Generative Bayesian Image Super Resolution With Natural Image Prior

Haichao Zhang, *Student Member, IEEE*, Yanning Zhang, *Senior Member, IEEE*, Haisen Li, and Thomas S. Huang, *Life Fellow, IEEE*

***Abstract*— We propose a new single image super resolution (SR) algorithm via Bayesian modeling with a natural image prior modeled by a high-order Markov random field (MRF). SR is one of the long-standing and active topics in image processing community. It is of great use in many practical applications, such as astronomical observation, medical imaging, and the adaptation of low-resolution contents onto high-resolution displays. One category of the conventional approaches for image SR is formulating the problem with Bayesian modeling techniques and then obtaining its maximum-a-posteriori solution, which actually boils down to a regularized regression task. Although straightforward, this approach cannot exploit the full potential offered by the probabilistic modeling, as only the posterior mode is sought. On the other hand, current Bayesian SR approaches using the posterior mean estimation typically use very simple prior models for natural images to ensure the computational tractability. In this paper, we present a Bayesian image SR approach with a flexible high-order MRF model as the prior for natural images. The minimum mean square error (MMSE) criteria are used for estimating the HR image. A Markov chain Monte Carlo-based sampling algorithm is presented for obtaining the MMSE solution. The proposed method cannot only enjoy the benefits offered by the flexible prior, but also has the advantage of making use of the probabilistic modeling to perform a posterior mean estimation, thus is less sensitive to the local minima problem as the MAP solution. Experimental results indicate that the proposed method can generate competitive or better results than *state-of-the-art* SR algorithms.**

***Index Terms*— Bayesian minimum mean square error estimation, field-of-experts, Markov chain Monte Carlo (MCMC), Markov random field, natural image statistics, super resolution (SR).**

I. Introduction

SUPER RESOLUTION (SR) is a very active research topic in the image processing community. In practice, many tasks, such as astronomical observation and medical diagnostic rely on high quality images for reliable and accurate analysis as well as prediction. However, in many practical situations, due to the inherent limitations of the optical system or other factors, the observed images are often of low resolution, thus limiting the subsequent tasks based on them. Image SR aims to estimate a high-resolution (HR) image from a single or a set of low-resolution (LR) observations [1], which is termed as single image SR and multiframe SR, respectively. SR has a long history and is still a very active field [1], with many new techniques emerging [2]–[5].

Conventional multiframe SR methods are reconstruction-based methods. They typically perform motion estimation followed by frame fusion [1], which relies on the complementary information contained in different frames for resolution enhancement. The success of this kind of approach relies heavily on the accuracy of the motion estimation procedure as well as the later frame fusion scheme. Furthermore, such reconstruction-based methods are inherently limited to the case of small zooming factors [6]. For single image SR, the simplest methods are the interpolation-based approaches, e.g., bicubic interpolation. However, such analytical interpolation-based methods do not exploit the underlying structures in natural images, such as edges, thus usually blurring the fine details and introducing artifacts in the interpolated results. More advanced interpolation-based methods take the underlying structures of the image into consideration during interpolation, e.g., [7] and [8], thus improving the interpolation quality by adjusting the interpolation scheme according to the latent structures. Recently, methods exploiting the natural image priors for SR have been proposed in the literature [2], [5], [9], [10]. Freeman *et al.* exploited the patch similarity prior with the help of a large training image set for SR [2]. Tappen *et al.* developed a Markov random field-based SR algorithm utilizing the sparse derivative prior of natural images [9]. Yang *et al.* proposed an SR approach based on the sparse representation prior of image patches with respect to a properly chosen dictionary [5]. Kim *et al.* developed a regression-based SR algorithm based on kernel regression and with a sparsity prior of natural images [10]. Apart from the sparsity prior, SR methods exploiting other properties of natural images, such as self-similarities and local/non-local regularities have also been proposed in the literature [3], [4]. Another property of natural images is the edge statistics, which have been exploited in [11]–[13] as a prior for SR. While exploiting the natural image statistical properties as a prior, these methods typically formulate the problem within the Bayesian framework and

Manuscript received December 23, 2011; revised March 21, 2012; accepted May 1, 2012. Date of publication May 15, 2012; date of current version August 22, 2012. This work was supported in part by the National Natural Science Foundation of China, under Grant 60872145 and Grant 60903126, the National High Technology Research and Development Program of China, under Grant 2009AA01Z315, and the U.S. Army Research Laboratory and Army Research Office, under Grant W911NF-09-1-0383. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Pascal Frossard.

H. Zhang, Y. Zhang, and H. Li are with the School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China (e-mail: hczhang@mail.nwpu.edu.cn; ynzhang@nwpu.edu.cn; haisenli@mail.nwpu.edu.cn).

T. S. Huang is with the Department of Electrical and Computer Engineering, Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: huang@ifp.uiuc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2012.2199330

then estimate the final HR image with a maximum a posteriori (MAP) estimation approach.[1] Although it is demonstrated effective in some applications, this kind of approach discards the further potential enabled by the Bayesian probabilistic modeling. Another limitation is that the regularization term is typically designed empirically *a priori*, thus is not guaranteed to be compatible with the image statistics. There are indeed some Bayesian SR approaches in the literature that do not use MAP solution [14]–[17]; however, they typically use very simple image priors to ensure the tractability of computation, thus limiting their performance. An SR approach using posterior-based sampling is proposed in [14]. This method is limited by the fact that it uses a very simple Ising model as its prior [14]. Multiframe SR methods based on variational Bayes (VB) technique have been developed in the past [15]–[17]. In [15], the authors used the total variation (TV) model as the prior, which prefers piecewise linear images thus is not quite adequate for general natural images. In [16], the authors used a casual Gaussian MRF model as the prior, which includes latent variables indicating the dependency of two adjacent pixels. Although the model has taken the edge structure of natural images into consideration, it is not flexible enough to capture the statistics of natural images well. To sum up, the conventional SR approaches using regularized regression are limited by the fact that they do not exploit the full potential offered by the probabilistic modeling, although various kinds of regularization terms have been designed. On the other hand, for the previously developed Bayesian SR approach, the prior is typically limited to very simple models to ensure the tractability in computation, thus limiting their estimation quality. Furthermore, VB technique is typically used for the posterior mean estimation in those approaches, which actually compromises the modeling accuracy.

In this paper, we propose a Bayesian SR algorithm with natural image statistics using generative scheme for HR image estimation via sampling. Our approach is different in several ways: 1) the proposed method is a fully Bayesian approach, which can incorporate the *a priori* knowledge on the latent HR image as well as other uncertainties, such as the noise level into the framework in a natural way; 2) a more advanced prior is used for capturing the HR image statistics rather than other very simple priors; specifically, a high-order MRF model is used for modeling the natural image statistics, which, at the same time, enables generative sampling from it; and 3) the Bayesian MMSE criteria are applied for estimating the HR image rather than the MAP criteria. The MMSE approach has several advantages over MAP. It does not require ad-hoc modifications as MAP to achieve desirable restoration performance. Also, it is less sensitive to the local minima in the solution space than MAP, especially when heavy-tailed priors are applied.

The rest of this paper is organized as follows. We first give a review of some related previous works briefly in Section II. Then, we discuss different prior models for natural images in Section III. We introduce the generative Bayesian SR model and present an effective SR algorithm in Section IV. Experiments are conducted in Section V and the results are compared with several *state-of-the-art* algorithms. We conclude this paper in Section VI.

[1]Some of the approaches may be formulated as a regularized regression problem, the solution to which is related to the MAP solution to the corresponding Bayesian formulation. See Section II for more details.

II. Previous Works

We review some of the related previous works in this section, which will lay the foundation for the deviation of our approach later. Mathematically, the LR image formation process is usually modeled as the convolution of the latent HR image with a low-pass filter, which models the low passing property of the imaging system followed with downsampling

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{H}\boldsymbol{x} + \boldsymbol{\epsilon} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{y} \in \mathbb{R}^n$ is the vector representation of the HR and LR image, respectively, obtained by lexicographical ordering the images. $\boldsymbol{H} \in \mathbb{R}^{m\times m}$ is the matrix corresponding to the blurring process and $\boldsymbol{D} \in \mathbb{R}^{n\times m}$ $(n < m)$ is the matrix representing the downsampling operator. $\boldsymbol{\epsilon} \in \mathbb{R}^n$ is a noise term. The task of SR is to estimate the underlying HR image $\boldsymbol{x}$ given only the LR observation $\boldsymbol{y}$. In the sequel, we review briefly some related works in the literature.

As the task of SR is to estimate a HR image from the LR observation, the number of unknowns m is larger than that of the knowns n, thus the problem is ill-posed. Therefore, regularization techniques based on the *a priori* knowledge are required to alleviate the ill-posedness of this problem. Typically, the task of SR is formulated as a regularized least-square regression problem. Mathematically

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{H}\boldsymbol{x}\|_2^2 + \lambda R(\boldsymbol{x}) \tag{2}$$

where the first term is the fidelity term reflecting the fact that the unknown HR image $\boldsymbol{x}$ should generate observation that is similar to $\boldsymbol{y}$ after passing it through the observation system. The second term, $R(\cdot)$, is a regularization term on the desired HR image to alleviate the ill-posedness of the SR problem. λ is a parameter balancing the contributions of these two terms. This is a very popular model and has been adopted in many recent SR and other related works [12], [13], [18], [19]. Following this direction, most of the current works focus on designing different formulations for the regularization term $R(\cdot)$. The most classical choice for $R(\cdot)$ would be the squared ℓ_2-norm, which is based on the assumption that the energy of the HR image should not be too large. Because of its well known blurring effect for the ℓ_2-norm-based regularization, some algorithms apply the ℓ_2-norm in the gradient domain. Recently, many approaches exploiting the sparsity property of natural images are developed [5], [9], [10], [12], [13], [19]. A general approach for designing $R(\cdot)$ utilizing the sparsity property is to use the sparsity property in the transform domain (e.g., wavelet domain, edge domain) as regularization [12], [13]. A novel regularization term called "Softcuts" has been developed in [12], by exploiting the sparse property of the edge profiles. In [13], Sun *et al*. use the sparsity prior of the edges in images by first estimating the gradient of the HR image via transforming that of the LR image with the learned relationship

between the LR and HR gradients. Then the regularization term $R(\cdot)$ is designed as the distance between the gradients of the unknown latent HR image and the estimated HR gradients. A sparse representation (coding) based SR (ScSR) method has been proposed in [5] based on the property that natural image patches have a sparse representation with respect to a properly chosen dictionary. Specifically, the ScSR method first recovers the sparse representation coefficients from the LR patches with respect to a LR dictionary, and then reconstructs a HR patch with the recovered representation coefficients and a HR dictionary, which is trained jointly with the LR dictionary. To further enhance the texture details in the estimated HR images, several approaches have been proposed by using texture prior offered by domain specific examples [19]–[21].

