SPARSITY-DRIVEN RECONSTRUCTION OF ℓ_∞-DECODED IMAGES

Yuanman Li and Jiantao Zhou

Department of Computer and Information Science
Faculty of Science and Technology, University of Macau
Emails: yuanmanx.li@gmail.com and jtzhou@umac.mo

ABSTRACT

In this paper, we propose a sparsity-driven restoration technique to improve the coding performance of the ℓ_∞-decoded images. This is achieved by incorporating a ℓ_1 minimization term into a ℓ_2 optimization framework, where the weighting vectors balancing the relative contribution of each term are appropriately determined. The ℓ_∞ constraints inherent to ℓ_∞-constrained predictive coding are also included to narrow the solution space, leading to more accurate estimation. Experimental results show that our proposed scheme significantly improves the ℓ_2 performance of the ℓ_∞-decoded images, while still preserving a tight error bound on every single pixel. In addition, when comparing with the existing scheme of restoring the ℓ_∞-decoded images, the PSNR gain can be up to 1 dB.

Index Terms— Image restoration, sparse representation, ℓ_∞-constrained predictive coding.

1. INTRODUCTION

Compression of images to reduce their storage and transmission bandwidth requirements is of great interest in many critical applications such as medical imaging, remote sensing and precise engineering. Although lossless compression methods are preferable in these scenarios, the achievable compression ratios are rather modest ranging from 2-3, depending on the image contents and imaging modalities. To obtain higher compression, certain amount of distortion must be tolerated in the reconstructed images.

Several methods for lossy image compression are available such as the transform-based codec JPEG 2000 and the DPCM-based approaches (or called ℓ_∞-constrained predictive coding). When the bit rate gets sufficiently high, the latter type of compression methods outperform the former ones in terms of ℓ_2 performance, while still guaranteeing a tight bound on every single pixel. Such rate-distortion behavior of the DPCM-based approaches makes them better candidates for those critical scenarios where high fidelity of the reconstructed images is required. There are a number of DPCM-based compression schemes proposed in the literature [1–3].

However, the ℓ_∞-constrained image coding technique becomes inferior in the ℓ_2 distortion metric if the bit rate decreases significantly from that of the lossless case. Wu and Bao attempted to improve the ℓ_2 performance of ℓ_∞-constrained coding by studying the adverse effect of residue quantization on the robustness of the predictor [3]. But the coding gain of their in-loop bias cancelation approach is rather limited (around 0.2 dB), and the encoding complexity is dramatically increased. Recently, Zhou *et. al* proposed a ℓ_2-based soft-decoding approach of ℓ_∞-decoded images, shifting the task of improving the coding performance from the encoder to the decoder [4]. In their regularization framework, the Piecewise AutoRegressive (PAR) model is utilized as the image prior domain knowledge. Though the method of [4] achieved significant performance gain, there still exists room for further improvement. In this paper, we demonstrate that the ℓ_2 performance can be further enhanced by considering the sparsity prior, which holds for any natural images. We design a sparsity-driven restoration technique to improve the coding performance of the ℓ_∞-decoded images. This is achieved by incorporating a ℓ_1 minimization term into a ℓ_2 optimization framework, where the weighting vectors balancing the relative contribution of each term are appropriately determined. The ℓ_∞ constraints inherent to ℓ_∞-constrained predictive coding are also included to narrow the solution space, leading to more accurate estimation. Experimental results show that our proposed scheme significantly improves the ℓ_2 performance of the ℓ_∞-decoded images, while still preserving a tight error bound on every single pixel. In addition, when comparing with the existing scheme of restoring the ℓ_∞-decoded images, our PSNR gain can be up to 1 dB.

The rest of this paper is organized as follows. In the next section, we briefly review the work of [4], and point out the weakness of neglecting the sparsity prior. Section 3-4 present our proposed sparsity-driven restoration framework, where the weighting vectors balancing the contribution of different terms are appropriately determined. Section 5 gives the experimental results on several satellite and medical images, demonstrating the effectiveness of our proposed tech-

This work was supported by the Macau Science and Technology Development Fund under grant FDCT/009/2013/A1 and the Research Committee at University of Macau under grant SRG023-FST13-ZJT, MYRG2014-00031-FST, and MRG021/ZJT/2013/FST.

nique. We finally conclude in Section 6.

