PATCH-BASED IMAGE INTERPOLATION: ALGORITHMS AND APPLICATIONS

Xin Li

Lane Dept. of Computer Science and Electrical Engineering
West Virginia University, Morgantown WV26506-6109

ABSTRACT

Motivated by recent advances in adaptive sparse representations and nonlocal image modeling, we propose a patch-based image interpolation algorithm under a set theoretic framework. Our algorithm alternates the projection onto two convex sets: one is given by the observation data and the other defined by a sparsity-based nonlocal prior similar to BM3D. In order to optimize the design of observation constraint set, we propose to address the issue of sampling pattern and model it by a spatial point process. A Monte-Carlo based algorithm is proposed to optimize the randomness of sampling patterns to better approximate homogeneous Poisson process. Extensive experimental results in image interpolation and coding applications are reported to demonstrate the potential of the proposed algorithms.

Index Terms—image interpolation, patch-based models, spatial point process, Monte-Carlo method.

I. INTRODUCTION

Image interpolation refers to the reconstruction of a plausible image from incomplete data (e.g., pixels at a subset of sampling lattice). Like other inverse problems, image prior plays a critical role in interpolation algorithms. Existing image interpolation algorithms are mostly local - for example, B-spline interpolation [1] uses cubic splines to fit local intensity functions; subpixel edge location was proposed in [2] to render sharper edges; local covariance structures can also be used to locally guide the linear interpolation [3]. Most recently, several existing ideas haven been unified under a kernel regression framework [4].

In recent years, nonlocal image priors have attracted increasingly more attention and found many successful applications (e.g., texture synthesis [5], image inpainting [6], image denoising [7], [8]). Nonlocal image interpolation has also appeared in the literature (e.g., [9], [10], [11]). However, there still lacks a systematic framework for understanding and analyzing the benefit of nonlocal priors for image interpolation. Given the nature of interpolation problem (a fundamental task in digital image processing), it is our belief that a better solution to image interpolation could shed useful insight to our understanding of other problems.

In this paper, we present a set theoretic framework for incorporating patch-based nonlocal prior into image interpolation. Based on previous work [12], we propose to cast image interpolation under the maximum a posterior (MAP) framework which admits a set theoretic implementation. The likelihood and prior terms respectively give rise to observation constraint and sparsity constraint. In particular, we note that the sparsity constraint can be enforced either locally (e.g., overcomplete DCT [13] or steerable pyramids [14]) or nonlocally (e.g., BM3D-based [11]). A strategy of progressive decreasing thresholds proposed in [15] is adopted to find the solution to the fixed-point equation characterizing the sparsity constraint. Issues of initial threshold selection and stopping criterion are briefly addressed as well.

In addition to the prior term, the second contribution of this work is on the likelihood term related to sampling pattern (the subset of sampling lattice). We propose to model sampling pattern by a spatial point process [16], [17]. Based on previous works of randomness measures [18], we have developed a Monte-Carlo based algorithm to optimize the randomness of sampling patterns and their benefits on image interpolation are justified by experimental results. The flexibility of designing optimal sampling pattern suggests an interesting application into compressive imaging which has potential in wireless sensor networks. A random-sampling protocol is shared by encoder and decoder, which eliminates the need of transmitting any bits to resolve location uncertainty. Such distributed image coding system has achieved highly encouraging performance when compared with conventional wavelet image coders such as SPIHT [19].

II. PATCH-BASED IMAGE INTERPOLATION

We use $\mathbf{x}, \mathbf{y}$ to denote unobservable image (target of reconstruction) and observation data respectively. For a binary sampling pattern $\mathbf{S}$, we simply have $y(m,n) = x(m,n)S(m,n)$ where (m,n) are the spatial coordinates. Under a Bayesian framework, the interpolation problem can be formulated as the MAP estimation, i.e.,

$$p(\mathbf{x}|\mathbf{y}, \mathcal{H}) = \frac{p(\mathbf{y}|\mathbf{x}, \mathcal{H})p(\mathbf{x}|\mathcal{H})}{p(\mathbf{y}|\mathcal{H})} \qquad (1)$$

where $\mathcal{H}$ denotes a model we assume (local vs. nonlocal), $p(\mathbf{y}|\mathbf{x}, \mathcal{H})$ is the likelihood term (it is an indicator function $1_{\{y(m,n)=x(m,n)\}}$) and $p(\mathbf{x}|\mathcal{H})$ is the prior term that plays the role of regularizing ill-posed problems. One popular choice of the prior term is so-called sparse constraint [20], [13] - i.e., only a small fraction of transform coefficients are significant. A great deal of effort has been spent on constructing good transforms $\mathcal{T}$ with maximal sparsity. If we use $X = \mathcal{T}\mathbf{x}$ to denote the transform coefficients, the sparsity constraint can be enforced by thresholding operator $G[X|\delta] = X \cdot 1_{\{|X|>\delta\}}$.