The regularized SR (2) is related to a probabilistic model as follows:

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})} \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}) \propto \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{DHx}\|_2^2}{\sigma^2}\right)\exp\left(-\frac{R(\boldsymbol{x})}{\eta^2}\right) \quad (3)$$

where σ denotes the standard deviation of the noise and η is a scale parameter for the prior distribution of $\boldsymbol{x}$. At this point, it is easy to see that the solution to (2) corresponds to the MAP solution to (3) due to the following relations:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{y}) \Leftrightarrow \arg\min_{\boldsymbol{x}} -\log p(\boldsymbol{x}|\boldsymbol{y}) \quad (4)$$

which corresponds to the mode of the posterior distribution for the HR image $\boldsymbol{x}$. The regularization parameter λ in (2) is related to σ and η in (3) by $\lambda = (\sigma^2/\eta^2)$.

Because of this relation, many previous works first model the SR problem in a probabilistic framework as (3) and then estimate the HR image by minimizing the negative logarithm of it, which actually boils down to the regularized regression model as shown in (2) [13], [19]. To solve (2), gradient-based optimization schemes are typically used [13], [19]. While some encouraging results can be obtained, the MAP approach could not exploit the full potential offered by the probabilistic modeling, as only the posterior mode is sought for solution, discarding much information offered by the probabilistic model. Even worse, it is prone to getting stuck in a local minima when the prior model is a heavy-tailed distribution, which actually leads to a non-convex optimization problem, thus there is no guarantee of obtaining even the posterior mode.

III. Prior Models Based on Natural Image Statistics

The statistics of natural images have received great attention of researchers from different communities for both understanding the human visual system and designing more effective processing algorithms [22]. Natural image statistics modeling has a fundamental relation with the field of image processing. Natural images are different from random noise image in that they exhibit structures. Examples are local regularities, such as edges [8] and self-similarities [3], [4]. As a result, the natural images only occupy a tiny fraction of the whole space of all the images generated by all the possibilities enables by the pixel value combinations. Natural images have strong statistical properties known as natural image statistics [23]–[25]. One of the most well known properties is that the natural images exhibit heavy-tailed distribution when applying derivative filters onto them. This can be explained intuitively as follows: on one hand, natural images are locally smooth, therefore, the local difference will be small, thus the marginal distribution will decrease faster than the Gaussian distribution; on the other hand, natural images have many structures such as edges, where the derivative response can be large, which contributes to the heavier tails than the Gaussian distribution. This prior enjoys wide range of applications, including image denoising [24], [25], deblurring [26] and SR [10]. Apart from this heavy-tailed marginal statistics, natural images also exhibit several other properties, such as scale invariance, which states that the natural images exhibit similar heavy-tailed distribution at different scales [23]. Another property is the joint statistics, meaning that the neighboring pixels in the natural images exhibit high dependency [25].

There are many works on image priors and it is still a very active research topic. Gaussian model applied onto the derivatives of images is the most classical and widely used one due to its simplicity

$$p(\boldsymbol{x}) \propto \exp\left(-\frac{\|\nabla\boldsymbol{x}\|_2^2}{\eta^2}\right) \quad (5)$$

where $\nabla\boldsymbol{x}$ denotes the gradients of image $\boldsymbol{x}$. Gaussian prior is related to the Tikhonov regularization in the inverse problem theory. Although it has the advantage of closed-form solution, it cannot generate satisfying solutions in most of the cases as it typically smoothens the image too much. To preserve the edge structure, Laplacian prior is later used as image prior and is defined as

$$p(\boldsymbol{x}) \propto \exp\left(-\frac{\|\nabla\boldsymbol{x}\|_1}{\eta^2}\right). \quad (6)$$

Laplacian prior is related to the TV model derived in the partial differential equation field [27], which has been proved to keep the image discontinuities better. It is also related to the ℓ_1-norm-based regularization, which is well known for its ability in promoting sparsity for the solution. Although the TV prior can preserve the edges, it cannot capture natural image statistics well enough, as the resulting images are typically piecewise linear, which is desirable for cartoon-like images, but not for natural images. Natural images follow a distribution with heavier tails than Gaussian and Laplacian distribution, which has been realized and used by many researchers recently [26], [28]. To overcome this limitation, some researchers proposed to use the hyper-Laplacian sparse prior for natural images, which has been successfully applied in many fields [26], [28]

$$p(\boldsymbol{x}) \propto \exp\left(-\frac{\|\nabla\boldsymbol{x}\|_\alpha}{\eta^2}\right) \quad (7)$$

where $\|\boldsymbol{x}\|_\alpha$ is defined as $\|\boldsymbol{x}\|_\alpha = \sum_i |\boldsymbol{x}(i)|^\alpha$. α here is the parameter controlling the sparseness of the desired natural images. For natural images, the response with respect to the

learned sparse filters is sparse, i.e., with a large number of zero elements and a few responds with large absolute values; therefore, α is typically chosen as $0.5 \leq \alpha \leq 0.8$ in the literature [28].

An elegant formulation for unifying the above mentioned probabilistic models of natural images is MRF [25]. MRF is a type of undirected graph, where the node and edge of this graph is defined differently according to different MRF configurations. In MRF, the probability of a whole image $\boldsymbol{x}$ is defined as follows based on the computation of local cliques

$$p(\boldsymbol{x}) = \frac{1}{Z} \prod_{c \in \zeta} f(\boldsymbol{x}_{[c]}) \tag{8}$$

where ζ denotes the set of all pixel locations of the image and c refers to a particular location. $\boldsymbol{x}_{[c]}$ is the vector made up by the local neighborhood at c for the image $\boldsymbol{x}$, which is also referred to as a clique. $f(\cdot)$ is a potential function, which takes a clique as input and outputs the potential or energy for that clique. Z is the partition function to ensure $p(\boldsymbol{x})$ to be integrated to one. One typical MRF model is the so called pairwise MRF. In this model, each pixel in the image is defined as a node. Each pixel is connected with its directed neighbors.

The choice of the potential function $f(\cdot)$ is crucial for the modeling quality. The conventional choice is the Gaussian function, as in (5). With the recent development in natural image statistics, it is demonstrated that the potential function defined as a heavy-tailed function over the response of derivative filters on the clique $\boldsymbol{x}_c$ can better capture the image statistics [25]. Motivated by the property of heavy-tailed marginal distribution of natural images, to improve the modeling quality, researchers suggest to use heavy-tailed potential functions, such as Laplacian and hyper-Laplacian as in (6) and (7) and Student-t distribution as used in [25].

In the terminology of MRF, (5)–(7) are pairwise MRF, as only two neighboring pixels are involved in each clique, due to the usage of the first-order derivative filters. However, the modeling power of pairwise MRF is limited, as the pixels in natural images also exhibit long range relations; therefore, high-order MRF model is developed in some recent work [25]. One challenge in the case of high-order MRF is that the proper potential function is hard to define in the high-dimensional space, as the size of a clique has been increased. The recently introduced field-of-experts (FoE) model formulates the potential on clique c as a product of several functions defined on different low-dimensional spaces, respectively [25]

$$f(\boldsymbol{x}_{[c]}) = \prod_{l=1}^{L} \phi\Big((\boldsymbol{k}_l * \boldsymbol{x})_c; \Theta_l\Big) = \prod_{l=1}^{L} \phi\left((\boldsymbol{K}_l \boldsymbol{x})_c; \Theta_l\right) \tag{9}$$

which is known as the product-of-expert (PoE) model in the literature [29]. $(\boldsymbol{k}_l * \boldsymbol{x})_c$ refers to the c-th pixel in the filtered image by $\boldsymbol{k}_l$. $\mathcal{K} = \{\boldsymbol{k}_l\}_{l=1}^{L}$ is a set of derivative filters, either designed under some criteria or learned from the data [24], [25]. $\boldsymbol{K}_l \in \mathbb{R}^{m \times m}$ is the matrix representation corresponding to the derivative filter $\boldsymbol{k}_l$, with the following relation: $\boldsymbol{K}_l \boldsymbol{x} = \boldsymbol{k}_l * \boldsymbol{x}$. The clique $\boldsymbol{x}_{[c]}$ in this model is actually the local patch covered by the filter $\boldsymbol{k}_l$ ($l \in \{1, \ldots, L\}$) at c. The model in (9) consists of L expert functions to model the high-dimensional distribution by using the products of them. $\{\Theta_l\}_{l=1}^{L}$ is the parameter set associated with the PoE.

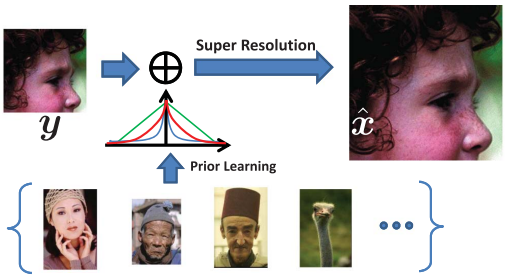

Fig. 1. Generative Bayesian SR framework. In the proposed SR framework, we first learn a distribution from the training images with a flexible high-order MRF model, with Gaussian scale mixture (GSM) model as its potentials. The learned distribution is then used as the prior for the underlying HR image to be estimated. The HR images are estimated under the MMSE criteria via sampling the posterior. In this framework, we are actually "transferring" the prior learned from training images to the unknown HR image.

