Gradient Profile Prior and Its Applications in Image Super-Resolution and Enhancement

Jian Sun, *Member, IEEE*, Jian Sun, Zongben Xu, and Heung-Yeung Shum, *Fellow, IEEE*

***Abstract*—In this paper, we propose a novel generic image prior—gradient profile prior, which implies the prior knowledge of natural image gradients. In this prior, the image gradients are represented by gradient profiles, which are 1-D profiles of gradient magnitudes perpendicular to image structures. We model the gradient profiles by a parametric gradient profile model. Using this model, the prior knowledge of the gradient profiles are learned from a large collection of natural images, which are called gradient profile prior. Based on this prior, we propose a gradient field transformation to constrain the gradient fields of the high resolution image and the enhanced image when performing single image super-resolution and sharpness enhancement. With this simple but very effective approach, we are able to produce state-of-the-art results. The reconstructed high resolution images or the enhanced images are sharp while have rare ringing or jaggy artifacts.**

***Index Terms*—Gradient field transformation, gradient profile prior, image enhancement, natural image statistics, super-resolution.**

I. Introduction

THERE HAS been much attention in modeling natural image prior in recent years [1]–[3]. Image prior has been widely investigated in super-resolution [4]–[7], denoising [2], [8], inpainting [9], deblur [10], [11], etc. Due to the ill-posedness of these problems, natural image prior is crucial to derive the reasonable results consistent with the natural image statistics.

The most traditional category of image prior is the edge smoothness prior, which generally models the smoothness of image edges. The edge smoothness prior has nice explanation in variational framework and partial differential equation (PDE) theory. This category of prior generally regularizes the first or higher order derivatives of image which implies the smoothness of image edges. Then a PDE is derived to minimize the related energy functional in the variational framework [12]–[14]. In recent years, sparsity deduced priors [15]–[17] were widely investigated, in which the image or signal is assumed to be represented by the sparse linear combination of over-complete dictionary of basis, and successfully applied to the problems of super-resolution, denoising and etc.

Another category of image prior is the generic image prior, which generally learns the non-Gaussian marginal distributions of the filter responses of natural images. Laplacian distribution [18], generalized Gaussian distribution (GGD) [18]–[20] and Gaussian scale mixture distribution [3] are proposed to model the image statistics. Steerable random fields model [8] improves the performance of image prior by considering the local image structures, in which the distribution of image gradients aligned with and orthogonal to local image structures are learned. In [2], [1], [21], filter bank and the distribution of filter responses over the filter bank are jointly learned in the framework of Markov random field (MRF) model. Especially, fields of experts (FoE) model [2] learns the translation invariant image statistics from the training set of image overlapping patches. This model has achieved excellent results for low level vision problems.

However, these pervious generic image priors only concern the marginal distribution of image filter responses (e.g., the gradients) over the whole image, and their spatial layouts are discarded [18], [3] or weakly incorporated [2], [8]. In this work, we will focus on the investigation of the spatial layouts of natural image gradients. We describe the spatial layout of image gradients by gradient profiles, in which each gradient profile is defined as the 1-D profile of gradient magnitudes along the gradient direction passing through an edge pixel. We parametrically model the gradient profile by GGD function, which is called gradient profile model. It describes the shape and sharpness of gradient profiles in natural image. Based on this model, we learn the regularities of gradient profiles from a large training set of natural images. We observe that, firstly, the shape statistics of the gradient profiles is stable to the image resolution. Second, the distribution of sharpness in natural images can be well modeled by γ-distribution. Thirdly, the sharpness of gradient profiles in high resolution (HR) image is statistically dependent on the sharpness of the corresponding gradient profiles in low resolution (LR) image. These prior knowledge of natural image gradient profiles are called gradient profile prior in this paper.

Based on the gradient profile prior, gradient field transformation is proposed to transform the observed gradient field to the target gradient field through mapping the shape and sharpness of the observed gradient profiles. The transformed gradient field provides an effective gradient domain constraint on the target image. As the shape parameter is stable to resolution, the remaining problem is to investigate the sharpness dependency between the observed image and the target image for specific application.

Manuscript received January 27, 2010; revised June 01, 2010, September 16, 2010; accepted November 10, 2010. Date of publication November 29, 2010; date of current version May 18, 2011. This work was supported by the National Natural Science Foundation of China under Grant 61003144 and National Key Fundamental Research Program of China under Grant 2007CB311002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xiaolin Wu.

J. Sun (first author) and Z. Xu are with the School of Science, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: jiansun@mail.xjtu.edu.cn; zbxu@mail.xjtu.edu.cn).

J. Sun (second author) and H.-Y. Shum are with Microsoft Research Asia, Beijing, 100084, China (e-mail: jiansun@microsoft.com; hshum@microsoft.com).

Digital Object Identifier 10.1109/TIP.2010.2095871

We apply the gradient profile prior to the problems of single image super-resolution and enhancement. Single image super-resolution [4]–[7], [22] is to estimate a sharp HR image with minimal artifacts (e.g., jaggy or ringing artifacts) from a single LR image. To design a good single image super-resolution algorithm, the essential issue is how to apply a good image prior or constraint on the HR image due to the ill-posedness of super-resolution. Using the relationship of sharpness between HR image and LR image, we infer the HR gradient field from the LR image by gradient field transformation. Then the estimated HR gradient field imposes a gradient domain constraint for the HR image. By coupling the gradient domain constraint and the image reconstruction constraint, the inferred HR image is sharp and rarely has jaggy or ringing artifacts along high frequency structures.

For the application of image enhancement, we apply the gradient field transformation to estimate the sharpened gradient field, in which the sharpness mapping function from the blurry image to the sharpened image should be given. Then the sharpened image can be recovered from the transformed gradient field. We design two methods to derive the sharpness mapping function. First, a parametric sharpness mapping function is designed for enhancing the sharpness of blurry image, in which a single parameter is free to be tuned by the user. Second, sharpness transfer mapping function is designed by transferring the sharpness distribution of a given sharp image or natural images to the blurry image.