An alternative to manually *selecting* a threshold is to use a collection of progressively decreasing thresholds $\delta_1 > \delta_2 > , ..., > \delta_K = 0$ [15]. Such strategy of progressive thresholding can be interpreted under a set theoretic framework, which makes it more widely applicable as a sparsity-enforcing tool. Referring to Fig. 1, we use a two-pixel image to illustrate how progressive thresholding works. The target image $\mathbf{x}$ is assumed to at the intersection of C_{obs} (observation constraint specified by the likelihood term) and an oracle manifold $\mathcal{M}_I : \{\mathbf{x}|\mathbf{x} = \mathcal{B}[\mathbf{x}]\}$ (prior term unknown to us) where $\mathcal{B} = T^{-1}GT$ is the nonlinear thresholding operator in the transform domain. With the knowledge of oracle threshold δ_{opt}, we can approximate $\mathcal{M}_I$ is by starting from some region (defined by initial threshold δ_1) and gradually increasing the search range by lowering the
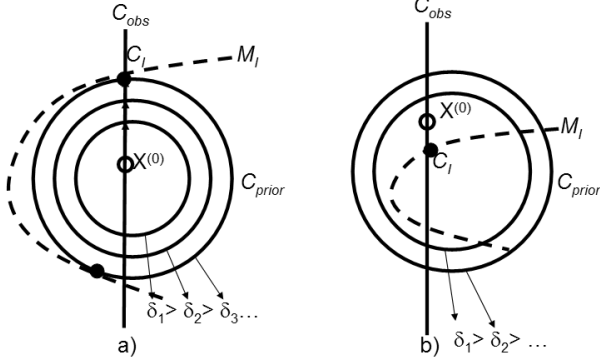
Figure 1. Toy-example illustration of POCS with progressively decreasing thresholds: C_{obs} (a line) and C_{prior} (a disk) respectively denote the likelihood and prior constraints for a two-pixel image; targeted image is located at the intersection which represents a local approximation of the unknown manifold M_I. a) as threshold δ_k approaches zero, the support of prior constraint set C_{prior} increases, which implies a larger search space; b) an over-shooting initial threshold δ_1 would not push $\mathbf{x}^{(k)}$ towards M_I because the search space is too large.

threshold. The catch is: if $\mathbf{x} \in C_I$ is a fixed point on $\mathcal{M}_I$ for some threshold δ_k, then it will remain a fixed point for all thresholds larger than δ_k (i.e., $\mathbf{x}^{(k)}$ will not diverge as the threshold further decreases).

The contraction mapping property of $\mathcal{B}$ has been established in [13]. The generality of such fixed-point perspective allows us to enforce sparsity in both local and nonlocal models. For example, instead of using overcomplete DCT or wavelet transforms, we can enforce the sparsity constraint with respect to packed similar patches as BM3D does [8], [11]. One pitfall with the progressive thresholding strategy lies in the choice of initial threshold δ_1. If the initial threshold δ_1 is too small (i.e., overfitting - C_{prior} would contain not only $\mathcal{M}_I$ but also less meaningful images), projection onto C_{prior} will move little because the constraint is too weak to drag $\mathbf{x}^{(k)}$ toward $\mathcal{M}_I$. On the other hand, if the initial threshold δ_1 is too large (underfitting - C_{prior} might not intersect with $\mathcal{M}_I$ at all), it will take a long time to converge and alternating projections in early iterations would be a waste of computations. In practice, we have found the following heuristic rule to work reasonably well - δ_1 is set to be the average of all local variances calculated on a block-by-block basis. The complete patch-based image interpolation algorithm is summarized as follows (note that it is conceptually similar to [11]).

Algorithm 1. Patch-based Nonlocal Image Interpolation
Input: observation data $\mathbf{y}$ and sampling pattern $\mathbf{S}$;
Output: reconstructed image $\hat{\mathbf{x}}$
• Initialization: obtain $\hat{\mathbf{x}}^{(0)}$ by a local interpolation method and calculate $\delta_1 = lv_{avg}$;
• Main loop: start with $k = 1$, set $\delta_k = lv_{avg} - (k-1)\Delta$
   Projection onto prior constraint set: apply BM3D filtering to $\hat{\mathbf{x}}^{(k)}$ with threshold δ_k;
   Projection onto observation constraint set: $\hat{x}^{(k)}(m,n) = y(m,n)$ for $\{(m,n)|S(m,n) = 1\}$;
   If $||\hat{\mathbf{x}}^{(k)} - \hat{\mathbf{x}}^{(k-1)}|| < \epsilon$, stop the iteration; otherwise $k = k + 1$.