In FoE, the filters are learned from the training data [25]. The advantage of learned filters over pre-specified ones (such as $\{\boldsymbol{k}_1 = [-1, 1], \boldsymbol{k}_2 = \boldsymbol{k}_1^T\}$) is that the learned filters have the potential to capture the statistics of natural images better than the simple horizontal and vertical derivative filters. Note that the prior can also be made adaptive to the content of the specific image by allowing its parameters to change according to the content [30]. One challenge in learning the MRF model is that the normalization term Z is typically intractable to evaluate, which poses a great difficulty for the learning process. To handle this challenge, several schemes, such as contrastive divergence [29] and score matching [31] are proposed in the literature.

IV. Natural Image Statistics-Based Generative Bayesian Image SR

As mentioned above, the conventional SR approaches using regularized regression cannot exploit the full potential of the probabilistic model and are also often faced with non-convex optimization problem when heavy-tailed sparse priors are used, thus is prone to be trapped in a local minima; on the other hand, previously proposed Bayesian SR approach[2] only used very simple prior model for natural images due to the limitation in computation [15]–[17]. In this section, we present a Bayesian SR approach and estimate the HR image under the MMSE criteria from the posterior, with more advanced prior model learned from natural images. The framework of the proposed approach is illustrated in Fig. 1. In the sequel, we first formulate the SR problem in the Bayesian framework and then present an effective algorithm for generating the SR image. In terms of Bayesian inference, the latent SR image is estimated via the posterior distribution $p(\boldsymbol{x}|\boldsymbol{y}; \Theta)$. Using the

[2]By Bayesian SR, we refer to those approaches that use the MMSE estimation criteria, rather than those using the MAP criteria, even though they are formulated in a Bayesian framework.

Bayes rule, we can get

$$p(\boldsymbol{x}|\boldsymbol{y};\Theta)=\frac{p(\boldsymbol{y}|\boldsymbol{x};\Theta)p(\boldsymbol{x})}{p(\boldsymbol{y})} \tag{10}$$

therefore, the task for the Bayes modeling includes the configuration of the likelihood $p(\boldsymbol{y}|\boldsymbol{x};\Theta)$ and the prior $p(\boldsymbol{x})$, which is presented in detail in the sequel.

A. Bayesian Image SR Formulation

In this section, we first present the likelihood and prior model, and then derive the posterior for the latent HR image, from which the posterior mean is computed as the HR estimation.

1) Likelihood: The noise term $\boldsymbol{\epsilon}$ is assumed to be known only up to the statistical property, i.e., its statistical distribution. The statistical distribution of the noise is application dependent. In many situations, the noise term can be modeled with Gaussian distribution. In this paper, we also use a Gaussian distribution for modeling the statistics of noise. Assuming the elements in the noise term $\boldsymbol{\epsilon}$ follow *i.i.d.* Gaussian distribution and the standard deviation is σ, then we have

$$\boldsymbol{\epsilon}\sim\mathcal{N}(\boldsymbol{0},\sigma^2\boldsymbol{I}). \tag{11}$$

With (11) and the observation model defined in (1), we can derive the following likelihood model:

$$p(\boldsymbol{y}|\boldsymbol{x})=\mathcal{N}(\boldsymbol{y};\boldsymbol{DHx},\sigma^2\boldsymbol{I}) \tag{12}$$

where $\boldsymbol{I}\in\mathbb{R}^{n\times n}$ is the identity matrix. The noise level σ^2 is not assumed to be known. Rather, we treat it as a random variable and place a hyperprior on it as shown later.

2) Prior Model for Image: The definition of the prior model for the HR image is crucial for the final HR estimation quality. The design of such a prior should reflect the desired property of natural HR images. The Gaussian prior reflects the smoothness property of natural images. However, this prior typically causes blurring effects for the structures in natural images, as it fails in capturing the heavy-tailed distribution of the natural image statistics. Also, the conventional first-order MRF model for natural images is limited by the pairwise interaction, as it is not flexible enough for capturing the image statistics. The recently developed FoE model is a high-order MRF model that can capture the statistics of natural images better by learning the high-order filters from the data [25]. Therefore, in this paper, we use a high-order FoE model as the prior for the underlying HR image $\boldsymbol{x}$

$$p(\boldsymbol{x})=\frac{1}{Z}\prod_{c\in\zeta}f(\boldsymbol{x}_{[c]})=\frac{1}{Z}\prod_{c\in\zeta}\left(\prod_{l=1}^{L}\phi\big((\boldsymbol{k}_l*\boldsymbol{x})_c;\Theta_l\big)\right) \tag{13}$$

where GSMs model (14) is used as the expert function $\phi(\cdot)$ following [24], [32], rather than Student-t distribution as used in [25], and the Gaussian, Laplacian, and hyper-Laplacian functions used in (5)–(7), respectively. The advantage of using such a model is two-fold: first, it is very flexible and can be adapted easily for capturing the natural image statistics; second, this model facilitates the generative sampling, which is required for the HR image estimation. Specifically, the expert function is defined as

$$\phi\big((\boldsymbol{k}_l*\boldsymbol{x})_c;\Theta_l\big)=\sum_{j=1}^{J}\beta_{lj}\mathcal{N}\left((\boldsymbol{k}_l*\boldsymbol{x})_c;0,\frac{\eta_l^2}{s_j}\right) \tag{14}$$

where β_{lj} is the normalized weight for the j-th component in the Gaussian mixture for modeling the responds of the l-th filter. η_l^2 is the base variance for the J Gaussian components in the l-th expert function. s_j is the scale parameter for the j-th component. Θ_l denotes the set of parameters $\Theta_l=\{\boldsymbol{k}_l,\{\beta_{lj}\}_{j=1}^{J},\{s_j\}_{j=1}^{J}\}$, which can be learned from the training data [25], [32]. The model in (14) is actually a 1-D Gaussian mixture model with zero mean and is very flexible to be adapted for capturing the statistics of natural images [24], [32]. Moreover, as this prior model is a Gaussian Mixture, we can perform sampling from it very efficiently.

3) Hyperpriors on Hyperparameters: We do not assume the noise level to be known in our model. Rather, we treat the noise level as a random variable and perform estimation for it jointly with other random variables. Specifically, to model this random variable, we use a Gamma distribution as its hyperprior

$$p(\tau)=\mathcal{G}(\tau;a^o,b^o)=\frac{1}{\Gamma(a^o)}(b^o)^{a^o}\tau^{a^o-1}e^{-b^o\tau} \tag{15}$$

where $\tau=\sigma^{-2}$ which is also referred to as precision. a^o and b^o are non-negative parameters controlling the shape and scale of the Gamma distribution, respectively. The reason of using a Gamma distribution as the hyperprior for the precision is that it is the conjugate prior for the precision of a Gaussian distribution, i.e., the posterior will have the same functional form as the prior distribution [33].

4) Posterior: By substituting (12), (13), and (15) into (10), we can derive the following posterior for the latent HR image $\boldsymbol{x}$ as well as the precision τ:

$$\begin{aligned}p(\boldsymbol{x},\tau|\boldsymbol{y};\Theta)&\propto p(\boldsymbol{y}|\boldsymbol{x},\tau;\Theta)p(\boldsymbol{x})p(\tau)\\&\propto\mathcal{N}(\boldsymbol{y};\boldsymbol{DHx},\tau^{-1}\boldsymbol{I})\cdot\prod_{c\in\zeta}f(\boldsymbol{x}_c)\cdot\mathcal{G}(\tau;a^o,b^o).\end{aligned} \tag{16}$$

In this paper, we estimate the HR image with the Bayesian posterior mean, which is the solution to the following MMSE estimation problem:

$$\hat{\boldsymbol{x}}=\arg\min_{\bar{\boldsymbol{x}}}\int\|\bar{\boldsymbol{x}}-\boldsymbol{x}\|^2p(\boldsymbol{x}|\boldsymbol{y};\Theta)d\boldsymbol{x}=E(\boldsymbol{x}|\boldsymbol{y}). \tag{17}$$

It is typically intractable to solve (17) directly in the general case. In this paper, we use an approximate approach, by using samples drawn from the posterior distribution. Denoting the set of samples for the HR image as $\{\tilde{\boldsymbol{x}}_i\}_{i=1}^{M}$, then (17) can be approximated as

$$\hat{\boldsymbol{x}}=E(\boldsymbol{x}|\boldsymbol{y})\approx\frac{1}{M}\sum_{i=1}^{M}\tilde{\boldsymbol{x}}_i. \tag{18}$$

The evaluation of (18) requires drawing samples from the posterior distribution. In the sequel, we will present an effective scheme for achieving this goal

$$\begin{aligned}\boldsymbol{\Sigma} &= \boldsymbol{W}_z^{-1} \\ &= \left(\tau \boldsymbol{H}^\top \boldsymbol{D}^\top \boldsymbol{D}\boldsymbol{H} + \sum_{l=1}^{L} \boldsymbol{K}_l^\top \boldsymbol{Z}_l \boldsymbol{K}_l\right)^{-1} \\ &= \left(\begin{bmatrix}\boldsymbol{H}^\top \boldsymbol{D}^\top & \boldsymbol{K}_1^\top \cdots \boldsymbol{K}_L^\top\end{bmatrix} \begin{bmatrix}\tau \boldsymbol{I} & \cdots & \boldsymbol{0} \\ & \boldsymbol{Z}_1 & \vdots \\ \vdots & & \ddots \\ \boldsymbol{0} & \cdots & \boldsymbol{Z}_L\end{bmatrix} \begin{bmatrix}\boldsymbol{D}\boldsymbol{H} \\ \boldsymbol{K}_1 \\ \vdots \\ \boldsymbol{K}_L\end{bmatrix}\right)^{-1} \\ &= \left(\boldsymbol{K}^\top \boldsymbol{Z} \boldsymbol{K}\right)^{-1}. \end{aligned} \tag{19}$$

B. Generative Bayesian SR Algorithm

As it is typically intractable to get the mean of the posterior distribution, i.e., evaluating (17), in general, we resort to approximate schemes for solution. One type of approach is to approximate the posterior distribution with separable distribution under criteria such as Kullback–Leibler-divergence, thus transforming the problem to another easy-to-solve problem. A representative example is the VB method [33], which has been applied to image deblurring and compressive sensing recently [26], [34]. While approaches based on the VB technique are computationally efficient as it typically offers analytic formulas for updating the variables, it compromises the modeling accuracy for computational efficiency. Apart from approximating the original model with tractable ones as in VB, another line of solution is based on generative sampling from the posterior distribution. After obtaining enough samples from the posterior distribution, the empirical mean calculated from the generated samples is used as the estimation for the posterior mean.