Compared with the previous generic image prior, the gradient profile prior has the following advantages: 1) unlike generic smoothness prior and edge smoothness prior, the gradient profile prior forces the image gradient field close to the transformed gradient field rather than preferring the gradient field with lower magnitude in the smoothness prior. Therefore both small scale and large scale details can be well recovered in the enhanced image; 2) the common artifacts in super-resolution or enhancement, such as ringing artifacts, can be well avoided by working in the gradient domain. A successful application of generic image prior is the work on camera shake removal proposed by Fergus *et al.* [10]. In that work, the marginal distribution of image gradients fitted by Gaussian mixture model is utilized to regularize the motion-deblurred image and produces the state-of-the-art results. The key difference with our proposed gradient profile prior is that, the marginal distribution in [10] concentrates on the heavy-tailed marginal distribution of gradients over the whole image and discards the spatial layout of gradients.

This paper is organized as follows. In Section II, we present the gradient profile model and the gradient profile prior learned from natural images. In Section III, we propose the gradient field transformation to enhance gradient field. In Section IV and V, we apply the gradient field transformation to single image super-resolution and enhancement respectively. We discuss and conclude this work in Section VI.

A. Related Work of Single Image Super-Resolution

There have been much work on single image super-resolution in recent years. They are classified into three categories of approach: interpolation based approach, reconstruction based approach, and learning based approach. The interpolation based approach [23]–[25] estimate the high-resolution image by interpolating the unknown pixels based on the surrounding known pixels. Recently, sophisticated interpolation methods are proposed using the sparse mixing estimation [15] and adaptive 2-D autoregressive modeling [26]. The interpolation based approach tends to blur the high frequency details if the low resolution image is generated with anti-aliasing operation on the high resolution image. The reconstruction based approach [27]–[30] enforces a reconstruction constraint which requires that the smoothed and down-sampled version of the HR image should be close to the LR image. Back-projection [27] is a typical reconstruction based method. It introduces jaggy or ringing artifacts around edges because no regularization is imposed. Image prior is necessary to regularize the reconstruction constraint to reduce these artifacts. The learning based approach [31], [32], [5], [33], [4], [28], [34]–[36], [17] "hallucinates" high frequency details from a training set of HR/LR image pairs. Primal sketches (e.g., edges, ridges and corners) are hallucinated in [5] because human eye is more sensitive to these features for recognition. The learning based approach highly relies on the similarity between the training set and the test set. It is still unclear how many training examples are sufficient for the generic images. Recently, Fattal [34] proposed an edge statistics (EFCM) of image gradients, which is learned from HR and LR image pairs, for image upsampling. EFCM is the statistics of the local continuity measures in the HR image conditional on the 3-D edge features in the LR image. The edge statistics are imposed on the HR image by constraining the local continuity measures searched from a learned 5-D EFCM table given the 3-D edge features of LR image. It has achieved convincingly excellent results for single image super-resolution. We will explicitly compare with the Fattal's work in the Section VI-A.

B. Related Work of Image Enhancement

The most popular approach for image enhancement is the unsharp masking (UM) technique [37]. In this approach, the observed image is firstly blurred by a low-pass filter, then subtract the blurred image from the observed image, and add a fraction of the difference back to the observed image. The unsharp masking approach can well recover image sharpness, however it tends to introduce ringing artifacts around image structures.

Another category of edge enhancement approaches is shock filter [38], [39]. It enhances image edges by solving an inverse partial differential equation. The shock filter generally enhances the image edges detected by the edge detector which is implicitly defined in the equation. Osher and Rudin [40] proposed to sharpen a blurry image I_0 by $(\partial I(\cdot,t))/(\partial t) = -\mathrm{sign}(\triangle I)|\nabla I|$, where $I(\cdot,0) = I_0(\cdot)$, and $\mathrm{sign}(\triangle I)$ can be considered as an edge detector. The pixels satisfying $\triangle I = 0$ are edge pixels. This equation governs a dilation process when $\triangle I > 0$ and an erosion process when $\triangle I < 0$. As the equation evolves, the enhanced image tends to be piecewise constant due to the morphological operations.

More types of shock filters are designed to enhance image edges detected by differently defined edge detectors [41], [39], [42]. The shock filters are capable of producing highly sharp enhancement results, however it tends to sharpen all the detected edges to be sharp discontinuities independent of their original sharpness. Our proposed edge enhancement approach is able to adaptively sharpen the image edges based on their original sharpness. Therefore the enhancement results are more reasonable to human perception.

II. Gradient Profile Prior

In this section, we will characterize the image gradient field by the gradient profiles centered at the zero-crossing pixels in the gradient domain. The regularities of gradient profiles in natural image, which is called gradient profile prior, will be learned from a large training set of natural images.

A. Gradient Profile

Denote the image gradient as $\nabla I = (\partial_x I, \partial_y I) = m \cdot \overrightarrow{N}$, where $m = \sqrt{(\partial_x I)^2 + (\partial_y I)^2}$ is the gradient magnitude and $\overrightarrow{N} = \arctan((\partial_y I)/(\partial_x I))$ is the gradient direction. In the gradient field, we denote the zero-crossing pixel, which is the local maximum on its gradient direction, as an edge pixel. To obtain the edge pixels, we firstly convolve the image by discrete gradient operator $[-(1/2), 0, (1/2)]$ and $[-(1/2), 0, (1/2)]^T$ to obtain $\partial_x I$ and $\partial_y I$ respectively. Then edge pixels are detected by nonmaximum suppression.