III. MODELING SAMPLING PATTERN VIA SPATIAL POINT PROCESS

Now, we turn our attention to the likelihood term - in particular, the design of sampling pattern $\mathbf{S}$. In the literature of spatial statistics [16], sampling pattern is viewed as the realization of a spatial point process. Let $\{\mathbf{s}_i\}$ be the location of events in R^d (we only consider $d = 2$ here). A spatial point process $\mathcal{S}$ on R^2 is a measurable mapping defined by $\phi(S) = \sum_i 1_{\mathbf{s}_i \in S}$ (i.e., the number of events in the realization S). The complete spatial randomness of a point process is characterized by homogeneous Poisson process [17]. For any infinitesimal region $d\mathbf{s} \subset R^2$, the point process $\mathcal{S}$ is a homogeneous Poisson process if 1) $\lim_{|d\mathbf{s}| \to 0} \frac{1 - P(\phi(d\mathbf{s}) = 0)}{|d\mathbf{s}|} = \lambda$; 2) $\lim_{|d\mathbf{s}| \to 0} \frac{P(\phi(d\mathbf{s}) = 1)}{|d\mathbf{s}|} = \lambda$; 3) $\phi(d\mathbf{s}_1), \phi(d\mathbf{s}_2), ...$ are statistically independent for any disjoint sequence of regions $d\mathbf{s}_1, d\mathbf{s}_2, ...$; where $|B|$ denotes the volume of set $B \subset R^d$ and λ is the intensity parameter. It is known that the superposition of two independent Poisson process with parameters λ_1, λ_2 is a Poisson process with parameter $\lambda_1 + \lambda_2$.

There are several classes of measures for quantifying the randomness of a spatial point pattern. The simplest class is quadrat-based - i.e., taking any subset $Q \subset R^2$ of arbitrary shape at random location, the expected counting measure $\phi(Q)$ should approach its mean value λ. By repeating such test for a large number of times (i.e., using $\{Q_i\}$), we can collect a sequence of countering measures $\{\phi(Q_i)\}$. Its deviation from the mean (i.e., the variance $v(\mathbf{S}) = E[(\phi(Q) - \lambda)^2]$) can be used as an indicator of randomness. Other measures (e.g., distance-based) also exist though often involve higher computational complexity due to the need of finding nearest neighbors.

With the definition of randomness measure, a systematic solution is to simulate spatial point processes via Monte-Carlo method [21], [22]. Starting with any initial configuration S_0, one can randomly choose a pair of pixels with opposite attributes to flip. If the randomness measure $v(\mathbf{S})$ decreases, accept the new configuration; otherwise accept the new configuration with a probability of $exp(-(v(\mathbf{S}^{new}) - v(\mathbf{S}^{old})))$. Such Metropolis-like algorithm can be viewed as an iterative optimization procedure of maximizing the randomness of a spatial point process and guarantees to reach a local optimum. To speed up the convergence, heuristics-based solution can be used as the initial configuration. The complete description of Monte-Carlo based algorithm for sampling pattern randomization is given below.

Algorithm 2. Sampling Pattern Randomness Maximization via Monte-Carlo Method
Input: Arbitrary spatial point process $\mathbf{S}_0$;
Output: Randomized spatial point process $\mathbf{S}_\infty$;
• Initialization: calculate the randomness measure for initial configuration $v(\mathbf{S}_0)$;
• Main loop: for $k = 0, 1, 2, ...$
   Randomly pick one location (r_i, c_i) from $\mathbf{S}_k$ and another (r_j, c_j) from $\mathbf{S}_k^c$;
   Exchange (r_i, c_i) with (r_j, c_j) to generate a new proposal configuration $\mathbf{S}_p$;
   Generate a random number a uniformly distributed over the interval $[0, 1]$;
   If $v(\mathbf{S}_p) < v(\mathbf{S}_k)$ or $a > exp(v(\mathbf{S}_k) - v(\mathbf{S}_p))$, let $\mathbf{S}_{k+1} = \mathbf{S}_p$; otherwise $\mathbf{S}_{k+1} = \mathbf{S}_k$.