In this paper, we also use a generative scheme to solve the SR problem by taking advantage of the probability modeling. However, directly drawing samples from the posterior (16) is hard. We follow [32] and [35] and use an auxiliary variable scheme for sampling. By introducing auxiliary variable $\boldsymbol{z} \in \mathbb{R}^{m \times L}$ as the matrix consisting the indicator variables for the scales in the GSM model for each expert at each location in the image $\boldsymbol{x}$, we can obtain

$$p(\boldsymbol{x}, \boldsymbol{z}|\Theta) = \frac{1}{Z(\Theta)} \prod_{c \in \zeta} \prod_{l=1}^{L} p(z_{cl}) \mathcal{N}\left((\boldsymbol{k}_l * \boldsymbol{x})_c; 0, \frac{\eta_l^2}{s_{z_{cl}}}\right) \tag{20}$$

where $p(z_{cl}) = \beta_{l z_{cl}}$. This can be verified easily corresponding to $p(\boldsymbol{x}|\Theta)$ when $\boldsymbol{z}$ is marginalized out. Note that from (20), the conditional distribution of $\boldsymbol{x}$ given $\boldsymbol{z}$ is a Gaussian distribution. Substituting (20) into (16), we can get a posterior distribution $p(\boldsymbol{x}, \boldsymbol{z}, \tau | \boldsymbol{y})$ as

$$\begin{aligned} p(\boldsymbol{x}, \boldsymbol{z}, \tau | \boldsymbol{y}) \propto\ & \mathcal{N}(\boldsymbol{y}; \boldsymbol{D}\boldsymbol{H}\boldsymbol{x}, \tau^{-1}\boldsymbol{I}) \\ & \cdot \prod_{c \in \zeta} \prod_{l=1}^{L} p(z_{cl}) \mathcal{N}\left((\boldsymbol{k}_l * \boldsymbol{x})_c; 0, \frac{\eta_l^2}{s_{z_{cl}}}\right) \end{aligned} \tag{21}$$

where we have omitted Θ here and in the following to keep the notation uncluttered without the danger of confusion. With MCMC scheme, we sample $\boldsymbol{z}$, τ, and $\boldsymbol{x}$ alternatively, which is detailed in the sequel.

1) Sample $\boldsymbol{z}$: Since the scale indices (i.e., the elements in $\boldsymbol{z}$) are conditionally independent, given the image $\boldsymbol{x}$, we can obtain the conditional distribution for z_{cl} as

$$p(z_{cl}|\boldsymbol{x}, \boldsymbol{y}) \propto p(z_{cl}) \cdot \mathcal{N}\left((\boldsymbol{k}_l * \boldsymbol{x})_c; 0, \frac{\eta_l^2}{s_{z_{cl}}}\right). \tag{22}$$

Therefore, $\boldsymbol{z}$ can be sampled easily with (22).

2) Sample τ: The conditional distribution for τ is

$$\begin{aligned} p(\tau|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{y}) &\propto \mathcal{N}(\boldsymbol{y}; \boldsymbol{D}\boldsymbol{H}\boldsymbol{x}, \tau^{-1}\boldsymbol{I}) \mathcal{G}(\tau; a^o, b^o) \\ &\propto \tau^{\frac{n}{2}} \exp\left(-\tau \frac{\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{H}\boldsymbol{x}\|_2^2}{2}\right) \cdot \tau^{a^o - 1} \exp(-\tau b^o) \\ &\propto \mathcal{G}\left(\tau; \frac{n}{2} + a^o, \frac{\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{H}\boldsymbol{x}\|_2^2}{2} + b^o\right) \end{aligned} \tag{23}$$

from which we can sample for the precision easily.

3) Sample $\boldsymbol{x}$: The conditional distribution for $\boldsymbol{x}$ is a multivariant Gaussian distribution as follows:

$$\begin{aligned} p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{y}, \tau) &\propto \mathcal{N}\left(\boldsymbol{y}; \boldsymbol{D}\boldsymbol{H}\boldsymbol{x}, \tau^{-1}\boldsymbol{I}\right) \mathcal{N}\left(\boldsymbol{x}; \boldsymbol{0}, \boldsymbol{V}_z^{-1}\right) \\ &\propto \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \end{aligned} \tag{24}$$

where the mean vector is $\boldsymbol{\mu} = \tau \boldsymbol{W}_z^{-1} \boldsymbol{H}^\top \boldsymbol{D}^\top \boldsymbol{y}$ and the covariance matrix $\boldsymbol{\Sigma}$ is given in (19). We denote $\boldsymbol{V}_z = \sum_{l=1}^{L} \boldsymbol{K}_l^\top \boldsymbol{Z}_l \boldsymbol{K}_l$, $\boldsymbol{W}_z = \tau \boldsymbol{H}^\top \boldsymbol{D}^\top \boldsymbol{D}\boldsymbol{H} + \boldsymbol{V}_z$ and $\boldsymbol{Z}_l = \text{diag}\{(s_{z_{cl}}/\eta_l^2)\}$. Recall that $\boldsymbol{K}_l$ is the matrix form corresponding to the filter $\boldsymbol{k}_l$. The difficulty for sampling from (24) is that the covariance matrix $\boldsymbol{\Sigma}$ is huge for typical image size, thus precluding the usage of the standard eigen-decomposition based sampling approach. We address this difficulty based on recent progress on high-dimensional data sampling [32], [36]–[38]. Specifically, to generate a sample $\boldsymbol{x}$ from $p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{y}, \tau)$, we first decompose it into two parts

$$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v} \tag{25}$$

where each part is obtained by solving the following two equations, respectively. For the $\boldsymbol{u}$-component, we compute it via solving

$$\boldsymbol{W}_z \boldsymbol{u} = \tau \boldsymbol{H}^\top \boldsymbol{D}^\top \boldsymbol{y}. \tag{26}$$

For obtaining the $\boldsymbol{v}$-component, we first draw a sample $\boldsymbol{r} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, and then solve the following equation for $\boldsymbol{v}$:

$$\boldsymbol{K}^\top \boldsymbol{Z} \boldsymbol{K} \boldsymbol{v} = \boldsymbol{K}^\top \sqrt{\boldsymbol{Z}} \boldsymbol{r} \tag{27}$$

where $\boldsymbol{K}$ and $\boldsymbol{Z}$ are defined as in (19). The sampling approach used above is closely related to the "Perturb-and-MAP" scheme recently developed for sampling Gaussian MRF models [36]. Both (26) and (27) can be solved efficiently with conjugate gradient (CG) method. Note that it is unnecessary to construct $\boldsymbol{K}$, $\boldsymbol{Z}$, and $\boldsymbol{W}_z$ explicitly, as only the ability to evaluate a vector with respect to them is required in the CG algorithm. The standard random sampling approach for multivariant Gaussian using eigen-decomposition, such as

Algorithm 1 Generative Bayesian SR

1: **Input:** LR observation $\boldsymbol{y}$, zooming factor r, number of samples M, burn in step length N.
2: **Initialize:** set up $\boldsymbol{D}$ and $\boldsymbol{H}$ according to r and set the initial HR estimation as $\tilde{\boldsymbol{x}}_0 = \boldsymbol{H}^\top \boldsymbol{D}^\top \boldsymbol{y}$.
3: **For** $i = 1, 2, \ldots, M$, do:
  1) sample $\tilde{z}_i \sim p(z|\tilde{\boldsymbol{x}}_{i-1}, \boldsymbol{y}, \boldsymbol{D}, \boldsymbol{H})$ via (22);
  2) sample $\tilde{\tau}_i \sim p(\tau|\tilde{\boldsymbol{x}}_{i-1}, \tilde{z}_i, \boldsymbol{y}, \boldsymbol{D}, \boldsymbol{H})$ via (23);
  3) sample $\tilde{\boldsymbol{x}}_i \sim p(\boldsymbol{x}|\tilde{z}_i, \tilde{\tau}_i, \boldsymbol{y}, \boldsymbol{D}, \boldsymbol{H})$:
    a) solve for $\tilde{\boldsymbol{u}}_i$ via (27);
    b) solve for $\tilde{\boldsymbol{v}}_i$ via (26);
    c) obtain a sample via $\tilde{\boldsymbol{x}}_i = \tilde{\boldsymbol{u}} + \tilde{\boldsymbol{v}}$.
4: **End**
5: **Perform** SR image estimation via (28).
6: **Output:** SR estimation $\hat{\boldsymbol{x}}$.

Cholesky decomposition has time and space complexity as $\mathcal{O}(w^2 m)$ and $\mathcal{O}(wm)$, respectively, where w denotes the bandwidth of the system matrix in (26) and (27), thus is expensive as the bandwidth w increases linearly with the size of the image and the spatial extent of the filters. By the CG method, the typical time complexity is as low as $\mathcal{O}(m)$ or $\mathcal{O}(m \log m)$ per iteration while the space complexity is $\mathcal{O}(m)$, which is more efficient than the eigen-decomposition based method. By using the property of a Gaussian random vector under a linear transformation, we can easily verify that a sample $\boldsymbol{x}$ generated according to the above described approach is indeed a sample drawn from the Gaussian distribution given in (24).