Fig. 1(a) is two image blocks containing two edges with different sharpness. Fig. 1(b) is corresponding gradient magnitude (normalized and inverted) maps. As shown in Fig. 1(b), starting from edge pixel x_0, we trace a path $p(x_0)$ along the gradient directions (two-sides) pixel by pixel until the gradient magnitude does not decrease anymore. We call the curve of gradient magnitudes along the 1-D path $p(x_0)$ as gradient profile. Fig. 1(c) is two illustrated gradient profiles. For subpixel accuracy when tracing the gradient profile, linear interpolation is used to estimate the gradient of the interpolated pixels with noninteger spatial coordinates as shown in Fig. 1(d).

To analyze the gradient profile, we first consider an image with a straight step edge blurred by a blur kernel $K(x, y)$. Without loss of generality, we assume the step edge is aligned with the y-axis. Then we can prove the following theorem.

Theorem 1: Given an ideal image $I(x, y)$ with a straight step edge aligned with y-axis, and the image is blurred by a blur kernel $K(x, y)$. Then the normalized gradient profile passing through the edge is $\int K(x, y)\mathrm{d}y$, which is the integral of the blur kernel along the edge direction (i.e., y-axis).

Proof: The image can be formulated as $I(x, y) = K(x, y) * [(u - v)H(x) + v]$, where u and v are gray-scale values on the right and left sides of the edge, $H(x)$ is the Heaviside function which equals to 1 when $x \geq 0$ and equals to 0 when $x < 0$, and $K(x, y)$ is the blur kernel. Then the gradient of image $I(x, y)$ is

$$\begin{aligned}\nabla I(x, y) &= K(x, y) * \nabla[(u - v)H(x) + v] \\ &= K(x, y) * \{\partial_x[(u - v)H(x) + v], 0\} \\ &= \{(u - v)K(x, y) * \delta(x), 0\} \\ &= \left\{(u - v)\int K(x, y)dy, 0\right\}\end{aligned}$$

and the gradient magnitude is $|u - v| \int K(x, y)dy$. Obviously, the gradient profile is a 1-D straight line orthogonal to the edge direction (i.e., y-axis) in this case. Therefore, the normalized gradient profile is the integral of blur kernel along the edge direction, i.e., $\int K(x, y)dy$.

This theorem tells us that the gradient profiles of a straight step edge blurred by a kernel $K(x, y)$ are determined by the blur kernel itself. If the kernel $K(x, y)$ is a Gaussian kernel $G(x, y; \sigma)$ with bandwidth σ, then the gradient profiles will be 1-D Gaussian kernel $G(x; \sigma)$. However, the edges in natural images are complicated (e.g., step edge, ridge edge, corner edge, etc.), which are not necessarily the straight and step edges. And the blurs are commonly caused by out-of-focus, motions, shadows, etc., which are hard to be explicitly modeled. Therefore it is unrealistic to theoretically model the gradient profiles of natural image covering all these factors. We will investigate the overall regularities of gradient profiles through learning from the natural images.

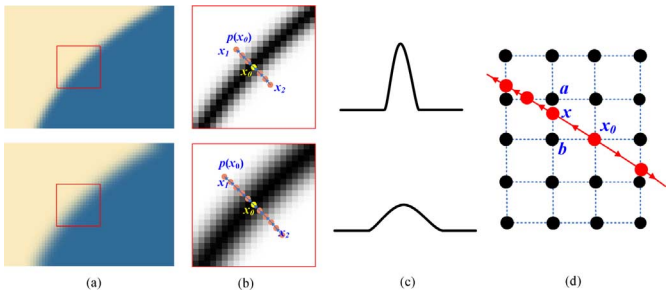

Fig. 1. Gradient profile. (a) Two edges with different sharpness. (b) Gradient maps (normalized and inverted magnitudes) of two rectangular regions in (a). $p(x_0)$ is a 1-D path passing through the edge pixel x_0, by tracing along gradient directions (two sides) pixel by pixel until the gradient magnitude does not decrease at x_1 and x_2. (c) Gradient profiles are curves of gradient magnitudes along the $p(x_0)$. (d) Subpixel technique is used to trace the curve of gradient profile. The gradient of pixel x is interpolated by its two nearest pixels (a and b) on the image grid.

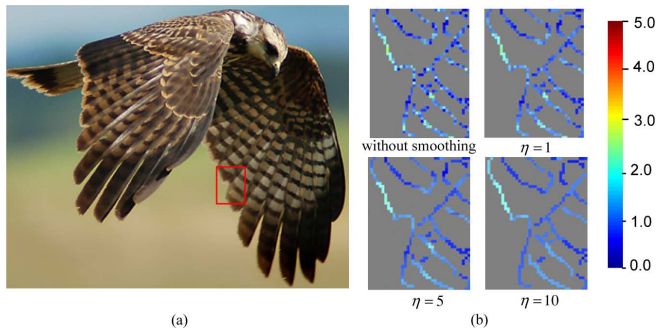

Fig. 2. Estimate the sharpness of gradient profiles. The sharpness of gradient profiles are shown by colors at the edge pixels indexed by the color bars in the rightmost part of the figure. The four sharpness maps in (b) show the sharpness of gradient profiles without global optimization, and with global optimization using $\eta = 1, 5, 10$. The sharpness map with $\eta = 5$ is smooth along edges while preserving the sharpness discontinuity when the gradients change significantly. (a) Image. (b) Sharpness estimation for edges in rectangle region of (a).

