Hindawi Publishing Corporation
EURASIP Journal on Image and Video Processing
Volume 2007, Article ID 41516, 12 pages
doi:10.1155/2007/41516

Research Article

Image Resolution Enhancement via Data-Driven Parametric Models in the Wavelet Space

Xin Li

Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506-6109, USA

Received 11 August 2006; Revised 29 December 2006; Accepted 9 January 2007

Recommended by James E. Fowler

We present a data-driven, project-based algorithm which enhances image resolution by extrapolating high-band wavelet coefficients. High-resolution images are reconstructed by alternating the projections onto two constraint sets: the *observation* constraint defined by the given low-resolution image and the *prior* constraint derived from the training data at the high resolution (HR). Two types of prior constraints are considered: spatially homogeneous constraint suitable for texture images and patch-based inhomogeneous one for generic images. A probabilistic fusion strategy is developed for combining reconstructed HR patches when overlapping (redundancy) is present. It is argued that objective fidelity measure is important to evaluate the performance of resolution enhancement techniques and the role of antialiasing filter should be properly addressed. Experimental results are reported to show that our projection-based approach can achieve both good subjective and objective performance especially for the class of texture images.

1. INTRODUCTION

Depending on the presence of antialiasing filer, there are two ways of formulating the resolution enhancement problem for still images—that is, how to obtain a high-resolution (HR) image from its low-resolution (LR) version? When no antialiasing filter is used (see Figure 1(a)), we might use classical linear interpolation [1], edge-sensitive filter [2], directional interpolation [3], POCS-based interpolation [4], or edge-directed interpolation schemes [5, 6]. When antialiasing filter is involved (see Figure 1(b)), resolution enhancement is twisted with contrast enhancement by deblurring which is an ill-posed problem itself [7].

When antialiasing filter takes the form of lowpass filter in wavelet transforms (WT) [8], there are a flurry of works [9–17] which transform the problem of resolution enhancement in the spatial domain to the problem of high-band extrapolation in the wavelet space. The apparent advantages of wavelet-based approaches include numerical stability and potential leverage into image coding applications (e.g., [18]). However, one tricky issue lies in the performance evaluation of resolution enhancement techniques—should we use subjective quality of high-resolution (HR) images or objective fidelity such as mean-square errors (MSE)?

The difficulty with the subjective option lies in that it opens the door to allow various contrast enhancement techniques as a postprocessing step after resolution enhancement. Both linear (e.g., [19]) and nonlinear (e.g., [20]) techniques have been proposed in the literature for sharpening reconstructed HR images. We note that *contrast* and *resolution* are two separate issues related to visual quality of still images. Tangling them together will only make the problem formulation less clean because it makes a fair comparison more difficult—that is, whether quality improvement comes from resolution enhancement or contrast enhancement? Therefore, we argue that subjective quality should not be used alone in the assessment of resolution enhancement schemes. Moreover, objective fidelity such as MSE can measure the closeness of computational approaches to the more cost-demanding optics-based solutions, which is supplementary to subjective quality indexes.

However, MSE-based performance comparison could be misleading if the role of antialiasing filter is not properly accounted. For example, in the presence of antialiasing filter, bilinear or bicubic interpolation would not be appropriate benchmark unless the knowledge of antialiasing filter is exploited by the reconstruction algorithm. To see this more clearly, we can envision a "lazy" scheme which simply pads

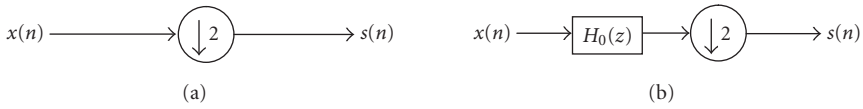
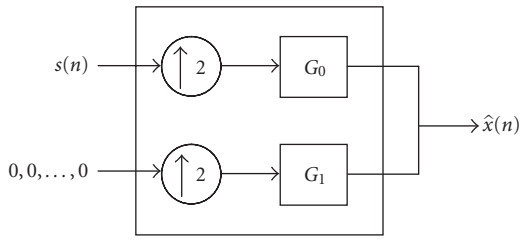


Figure 1: Two ways of formulating the resolution enhancement problem in 1D (2D generalization is straightforward): (a) without antialiasing filter; (b) with antialiasing filter H_0 (lowpass filter in wavelet transforms).



	Carey et al.'s scheme △PSNR	Lazy scheme △PSNR
Lena	4.2	4.9
Mandrill	0.2	1.3
Peppers	2.7	4.3

(b)

Figure 2: (a) Diagram of lazy scheme (padding zeros to high band); (b) comparison of PSNR gains (dB) over bicubic between [10] and lazy scheme for three USC test images. Note that zero-padding-based lazy scheme achieves even higher PSNR values than more sophisticated scheme [10].

zeros into the three high bands before doing inverse WT (refer to Figure 2(a)). Figure 2(b) shows the PSNR gain of lazy scheme over bicubic interpolation—note that the impressive gain is not due to the ingeniousness of the lazy scheme itself but an unfair comparison because bicubic interpolation does not make use of the antialiasing filter at all. Unfortunately, such subtle difference caused by antialiasing filter appears to be largely ignored in the literature [10–15] which use bilinear/bicubic interpolation as the benchmark.

In this paper, we propose a data-driven, projection-based approach toward resolution enhancement by extrapolating high-band wavelet coefficients. Our work is built upon parametric wavelet-based texture synthesis [21] and nonparametric example-based superresolution (SR) [22]. Similar to [22], we also assume the availability of some HR images as the training data; however, our extrapolation method is based on the parametric model proposed in [21]. Since parametric texture models [21] cannot be directly used for resolution enhancement of generic images due to their inhomogeneity, we propose to use [22] as a preprocessing step of preparing HR training patches to drive parametric models. Moreover, to reduce the artifacts introduced by patch-based representations, we propose a strategy of probabilistically fusing the overlapped patches synthesized at the HR, which can be viewed as the extension of averaging strategy adopted by [22].

The rest of the paper is structured as follows. In Section 2, we briefly cover the background and motivation behind our approach. In Section 3, we present a basic extension of synthesis technique [21] for resolution enhancement of spatially homogeneous textures. In Section 4, we generalize our new resolution enhancement into the spatially inhomogeneous case by introducing patch-based representation and weighted linear fusion. Experimental results are reported in Section 5 to demonstrate the performance of our schemes and we make final concluding remarks in Section 6.

2. PROBLEM FORMULATION AND MOTIVATION

In wavelet-space extrapolation, the objective is to obtain an estimation of high-band coefficients $\hat{d}(n)$ from $s(n)$ (refer to Figure 3). Due to aliasing introduced by the down-sampling operator, such inter-band prediction (note its difference from interscale prediction in wavelet-based image coding [18]) is not expected to work unless we impose some constraints on the original HR signal $x(n)$. For example, it is well known that in 1D scenario, the way that extrema points of isolated singularities propagate across the scales can be characterized by local Lipschitz regularity [23]. Many previous wavelet-based interpolation schemes (e.g., [9, 10]) are based on such observation.