We alternatively sample z, τ, and $\boldsymbol{x}$ with (22)–(24), generating a sample set $\{\tilde{\boldsymbol{x}}_i, \tilde{\tau}_i, \tilde{z}_i\}_{i=1}^M$. As there is a burn in process, only the samples after a certain number of iterations are considered as drawn from the target joint distribution. We use N to denote the number of steps for the burn in process. Therefore, the estimated HR image can be calculated via simple averaging scheme as follows:

$$\hat{\boldsymbol{x}} = \frac{1}{M-N} \sum_{i=N+1}^{M} \tilde{\boldsymbol{x}}_i. \tag{28}$$

Alternatively, we can also estimate the HR image via the Rao–Blackwellized estimator by averaging over the conditional expectations as suggested in [36], which can achieve the same performance with less iterations

$$\hat{\boldsymbol{x}} = \frac{1}{M-N} \sum_{i=N+1}^{M} \tilde{\boldsymbol{u}}_i \tag{29}$$

where $\tilde{\boldsymbol{u}}_i$ is the $\boldsymbol{u}$-component for the sample $\tilde{\boldsymbol{x}}_i$. This scheme is used in the current work. Finally, we will estimate the noise level. From (23), we have the conditional expectation for τ as

$$\tilde{\tau}_i = \frac{n + 2a^o}{\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{H}\tilde{\boldsymbol{x}}_i\|_2^2 + 2b^o}. \tag{30}$$

Recall the relationship $\tau = \sigma^{-2}$, we can estimate the noise level as follows:

$$\hat{\sigma} = \frac{1}{M-N} \sum_{i=N+1}^{M} \tilde{\tau}_i^{-\frac{1}{2}} = \sqrt{\frac{\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{H}\tilde{\boldsymbol{x}}_i\|_2^2 + 2b^o}{n + 2a^o}} \tag{31}$$

which again uses the Rao–Blackwellized estimation scheme [36].

Note that in (31), the prior knowledge on the noise level is incorporated into the final noise estimation via the shape and scale parameter a^o and b^o of the hyperprior. If we have no knowledge available on the noise level, we can simply use a uninformative prior by setting $a^o = 1$ and $b^o = 0$.

The overall procedure of the proposed method is summarized in Algorithm 1. Note that when the downsampling operator $\boldsymbol{D}$ is an identity matrix, our model reduces naturally to the deblurring method proposed in [39]. Therefore, our method can be recognized as a generalization of the method in [39]. Finally, we want to make some comments on the proposed method and the sparse representation-based SR method [5], which is a *state-of-the-art* SR algorithm. The ScSR method proposed in [5] is actually a *synthesis* based SR approach, which first recovers the sparse representations with the LR patch and then uses them to synthesize the HR patch with a corresponding HR dictionary. On the other hand, our method is a *generative* SR approach, which uses flexible prior models to exploit the statistics of natural images and takes advantage of the Bayesian probabilistic modeling via sampling. As a coupled dictionary pair is used for synthesizing the HR image in the ScSR method, the training process has to be performed again once the zooming factor is changed. The proposed method, however, does not have this limitation and thus is more flexible.

C. Implementation Details

In this section, we make some comments on the implementation details. The low-passing filter corresponding to $\boldsymbol{H}$ is modeled as a Gaussian filer with standard deviation of 2 when the zooming factor is 3. The size of the Gaussian filer is set as 7. The filtering process is implemented via FFT in the frequency domain. The down sampling operator $\boldsymbol{D}$ is implemented via subsampling the image for every r rows and columns for the zooming factor of r. Eight experts (thus $L = 8$) are multiplied together to form the energy potential $f(\cdot)$, which is a PoE model. Each expert is associated with one of the eight different filters. GSM model with $J = 15$ components is used as the expert function. All the 15 components share a common base variance parameter $\eta_l = 2e^{-3}$ and each has a scale parameter as $s_1 \sim s_{15} = \exp\{-9, -7, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 7, 9\}$. In our implementation, we use filters $\{\boldsymbol{k}_l\}_{l=1}^L$ with size of 3×3. The filters $\{\boldsymbol{k}_l\}_{l=1}^L$ as well as the mixing weights for each GSM $\{\beta_{lj}\}_{l,j}$ are learned from the training images using the contrastive divergence algorithm [29]. Specifically, we use the implementation given by [32] which is trained on the Berkeley segmentation database [40]. Note that the training dataset does not include any test images used for comparison in the experimental section. The learned filters, the experts, and the combination weights in the GSM model are shown in Fig. 2. The convergence condition is a common issue to be considered for MCMC-based algorithms. As to the proposed approach, we observe empirically that overall, the performance of the algorithm increases with the increase of the number

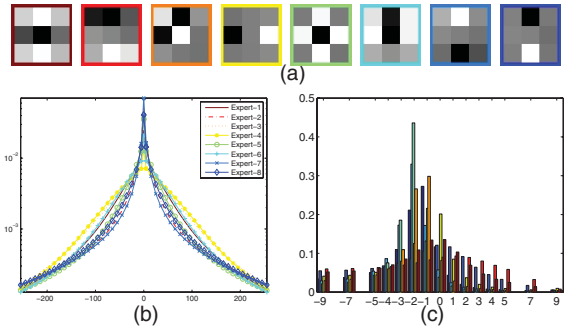

Fig. 2. Illustration of the learned prior for natural images [32]. (a) Learned filters. (b) Learned expert functions, each of which is a GSM model with learned weights, as shown in (c). x-axis in (c) denotes $\log s$ while the y-axis denotes the weight β. The plot in (b) and bar in (c) corresponds to the filter in (a) with the same color.

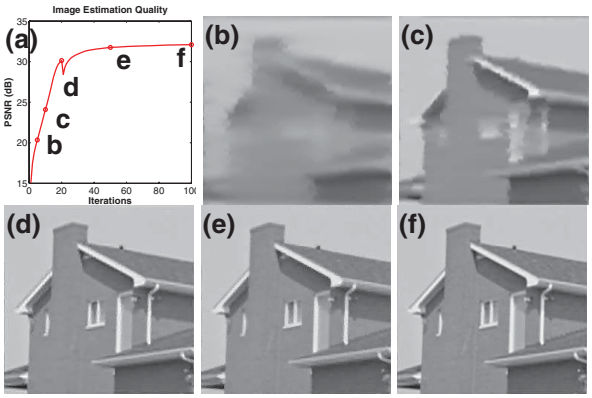

Fig. 3. Illustration of the SR ($\times 3$) estimation evolution. (a) PSNR curve with (b)–(f) intermediate SR estimation results for the "House" image.

of iterations [see Fig. 3 (a)] and converges after a sufficient number of iterations. Empirically, we perform the sampling process for $M = 100$ times and then treat the first $N = 20$ steps as burn in process. In the case of noisy image SR, we set $M = 200$.

"house"

V. Experiments and Results

In this section, we conduct several experiments to evaluate the effectiveness of the proposed method, both qualitatively and quantitatively. We first give an illustrative example to demonstrate some basic properties of the proposed algorithm. Then we compare the SR estimation results with several classical as well as *state-of-the-art* SR methods, on several standard test images as well as real-world color images. Finally, the robustness of the proposed SR method with respect to noise is also investigated. Peak signal to noise ratio (PSNR) is adopted as an evaluation metric, which is defined as

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\frac{1}{m}\sum_i (\hat{\boldsymbol{x}} - \boldsymbol{x})_i^2}$$

where $\hat{\boldsymbol{x}}$ denotes the estimated SR image and $\boldsymbol{x}$ denotes the original HR image. m is the total number of pixels in $\boldsymbol{x}$. The structural SIMilarity (SSIM) index is used as another objective measure, which aligns better with the human perception than PSNR [41].

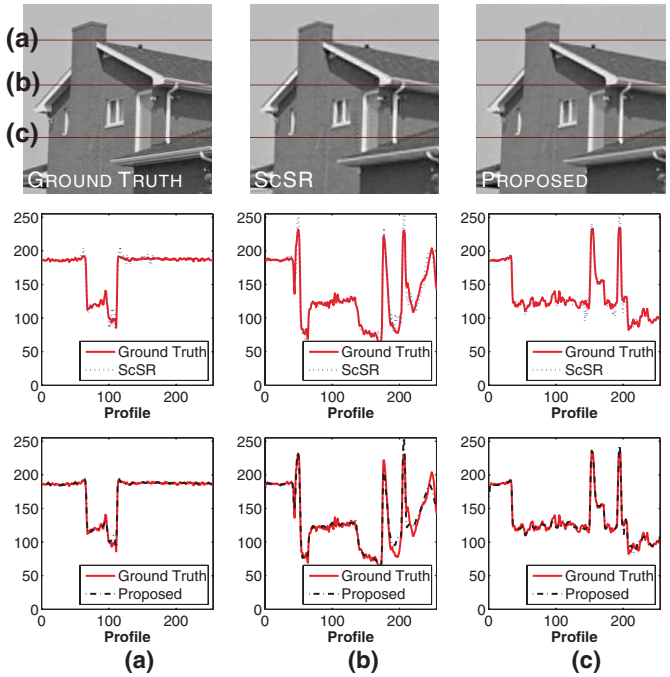

Fig. 4. Comparison of several horizontal profiles of: the ground-truth image, SR results from the ScSR method [5], and the proposed method ($\times 3$). (a)–(c) Three different horizontal profiles.

A. Illustrative Example

We first give an illustrative example in Fig. 3 to demonstrate the estimation process of the proposed algorithm with the "House" image, with the zooming factor as 3. In Fig. 3, the plot in (a) is the PSNR curve with respect to the iterations (i.e., the sampling process). In Fig. 3(b)–(f), we show the intermediate estimation results when the iteration number is $t \in \{5, 10, 20, 50, 100\}$. It can be observed from Fig. 3 that for the proposed algorithm, the estimation quality improves as the number of iteration increases, generating more and more accurate SR estimation result. The evolution curve with such kind of pattern has been observed for all the experiments in this paper. Fig. 4 shows the comparison of three horizontal profiles [denoted as (a)–(c)] from: the ground truth HR image, the HR estimation from ScSR method [5], and the HR result from the proposed method. It is observed from the figure that the profiles from ScSR may have large amplitude differences near the edges, which is manifested as the visible artifacts in the image; the profiles of the SR result from the proposed method, on the other hand, do not suffer from this problem, which indicates the benefits of using a global natural image prior for images. The experimental results shown in the following sections will further verity the effectiveness of the proposed method.