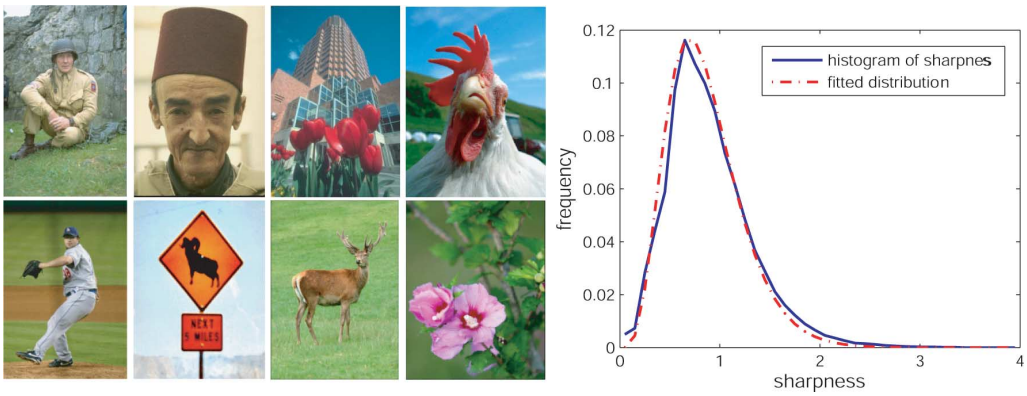

Fig. 3. Profile sharpness distribution of natural images. The histogram of profile sharpness in natural images is shown in the right subfigure. It is shown that the sharpness of natural images are dominantly distributed around 0.70, and heavy-tailed distributed when the sharpness is increasing. The γ-distribution with parameters $\theta = 0.16$ and $s = 5.35$ can well fit the sharpness distribution.

B. Gradient Profile Sharpness

Now we measure the sharpness of gradient profile. As shown in Fig. 1(b), (c), the gradient profile in blurry image has wider nonzero spatial scattering than the gradient profile in sharp image. We measure the sharpness of gradient profile by its spatial scattering which can be well modeled by the square root of the variance (second moment) of gradient profile

$$\sigma(p(x_0)) = \sqrt{\sum_{x \in p(x_0)} \frac{m(x)}{M(x_0)} {d_c}^2(x, x_0)} \qquad (1)$$

where $M(x_0) = \sum_{s \in p(x_0)} m(s)$ and $d_c(x, x_0)$ is the curve length of the gradient profile between x and x_0. Obviously, the sharper image gradient profile, the smaller the square root of variance σ is. We call this measure as the profile sharpness.

Profile Sharpness Estimation: Individually estimating the sharpness for each gradient profile is not robust. To have a better estimation, we apply a global optimization to enforce the sharpness consistency of neighboring gradient profiles as follows.

First, we construct a graph over all the gradient profiles, and the neighboring gradient profiles of a profile $p(x_i)$ passing through edge pixel x_i are defined as the profiles with edge pixels within a predefined distance (five pixels in this paper) to x_i. Then we minimize the following energy to estimate the sharpness σ_i for each profile $p(x_i)$ with global optimization

$$\begin{aligned} \{\sigma_i^*\} &= \text{argmin}_{\{\sigma_i\}} E(\{\sigma_i\}) \\ &= \sum_i \left[(\sigma_i - \hat{\sigma}_i)^2 + \eta \sum_{j \in N(i)} w_{i,j} \cdot (\sigma_i - \sigma_j)^2 \right] \end{aligned} \qquad (2)$$

where $\hat{\sigma}_i$ is individually estimated using (1) for profile $p(x_i)$. The edge weight $w_{i,j}$ for neighboring profiles $p(x_i)$ and $p(x_j)$ is defined as $w_{i,j} = \exp(-\zeta_1 \cdot |\nabla I(x_i) - \nabla I(x_j)|^2 - \zeta_2 \cdot d_e(x_i, x_j)^2)$, in which $d_e(\cdot, \cdot)$ implies the Euclidean distance. In our implementation, ζ_1 and ζ_2 are set to 0.16 and 0.08, which impose larger importance on the gradient similarity. By the definition of weight, the gradient profiles with similar spatial position and gradients on their edge pixels are enforced to have similar sharpness. The parameter η controls the strength of sharpness smoothness, which is set to 5 to achieve smooth profile sharpness while preserving the sharpness discontinuity when the gradient profiles change significantly. Please refer to Fig. 2 for the effect of parameter η. The energy function can be effectively minimized because it is a Gaussian MRF model.

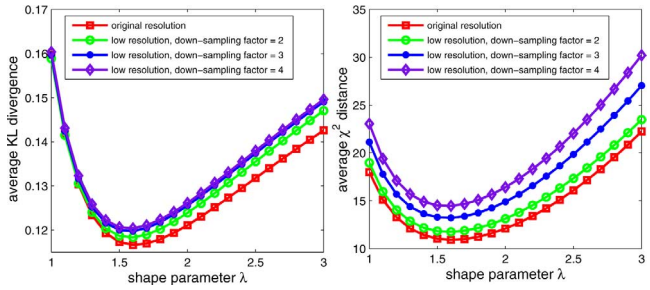

Fig. 4. Average KL divergences and χ^2 distances between the fitted gradient profiles and 1 million gradient profiles by varying the shape parameter λ. The optimal λ is near 1.6 on four data sets with different resolutions.

Profile Sharpness Distribution of Natural Images: We now investigate the profile sharpness distribution of natural images. Generally, the natural images are captured by focusing on the foreground layer in which the objects are sharp and textured, while the objects in the out-of-focus layers are blurry. Therefore, the sharpness values of gradient profiles in natural images should be dominantly distributed around a smaller value which implies higher sharpness in appearance. Fig. 3 illustrates the profile sharpness distribution of natural images in the left subfigure of Fig. 3. It is shown that the sharpness of gradient profiles are dominantly distributed around 0.7, and heavy-tailed distributed with the increase of sharpness value. We also find that the sharpness distribution of natural images can be well fitted by γ- distribution with density function $g_\gamma(\sigma) = \sigma^{s-1}(\exp(-\sigma/\theta))/(\Gamma(k)\theta^s)$, in which σ is the profile sharpness and always nonnegative, $s > 0$ and $\theta > 0$ are parameters of this distribution. For the empirical distribution