However, there are caveats with the above observation. First, aliasing introduced by the down-sampling operator adds phase ambiguity to the extrapolation problem. That is, the extrema points across the scales cannot be exactly located due to the phase uncertainty. Additional constraints are required to help partially resolve such ambiguity. Such issue was insightfully pointed out by the authors of [9, 16], but the success has been limited to subjective quality improvement so far. In fact, if such ambiguity is not properly resolved, the predicted high-frequency band is often no better than zero-padding in the lazy scheme (i.e., lower MSE cannot be achieved). Second and more importantly, the problem of inter-band prediction becomes dramatically more difficult in 2D scenario due to the increased complexity of modeling image signals in the wavelet space. The diversity of image structures in generic images (e.g., edges, textures, etc.) dramatically increases the difficulty of the extrapolation task.

The motivation behind our attack is largely based on the existing parametric models [21] for texture synthesis in the wavelet space. However, we face two obstacles while applying parametric models into resolution enhancement: *aliasing* and *inhomogeneity*. Aliasing makes the parameter extraction

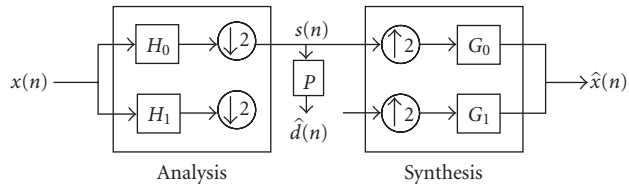


Figure 3: Problem formulation in 1D scenario: in wavelet-based interpolation, interscale prediction is designed to predict high-band coefficients from the low-band ones at the same scale.

nontrivial (essentially a missing data problem) and inhomogeneity calls for spatially varying (or localized) models. To overcome those difficulties, we borrow ideas from data-driven or example-based superresolution (SR) [22] to make the problem tractable. Assuming the availability of some correlated HR images as training data, we propose to use nonparametric sampling [22] to first generate initial HR patches, then use them to drive the parametric model to synthesize intermediate HR patches and lastly obtain the final HR patches via probabilistic fusion.

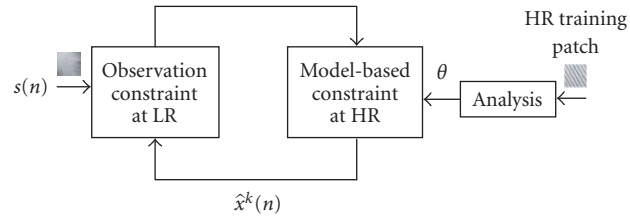


Figure 4: Resolution enhancement of textures: HR image is obtained by alternating the projection onto two constraint sets.

3. RESOLUTION ENHANCEMENT OF TEXTURE IMAGES

In this work, we have adopted a definition of textures in the narrow sense—that is, textures are modeled by a homogeneous (stationary) random field. Homogeneity refers to that the probability distribution function (pdf) is independent of the spatial position. Statistical modeling of textures has been extensively studied in the literature (see [24–26]). In recent years, multiscale approaches toward texture analysis and synthesis have also received more and more attention (e.g., [21, 27–29]). Both parametric and nonparametric models have been developed and demonstrated visually appealing synthesis results. Among them, parametric models in the wavelet space [21] are adopted as the foundation for this work.

Resolution enhancement, unlike synthesis, addresses a new dimension of challenge due to aliasing introduced by the down-sampling operation. Depending on the choice of antialiasing filter and the spectral distribution of texture images, we might observe significant visual difference between LR and HR pairs due to spatial aliasing. Even when aliasing does not dramatically change the visual appearance, HR image reconstructed by the lazy scheme often appears blurred due to the knock down of high-frequency coefficients. In previous works on wavelet-based interpolation such as [30], no experimental results are reported for texture images. According to [10], the PSNR gain of wavelet-based interpolation over bilinear/bicubic is almost unnoticeable for *mandrill* image which contains abundant texture regions.

In view of the difficulty with finding a universal prior constraint for textures, we propose to make additional assumption that some HR training patches are available (refer to Figure 5(a)). It is believed that such training data are necessary for resolution enhancement of textures because the problem is ill-posed (i.e., two HR images corresponding to the same LR data can be visually different). However, the size of training patch could be small since its role is to resolve the ambiguity among multiple solutions caused by aliasing. Specifically, we propose to combine patch-based prior constraint with observation data constraint (i.e., the low-low band in the wavelet space is specified by the given LR image) and reconstruct HR images by alternating projections (refer to Figure 4).

Various statistical models developed for texture synthesis (e.g., [21, 27, 28]) can be used to derive the prior constraint sets. Since the parametric model developed in [21] is projection-based and computationally efficient, we can easily build our resolution enhancement algorithm upon it. In [21], four types of statistical constraints (SC), namely, marginal statistics, raw coefficient correlation, coefficient magnitude statistics, and cross-scale phase statistics, are sequentially enforced to iteratively adjust complex high-band coefficients (we denote it by projection operator $P_{\text{sc}}[x]$). Mathematical details on adjustment of constraints can be found in the appendix of [21]. The implementation of projection onto observation constraint ($P_{\text{obs}}[x]$) is trivial—we simply replace the low-low band of x in the wavelet space by the given LR image (the MSE of low-low band is denoted by MSE_{LL}). By alternatively applying model-based prior constraint and data-driven observation constraint to high-band and low-band coefficients, we have the following algorithm.

Like any iterative schemes, starting point and stopping criterion are important to the performance of Algorithm 1. We have found that Algorithm 1 is reasonably robust to the starting point ($\hat{x}^0$) (one example can be found in Figure 10). We also note that unlike existing projection onto convex set (POCS) based algorithms [31], convergence is not a necessary condition even though we have found that MSE_{LL} often drops rapidly in the first few iterations and then goes saturated (refer to Figure 6(b)). In fact, as pointed out in [21], the convexity of constraint sets defined by parametric texture

(i) Initialization: extract the parameter set Θ from the training patch and obtain HR image $\hat{x}^0$ by lazy scheme or example-based SR [22].

(ii) Iterations: alternate the following two projections.

(1) Projection onto prior constraint set: sequentially run the projection onto four statistical constraint sets to modify the HR image

$$\hat{x}^{n+1} = P_{\text{sc}}[\hat{x}^n \mid \Theta]. \quad (1)$$

(2) Projection onto observation constraint set:

$$\hat{x}^{n+2} = P_{\text{obs}}[\hat{x}^{n+1}]. \quad (2)$$

(iii) Termination: if MSE_{LL} keeps decreasing, continue the iteration; otherwise stop.

Algorithm 1: Project-based resolution enhancement for textures.

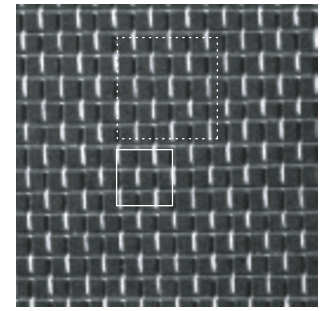
model is often unknown. However, in the application of resolution enhancement, our projection-based algorithm can be stopped by checking MSE_{LL} because it is correlated with the MSE of reconstructed HR image as shown in Figure 6. Despite the lack of theoretical justification, such empirical stopping criterion works fairly well in practice.