B. SR Experiment

To further verify the effectiveness of the proposed method, we compare the SR performance of the proposed method with several algorithms: Nearest neighbor (NN) interpolation, bicubic interpolation (BI), fast SR method [18], ScSR method [5],[3] as well as another variational Bayesian-based

[3]For the ScSR method, the dictionary size is 1024 and the regularization parameter is set as $\lambda = 0.1$, which is the same setting as in [5].

Fig. 5. Single image super resolution results for "Peppers," "Cameraman" and "Lena" images ($\times 3$) with different algorithms: NN interpolation, fast SR method [18], sparse representation based SR method [5] and the proposed SR method. The results evaluated in terms of PSNR and SSIM are reported in Table I.

SR method [15].[4] In this experiment, we use the standard test images in the noise-free setting for evaluation and use PSNR and SSIM as the objective measures. The zooming factor in this experiment is $r = 3$. The SR results in terms of PSNR and SSIM are summarized in Table I. As can be seen from Table I, the proposed method can generate SR images with better quality in terms of PSNR and SSIM than the fast SR method [18] and the ScSR method [5] on all the testing images. This clearly demonstrates the effectiveness of the proposed method. To further compare the SR results visually, we show some of the SR results from different algorithms in Fig. 5. As can be seen from Fig. 5, the fast SR method tends to removing the fine details from the image and the SR results exhibit cartoon-like artifacts. Furthermore, it is noted from Fig. 5 that the SR results from the ScSR method have some artifacts along the edge structures (the zoomed patches for "Peppers" and "Cameraman" images in Fig. 5). This may result from the inconsistency in recovering the HR patches from the neighboring LR patches. On the other hand, the SR image from the proposed method does not suffer from these artifacts, due to the usage of the global statistical model as well as the MMSE estimation criteria for the latent HR images. This clearly demonstrates the benefit of using a global statistical model as a prior based on natural image statistics, combined with Bayesian posterior mean estimation criteria.

[4]Codes are available at http://decsai.ugr.es/pi/superresolution/software.html.

TABLE I
SR ($\times 3$) RESULT COMPARISON OF ESTIMATION QUALITY

Methods		NN	BI	Fast [18]	ScSR [5]	Proposed
House	PSNR	24.90	25.62	28.72	30.80	**32.06**
	SSIM	0.763	0.788	0.837	0.863	**0.897**
Peppers	PSNR	21.59	22.18	24.38	25.52	**25.88**
	SSIM	0.727	0.775	0.841	0.863	**0.910**
Cameraman	PSNR	21.71	22.12	24.64	25.79	**26.42**
	SSIM	0.695	0.714	0.774	0.810	**0.847**
Barbara	PSNR	23.07	23.43	24.58	25.24	**25.62**
	SSIM	0.618	0.649	0.683	0.732	**0.753**
Lena	PSNR	25.91	26.62	29.59	31.89	**33.25**
	SSIM	0.769	0.802	0.833	0.888	**0.911**
Boat	PSNR	24.14	24.63	26.70	28.25	**29.47**
	SSIM	0.659	0.687	0.729	0.815	**0.844**
Bill	PSNR	25.88	26.41	27.72	30.15	**30.44**
	SSIM	0.654	0.686	0.714	0.811	**0.831**
Couple	PSNR	24.16	24.60	26.34	27.69	**28.53**
	SSIM	0.621	0.649	0.698	0.788	**0.811**

To further verify the effectiveness of the proposed method, we compare its SR performance with more SR algorithms, including the Kernel ridge (KRR) regression SR method [10] and the variational Bayesian SR (VBSR) method [15], with

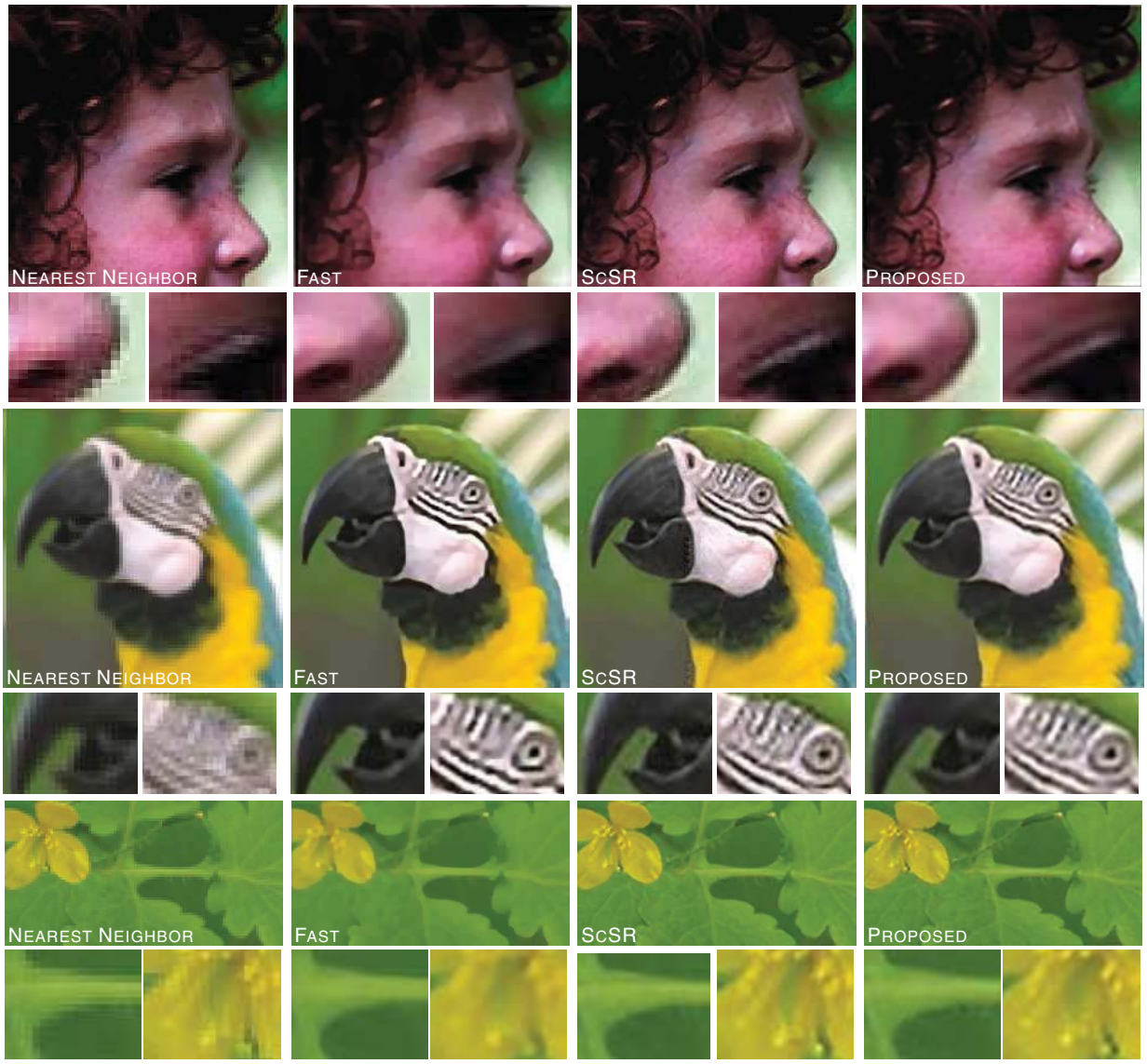

Fig. 6. Single image SR (×3) results for "Girl," "Parrot," and "Flower" images (×3). PSNR and SSIM results for the image "Flower" are given in brackets. Left to right: NN interpolation (32.95, 0.849), fast SR method [18] (35.81, 0.890), sparse representation-based SR method [5] (37.60, 0.923), and the proposed SR method (38.07, 0.927).

zooming factor of 4.[5] The KRR method achieves SR via patch-based sparse kernel regression for capturing the mapping between LR and HR patches. The VBSR method uses a simple TV prior model and estimates the HR image with posterior mean via VB approximation. As in the current implementation of the VBSR method, the matrices corresponding to the low-pass filtering as well as the down-sampling operator are constructed explicitly, thus it can only handle small images. Therefore, we crop image parts of size 192×192 from the standard test images and summarize the results in Table II. As can be seen from Table II, the KRR method performs better than the fast SR method in general, and the VBSR method overall performs similar to the ScSR method and the KRR method. Finally, the proposed method outperforms the VBSR method and all the other methods constantly in terms of both PSNR and SSIM, which demonstrates the superiority of the proposed method. The average computational time per iteration for different algorithms is also reported in Table II.[6] The typical number of iteration is ten for the fast SR method, 30 for VBSR, and 100 for the proposed method. The other algorithms in Table II are "one-shot" methods that require no iteration.

We further evaluate the proposed SR method on the task of color image SR. For color image SR, we only perform SR for the intensity channel in the YCbCr color space and use BI for the other two channels. The reason is that human eyes are more sensitive to the intensity changes of the image than to the changes in other channels, such as saturation and hue. This scheme has been practiced already in the previous SR works [5]. The SR results on color images with the zooming factor of 3 are shown in Fig. 6. The PSNR and SSIM results accompanying these images are also

[5]As both the KRR and VBSR methods can only handle SR tasks with even zooming factor, they are not compared in Table I.

[6]The computational time is measured on a Laptop with Intel Core2 Duo P8600 (2.4 GHz) CPU and 1 GB RAM.