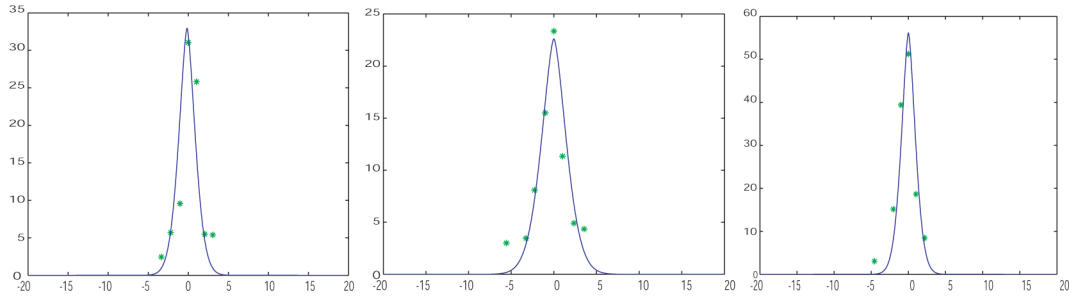

Fig. 5. Examples of fitted gradient profiles using GGD with shape parameter 1.6. X-axis denotes the distance or minus distance of pixels on the two sides of gradient profile to the edge pixel. Y-axis denotes the gradient magnitude. The scattered points are the observed gradient magnitudes of pixels on the gradient profiles, and the continuous curves denote the fitted gradient profiles.

shown in Fig. 3, the γ-distribution with parameters $s = 5.35$ and $\theta = 0.16$ well fits this sharpness distribution.

C. Gradient Profile Model

Next, we investigate the regularities of gradient profiles in natural images. We parametrically describe the gradient profiles by the density function of a generalized exponential distribution, i.e., GGD [43], which is defined as

$$g(x;\sigma,\lambda) = \frac{\lambda\alpha(\lambda)}{2\sigma\Gamma\left(\frac{1}{\lambda}\right)} \exp\left\{-\left[\alpha(\lambda)\left|\frac{x}{\sigma}\right|\right]^{\lambda}\right\} \tag{3}$$

where $\Gamma(\cdot)$ is gamma function and $\alpha(\lambda) = \sqrt{\Gamma(3/\lambda)/\Gamma(1/\lambda)}$ is the scaling factor which makes the second moment of GGD equal to σ^2. Therefore, σ can be directly estimated as the second moment of the profile, i.e., the profile sharpness. λ is the shape parameter which controls the overall shape of the distribution. The distribution function $g(x;\sigma,\lambda)$ is Gaussian distribution function if $\lambda = 2$, and a Laplacian distribution function if $\lambda = 1$. This parametric model of (6) for image gradient profiles is called gradient profile model in this paper.

GGD has been widely used to model the statistical distribution of filter responses of natural images [19], [20]. Different from these work, the density function of GGD is utilized to fit the curves of gradient profiles in natural images. To fit the gradient profiles in natural images, we collect an image set containing 1000 natural images downloaded from professional photography forums. All images are in the original resolution without down-sampling or up-sampling. For each image, we randomly select 1000 gradient profiles to construct a data set Ω_1 which consists of 1 million gradient profiles. We also construct other three profile data sets Ω_2, Ω_3 and Ω_4 from the down-sampled versions of the original resolution images with the down-sampling factors of 2, 3, and 4.

We use Kullback-Leibler (KL) divergence and χ^2 distance to measure the fitting error. The optimal λ^* is obtained by minimizing the average fitting error over the training set of gradient profiles Ω (e.g., Ω_i, $i = 1, 2, 3$, or 4)

$$\lambda^* = \text{argmin}_{\lambda}\left\{\frac{1}{|\Omega|}\sum_{p\in\Omega}\text{Err}(p, g(\cdot;\sigma_p,\lambda))\right\} \tag{4}$$

where σ_p is the sharpness of one gradient profile p in the set Ω, which is estimated in Section II-B. Using Kullback-Leibler (KL) divergence as the error measurement, the fitting error between gradient profile p and the gradient profile model g is defined as

$$\text{KL}(g(\cdot;\sigma_p,\lambda), p) = \sum_{x\in p}\hat{g}(x;\sigma_p,\lambda)\log\frac{\hat{g}(x;\sigma_p,\lambda)}{\hat{m}(x)} \tag{5}$$

where $\hat{g}(x;\sigma_p,\lambda) = (g(x;\sigma_p,\lambda))/(\sum_{s\in p} g(s;\sigma_p,\lambda))$, $\hat{m}(x) = (m(x))/(\sum_{s\in p} m(s))$ are the normalized gradient profile model and gradient profile, x and s denote the pixels on the gradient profile p, and $m(x)$ is the gradient magnitude of pixel x. Using χ^2 distance as the error measurement, the fitting error between the gradient profile p and the gradient profile model g is defined as

$$\chi^2(g(\cdot;\sigma_p,\lambda), p) = \sum_{x\in p}\frac{[m(x) - E(x)]^2}{E(x)} \tag{6}$$

where $E(x) = (g(x;\sigma_p,\lambda))/(\sum_{s\in p} g(s;\sigma_p,\lambda)) \cdot \sum_{s\in p} m(s)$.

Due to the high nonlinearity of (4) combining with (5) or (6), we resort to the direct search of λ minimizing the average fitting error over the training set. We compute the average KL divergences and χ^2 distances on four profile sets Ω_1, Ω_2, Ω_3, and Ω_4 by varying the shape parameter λ, as shown in Fig. 4. We observe that the optimal shape parameter is about 1.6 for all down-sampling factors and two types of error measurements. The shape parameter λ is stable across different resolutions, which means that the shape parameter of gradient profile model is resolution independent in natural images.