4. RESOLUTION ENHANCEMENT OF GENERIC IMAGES

Generic photographic images contain a variety of singularities including edges, textures, and so on. The diversity of singularities suggests that image source cannot be modeled by a globally stationary (homogeneous) process. A natural strategy of handling nonstationary process is via spatial localization—that is, to view an image as the composition of overlapping patches [22] (refer to Figure 5(b)). Such patch-based representation has led to many state-of-the-art image processing algorithms in both spatial and wavelet domains. Using patch-based representation, we decompose resolution enhancement of generic images into two subproblems: (1) how to enhance the resolution of a single patch? (2) How to combine the enhancement results obtained for overlapped patches? The first can be solved by Algorithm 1 except the generation of HR training patch; the second is related to the issue of global consistency due to the locality assumption of patches. We will study these two problems, respectively, next.

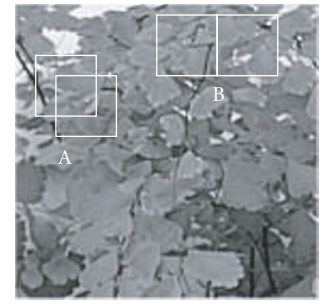
4.1. Single-patch resolution enhancement

Since generic images do not satisfy the assumption of global homogeneity, HR training patches have to be made spatially adaptive. Unlike texture images, how to generate an appropriate HR training patch is nontrivial due to the *location* uncertainty. In texture images, an HR patch of any location is arguably useable because of the homogeneity constraint (we will illustrate this in Figure 10). However, such flexibility



Testing patch
Training patch

(a)



A: Overlapping patches
B: Nonoverlapping patches

(b)

Figure 5: (a) Training patch and test patch in texture images; (b) overlapping and nonoverlapping patches in generic images.

does not hold for generic images any more—since the conditional probability distribution becomes a function of location, additional uncertainty needs to be resolved in the generation of HR training patches.

One solution to resolve such location uncertainty is through nonparametric sampling [22, 32]. In nonparametric sampling, patches with similar photometric patterns are clustered and new patch can be synthesized by sampling the empirical distribution. Such strategy cannot be directly applied here because the target to approximate is an LR patch and the population to draw from is the collection of HR patches. However, we can modify the distance metric in nonparametric sampling to accommodate such resolution discrepancy, that is,

$$d[\mathbf{x}_l, \mathbf{y}_h] = \left|\left|\mathbf{x}_l - DH(\mathbf{y}_h)\right|\right|_{L_2}, \quad (3)$$

where D, H denotes the down-sampling operation and convolution with antialiasing filter, respectively. When antialiasing filter H is the same as the lowpass filter of WT,

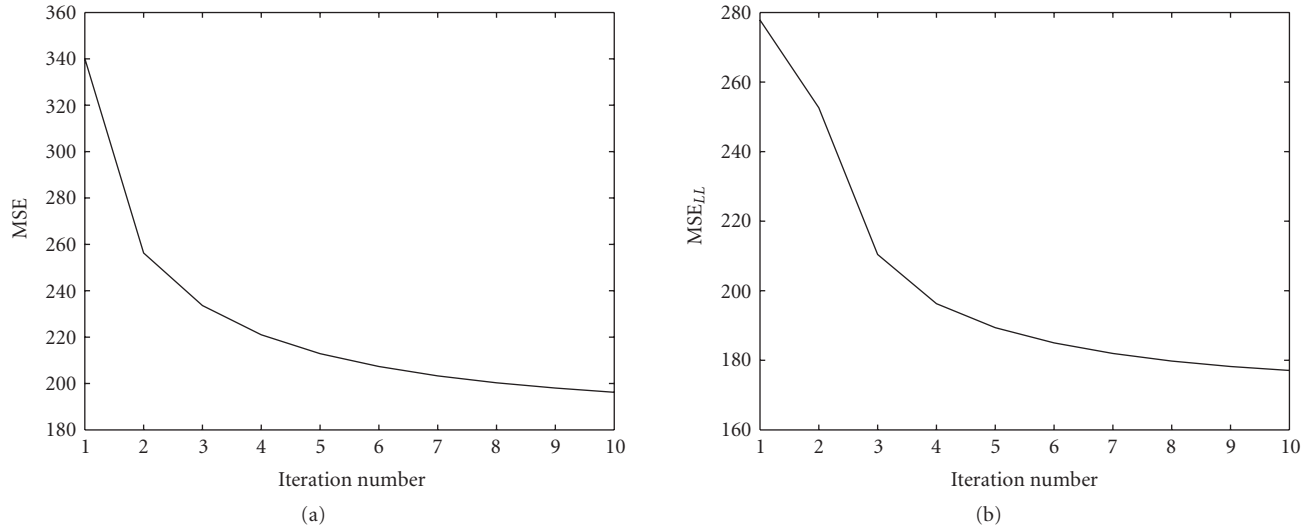


Figure 6: The behavior of iterative Algorithm 1: (a) MSE of reconstructed HR image; (b) MSE of low-low band MSE_{LL}. Note that they are highly correlated which empirically justifies the stopping criterion based on MSE_{LL}.

example-based superresolution [22] offers a convenient implementation of generating HR training patch.

Unlike [22], nonparametric sampling is used here to generate the initial rather than the final result. This is because although nonparametric sampling often produces perceptually appealing results, they do not necessarily have small L_2 distance to the ground truth. Therefore, we propose to use the outcome of nonparametric sampling as the training HR patch to drive the parametric texture model, as shown in Figure 7. Meantime, due to the descriptive nature of parametric texture models, synthesized images might have similar statistical properties such as marginal or joint pdf but large L_2 distance to the original. Such weakness with parametric models can be alleviated by defining a new prior constraint projection operator P'_{sc}

$$\hat{x}^{k+1} = P'_{\text{sc}}\left[\hat{x}^k\right] = \frac{P_{\text{sc}}\left[\hat{x}^k\right] + \hat{x}^0}{2}. \tag{4}$$

Such modification can be viewed as adding a bounded variation constraint enforcing the initial condition $\hat{x}^0$.

Such combination of nonparametric and parametric sampling is important to achieve good performance in terms of both subjective quality and objective fidelity. On one hand, it extends the parametric texture model [21] by introducing nonparametric sampling to generate training patches required at the HR. Despite being conceptually simple, such extension effectively overcomes the difficulty of resolution discrepancy and handles inhomogeneity in generic images. On the other hand, our combined scheme is more robust to training data than example-based SR [22]. This is because parametric texture model can tolerate some errors in the initial estimate as long as they do not significantly change the four types of statistical constraints.

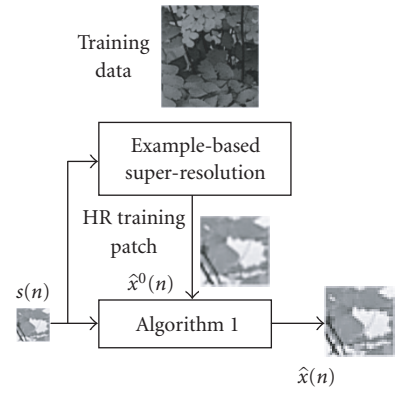


Figure 7: Algorithm 2 for resolution enhancement of a single patch (example-based SR provides an initial result to drive the parametric texture model).