TABLE II
SR ($\times 4$) QUALITY COMPARISON ($\sigma = 1$)

Methods		NN	BI	Fast [18]	KRR [10]	ScSR [5]	VBSR [15]	Proposed
House	PSNR	22.24	23.14	26.79	27.60	26.86	27.00	**28.20**
	SSIM	0.627	0.671	0.771	0.785	0.742	0.767	**0.802**
Peppers	PSNR	21.47	22.20	25.36	25.95	25.61	25.36	**26.90**
	SSIM	0.647	0.711	0.786	0.821	0.764	0.807	**0.837**
Cameraman	PSNR	18.35	18.91	21.06	21.35	21.17	21.73	**21.89**
	SSIM	0.695	0.714	0.742	0.758	0.695	0.810	**0.847**
Barbara	PSNR	21.91	21.50	24.12	24.75	24.62	24.12	**24.94**
	SSIM	0.515	0.582	0.640	0.696	0.654	0.699	**0.703**
Lena	PSNR	21.68	22.41	26.66	26.88	27.17	27.62	**28.32**
	SSIM	0.624	0.678	0.766	0.807	0.760	0.814	**0.825**
Boat	PSNR	19.52	20.17	22.47	23.27	22.94	23.71	**23.92**
	SSIM	0.659	0.687	0.614	0.652	0.598	0.815	**0.844**
Hill	PSNR	24.56	25.17	26.94	27.75	24.92	27.10	**27.94**
	SSIM	0.540	0.581	0.623	0.666	0.669	0.675	**0.680**
Couple	PSNR	21.98	22.41	24.08	24.79	25.17	25.59	**27.94**
	SSIM	0.524	0.564	0.655	0.697	0.641	0.707	**0.718**
Time (s)		0.003	0.005	0.9	23.4	5.4	4.5	14.5

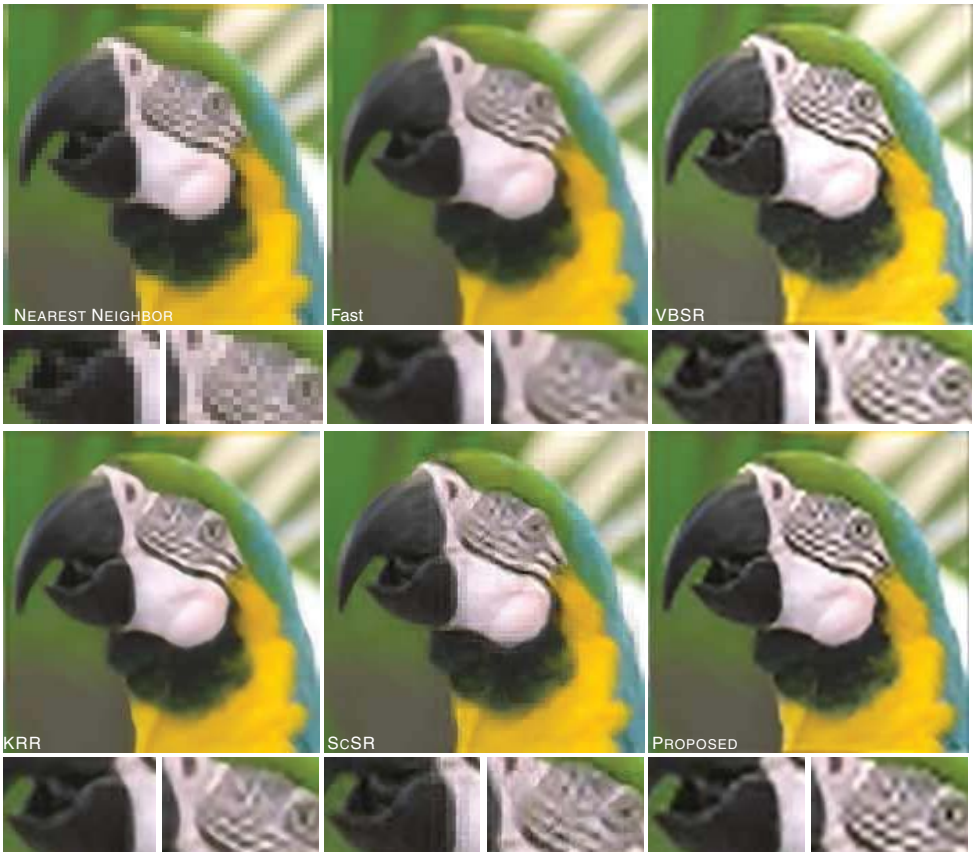

Fig. 7. Single image SR results with different algorithms ($\times 4$). The results are generated with: NN interpolation (22.64, 0.699), Shan *et al.* fast SR method [18] (24.09, 0.775), VBSR [15] (23.33, 0.823), Kim *et al.* kernel regression-based method (24.22, 0.818) [10], Yang *et al.* ScSR method (25.47, 0.798) [5], and the proposed method (26.22, 0.847).

reported. As can be seen from Fig. 6, the proposed method can generate SR images visually comparable to the sparse representation-based method, which are much better than the fast SR method. Detailed inspections also reveal that the SR results from the proposed method have less artifacts than those from the ScSR method (the zoomed patches for the *nose* and *eye* parts of the "Girl" image in Fig. 6). Furthermore, the quantitative evaluation also indicates that the SR results from the proposed method have better quality in terms of PSNR and SSIM.

Fig. 8. Single image SR results with different algorithms (×4). The results are generated with: NN interpolation (26.05, 0.741), BI (26.92, 0.785), Freeman *et al.* example based SR method (25.38, 0.729), Kim *et al.* kernel regression-based method (31.22, 0.863) [10], Fattal *et al.* edge statistics-based SR method (25.55, 0.751) [11], Glasner *et al.* method (25.28, 0.741) [3], Shan *et al.* method (24.43, 0.727) [18], and the proposed method (31.80, 0.885).

More color image SR results are shown in Figs. 7 and 8, with the zooming factor of 4. We compare the SR result from our proposed method with the results from several SR methods in the literature: the results from Freeman's example-based SR method [2], Kim's kernel regression method [10], Glasner's method [3], Shan *et al.* fast SR method [18], Yang *et al.* sparse representation-based SR method [5] as well as the VB-based SR method [15]. From Fig. 7, we can see that the fast SR method again over-smoothens the HR image, and the ScSR method generates HR images with artifacts along the edges. The HR image from the VBSR method suffers from ringing artifacts. The proposed method, on the other hand, generates HR image with less artifacts and with desirable quality both visually and in terms of PSNR and SSIM. More SR algorithms are compared in Fig. 8. It can be observed from Fig. 8 that the SR results from Freeman's example-based SR method and Glasner's method although have visual resolution improvement, they actually hallcinating some HR information that is not fidelity to the original image, thus hindering the improvement in terms of PSNR and SSIM. Fattal's method and Shan's method appear to suffer from cartoon-like artifacts. The proposed method is visually much better than results via NN and BI; also, it has less artifacts compared to the result from example-based SR method [2], while appears similar to result from Kim's kernel regression method [10], and achieves the best performance in terms of PSNR and SSIM. The results on both gray and color image SR, all verified the effectiveness of the proposed method.

C. Noisy Image SR

In real-world SR tasks, the observed LR images are rarely noise-free, but are often contaminated by noise. Therefore, it is important for the SR algorithm to be robust to the noise contained in the LR observations. In this section, we evaluate the effectiveness of the proposed approach in the case of noisy SR situation, i.e., the given input is both of low in resolution and is contaminated by noise. The SR estimation results under the noise levels $\sigma \in \{0, 1, 2, 4\}$ are reported in Table III. The bar plots for the evaluated algorithms are shown in Fig. 9. The PSNR and SSIM performance shown in Fig. 9 is the average performance over the test images. As can been seen from Table III and Fig. 9, the proposed method performs better than the other methods when the noise level is low; as the noise level increases, the performance in terms of PSNR and SSIM measures of all the SR algorithms decreases. In this case, the performance of the proposed method is still better than the other algorithms in general in terms of SSIM and is comparable to ScSR in terms of PSNR, indicating its robustness to the noise in the LR images and its applicability to general SR scenarios. It is worthwhile to point out that for the ScSR method, it is given the ground truth noise level, and

TABLE III
NOISY IMAGE SR (×3) RESULT COMPARISON OF ESTIMATION QUALITY

σ	Methods		NN	BI	Fast [18]	ScSR [5]	Proposed
0	House	PSNR	24.90	25.62	28.72	30.80	**32.06**
		SSIM	0.763	0.788	0.837	0.863	**0.897**
	Peppers	PSNR	21.59	22.18	24.38	25.52	**25.88**
		SSIM	0.727	0.775	0.841	0.863	**0.910**
	Cameraman	PSNR	21.71	22.12	24.64	25.79	**26.42**
		SSIM	0.695	0.714	0.774	0.810	**0.847**
1	House	PSNR	24.88	25.60	28.88	30.67	**31.24**
		SSIM	0.757	0.784	0.844	0.854	**0.876**
	Peppers	PSNR	21.58	22.17	24.53	25.49	**25.65**
		SSIM	0.723	0.772	0.851	0.858	**0.886**
	Cameraman	PSNR	21.70	22.15	24.73	25.75	**26.09**
		SSIM	0.689	0.710	0.778	0.801	**0.823**
2	House	PSNR	24.81	25.55	28.76	**30.47**	**30.47**
		SSIM	0.738	0.773	0.840	0.834	**0.858**
	Peppers	PSNR	21.55	22.15	24.49	**25.36**	25.23
		SSIM	0.710	0.764	0.847	0.844	**0.860**
	Cameraman	PSNR	21.66	22.13	24.70	25.63	**25.67**
		SSIM	0.671	0.700	0.775	0.779	**0.802**
3	House	PSNR	24.57	25.36	28.41	**29.38**	29.33
		SSIM	0.678	0.734	0.823	0.798	**0.832**
	Peppers	PSNR	21.44	22.06	24.29	**24.65**	24.54
		SSIM	0.667	0.736	**0.830**	0.803	0.823
	Cameraman	PSNR	21.54	22.03	24.52	24.88	**24.94**
		SSIM	0.612	0.662	0.757	0.734	**0.770**

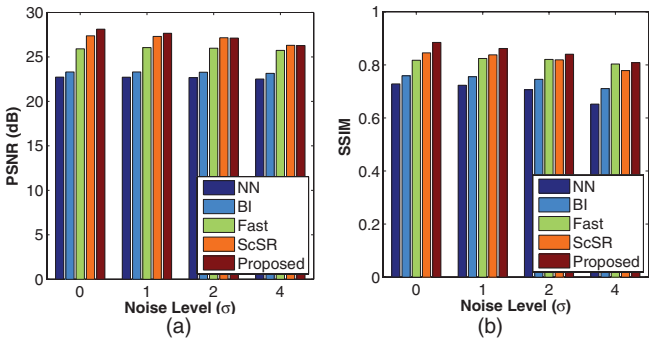

Fig. 9. Noisy image SR results. (a) PSNR bar plot and (b) SSIM bar plot with respect to increasing noise level ($\sigma \in \{0, 1, 2, 4\}$).

its regularization weight has to be adjusted according to the noise level.[7] The proposed method, on the other hand, does not require the noise level to be known, but can infer it from the noise LR image itself, while generating results comparable or better than those from ScSR, demonstrating its effectiveness for real-world SR tasks.