Fig. 5 illustrates three examples of observed gradient profiles fitted by GGD with shape parameter 1.6. To verify whether the parameter $\lambda = 1.6$ is independent on our collected data or not, we repeat the above experiments on two different image sources. One is 500 images randomly downloaded from Flickr image site. The other is 500 images from a home photo gallery taken with four different digital cameras. Again, the obtained optimal shape parameters are stable and between 1.55 and 1.65, which means the GGD function with $\lambda = 1.6$ is a good generic model for the natural image gradient profiles and independent on the image resolution. Based on this nice property, we only need to study the relationship of the profile sharpness σ for specific application.

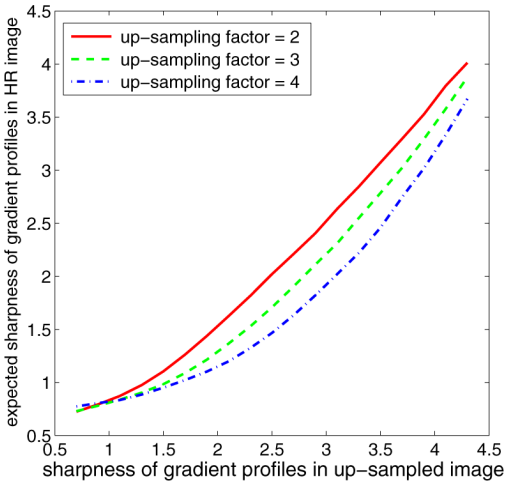

Fig. 6. Expected sharpness of the gradient profiles in HR image with respect to sharpness of the corresponding profiles in up-sampled image.

D. Relationship of Profile Sharpness Between HR Image and LR Image

We now investigate the relationship of profile sharpness between HR image and LR image, which is important for the application of image super-resolution. Similar to the previous methods [31], [5], [34], we study the relationship of gradient profile sharpness between the up-sampled image I_l^u and the HR image I_h, in order to avoid the shifting problem of the zero-crossing pixels in scale space [44]. In our implementation, the up-sampled image I_l^u is the bicubic interpolation[1] of the LR image I_l.

For each gradient profile in the up-sampled image I_l^u, we extract its corresponding gradient profile in the HR image I_h. Because the edge pixels are not exactly aligned in two images, we find the best matched edge pixels by measuring the distance and direction. For each edge pixel x_l in I_l^u, the best matched edge pixel x_h in I_h is found by $x_h = \text{argmin}_{x \in \mathcal{N}(x_l)}\{\|x - x_l\| + 2\|\overrightarrow{N}(x) - \overrightarrow{N}(x_l)\|\}$ where $\mathcal{N}(x_l)$ is the 5×5 neighbors of x_l in the HR image.

To compute the statistical dependency of profile sharpness between HR image and LR image, we quantize the sharpness σ into a number of bins. The width of bin is 0.1. For all LR gradient profiles whose sharpness value falls in the same bin, we calculate the expectation of sharpness of the corresponding HR gradient profiles. Fig. 6 shows three fitted curves of computed expectations for the up-sampling factors of 2, 3, and 4. X-axis is the sharpness of the (up-sampled) LR gradient profile and Y-axis is the expected sharpness of the corresponding HR gradient profile.

There are two basic observations from Fig. 6: 1) the HR gradient profile is sharper than the LR gradient profile because the bicubic interpolation blurs the profile; 2) the higher the up-sampling factor, the larger the sharpness difference between the HR gradient profile and the LR gradient profile is. Notice that three curves converge together when the sharpness is below 1.0 in Fig. 6. One possible reason is due to the inaccuracy of our sharpness estimation. The sharpness estimation for the small scale edge is sensitive to the noise. Also, the introduced image aliasing in the LR image by down-sampling may result in over-estimated sharpness.

[1]Note that the statistic of shape parameter λ in the up-sampled image may be slightly influenced by the bicubic interpolation. However, we found that the optimal λ value for the up-sampled image is still stable. They are 1.63, 1.68, and 1.69 for the up-sampling factors of 2, 3, and 4 on our data sets.

E. Summary of Gradient Profile Prior

In summary, we have learned the following prior knowledge of the gradient fields of natural images based on the gradient profile model : 1) the shape parameter of the gradient profiles in the natural image is close to the value 1.6; 2) the sharpness distribution of the natural image can be well fitted by a γ-distribution with learned parameters; 3) the sharpness relationship of gradient profiles between two resolutions follows the statistical dependency learned in Section II-D. These prior knowledge of the gradient field of natural image are called gradient profile prior in this paper.

III. Gradient Field Transformation

Based on gradient profile model and gradient profile prior, we propose the gradient field transformation to estimate the gradient field of the target image in image super-resolution and enhancement. The transformed gradient field is derived by transforming the gradient profiles of the observed image. We also propose the approach to reconstruct the image from the transformed gradient field. Please refer to the electronic versions of the color figures in the following sections for better illustration.

A. Gradient Field Transformation

First, we study how to transform a gradient profile p_s with parameters $\{\lambda_s, \sigma_s\}$ in the observed gradient field to the gradient profile p_t with parameters $\{\lambda_t, \sigma_t\}$ in the target gradient field. Obviously, the gradient profile p_t can be derived through multiplying p_s by the ratio between p_t and p_s. Based on the gradient profile model in (3), the transform ratio between gradient profiles p_t and p_s is

$$\begin{aligned} r(x) &= \frac{g(x;\sigma_t,\lambda_t)}{g(x;\sigma_s,\lambda_s)} \\ &= c \cdot \exp\left\{ -\left(\frac{\alpha(\lambda_t)\cdot|d_c(x;x_0)|}{\sigma_t}\right)^{\lambda_t} \right. \\ &\quad \left. + \left(\frac{\alpha(\lambda_s)\cdot|d_c(x;x_0)|}{\sigma_s}\right)^{\lambda_s} \right\} \end{aligned} \tag{7}$$

where $c = (\lambda_t \cdot \alpha(\lambda_t) \cdot \sigma_s \cdot \Gamma(1/\lambda_s))/(\lambda_s \cdot \alpha(\lambda_s) \cdot \sigma_t \cdot \Gamma(1/\lambda_t))$, x and x_0 are a pixel and the edge pixel of the gradient profile (p_s or p_t), and $d_c(x;x_0)$ is the curve distance between x and x_0 along the gradient profile. Thus, the gradient profile p_t is estimated by multiplying the gradient profile p_s by the transform ratio r.