4.2. Bayesian fusion of overlapped HR patches

When patches overlap with each other, a pixel might be included into multiple patches and therefore the pixel can have more than one HR synthesized result (refer to Figure 5(b)). Such redundancy is the outcome of spatial localization—although it effectively reduces the dimensionality, the potential inconsistency across patches arises. For instance, how to consolidate the multiple synthesis results generated by overlapping patches is related to the enforcement of global consistency. In example-based SR [22], multiple HR versions are simply averaged to produce the final result. Although averaging represents the simplest way of enforcing global consistency across patches, its optimality is questionable especially due to the ignorance of the impact of location (i.e., whether a pixel is at the center or at the border of a patch) on the fusion performance. We propose to formulate such

patch-based fusion problem under a Bayesian framework and derive a closed-form solution as follows.

Using patch-based representation, we adopt the following probability model for each pixel:

$$p(x) = \int p(x,z)dz = \int p(x \mid z)p(z)dz, \tag{5}$$

where the new random variable z denotes the *location* of pixel x in the patch. Given a set of HR reconstruction results $\mathbf{y} = [y_1, \ldots, y_k, \ldots, y_N]$ (k is the discretized version of location variable z, N is the total number of patches containing x), the Bayesian least-square estimator is

$$\begin{aligned} E[x \mid \mathbf{y}] &= \int x p(x \mid \mathbf{y})dx \\ &= \int x p(x,z \mid \mathbf{y})dx\,dz \\ &= \int x p(x \mid z,\mathbf{y}) p(z \mid \mathbf{y})dx\,dz \\ &= \int p(z \mid \mathbf{y}) E[x \mid z,\mathbf{y}]dz. \end{aligned} \tag{6}$$

Note that when z is given (i.e., the indexing k of HR patch y_k), we have $E[x \mid k,\mathbf{y}] = y_k$ and (6) boils down to

$$\hat{x} = E[x \mid \mathbf{y}] = \sum_{k=1}^{N} w_k y_k, \tag{7}$$

where $w_k = p(k \mid y_k)$ is the weighting coefficient for the kth patch. To determine w_k, we use Bayesian rule

$$p(k \mid y_k) = \frac{p(y_k \mid k)p(k)}{\sum_k p(y_k \mid k)p(k)}, \tag{8}$$

where likelihood function $p(y_k \mid k)$ (the likelihood of pixel x belonging to the kth patch) can be approximated by a Gaussian distribution of $\exp(-e^2/K)$ where $e = d[\mathbf{x_l}, \mathbf{y_h}]$ as defined in (3) indicates how well the observation constraint is satisfied and K is a normalizing constant as used in bilateral filter [33]. Currently, we adopt a uniform prior $p(k) = 1/N$ for the simplicity but more sophisticated form such as Gaussian can also be used.

Combining single-patch resolution enhancement and Bayesian fusion, we obtain the following algorithm of resolution enhancement for generic images.

We note that the above Bayesian fusion degenerates into simple averaging across overlapping patches [22] when the likelihood function is approximately independent of locations (i.e., all coefficients in (7) have the same weights). The characteristics of likelihood function depend on the size of patches as well as their overlapping ratio. As we will see from the experimental results next, even simple averaging can significantly improve the objective performance due to the exploitation of the diversity provided by overlapping patches. The only penalty is the increased computational complexity which is approximately proportional to the redundancy ratio.

(i) Initialization: obtain HR training image $\hat{x}^0$ by example-based SR [22].
(ii) Iteration: for every patch $\mathbf{x_l}$ in the LR image, use the corresponding patch in $\hat{x}^0$ as the training patch and call Algorithm 1 to reconstruct the HR patch $\mathbf{y_h}$ and record the residue $d[\mathbf{x_l}, \mathbf{y_h}]$.
(iii) Fusion: calculate the final HR image by (7) and (8).

Algorithm 2: Patch-based resolution enhancement for generic images.

Table 1: Comparison of PSNR(dB) performance among lazy scheme, example-based SR, and Algorithm 1 for six texture images.

	Lazy scheme	Example-based SR	This work
$D6$	22.85	22.37	26.51
$D20$	23.22	22.05	25.27
$D21$	16.22	17.05	18.44
$D34$	23.84	25.47	28.04
$D49$	17.71	19.99	20.63
$D53$	24.43	25.08	26.94

5. EXPERIMENTAL RESULTS

In this section, we use experimental results to show that (1) for texture images, Algorithm 1 significantly outperforms lazy scheme and example-based SR [22] on both subjective and objective qualities; (2) for generic images, Algorithm 2 achieves arguably better subjective performance than lazy scheme and better objective performance than example-based SR [22]. The wavelet filter used in this work is Daubechies' 9-7 filter and resolution enhancement ratio is fixed to be two (i.e., one-level WT). Our implementation is based on several well-known toolboxes including WaveLab 8.5 for wavelet transforms, OpenTSTool for example-based SR [34], and MATLAB package for texture analysis/synthesis [21]. Test images and research codes accompanying this work will be made available at http://www.csee.wvu.edu/~xinl/demo/wt-interp.html.

5.1. Resolution enhancement of texture images

We have chosen six Brodatz texture images which approximately satisfy the homogeneity condition (see Figure 8) to test the performance of Algorithm 1. The training patch and testing patch are sized 128×128 and 64×64, respectively. The training patch driving the parametric texture model does not overlap with the testing patch for the reason of fairness (refer to Figure 5(a)). The benchmark includes lazy scheme and example-based SR [22] and MSE is calculated for nonborder pixels only (to eliminate potential bias introduced by varying boundary handling strategies in different schemes).

Table 1 includes the PSNR performance comparison among lazy scheme, example-based SR, and Algorithm 1. It

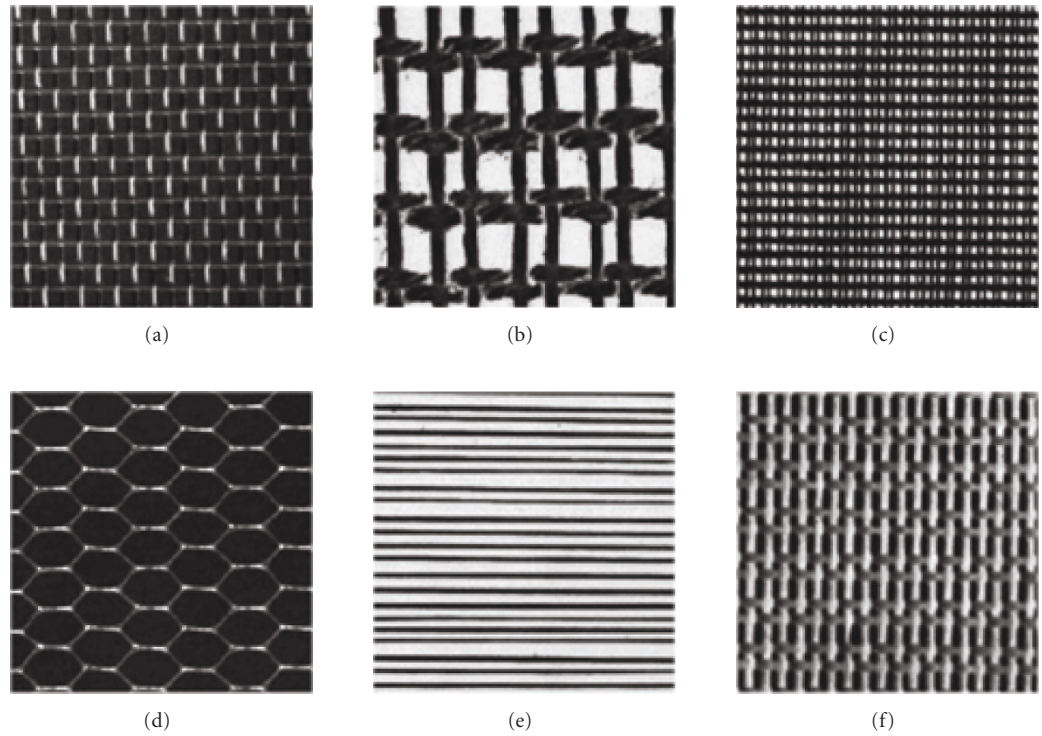


Figure 8: The collection of Brodatz texture images used in our experiments (left to right and top to bottom: *D*6, *D*20, *D*21, *D*34, *D*49, and *D*53).

