text{sgn}(x)$. Similarly, the $\mathbf{w}^{\text{v}}$ sub-problem can be solve by

$$\mathbf{w}^{\text{v}}=\mathbf{m}^{\text{v}}_{\Delta}+\text{shrink}(\tilde{\mathbf{w}}^{\text{v}}-\mathbf{m}^{\text{v}}_{\Delta},\frac{\sqrt{2}}{(\beta+\theta)\sigma^{\text{v}}_{\Delta}}). \tag{17}$$

When $\mathbf{w}^{\text{h}}$ and $\mathbf{w}^{\text{v}}$ are fixed, the $\mathbf{u}$ sub-problem can be rewritten as (18), which is a quadratic function. Since D^{h} and D^{v} are both convolution operators, the least square problem can be efficiently solved in the frequency domain. The solution is formulated in (19), where "$*$" denotes complex conjugacy and both the multiplication and the division are component-wise calculations.

Now we may conclude this section by summarizing the main procedure in Algorithm 1.

4. EXPERIMENTAL RESULTS

This section examines the performance of G-Cast utilizing NLGSR. TV based G-Cast and SoftCast are used for comparison. TV based restoration also serves as the first step of NLGSR to generate $\mathbf{u}^{\text{basic}}$. The transmission in SoftCast is performed twice and averaged at the receiver side so that they send the same amount of data as G-Cast does.

Since error can hardly be completely circumvented in the process of parameter estimation, plus the fact that gradient

$$\mathcal{L}_A(\mathbf{u}) = \frac{\gamma}{2}\|E \circ \mathcal{F}(\mathbf{u}) - (\mathbf{m} + \frac{\rho}{\gamma})\|_2^2 + \frac{\beta}{2}\left(\|D^{\mathrm{h}}\mathbf{u} - (\mathbf{w}^{\mathrm{h}} + \frac{\lambda^{\mathrm{h}}}{\beta})\|_2^2 + \|D^{\mathrm{v}}\mathbf{u} - (\mathbf{w}^{\mathrm{v}} + \frac{\lambda^{\mathrm{v}}}{\beta})\|_2^2\right) \tag{18}$$

$$\tilde{\mathbf{u}} = \mathcal{F}^{(-1)}\left(\frac{\mathcal{F}^*(D^{\mathrm{h}}) \circ \mathcal{F}(\mathbf{w}^{\mathrm{h}} + \frac{\lambda^{\mathrm{h}}}{\beta}) + \mathcal{F}^*(D^{\mathrm{v}}) \circ \mathcal{F}(\mathbf{w}^{\mathrm{v}} + \frac{\lambda^{\mathrm{v}}}{\beta}) + \frac{\gamma}{\beta}(\mathbf{m} + \frac{\rho}{\gamma})}{\mathcal{F}^*(D^{\mathrm{h}}) \circ \mathcal{F}(D^{\mathrm{h}}) + \mathcal{F}^*(D^{\mathrm{v}}) \circ \mathcal{F}(D^{\mathrm{v}}) + \frac{\gamma}{\beta} \cdot E}\right) \tag{19}$$

Algorithm 1: A Summary of NLGSR

Data: The received noisy horizontal gradient image $\mathbf{d}^{\mathrm{h}}$, noisy vertical gradient image $\mathbf{d}^{\mathrm{v}}$; $\mathbf{m}$, β, γ;
Result: Recovered image $\mathbf{u}^{\mathrm{final}}$.
initialization: Generate $\mathbf{u}^{\mathrm{basic}}$ from $\mathbf{d}^{\mathrm{h}}$ and $\mathbf{d}^{\mathrm{v}}$ using TV based restoration; $\mathbf{g}^{\mathrm{h}} = D^{\mathrm{h}}\mathbf{u}^{\mathrm{basic}}$, $\mathbf{g}^{\mathrm{v}} = D^{\mathrm{v}}\mathbf{u}^{\mathrm{basic}}$, $\mathbf{w}^{\mathrm{h}} = \mathbf{d}^{\mathrm{h}}$, $\mathbf{w}^{\mathrm{v}} = \mathbf{d}^{\mathrm{v}}$, $\lambda^{\mathrm{h}} = \lambda^{\mathrm{v}} = \rho = \mathbf{0}$;
Use block matching within $\mathbf{g}^{\mathrm{h}}$ and $\mathbf{g}^{\mathrm{v}}$, calculate $\mathbf{m}^{\mathrm{h}}_{\triangle}$, $\sigma^{\mathrm{h}}_{\triangle}$ and $\mathbf{m}^{\mathrm{v}}_{\triangle}$, $\sigma^{\mathrm{v}}_{\triangle}$ according to Eq. (5) and (6);

```
while Outer stopping criteria unsatisfied do
    while Inner stopping criteria unsatisfied do
        solve w^h-problem by computing Eq. (16);
        solve w^v-problem by computing Eq. (17);
        solve u-problem by computing Eq. (19);
    end
    Update multipliers λ^h, λ^v and ρ by Eq. (11);
    Choose β_(k+1) ≥ β_(k), γ_(k+1) ≥ γ_(k);
end
```