Generally, the parameters $\{\lambda_s, \sigma_s\}$ of the observed gradient profile can be computed from the observed image. However, the setting of parameters $\{\lambda_t, \sigma_t\}$ is nontrival, which should be based on the prior knowledge of the gradient profiles in the target image. The setting of parameters $\{\lambda_t, \sigma_t\}$ will be investigated for specific application in the following sections. Fig. 7

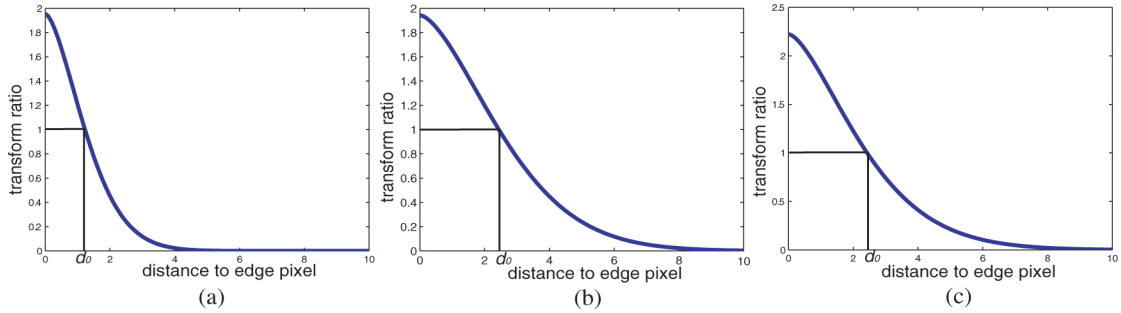

Fig. 7. Illustration of transform ratio in (7). (a) $\lambda_t = 1.6, \sigma_t = 1.0, \lambda_s = 1.6, \sigma_s = 2.0$. (b)$\lambda_t = 1.6, \sigma_t = 2.0, \lambda_s = 1.6, \sigma_s = 4.0$. (c)$\lambda_t = 1.6, \sigma_t = 2.0, \lambda_s = 2.0, \sigma_s = 4.0$. When the distance to edge pixel is smaller than d_0, the transform ratio is larger than 1, and otherwise it is smaller that 1. Multiplying gradient profile by the transform ratio, the gradient magnitude of pixel on the gradient profile within distance d_0 to the edge pixel will be increased, and otherwise it will be decreased.

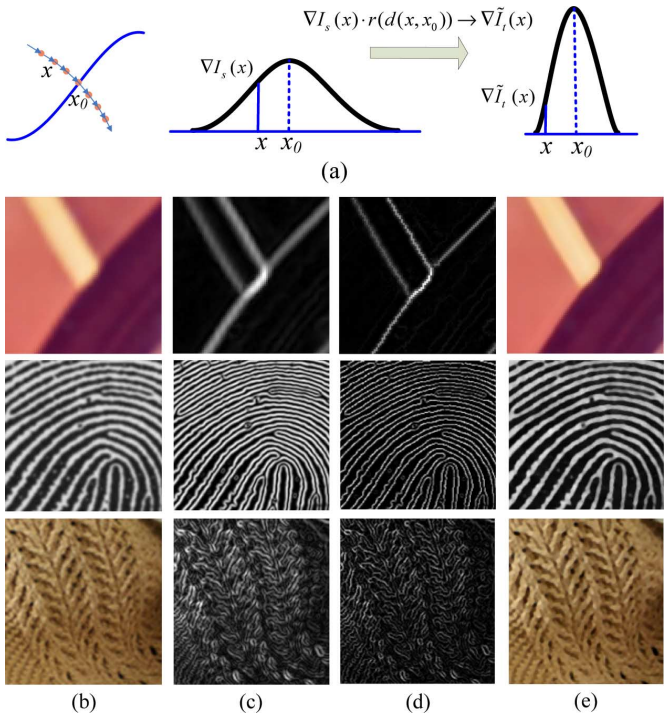

Fig. 8. Gradient field transformation (please refer to the electronic version for better illustration). (a) Left and middle subfigures illustrate a gradient profile passing through x and x_0. The gradient of x is transformed to its enhanced version (right) by multiplying a transform ratio r. (b) Observed images. (c) Gradient maps of images in (b). (d) Transformed gradient fields. (e) Reconstructed images by solving poisson equation. The gradient field transformation can well transform the step edges and edges around corners in the second row, ridge edges in the third row, and highly textured edges in the fourth row.

shows three examples of transform ratios with different parameter settings. Multiplying a gradient profile by the transform ratio, the gradient magnitudes of the pixels on the gradient profile within distance d_0 to the edge pixel will be increased, and otherwise will be decreased. In this way, the observed gradient profile is enhanced to be a sharper profile with smaller spatial scattering. Fig. 8(a) presents an illustrative example to show how the gradient profile is enhanced by multiplying the transform ratio.

Second, using the transform ratio defined in (7), we estimate the target gradient field by transforming the gradient profiles in the observed gradient field ∇I_s. Then the target gradient field ∇I_t is estimated by

$$\tilde{\nabla I_t}(x) = r(x) \cdot \nabla I_s(x) \tag{8}$$

where x is any pixel with nonzero gradient magnitude, and $r(x)$ is computed with respect to the gradient profile passing through x. In our implementation, to find the gradient profile passing through x, we trace from x along the direction (gradient direction or minus gradient direction) with increasing gradient magnitude until reach an edge pixel x_0 (in a threshold distance, e.g., 1 pixel), then adjust the gradient of x by (8).