2. SOFT-DECODING OF ℓ_∞-DECODED IMAGES [4]

In [4], Zhou *et. al.* suggested a soft decoding technique to re-estimate the original image $\mathbf{x}$ from the hard decision ℓ_∞-decoded image $\mathbf{y}$. This is formulated as the following inverse problem of ℓ_2 restoration of ℓ_∞-decoded images

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}}\{\|\mathbf{x}-\mathbf{x}_{\mathbb{M}}\|_2^2 + \lambda\|\mathbf{x}-\tilde{\mathbf{y}}\|_2^2\} \quad \text{subject to } \|\mathbf{x}-\mathbf{y}\|_\infty \le \tau \tag{1}$$

where

- $\mathbf{x}_{\mathbb{M}}$ denotes an estimate of $\mathbf{x}$ generated by an image model $\mathbb{M}$;
- $\tilde{\mathbf{y}}$ is a context-adaptive initially refined estimate of $\mathbf{x}$ based on $\mathbf{y}$ [4];
- $\|\mathbf{x}-\mathbf{x}_{\mathbb{M}}\|_2^2$ is the regularization term;
- $\|\mathbf{x}-\tilde{\mathbf{y}}\|_2^2$ is the fidelity term;
- λ is the regularization parameter to trade-off the fidelity to the received data and the image prior;
- The constraints inherent to the ℓ_∞-constrained coding $\|\mathbf{x}-\mathbf{y}\|_\infty \le \tau$ are to confine the solution space.

In the regularization term above, an image model $\mathbb{M}$ of 2-D piecewise autoregressive (PAR) was assumed

$$x_{i,j} = \sum_{(m,n)\in\mathcal{S}} a_{i,j}^{m,n} x_{i+m,j+n} + \zeta, \tag{2}$$

where $\mathcal{S}$ specifies the support of the PAR model relative to position (i,j), and ζ is a random perturbation. Here, the model parameter vectors $\mathbf{a}_{i,j}$'s are allowed to change from pixel to pixel. The PAR model parameter vectors $\mathbf{a}_{i,j}$'s are estimated adaptively for each spatial location (i,j) using samples from the initially refined image $\tilde{\mathbf{y}}$ in a moving window $\mathcal{W}_{i,j}$. Specifically, $\mathbf{a}_{i,j}$ can be determined by

$$\hat{\mathbf{a}}_{i,j} = \underset{\mathbf{a}}{\operatorname{argmin}} \sum_{(i,j)\in\mathcal{W}_{i,j}} \left(\tilde{y}_{i,j} - \sum_{(m,n)\in\mathcal{S}} a_{m,n}\tilde{y}_{i+m,j+n}\right)^2 \tag{3}$$

With the PAR model, $\mathbf{x}_{\mathbb{M}}$ can be written as $\mathbf{x}_{\mathbb{M}} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is the sparse zero-diagonal matrix consisting of model parameters, whose ith row is the parameter vector $\mathbf{a}_i^T$.

Therefore, (1) can be re-written into the following format

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}}\{\|\mathbf{x}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}-\tilde{\mathbf{y}}\|_2^2\} \quad \text{subject to } \|\mathbf{x}-\mathbf{y}\|_\infty \le \tau. \tag{4}$$

This is a constrained ℓ_2 minimization problem, which can be solved easily. Please refer to [4] for more details.

3. SPARSITY-DRIVEN APPROACH FOR RESTORATION OF ℓ_∞-DECODED IMAGES

The above framework of restoration of ℓ_∞-decoded images, though effective in improving the PSNR values of the resulting images, still has room to be further refined. As a regularization-based technique, it is of great importance to incorporate appropriate prior domain knowledge of natural images. One of the key factors leading to the success of the framework (4) is the utilization of the PAR model, which is capable of fitting a rich amount of image waveforms.

However, the sparsity prior, another type of important image prior, which have found tremendous applications in various image processing tasks, is unfortunately neglected. If we can unlock the power of sparsity, then the restoration performance can be expected to be significantly improved.

Mathematically, the sparse representation model assumes that a signal can be represented by $\mathbf{x} \approx \mathbf{D}\boldsymbol{\alpha}$, where $\mathbf{D}$ is an over-complete dictionary and the coding vector $\boldsymbol{\alpha}$ is sparse. The sparse coding problem can be generally formulated in a ℓ_1 minimization form $\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}}\{\|\mathbf{x}-\mathbf{D}\boldsymbol{\alpha}\|_2^2 + \gamma\cdot\|\boldsymbol{\alpha}\|_1\}$, where γ is the regularization parameter.