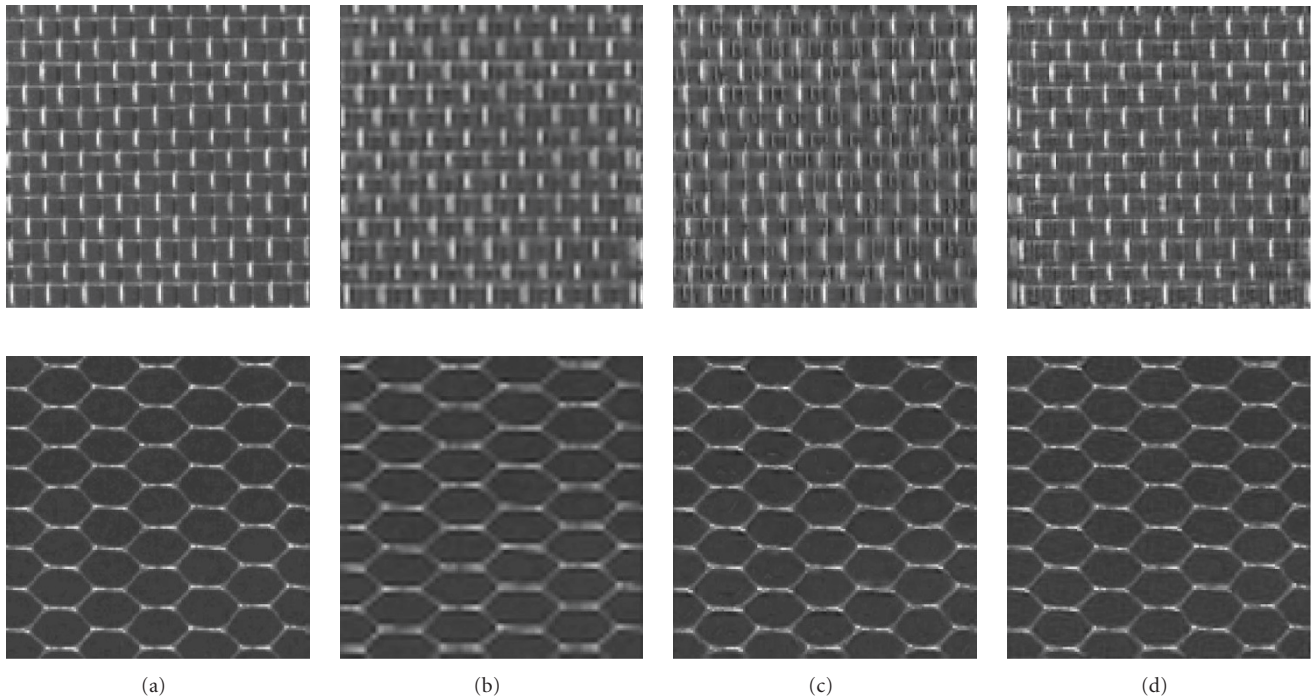
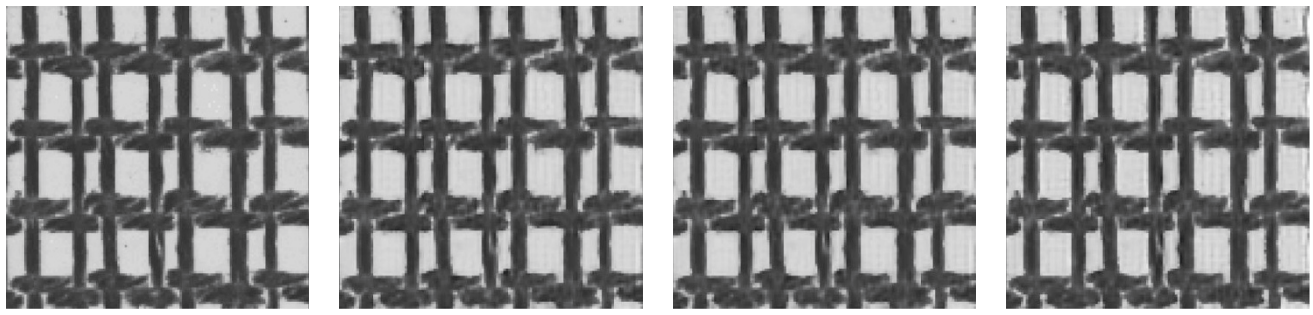
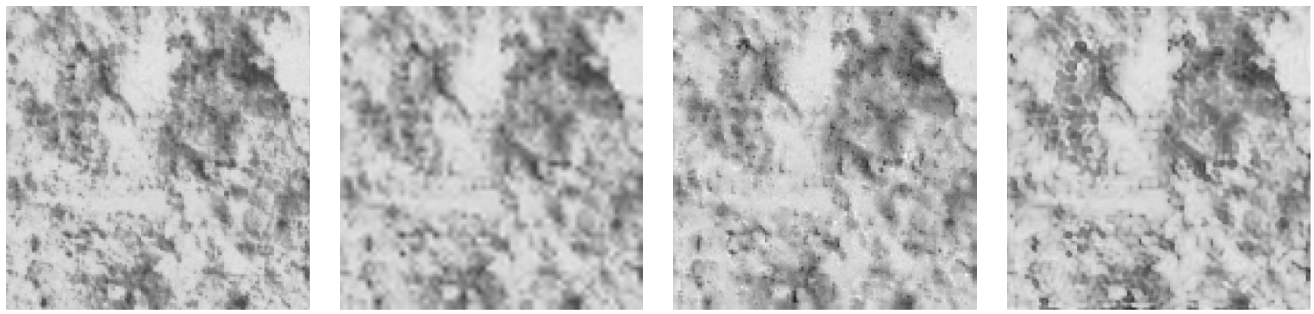


Figure 9: Performance comparison for *D*6 (top) and *D*34 (bottom): (a) original HR images; (b) reconstructed HR image by lazy scheme; (c) reconstructed HR image by example-based SR; (d) reconstructed HR image by Algorithm 1.



(a) (b) (c) (d)

Figure 10: Impact of training patch on the performance of Algorithm 1: (a) original $D20$ image; (b) reconstructed image by Algorithm 1 (PSNR = 25.27 dB); (b) reconstructed image by Algorithm 1 with a different starting point (PSNR = 25.32 dB); (d) reconstructed image by Algorithm 1 with a different training patch (PSNR = 23.79 dB).