VI. CONCLUSION

An effective generative Bayesian single image SR algorithm with natural image prior was proposed in this paper. The proposed method exploits the natural image statistics for image SR with a flexible high-order MRF model. Specifically, an FoE model with products of GSM potentials was used for learning the prior model from natural images, which was further incorporated into a fully Bayesian framework for image SR. The SR image was estimated under the MMSE criteria, rather than the conventional MAP criteria. In other words, the HR image was estimated with the posterior mean, rather than the posterior mode. To estimate the latent HR image, we take advantage of the recent development in efficient sampling of high-dimensional data for non-Gaussian models and develop an effective sampling scheme for approximating the MMSE estimation for the HR image. Experimental results under different settings compared with several *state-of-the-art* SR algorithms in the literature verified the effectiveness of the proposed method. As the proposed SR method is an MCMC sampling-based generative approach, it is generally not as fast as the SR approaches using the MAP solution. However, we believe that it is beneficial to pursue this direction, due to its flexibility in using advanced natural image priors in an elegant Bayesian model as well as its potentials in the SR task as shown by the experimental results. For future work, we would like to examine the possibility of combining the generative models with synthesize-based methods to further improve both the computational efficiency of the proposed approach and the estimation quality of the SR results.

[7] In our experiments, we adjust the regularization weight the same way as suggested in [5].

REFERENCES

[1] R. Y. Tsai and T. Huang, "Multiframe image restoration and registration," in *Proc. Adv. Comput. Vision Image Process.*, 1984, pp. 317–339.

[2] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example based super-resolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, Mar.–Apr. 2002.

[3] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. IEEE Int. Conf. Comput. Vision*, Oct. 2009, pp. 349–356.

[4] H. Zhang, J. Yang, Y. Zhang, and T. Huang, "Non-local kernel regression for image and video restoration," in *Proc. Eur. Conf. Comput. Vision*, 2010, pp. 566–579.

[5] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[6] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1167–1183, Sep. 2002.

[7] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[8] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[9] M. F. Tappen, B. C. Russell, and W. T. Freeman, "Exploiting the sparse derivative prior for super-resolution and image demosaicing," in *Proc. IEEE Workshop Stat. Comput. Theories Vision*, Oct. 2003, pp. 1–24.

[10] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, Jun. 2010.

[11] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 95–102, 2007.

[12] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, and A. K. Katsaggelos, "Softcuts: A soft edge smoothness prior for color image super-resolution," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 969–981, May 2009.

[13] J. Sun, J. Sun, Z. Xu, and H. Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, Jun. 2008, pp. 1–8.

[14] A. Mohebi and P. Fieguth, "Posterior sampling of scientific images," in *Proc. Int. Conf. Image Anal. Recognit.*, 2006, pp. 1–12.

[15] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian super resolution," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 984–999, Apr. 2011.

[16] T. Katsuki, A. Torii, and M. Inoue, "Posterior mean super-resolution with a causal Gaussian Markov random field prior," *IEEE Trans. Image Process.*, to be published.

[17] M. E. Tipping and C. M. Bishop, "Bayesian image super-resolution," in *Proc. Adv. Neural Inf. Process. Syst.*, 2005, pp. 1303–1310.

[18] Q. Shan, Z. Li, J. Jia, and C. K. Tang, "Fast image/video upsampling," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 1–5, 2008.

[19] Y. W. Tai, S. Liu, M. S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2400–2407.

[20] L. C. Pickup, S. J. Roberts, and A. Zisserman, "A sampled texture prior for image super-resolution," in *Proc. Adv. Neural Inf. Process. Syst.*, 2003, pp. 1587–1594.

[21] Y. HaCohen, R. Fattal, and D. Lischinski, "Image upsampling via texture hallucination," in *Proc. IEEE Int. Conf. Comput. Photography*, Mar. 2010, pp. 1–8.

[22] D. J. Field, "Relations between the statistics of natural images and the response profiles of cortical cells," *J. Opt. Soc. Amer.*, vol. 4, no. 12, pp. 2379–2394, 1987.

[23] M. Zontak and M. Irani, "Internal statistics of a single natural image," in *Proc. IEEE Conf. Comput. Vis. Patern Recognit.*, Jun. 2011, pp. 977–984.

[24] Y. Weiss and W.T. Freeman, "What makes a good model of natural images?" in *Proc. IEEE Conf. Comput. Vis. Patern Recognit.*, Jun. 2007, pp. 1–8.

[25] S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vis.*, vol. 82, no. 2, pp. 205–229, 2009.

[26] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph. (SIGGRAPH)*, vol. 25, pp. 787–794, Aug. 2006.

[27] S. Osher, M. Burger, and D. Goldfarb, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.

[28] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-laplacian priors," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1033–1041.

[29] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural Comput.*, vol. 14, no. 8, pp. 1771–1800, 2002.

[30] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman, "Image restoration by matching gradient distributions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 683–694, Apr. 2012.

[31] S. Lyu, "Interpretation and generalization of score matching," in *Proc. 25th Conf. Annu. Conf. Uncertainty Artif. Intell.*, 2009, pp. 359–366.

[32] U. Schmidt, Q. Gao, and S. Roth, "A generative perspective on MRFs in low-level vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1751–1758.

[33] J. Miskin and D. J. C. Mackay, "Ensemble learning for blind image separation and deconvolution," in *Advances in Independent Component Analysis*, M. Girolami, Ed. New York: Springer-Verlag, 2000.

[34] L. He, H. Chen, and L. Carin, "Tree-structured compressive sensing with variational Bayesian analysis," *IEEE Signal Process. Lett.*, vol. 17, no. 3, pp. 233–236, Mar. 2010.

[35] E. Levi, "Using natural image priors-maximizing or sampling?" M.S. thesis, School Comput. Sci. Eng., Hebrew Univ. Jerusalem, Jerusalem, Israel, 2009.

[36] G. Papandreou and A. Yuille, "Gaussian sampling by local perturbations," in *Proc. Adv. Neural Inf. Process. Syst.*, 2010, pp. 1858–1866.

[37] G. Papandreou and A. Yuille, "Perturb-and-map random fields: Using discrete optimization to learn and sample from energy models," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 193–200.

[38] F. Orieux, O. Féron, and J.-F. Giovannelli, "Efficient sampling of high-dimensional Gaussian fields: The non-stationary/non-sparse case," *CoRR*, vol. abs/1105.5887, 2011.

[39] U. Schmidt, K. Schelten, and S. Roth, "Bayesian deblurring with integrated noise estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2625–2632.

[40] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2. Jul. 2001, pp. 416–423.

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

Haichao Zhang (S'10) received the B.Eng. degree in computer science from Northwestern Polytechnical University, Xi'an, China, in 2007, where he is currently pursuing the Ph.D. degree with the School of Computer Science.

His current research interests include image processing, computer vision, machine learning, sparse representation and its applications in pattern recognition, and signal processing.

Dr. Zhang was a recipient of the Best Student Paper Award from the International Conference on Computer Vision in 2011 and the Honorable Mention Student Paper Award from the International Conference on Information Fusion in 2011.

Yanning Zhang (SM'12) received the B.Eng. degree from the Dalian University of Technology, Dalian, China, and the Ph.D. degree from the School of Marine Engineering, Northwestern Polytechnical University, Xi'an, China, in 1996.

She is currently a Professor and the Executive Dean with the School of Computer Science, Northwestern Polytechnical University. Her current research interests include computer vision and pattern recognition, image and video processing, and intelligent information processing.

Dr. Zhang was the Organization Chair of the Asian Conference on Computer Vision in 2009 and served as the Program Committee Chairs of several international conferences.

Haisen Li received the B.Eng. and M.Eng. degrees from Northwestern Polytechnical University, Xi'an, China, in 2007 and 2011, respectively, where he is currently pursuing the Ph.D. degree with the School of Computer Science.

His current research interests include image processing, sparse representation, and related problems.

Thomas S. Huang (LF'01) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, and the M.S. and D.Sc. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1956 and 1960, respectively.

He was a Faculty Member with the Department of Electrical Engineering, MIT, from 1963 to 1973, and a Faculty Member with the School of Electrical Engineering and the Director of its Laboratory for Information and Signal Processing, Purdue University, West Lafayette, IN, from 1973 to 1980. In 1980, he joined the University of Illinois at Urbana-Champaign, Urbana, where he is currently a W. L. Everitt Distinguished Professor of electrical and computer engineering, a Research Professor with the Coordinated Science Laboratory, and a Technology and Co-Chair of the major research theme Human-Computer Intelligent Interaction, Beckman Institute for Advanced Science, Urbana. He has published 21 books and over 600 papers in network theory, digital filtering, image processing, and computer vision. His current research interests include information technology, especially the transmission and processing of multidimensional signals.

Dr. Huang is a member of the National Academy of Engineering, Academia Sinica, China, and a Foreign Member of the Chinese Academies of Engineering and Sciences. He is a fellow of the International Association of Pattern Recognition and the Optical Society of America. He was a recipient of the Honda Lifetime Achievement Award, the IEEE Jack Kilby Signal Processing Medal, and the K.S. Fu Prize of the International Association for Pattern Recognition.