By introducing the sparsity prior, we propose a new framework of restoring ℓ_∞-decoded images

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}}\{\|\tilde{\mathbf{y}}-\mathbf{D}\boldsymbol{\alpha}\|_2^2 + \boldsymbol{\mu}\|(\mathbf{I}-\mathbf{A})\mathbf{D}\boldsymbol{\alpha}\|_2^2 + \boldsymbol{\nu}\|\boldsymbol{\alpha}\|_1\} \quad \text{subject to } \|\mathbf{D}\boldsymbol{\alpha}-\mathbf{y}\|_\infty \le \tau \tag{5}$$

To better characterize the locally stationary nature of image signal, we propose to solve the above optimization problem in a patch-by-patch manner. Let $\mathbf{x}_k = \mathbf{E}_k\mathbf{x}$ be the image patch of size $\sqrt{n}\times\sqrt{n}$ extracted at location k, where $\mathbf{E}_k$ serves as the extraction matrix. The extracted patches are allowed to be overlapped to reduce blockiness artifacts. In addition, similar to [5], we cluster the image patches of the initially refined image $\tilde{\mathbf{y}}$ into K clusters, and learn a PCA sub-dictionary for each cluster. Letting $\mathbf{D}_k$ be the learned sub-dictionary for patch $\mathbf{x}_k$, then $\mathbf{x}_k$ can be approximated by $\hat{\mathbf{x}}_k = \mathbf{D}_k\boldsymbol{\alpha}_k$, where $\boldsymbol{\alpha}_k$ is the sparse coding vector for $\mathbf{x}_k$. By fusing all $\hat{\mathbf{x}}_k$'s in the least-square (LS) sense, we can reconstruct the whole image by

$$\hat{\mathbf{x}} \approx \mathbf{D}\circ\boldsymbol{\alpha} = \left(\sum_k \mathbf{E}_k^T\mathbf{E}_k\right)^{-1}\sum_k\left(\mathbf{E}_k^T\mathbf{D}_k\boldsymbol{\alpha}_k\right). \tag{6}$$

By considering the patch-by-patch reconstruction, we reformulate (5) as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}}\Big\{\|\tilde{\mathbf{y}}-\mathbf{D}\circ\boldsymbol{\alpha}\|_2^2 + \boldsymbol{\mu}\|(\mathbf{I}-\mathbf{A})\mathbf{D}\circ\boldsymbol{\alpha}\|_2^2 + \boldsymbol{\nu}\sum_k\|\boldsymbol{\alpha}_k\|_1\Big\} \quad \text{subject to } \|\mathbf{D}\circ\boldsymbol{\alpha}-\mathbf{y}\|_\infty \le \tau \tag{7}$$

Table 1

PERFORMANCE COMPARISON FOR $\tau = 1$

Image	rate	CALIC		J2K		[4]		Proposed	
		PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$
Liver	2.265	49.89	1	48.16	4	49.99	2	50.63	2
Cloud	2.382	49.89	1	48.01	5	49.91	2	50.39	2
Factory	2.226	49.91	1	48.35	5	50.24	2	51.38	2
Stadium	3.255	49.86	1	47.56	5	49.90	2	50.27	2
Pentagon	3.193	48.87	1	45.94	6	49.89	2	50.01	2

Table 2

PERFORMANCE COMPARISON FOR $\tau = 3$

Image	rate	CALIC		J2K		[4]		Proposed	
		PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$
Liver	1.318	42.28	3	43.39	10	43.94	5	44.43	5
Cloud	1.384	42.25	3	43.13	9	43.68	5	43.89	5
Factory	1.373	42.30	3	42.99	11	44.35	5	45.25	5
Stadium	2.230	42.12	3	41.55	11	42.80	5	43.12	5
Pentagon	2.047	42.12	3	39.96	12	42.34	5	42.79	5