(a) (b) (c) (d)

Figure 11: Performance comparison for $D2$. From left to right: original HR image, reconstructed images by lazy scheme (PSNR = 25.00 dB), example-based SR (PSNR = 22.12 dB), and Algorithm 1 (PSNR = 23.06 dB).

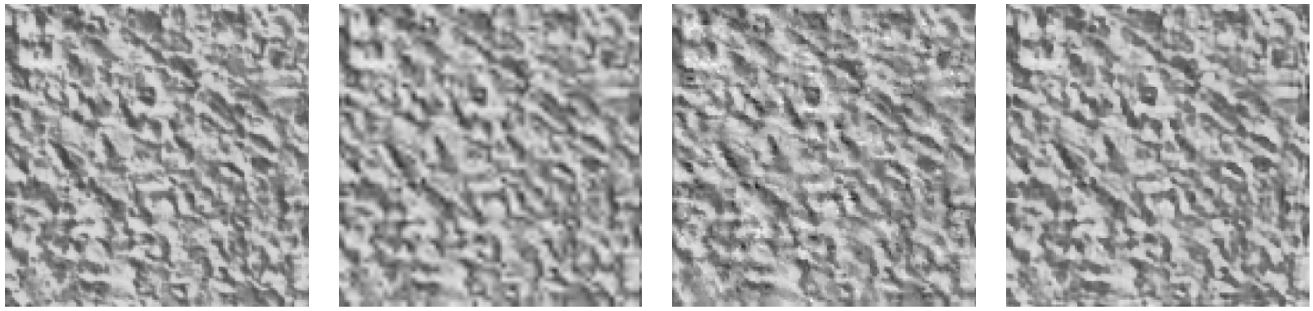
can be observed that Algorithm 1 uniformly outperforms lazy scheme and example-based SR by a large margin (0.7–4.1 dB) for the six test images. The most significant SNR improvement is observed for $D6$ and $D34$ which contain sharp contrast and highly regular texture patterns. Figure 9 compares the original HR image with the reconstructed HR images by three different schemes. It can be observed that Algorithm 1 driven by parametric texture model achieves the best visual quality among the three, lazy scheme suffers from blurred edges, and example-based SR introduces noticeable artifacts.

To illustrate the impact of starting point ($\hat{x}^0$) on reconstructed HR image, we test Algorithm 1 with two different initial settings: lazy scheme versus example-based SR. Figure 10 includes the comparison between reconstructed HR images by these two different starting points. It can be observed that the PSNR gap is negligible (0.05 dB), which suggests the insensitivity of Algorithm 1 to $\hat{x}^0$. To show how the choice of training patch affects the performance of Algorithm 1, we run it with two different training patches on $D20$. It can be seen from Figure 10 that although two training patches produce visually similar results, the gap on PSNR values of reconstructed HR images could be as large as 1.4 dB. Such finding is not surprising because it is widely known that MSE does not well correlate with the subjective quality of an image.

The discrepancy between subjective quality and objective fidelity becomes even more severe as texture patterns become more irregular (i.e., spatial homogeneity condition is less valid). To see this, we report the experimental results of Algorithm 1 for two other Brodatz texture images ($D2$ and $D4$) containing less periodic patterns (refer to Figures 11 and 12). Due to more complex texture patterns involved, we observe that the PSNR performance of Algorithm 1 falls behind lazy scheme (though still outperforms example-based SR). However, the subjective quality of HR images reconstructed by Algorithm 1 is convincingly better than that by lazy scheme especially in view of the improvements on edge sharpness. Therefore, we conclude that our Algorithm 1 achieves a better balance between subjective quality and objective fidelity than lazy scheme or example-based SR.

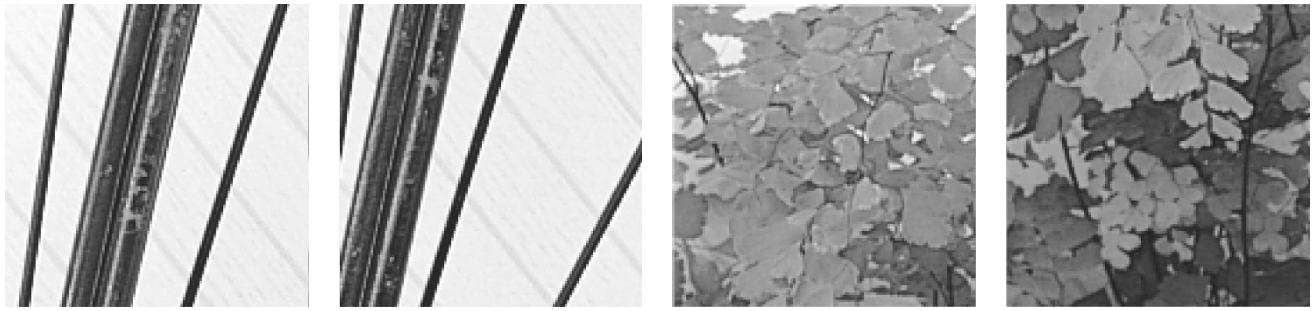
5.2. Resolution enhancement of generic images

The generic image for testing the proposed algorithms is chosen to be the JPEG2000 test image *bike* which contains a diversity of image structures. Due to its large size, we



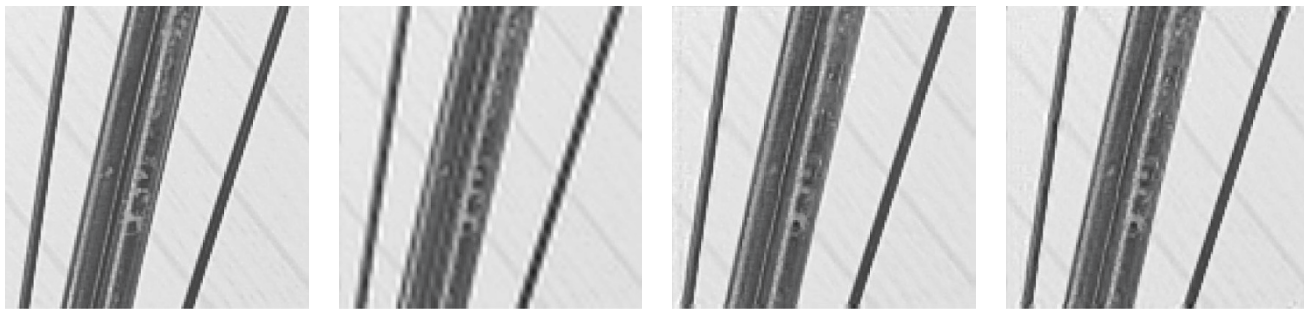
(a) (b) (c) (d)

Figure 12: Performance comparison for $D4$. From left to right: original HR image, reconstructed images by lazy scheme (PSNR = 22.23 dB), example-based SR (PSNR = 19.16 dB), and Algorithm 1 (PSNR = 21.39 dB).