IV. APPLICATION INTO COMPRESSIVE IMAGING

Conventional wisdom in designing image communication systems is to divide-and-conquer. For example, image acquisition,

compression and transmission are often handled by three independent components namely sensor (analog-digital conversion), source coding (redundancy reduction) and channel coding (redundancy addition). Such divide-and-conquer strategy is often suboptimal when other resource constraints (e.g., the cost of analog-digital conversion, transmission bandwidth, error rate and delay). There is a new trend of optimizing the system by jointly designing those components (e.g., joint sensing-processing, joint source-channel coding). Algorithms 1 and 2 presented in this work open the door to design a new compressive imaging system suitable for wireless sensor network applications as shown in Fig. 2 (such system can also be viewed as a distributed coding of image source since complexity has been shifted from encoder to decoder).

The most significant departure from the current practice is that only a small percentage of observation data at *random* locations need to be acquired and transmitted (therefore the complexity of sensor node is extremely low). Note that here observation data have a general meaning - e.g., pixel values in Alg. 1 can be replaced by Gabor filtered samples (accordingly the definition of observation constraint sets varies). We also note that the sampling pattern $\mathbf{S}$ (generated by Alg. 2) is assumed to be shared by both encoder and decoder. Therefore there is no need for transmitting any overhead related to location information (by contrast, conventional image coding systems strive to achieve the optimal balance between bits spent on coding location and intensity values). With an appropriately chosen intensity parameter λ, images are reconstructed from the quantized samples at the decoder by Alg. 1. In other words, our guiding principle has evolved from the traditional redundancy reduction to redundancy exploitation, which has been argued as a more plausible hypothesis for sensory processing by human visual cortexes [23].
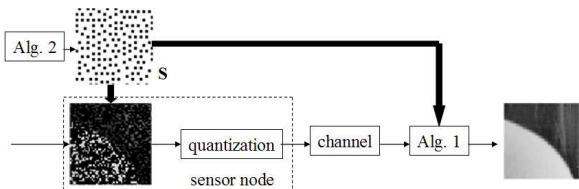

Figure 2. Compressive imaging system built upon Algorithms 1 and 2.

Our compressive imaging system has good error-resilient properties On one hand, since the task of redundancy exploitation is located at the decoder, image quality would have a graceful degradation in the presence of channel errors (damaged samples and missing samples are both exceptions to sparsity constraint which Alg. 1 attempts to correct). On the other hand, superposition property of Poisson process makes it convenient to generate multiple descriptions of an image [24], which can support joint reconstruction from an arbitrary collection of descriptions received by the decoder (thanks to the superposition property of Poisson process). We note that the diversity of descriptions can arise from both sampling locations (e.g., $\mathbf{S}_1 \cap \mathbf{S}_2 = \phi$) and filter responses (e.g., Gabor filters with different orientation and scale parameters resembling the rod/cone receptors in human retinas). To simplify the implementation, we have only considered the location-related diversity in our experiments (i.e., a delta-function filter is assumed).
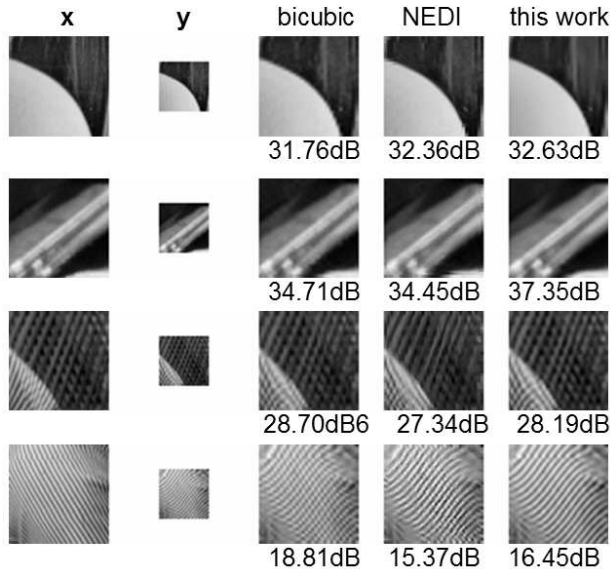

Figure 3. PSNR and subjective quality comparison among different reconstruction schemes for four test images: DT-based (starting point of Alg. 1), NEDI [3] and Alg. 1 (this work).