We also find that such sparsity-driven restoration could be further improved by taking the geometric structure of images into consideration. Dong *et. al* [6] proposed the concept of sparse coding noise (SCN), which is defined as the difference between the sparse coding vectors obtained from the original image patch $\mathbf{x}_k$ and the available degraded patch $\tilde{\mathbf{y}}_k$. It was shown that better restoration performance could be obtained by replacing the ℓ_1 minimization term in sparsity regularized restoration problem with a term corresponding to the SCN. To this end, we introduce the term suppressing the SCN into the restoration framework of ℓ_∞-decoded images

$$\begin{aligned}\hat{\boldsymbol{\alpha}} = \quad & \arg\min_{\boldsymbol{\alpha}} \Big\{ \|\tilde{\mathbf{y}} - \mathbf{D} \circ \boldsymbol{\alpha}\|_2^2 + \boldsymbol{\mu} \|(\mathbf{I} - \mathbf{A})\mathbf{D} \circ \boldsymbol{\alpha}\|_2^2 \\ & + \boldsymbol{\nu} \sum_k \|\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_k^*\|_1 \Big\} \\ & \text{subject to } \; \|\mathbf{D} \circ \boldsymbol{\alpha} - \mathbf{y}\|_\infty \leq \tau \end{aligned} \tag{8}$$

where

$$\boldsymbol{\alpha}_k^* = \sum_{q \in \Omega_k} \omega_{k,q} \boldsymbol{\alpha}_{k,q} \tag{9}$$

Here, Ω_k is the set of image patches similar to $\tilde{\mathbf{y}}_k$, $\boldsymbol{\alpha}_{k,q}$ is the spare coding vector for the qth block in Ω_k, and $\omega_{k,q}$'s denote the non-local weights.

4. ADAPTIVE WEIGHTING STRATEGY

A crucial task now is to appropriately determine the weighting vectors $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, which tradeoff the relative contribution of different regularization terms and the fidelity term. To this end, we first determine the variance of the fidelity term, and hence define the compression noise as

$$d \;=\; x - \tilde{y} \tag{10}$$

which is found experimentally to follow a Gaussian distribution. The variance of the compression noise can be estimated in a context-adaptive manner from the training images. For fixed context vector $\mathbf{C}$, we denote the conditional variance of the compression noise by $\mathrm{Var}(d|\mathbf{C})$. For a specific pixel j in the kth block, the conditional variance of the compression noise can be expressed by $\sigma_F^2(k,j) = \mathrm{Var}(d|\mathbf{C}_{k,j})$, where $\mathbf{C}_{k,j}$ is the associated context vector.

In addition, a byproduct of estimating the PAR model coefficient matrix $\mathbf{A}$ is the model fitting error $\sigma_A^2(k,j) = (\tilde{y}_{k,j} - \mathbf{a}_j^T \mathbf{y}_k)^2$, which can be treated as an estimate of the variance of the jth pixel in the current block k.

As pointed out in [6], the SCN follows a Laplacian distribution, whose variance can be straightforwardly estimated from training samples. For the jth pixel in the kth block, let $\sigma_S^2(k,j)$ be the corresponding variance of the SCN.

To reflect the variance of the fidelity term and the PAR regularization term, we set $\boldsymbol{\mu}$ as a diagonal matrix with diagonal elements being

$$\mu_{i,i} = \frac{\sigma_F^2(k,i)}{\sigma_A^2(k,i) + \varepsilon} \tag{11}$$

where ε is a small adjusting parameter.

Table 3

PERFORMANCE COMPARISON FOR $\tau = 5$

Image	rate	CALIC		J2K		[4]		Proposed	
		PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$
Liver	0.914	38.64	5	40.88	12	40.89	9	41.37	8
Cloud	0.938	38.61	5	40.64	14	40.64	9	40.79	8
Factory	1.056	38.48	5	40.19	14	41.10	8	42.09	8
Stadium	1.766	38.39	5	38.31	14	39.49	9	39.81	8
Pentagon	1.505	38.23	5	37.64	16	39.13	9	39.60	8

Table 4

PERFORMANCE COMPARISON FOR $\tau = 7$

Image	rate	CALIC		J2K		[4]		Proposed	
		PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$	PSNR	$\|e\|_\infty$
Liver	0.661	36.32	7	39.21	14	38.74	11	39.08	11
Cloud	0.676	36.30	7	39.12	15	38.53	12	38.71	11
Factory	0.828	35.91	7	37.93	23	38.77	11	39.73	11
Stadium	1.436	35.48	7	36.35	23	36.73	12	37.47	11
Pentagon	1.141	35.80	7	36.12	24	36.99	12	37.58	11