(a) (b) (c) (d)

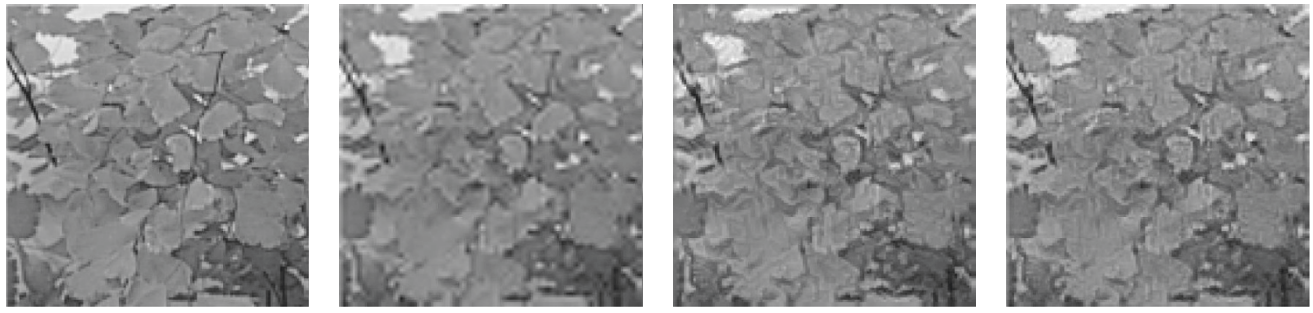
Figure 13: 128×128 portiones cropped out from the *bike* image. (a), (c) test data; (b), (d) training data.



(a) (b) (c) (d)

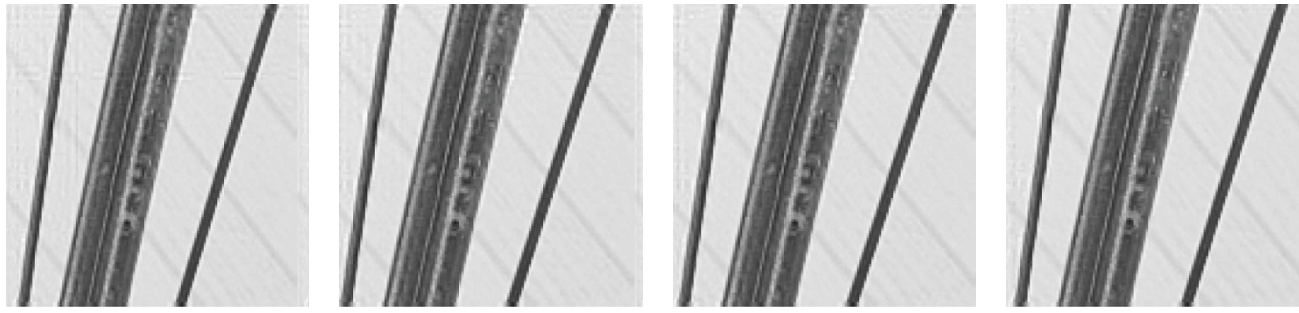
Figure 14: (a) Original *wheel* image; (b) reconstructed HR image by lazy scheme (PSNR = 21.86 dB); (c) reconstructed HR image by example-based SR (PSNR = 26.91 dB); (d) reconstructed HR image by Algorithm 1 (PSNR = 26.88 dB). Note that lazy scheme suffers from severe ringing artifacts around sharp edges.

crop out two 128×128 portions (called *wheel* and *leaves*) as the ground-truth HR images and their adjacent portions as the training data (refer to Figure 13). Figures 14 and 15 include the comparison between reconstructed HR images by lazy scheme, example-based SR, and our Algorithm 1 which can be viewed as a special case of Algorithm 2 with patch size being the same as the image size. It can be observed that Algorithm 1 achieves higher subjective quality than lazy scheme and comparable quality to example-based SR. The objective PSNR performance depends on the training data—for instance, significant positive gain (> 5 dB) is achieved for *wheel* (favorable training data) while the gain over lazy scheme becomes negative for *leaves* (unfavorable training data).



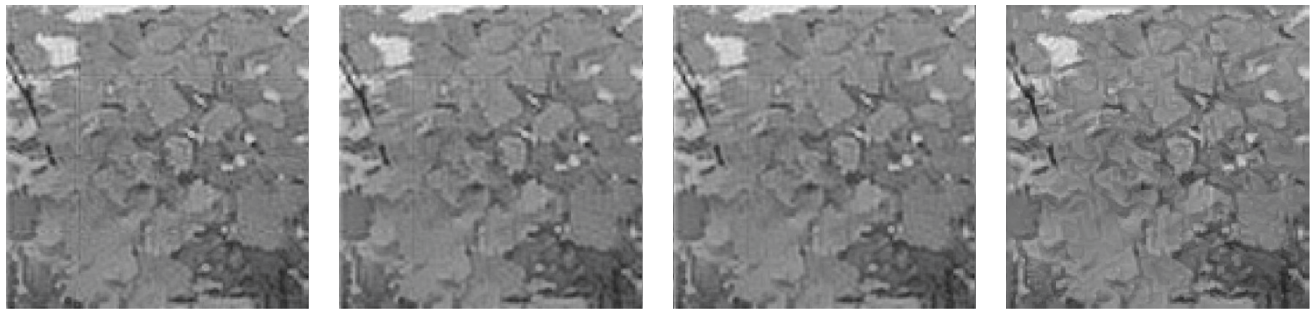
(a) (b) (c) (d)

Figure 15: (a) Original *leaves* image; (b) reconstructed HR image by lazy scheme (PSNR = 27.08 dB); (c) reconstructed HR image by example-based SR (PSNR = 24.31 dB); (d) reconstructed HR image by Algorithm 1 (PSNR = 25.13 dB). Note that despite lower PSNR value, our HR image appears sharper than the one by lazy scheme.



(a) (b) (c) (d)

Figure 16: Comparison of reconstructed *wheel* images: (a) Algorithm 2 with redundancy ratio of 1 (PSNR = 27.06 dB); (b) Algorithm 2 with redundancy ratio of 4 (PSNR = 27.55 dB); (c) Algorithm 2 with redundancy ratio of 16 (PSNR = 27.60 dB); (d) example-based SR [22] (PSNR = 27.23 dB).



(a) (b) (c) (d)

Figure 17: Comparison of reconstructed *leaves* images: (a) Algorithm 2 with redundancy ratio of 1 (PSNR = 25.73 dB); (b) Algorithm 2 with redundancy ratio of 4 (PSNR = 26.05 dB); (c) Algorithm 2 with redundancy ratio of 16 (PSNR = 26.09 dB); (d) example-based SR [22] (PSNR = 24.31 dB).

To test Algorithm 2, we have chosen a fixed patch size of 32×32 but different redundancy ratios. By increasing the overlapping ratio of adjacent patches from 0 to 1/2 and then 3/4, we observe that the redundancy ratio goes from 1 (nonoverlapping) to 4 and then 16. In our current implementation, we have adopted the averaging strategy in [22] instead of the Bayesian fusion formula in Section 4 (therefore, better performance is expected from nonuniform weighting). Figures 16 and 17 include the reconstructed HR images by Algorithm 2 with different redundancy ratios as well as the benchmark scheme [22]. It can be seen that PSNR improvement over no-fusion scheme is around 0.6–0.8 dB and noticeable suppression of artifacts around patch boundaries can be observed. Algorithm 2 with fusion strategy also

outperforms example-based SR [22] on PSNR performance due to the enforcement of observation and priori constraints by alternating projections.