V. EXPERIMENTAL RESULTS

V-A. Image Interpolation

The implementation of Algorithm 1 is based on the public availability of BM3D demo package[1]. The parameters in Algorithm 1 are fixed in all our experiments: $\Delta = 0.02, \epsilon = 0.001, B = 16, d = 4, N_{max} = 16, N_s = 32$ (the last four parameters are used by BM3D algorithm). Such choice has been found to achieve good results for a wide class of images. The MATLAB codes and sample images accompanying this paper can be found at http://www.csee.wvu.edu/~xinl/code/patch_recon.zip.

In the first experiment, we apply Alg. 1 to the interpolation of uniformly-sampled images (i.e., $\mathbf{y}$ is a down-sampled version of $\mathbf{x}$). The starting point is chosen to be bicubic interpolation and benchmark scheme also includes new edge-directed interpolation (NEDI) [3]. Fig. 3 shows the performance comparison among different schemes. It can be seen the benefit of exploiting nonlocal dependency is apparent for resolution enhancement of regular edges (due to their scale-invariant properties). The findings for textures are more interesting - for example, for the first texture image, spatial aliasing is less severe and even bicubic interpolation achieved good result (NEDI is not appropriate for this image due to its abundance of textures). Although patch-based nonlocal interpolation gives slightly lower PSNR than bicubic, we note that the gain remains positive when measured by SSIM metric (bicubic: 0.8803, ours: 0.8860).

For the second texture image, the reconstructed image does not look like the original (due to severe spatial aliasing) but still represents a perceptually plausible solution. This suggests that image manifold M_I could have multiple intersection points with constraint set C_{prior} - when the likelihood term is not sufficient to resolve the ambiguity, reconstructed image might appear drastically different from the original one (refer to the two solid dots in Fig. 1a). In Fig. 4, we have compared

[1]The MEX-based demo can be downloaded from http://www.cs.tut.fi/~foi/GCF-BM3D/
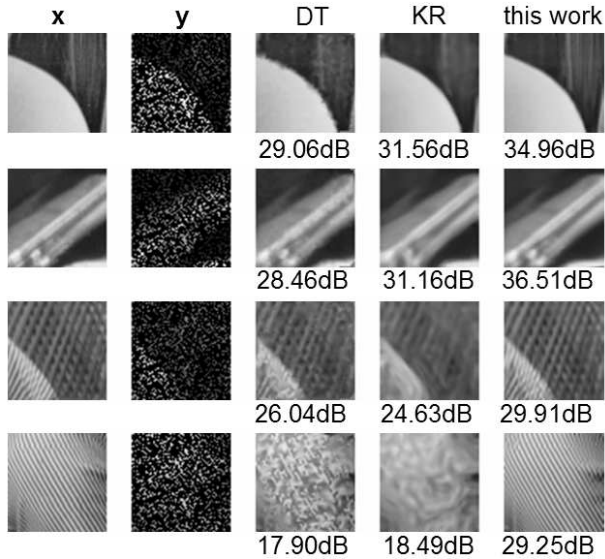
Figure 5. PSNR and subjective quality comparison among different reconstruction schemes for four test images: DT-based (starting point of Algo. 1), KR-based [4] and Algo. 1 (this work).

the results with three different sampling matrixes with the same amount of data (75% missing) but increasing amount of randomness: uniform sampling, Hilbert-scanning and pseudo-random sampling (no randomness optimization involved). The benefit of randomness in image reconstruction can already be clearly seen for this example (more justification comes later).

In our second experiment, we test the performance of Alg. 1 on interpolating nonuniformly-sampled images. Observation data $\mathbf{y}$ is obtained by randomly throwing away 75% pixels in an image. We use Delaunay-triangle (DT) based interpolation (implemented by MATLAB function $griddata$) to generate the starting point $\hat{\mathbf{x}}^{(0)}$. Three reconstruction schemes are compared: Delaunay-triangle (DT) based interpolation (starting point of Algorithm 1), kernel-regression (KR) based interpolation [4] and Alg. 1. Fig. 5 includes the comparison for four images: two edge-dominated and the other two texture-abundant. It can be observed that our work dramatically outperforms other competing schemes (the average PSNR gain is over $4dB$) especially on texture images whose nonlocal dependency can not be effectively characterized by local models.

Figs. 6 and 7 include the comparison of reconstructed images for two generic images: $house$ (edge-abundant) and $lena$ (mixture of regular edges and irregular textures). Similar to the experimental setup in [4], we have chosen sampling matrix S to randomly keep 15% pixels in the original image. We observe that the PSNR performance improvement for $lena$ is less impressive than that for $house$ due to the presence of irregular texture patterns. However, the visual quality of reconstructed image by patch-based nonlocal interpolation appears to be the highest (fewer artifacts). When measured by a recently proposed objective image quality metric named SSIM [25], we have found the PN scheme achieves the best results and the gain becomes even more noticeable.