Due to the fact that the compression noise can be modeled as a Gaussian distribution, we can determine the diagonal regularization matrix $\boldsymbol{\nu}$ in ML sense [6] by

$$\nu_{i,i} = \frac{2\sqrt{2}\sigma_F^2(k,i)}{\sigma_S(k,i) + \varepsilon} \tag{12}$$

After determining all the regularization parameters, the optimization problem of (8) can be expressed by

$$\begin{aligned}\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \Big\{ & \|\tilde{\mathbf{y}} - \mathbf{D} \circ \boldsymbol{\alpha}\|_2^2 + \|\sqrt{\boldsymbol{\mu}}(\mathbf{I} - \mathbf{A})\mathbf{D} \circ \boldsymbol{\alpha}\|_2^2 \\ & + \sum_k \|\sqrt{\boldsymbol{\nu}_k}(\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_k^*)\|_1 \Big\} \\ & \text{suject to } \|\mathbf{D} \circ \boldsymbol{\alpha} - \mathbf{y}\|_\infty \leq \rho\tau. \end{aligned} \tag{13}$$

where $\rho < 1$ is a shrinkage factor empirically found to be useful in improving both the ℓ_2 and ℓ_∞ performance. This problem can be solved iteratively by using the variable splitting technique [7, 8].

5. EXPERIMENTAL RESULTS

To demonstrate the performance of our proposed sparsity-driven method for restoring the ℓ_∞-decoded images, we compare it with the CALIC, the soft-decoding approach in [4], and the encoder-centralized codec JPEG 2000, under both ℓ_2 and ℓ_∞ criterion. For fair comparison, our method and the one in [4] are both initialized by the CALIC of ℓ_∞ error bounds $\tau = 1$, $\tau = 3$, $\tau = 5$ and $\tau = 7$, respectively. Tables 1 through 4 show the results performed on five test images.

Compared with the CALIC, the gain in PSNR can be up to 3.73 dB, which is achieved by the image Factory when $\tau = 7$. Due to the employment of the shrinkage parameter $\rho < 1$, the proposed method still grantees ℓ_∞ bounds less than 2τ.

When comparing with the JPEG 2000, the PSNR gain is quite remarkable when the bit rates are sufficiently high. For instance, our method achieves 4.07 dB higher PSNR values for image Pentagon when the bit rate is 3.193 bpp. As bit rate decreases, the gain becomes smaller, and our method losses the competition when the bit rate is below 0.9 bpp. However, the new method maintains much tighter ℓ_∞ error bounds.

In addition, the new sparsity-driven restoration technique obtains higher PSNR values over the soft-decoding method [4], for all test images under different τ. The maximum PSNR improvement is around 1 dB, which is observed in the image Factory for $\tau = 1$. Also, we notice that the ℓ_∞ performance is enhanced for several test images, especially for large τ.

6. CONCLUSION

By introducing a ℓ_1 norm term into a ℓ_2 restoration framework, we have suggested a sparsity-driven restoration technique to improve the coding performance of the ℓ_∞-decoded images. It has been shown that our scheme significantly improves the ℓ_2 performance of ℓ_∞-decoded images, while still preserving a tight error bound on every single pixel. When comparing with the existing methods, we have achieved significant PSNR gains and smaller ℓ_∞ error bounds.

7. REFERENCES

[1] K. Chen and T. Ramabadran, "Near-lossless compression of medical images through entropy-coded dpcm," *IEEE Trans. Med. Imag.*, vol. 13, no. 3, pp. 538–548, Sept. 1994.

[2] L. Ke and M. Marcellin, "Near-lossless image compression: minimum-entropy, constrained-error dpcm," *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 225–228, Feb. 1998.

[3] X. Wu and P. Bao, "ℓ_∞-constrained high-fidelity image compression via adaptive context modeling," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 536–542, Apr. 2000.

[4] J. Zhou, X. Wu, and L. Zhang, "ℓ_2 restoration of ℓ_∞ decoded images via soft-decision estimation," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4797–4807, Dec. 2012.

[5] W. Dong, D. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.

[6] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1620–1630, Apr. 2013.

[7] J. Bioucas-Dias and M. A. T. Figueiredo, "An iterative algorithm for linear inverse problems with compound regularizers," in *IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 685–688.

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