Finally, we want to report the experimental results on computational complexity. In our current nonoptimized MATLAB implementation, the running time of Algorithm 1 with 10 iterations is typically 30 seconds for reconstructing an HR image sized 128×128 on a Pentium-IV laptop (2.4 GHz and 512 M memory). The running time of Algorithm 2 depends on the redundancy ratio of patch-based representation (i.e., how much overlap is allowed from one patch to the next) as well as patch size. For 128×128 images, it takes around 2 minutes to run our Algorithm 2 with redundancy ratio of one and patch size of 32×32 (iteration number is 5). When the redundancy ratio is increased to 4 and 16, the running time becomes 4 minutes and 20 minutes, respectively. In view of PSNR results in Figures 16-17, we conclude that a modest redundancy ratio of 4 is preferred to achieve a good balance between the performance and the computational cost.

6. CONCLUDING REMARKS

In this paper, we present a data-driven, projection-based resolution enhancement scheme which extends the previous work of parametric texture models in the wavelet space. When both target HR data and training data are characterized by homogeneous textures, parametric models are used to define prior constraint and we show how the parametric texture model can be used as prior constraint along with observation constraint to derive an alternating projection-based HR image reconstruction algorithm. When both target HR data and training data are generic images, we propose to borrow the idea of nonparametric sampling and synthesize new training data to drive the parametric texture models. Using patch-based representation, we show how to probabilistically fuse the reconstruction results at HR. Experimental results have shown that our new schemes achieve a good balance between subjective quality and objective fidelity. The importance of using both subjective quality and objective fidelity in evaluating the performance of resolution enhancement is argued, which is expected to clarify some misunderstandings about wavelet-based approaches toward resolution enhancement in the literature.

ACKNOWLEDGMENT

The author wants to thank Dr. T. Q. Pham at Delft University of Technology for sharing his implementation of example-based SR [22].

REFERENCES

[1] H. C. Andrews and C. L. Patterson III, "Digital interpolation of discrete images," *IEEE Transactions on Computers*, vol. 25, no. 2, pp. 196–202, 1976.

[2] S. Carrato, G. Ramponi, and S. Marsi, "A simple edge-sensitive image interpolation filter," in *Proceedings of IEEE International Conference on Image Processing (ICIP '96)*, vol. 3, pp. 711–714, Lausanne, Switzerland, September 1996.

[3] 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.

[4] K. Ratakonda and N. Ahuja, "POCS based adaptive image magnification," in *Proceedings of IEEE International Conference on Image Processing (ICIP '98)*, vol. 3, pp. 203–207, Chicago, Ill, USA, October 1998.

[5] J. Allebach and P. W. Wong, "Edge-directed interpolation," in *Proceedings of IEEE International Conference on Image Processing (ICIP '96)*, vol. 3, pp. 707–710, Lausanne, Switzerland, September 1996.

[6] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1521–1527, 2001.

[7] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proceedings of the IEEE*, vol. 78, no. 5, pp. 856–883, 1990.

[8] G. Strang and T. Q. Nguyen, *Wavelets and Filterbanks*, Wellesley-Cambridge, Wellesley, Mass, USA, 1997.

[9] S. G. Chang, Z. Cvetkovic, and M. Vetterli, "Resolution enhancement of images using wavelet transform extrema extrapolation," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '95)*, vol. 4, pp. 2379–2382, Detroit, Mich, USA, May 1995.

[10] W. K. Carey, D. B. Chuang, and S. S. Hemami, "Regularity-preserving image interpolation," *IEEE Transactions on Image Processing*, vol. 8, no. 9, pp. 1293–1297, 1999.

[11] D. D. Muresan and T. W. Parks, "Prediction of image detail," in *Proceedings of IEEE International Conference on Image Processing (ICIP '00)*, vol. 2, pp. 323–326, Vancouver, BC, Canada, September 2000.

[12] Y. Zhu, S. C. Schwartz, and M. T. Orchard, "Wavelet domain image interpolation via statistical estimation," in *Proceedings of IEEE International Conference on Image Processing (ICIP '01)*, vol. 3, pp. 840–843, Thessaloniki, Greece, October 2001.

[13] K. Kinebuchi, D. D. Muresan, and T. W. Parks, "Image interpolation using wavelet-based hidden Markov trees," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '01)*, vol. 3, pp. 1957–1960, Salt Lake, Utah, USA, May 2001.

[14] D. H. Woo, I. K. Eom, and Y. S. Kim, "Image interpolation based on inter-scale dependency in wavelet domain," in *Proceedings of International Conference on Image Processing (ICIP '04)*, vol. 3, pp. 1687–1690, Singapore, October 2004.

[15] Y.-L. Huang, "Wavelet-based image interpolation using multilayer perceptrons," *Neural Computing and Applications*, vol. 14, no. 1, pp. 1–10, 2005.

[16] C.-L. Chang, X. Zhu, P. Ramanathan, and B. Girod, "Light field compression using disparity-compensated lifting and shape adaptation," *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 793–806, 2006.

[17] Y. Itoh, Y. Izumi, and Y. Tanaka, "Image enhancement based on estimation of high resolution component using wavelet transform," in *Proceedings of IEEE International Conference on Image Processing (ICIP '99)*, vol. 3, pp. 489–493, Kobe, Japan, October 1999.

[18] J. Liu and P. Moulin, "Information-theoretic analysis of inter-scale and intrascale dependencies between image wavelet coefficients," *IEEE Transactions on Image Processing*, vol. 10, no. 11, pp. 1647–1658, 2001.

[19] N. R. Shah and A. Zakhor, "Resolution enhancement of color video sequences," *IEEE Transactions on Image Processing*, vol. 8, no. 6, pp. 879–885, 1999.

[20] V. Caselles, J.-M. Morel, and C. Sbert, "An axiomatic approach to image interpolation," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 376–386, 1998.

[21] J. Portilla and E. P. Simoncelli, "Parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–71, 2000.

[22] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.

[23] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 2nd edition, 1999.

[24] B. Julesz, "Visual pattern discrimination," *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 84–92, 1962.

[25] H. Wechsler, "Texture analysis—a survey," *Signal Processing*, vol. 2, no. 3, pp. 271–282, 1980.

[26] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[27] D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis/synthesis," in *Proceedings of the 22nd Annual ACM Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '95)*, pp. 229–238, Los Angeles, Calif, USA, August 1995.

[28] S. C. Zhu, Y. Wu, and D. Mumford, "Filters, random fields and maximum entropy (frame): towards a unified theory for texture modeling," *International Journal of Computer Vision*, vol. 27, no. 2, pp. 107–126, 1998.

[29] J. S. de Bonet, "Multiresolution sampling procedure for analysis and synthesis of texture images," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*, pp. 361–368, Los Angeles, Calif, USA, August 1997.

[30] S. Chang, "Image interpolation using wavelet-based edge enhancement and texture analysis," M.Sc. thesis, University of California, Berkeley, Calif, USA, 1995.

[31] P. L. Combettes, "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, 1993.

[32] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, vol. 2, pp. 1033–1038, Kerkyra, Greece, September 1999.

[33] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the 6th IEEE International Conference on Computer Vision (ICCV '98)*, pp. 839–846, Bombay, India, January 1998.

[34] T. Q. Pham, L. J. van Vliet, and K. Schutte, "Resolution enhancement of low quality videos using a high-resolution frame," in *Visual Communications and Image Processing*, vol. 6077 of *Proceedings of SPIE*, San Jose, Calif, USA, January 2006.