V-B. Compressive Imaging

For compressive imaging application, we first demonstrate how optimizing the randomness of sampling pattern $\mathbf{S}$ by Alg. 2
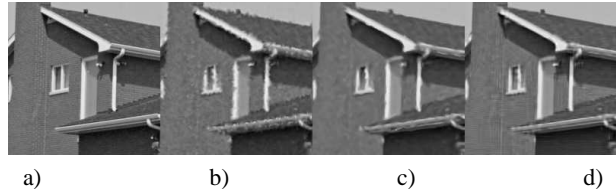

Figure 6. Subjective comparison among a) original $house$ image and reconstructed images by b) DT-based interpolation ($PSNR = 28.01dB, SSIM = 0.8212$), c) kernel-regression based scheme ($PSNR = 29.39dB, SSIM = 0.8343$) and d) this work ($PSNR = 30.96dB, SSIM = 0.8787$).
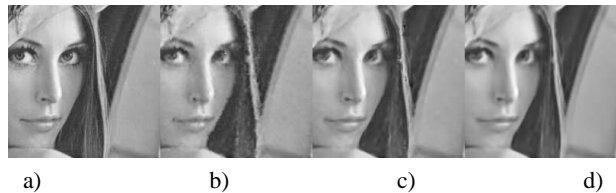

Figure 7. Subjective comparison among a) original $lena512$ image and reconstructed images by b) DT-based interpolation ($PSNR = 28.92dB, SSIM = 0.8443$), c) kernel-regression based scheme ($PSNR = 29.74dB, SSIM = 0.8485$) and d) this work ($PSNR = 30.14dB, SSIM = 0.8687$).

improves the performance of Alg. 1. The starting point of Alg. 2 ($\mathbf{S}_0$) is obtained by randomly choosing $(100 \cdot \lambda)\%$ locations from the sampling lattice of a given image as the location of events $\{\mathbf{s}_i\}$. It is known such strategy is only sufficient for 1D; for instance, the outcome in 2D often demonstrates structures (e.g., clumps, constellations as shown in Fig. 9a). Note that Alg. 2 can work any definition of randomness measure. In our implementation, we have adopted a combination of quadrat-based and distance-based measures.

To illustrate the behavior of Alg. 2, we have taken an example of $\lambda = 0.0625$ on a 64×64 lattice. Fig. 8 includes the evolution of randomness measure for the first 2000 iterations. The initial condition $\mathbf{S}_0$ and final result $\mathbf{S}_2000$ are displayed in Fig. 9a. Even visual inspection can easily verify that the sampling pattern after optimization appears more random. The PSNR gain brought by such randomness optimization is dramatic as can be seen from Fig. 9b. We note that such significant benefit of randomizing the sampling locations has clearly been harvested by human eyes as the result of evolution (refer to [26] for an image of spatially randomly-distributed rod/cone cells).

We have designed a batch of randomness-optimized sampling patterns with various λ's for image coding system of Fig. 2. The quantization in our system is the simplest uniform quantization with an adjustable stepsize Q (intensity values at sampled locations are simply treated as a 1D signal and fit to a finite mixture of Gaussian models [27]). In our preliminary study, we have fixed $Q = 8$ and compared the R-D performance with the popular wavelet-based SPIHT scheme [19]. Figs. 10 and 11 include the comparison between our system and SPIHT for two test images: $\lambda = 0.0625$ for the edge image and $\lambda = 0.25$ for the texture image. Both images are compressed at the same bit rate to have a fair comparison. The superiority of our system on texture image is apparent (this is not surprising because it is long known that localized wavelet bases are not optimal for representing textures). For edge image, our system
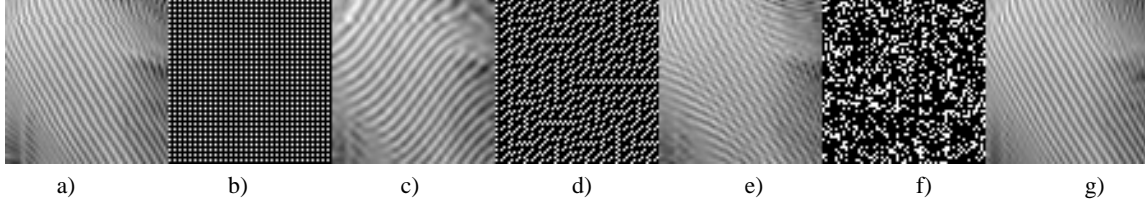
Figure 4. Illustration of how more randomness in sampling renders better reconstruction. a) original image; b),d),f) three different sampling patterns $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$; c),e),g) reconstructed images from $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$.
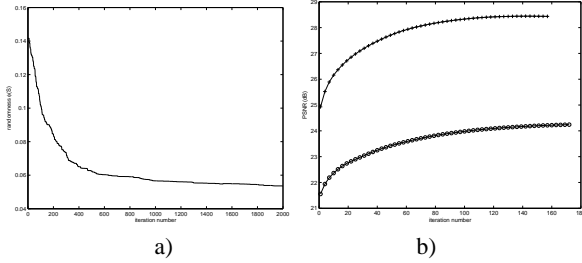