data may not strictly conform to Laplacian distribution, the estimated $\sigma^{\mathrm{h}}_{\triangle}$ and $\sigma^{\mathrm{v}}_{\triangle}$ are adjusted by a factor $\delta^{\mathrm{h}}_{\triangle}$, which is empirically chosen according to σ_n. The number of similar blocks K is set to 50 their block size L is set to 7×7. The dimension of low-frequency data is 8×8.

15 natural images are tested with CSNR ranging from 0dB to 15dB. As can be seen from Fig. 1 and Fig. 2, NLGSR has evident gain over TV measured by both SSIM and gradient signal-to-noise ratio (GSNR), while the TV based G-Cast outperforms SoftCast. Other tested images show similar results. The average GSNR of all the tested pictures shows that the gain of NLGSR over traditional TV ranges from 0.77dB to 0.34dB as CSNR varies from 0dB to 15dB, while the gain of TV based G-Cast over SoftCast falls in the interval of 1.95dB to 1.4dB. Despite that our proposed method is not optimized w.r.t. MSE, it shows competitive PSNR values compared with SoftCast, except for some specific cases like Barbara. The average PSNR shows a gain of up to 0.59dB.

Our ultimate goal is to improve perceptual quality. The reconstructed images of the three tested methods are shown in Fig. 4. The superiority of NLGSR based G-Cast is obvious.

5. CONCLUSION

Inspired by the observation that gradient structures are highly relevant to image quality assessment, gradient based SoftCast (G-Cast) has been designed for wireless image communication. This paper introduces a kind of perception oriented image reconstruction method, i.e., non-local gradient sparsity regularization (NLGSR), for the receiver of G-Cast. NLGSR utilizes the property of self-similarity in gradient domain, estimating the statistics of image gradient at any pixel from a group of non-locally searched gradient blocks most similar to the gradient block centered at the current location. These groups of blocks are used as the samples of the gradient distribution to be learnt. Augmented Lagrangian method is applied to solve the optimization problem. Experimental results show that G-Cast outperforms SoftCast in terms of visual quality, and NLGSR based G-Cast can substantially improve the reconstruction quality over traditional TV restoration.

6. 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.

[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] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[4] J. Zhu and N. Wang, "Image quality assessment by visual gradient similarity," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 919–933, March 2012.

[5] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1500–1512, April 2012.

[6] R. Xiong, H. Liu, S. Ma, X. Fan, F. Wu, and W. Gao, "G-cast: Gradient based image softcast for perception-friendly wireless visual communication," *IEEE Data Compression Conference (DCC'14)*, pp. 133–142, March 2014.

[7] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[8] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

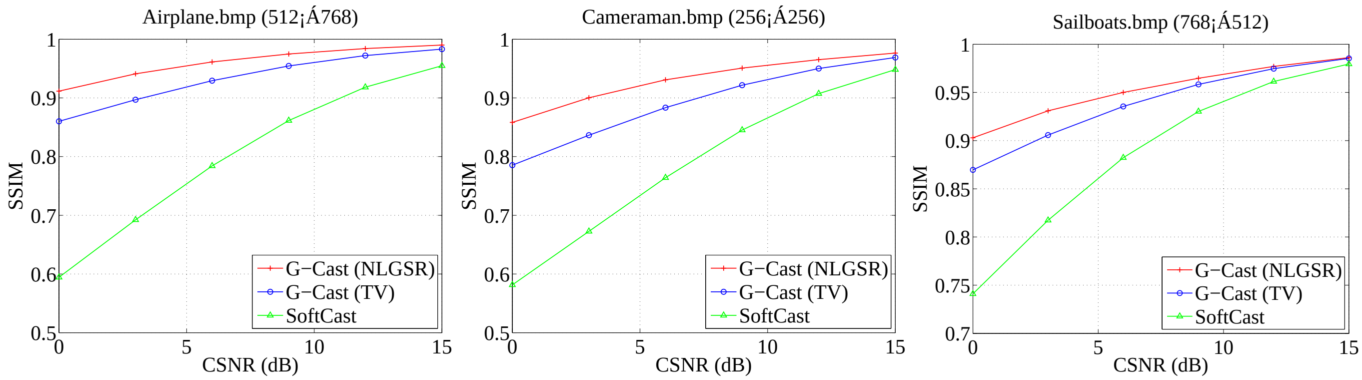

Fig. 1. SSIM comparison

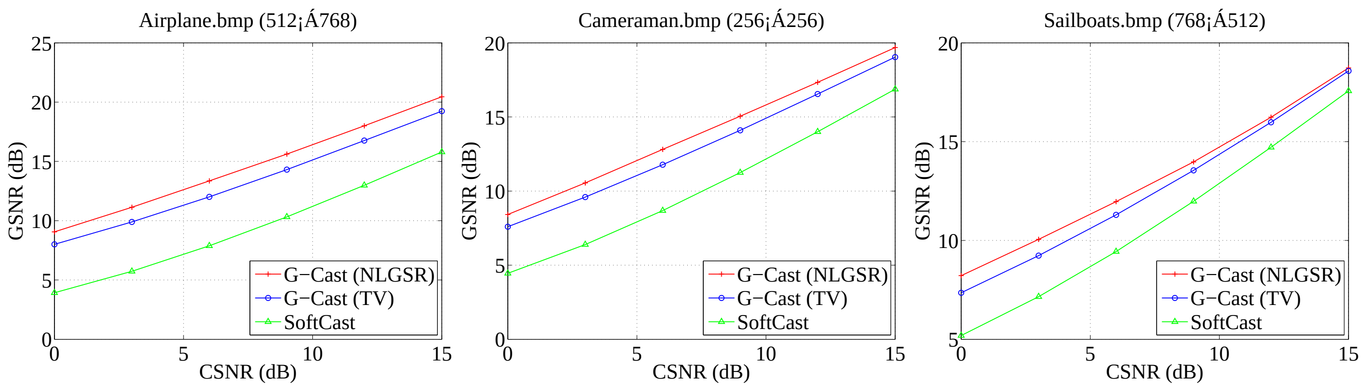

Fig. 2. GSNR comparison

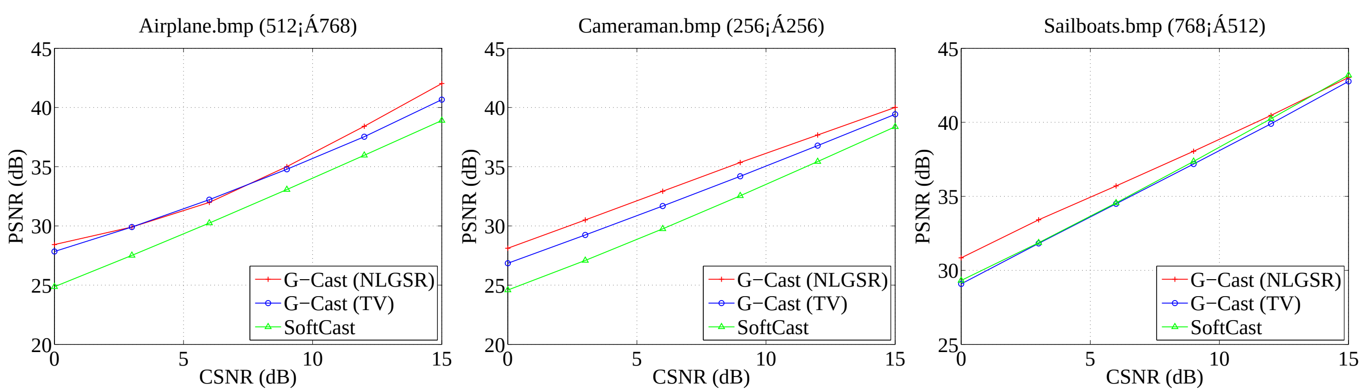

Fig. 3. PSNR comparison

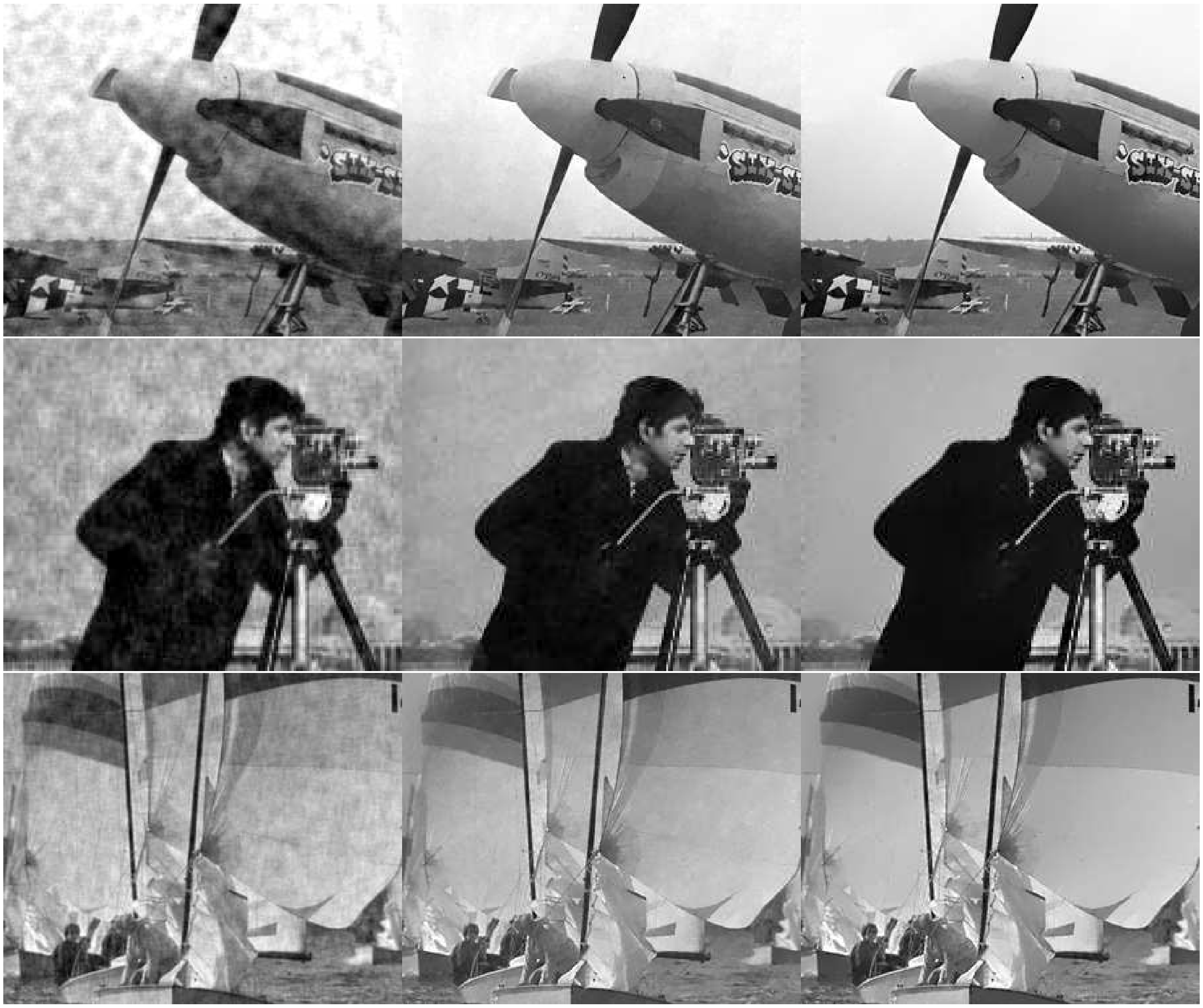

Fig. 4. Visual comparison of reconstructed images (CSNR=0dB). From left to right: SoftCast, G-Cast(TV), G-Cast(NLGSR). Please enlarge the figure for better comparison.

[9] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[10] A. Buades, B. Coll, and J.M. Morel, "A non-local algorithm for image denoising," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05)*, June 2005, vol. 2, pp. 60–65 vol. 2.

[11] M.V. Afonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2345–2356, Sept 2010.

[12] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[13] J.-F. Cai, S. Osher, and Z. Shen, "Split bregman methods and frame based image restoration," *Multiscale Modeling & Simulation*, vol. 8, no. 2, pp. 337–369, 2009.

[14] C. Li, W. Yin, and Y. Zhang, "Tval3: Tv minimization by augmented lagrangian and alternating direction algorithms," 2009.

[15] S. Chretien, "An alternating l_1 approach to the compressed sensing problem," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 181–184, Feb 2010.