Figure 8. a) Evolution of randomness measure $v(\mathbf{S})$ in Alg. 2 (2000 iterations are performed); b) PSNR performance comparison between two sampling patterns ('+' - after optimization; 'o' - before optimization).
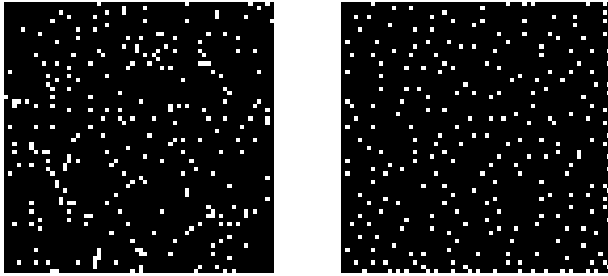

Figure 9. Random sampling pattern generation via Metropolis algorithm before (left) and after (right) optimization (the pattern on the right appears more random).
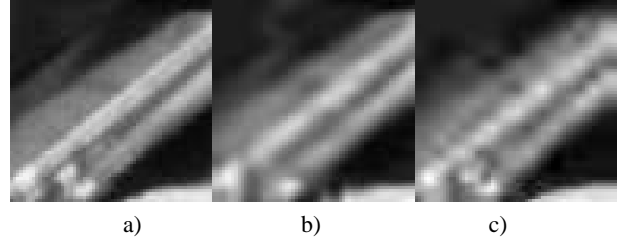

Figure 10. Performance comparison between ours and SPIHT on a regular edge image at the same rate of 0.21bpp: a) original; b) this work ($PSNR = 27.85dB, SSIM = 0.8750$); c) SPIHT ($PSNR = 28.82dB, SSIM = 0.8637$).
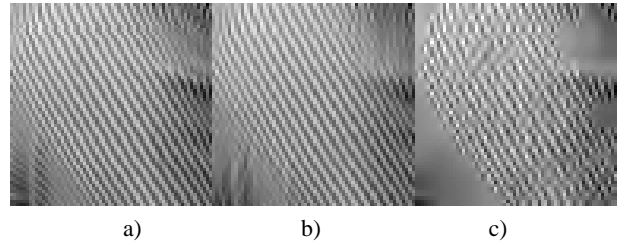

Figure 11. Performance comparison between ours and SPIHT on a regular texture image at the same rate of 0.81bpp: a) original; b) this work ($PSNR = 28.10dB, SSIM = 0.9182$); c) SPIHT ($PSNR = 22.98dB, SSIM = 0.7512$).

have achieved noticeably better subjective quality (e.g., fewer artifacts). Note that ours is 1dB worse than SPIHT in terms of PSNR but more favorable on SSIM metric which has shown to better correlate with subjective results.

Last, we want to demonstrate the error resilience of our coding system. To simulate the channel impairments, we assume only a subset of the samples at the locations $\{s_i\}$ are successfully received. If the error rate $p \geq 0$ denotes the percentage of lost samples, Alg. 1 at the decoder simply work with a Poisson process with intensity parameter $\lambda - p$. For each error rate, we generate 10 independent error patterns and take the average PSNR value as the final result. Fig. 12 shows the PSNR profile as p gradually increases and Fig. 13 contains several examples of decoded images at different error rates. The graceful degradation of both objective and subjective qualities justifies the insensitivity of our coding system.

VI. CONCLUSIONS

The contributions of this work are two-fold. First, two new algorithms related to patch-based image interpolation are pre-
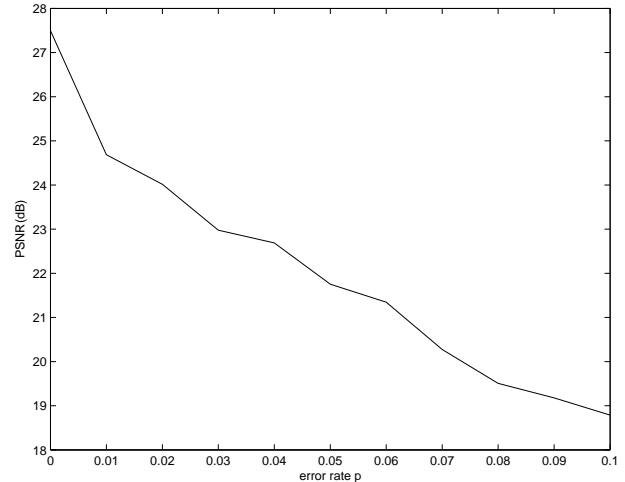

Figure 12. PSNR performance degradation of Alg. 1 in the presence of errors.
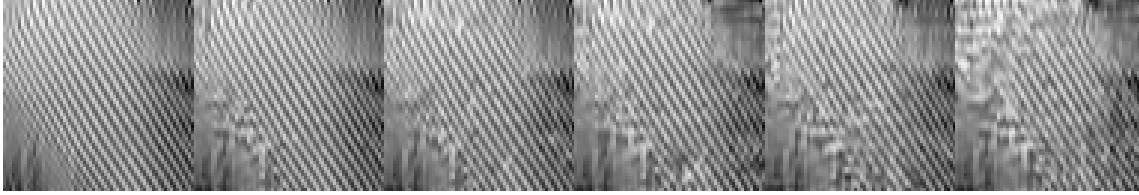
Figure 13. Decoded images at different error rates: from left to right, $p = 0, 0.02, 0.04, 0.06, 0.08, 0.10$.

sented: one improves previous works by enforcing a nonlocal sparsity-based constraint (prior optimization) and the other targets at increasing the randomness of sampling patterns (likelihood optimization). These two algorithms jointly produce a better solution to the sensing of image signals. Second, a new compressive imaging system is built on the proposed algorithms and its coding performance has shown to competitive with state-of-the-art wavelet coding algorithm SPIHT. The good error resilience property of this system is demonstrated by our preliminary experimental results.

VII. REFERENCES

[1] H. S. Hou and H. C. Andrews, "Cubic Splines for Image Interpolation and Digital Filtering," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 26, pp. 508–517, 1978.

[2] K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images." *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 285–295, 1995.

[3] X. Li and M. Orchard, "New edge directed interpolation," *IEEE Trans. on Image Processing*, vol. 10, pp. 1521–1527, 2001.

[4] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, 2007.

[5] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *International Conference on Computer Vision*, 1999, pp. 1033–1038.

[6] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, September 2004.

[7] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," *cvpr*, vol. 2, pp. 60–65, 2005.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Trans. on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[9] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, pp. 56–65, 2002.

[10] W. P. H.Q. Luong, A. Ledda, "Non-local interpolation," in *ICIP*, 2006, pp. 693–696.

[11] K. Egiazarian, A. Foi, and V. Katkovnik, "Compressed sensing image reconstruction via recursive spatially adaptive filtering," in *IEEE International Conference on Image Processing*, vol. 1, San Antonio, TX, USA, Sep 2007.

[12] P. Ishwar and P. Moulin, "On the equivalence of set-theoretic and maxent MAP estimation," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 51, no. 3, pp. 698–713, Mar 2003.

[13] O. G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part i: theory." *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 539–554, 2006.

[14] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, pp. 1338–1351, Nov 2003.

[15] O. G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part ii: adaptive algorithms." *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 555–571, 2006.

[16] B. D. Ripley, *Spatial statistics*. Wiley Series in Probability and Mathematical Statistics, 1981.

[17] N. Cressie, *Statistics for Spatial Data*. John Wiley&Sons, 1993.

[18] B. D. Ripley, "Tests of 'randomness' for spatial point patterns," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 41, no. 3, pp. 368–374, 1979.

[19] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243–250, 1996.

[20] B. Olshausen and D. Field, "Sparse coding with over-complete basis set: a strategy employed by v1?" *Vision Research*, vol. 37, pp. 3311–3325, 1997.

[21] N. Metropolis and S. Ulam, "The monte carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, Sep. 1949.

[22] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*. Springer Series in Statistics, 2001.

[23] H. Barlow, "Redundancy reduction revisited," *Network: Computation in Neural Systems*, vol. 12, pp. 241–253(13), 1 March 2001.

[24] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Transactions on Image Processing*, vol. 10, pp. 351–366, 2001.

[25] 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.

[26] D. H. Hubel, *Eye, Brain, and Vision*. Scientific American Library, No 22, 1988.

[27] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, 2002.