Gradient Field Transformation for Different Types of Edges: Natural images are composed of complicated edges, e.g., step edge, ridge edge and edge around corners or in highly textured region. In these cases, multiple gradient profiles may intersect at a certain pixel due to the sudden change of edge direction and the discrete image grid. To avoid the multiple enhancements for a single pixel at which multiple profiles intersect, we perform gradient field transformation for each pixel with nonzero gradient rather than each gradient profile. To enhance the gradient of each pixel by (8), the gradient profile passing through this pixel is traced using subpixel interpolation along the gradient direction [similar to Fig. 1(b)]. Therefore, a dense enhanced gradient field over all the nonzero gradient pixels is obtained by gradient profile transform. Fig. 8 presents three real examples with complicated step edge, ridge edge and edge around corners. Fig. 8(d) shows the enhanced gradient fields transformed from Fig. 8(c). As we can observe, the gradients of pixels along image structures are well enhanced by the gradient field transformation.

B. Image Reconstruction From the Transformed Gradient Field

Given the transformed gradient field, we now investigate how to reconstruct image from the transformed gradient field [45]. Assume I is the image to be reconstructed, the objective of image reconstruction is to constrain the gradient field of I as close to the transformed gradient field $\tilde{\nabla I_t}(x)$ as possible, i.e., to minimize the energy function $E(I; \tilde{\nabla I_t}(x)) = \int |\nabla I - \tilde{\nabla I_t}(x)|^2$. This energy can be minimized by solving the poisson equation $\mathrm{div}(\nabla I) = \mathrm{div}(\tilde{\nabla I_t})$. This equation can be discretized and solved efficiently by

Fig. 9. Effect of parameter β on the super-resolution results. The first image shows the original image, and the other images are the down-sampled and then up-sampled (3X) images with $\beta = 0, 0.01, 0.05, 0.10$. The value under each image is the Structural SIMilarity (SSIM) [47] between the super-resolution result and the original image.

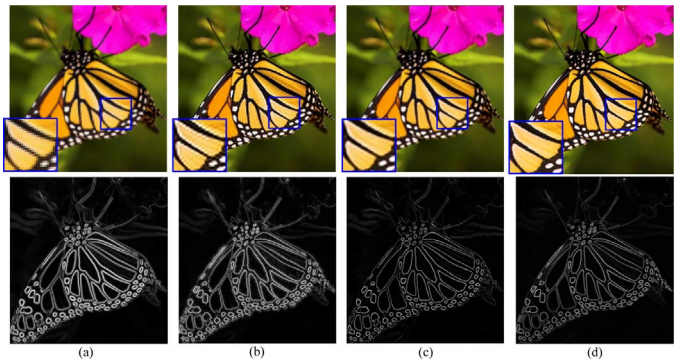

Fig. 10. HR image reconstruction (3X). (a) LR image (nearest neighbor interpolation) and gradient field of its up-sampled image (bicubic interpolation). (b) Result of back-projection and its gradient field. (c) Our result and transformed gradient field for HR image. (d) Ground truth image and its gradient field. Compared with the gradient field of result by back-projection, the transformed gradient field is much closer to the ground truth gradient field of HR image. Our reconstructed result has rare jaggy or ringing artifacts.

Gauss-Seidel iteration with over-relaxation method [46]. Fig. 8(e) shows the reconstructed images by solving poisson equation. The recovered images are sharp and with rare ringing artifacts.

IV. Gradient Profile Prior for Image Super-Resolution

In this section, we apply the gradient profile prior to image super-resolution. Given a LR image, the gradient profile prior learned from natural images provides prior knowledge of the gradient field for the HR image: 1) the shape parameter of the gradient profiles in the HR image is close to the value 1.6; 2) the sharpness relationship of gradient profiles between two resolutions follows the statistical dependency learned in Section II-D. Through performing gradient field transformation, the prior estimation of the HR gradient field will impose a gradient field constraint on the HR image.

A. Super-Resolution Model

Given the LR image I_l, we first transform the gradient field of the up-sampled image I_l^u to estimate the HR gradient field using the gradient field transformation defined in Section III-A. The shape parameters λ_s and λ_t are set to the learned values of the shape parameter in Section II-C. The sharpness σ_s is estimated from the observed gradient profile in image I_l^u and the sharpness σ_t is set as the expected sharpness of gradient profiles in HR image corresponding to σ_s using the relationship we learned in Section II-D.

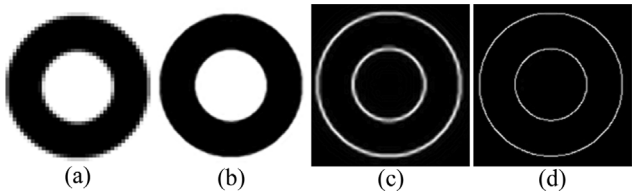

Fig. 11. Super-resolution on synthetic image (4X). (a) LR image (nearest neighbor interpolation). (b) Reconstructed HR image. (c) Gradient field of the up-sampled image (bicubic interpolation). (d) Transformed gradient field from (c).

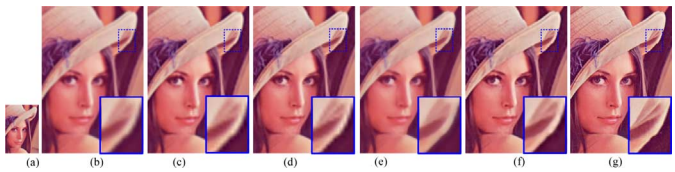

Fig. 12. Super-resolution comparison (3X). Gradient reconstruction is obtained by solving poisson equations on the transformed gradient field. Both of gradient reconstruction result (e) and our result (f) contain much less ringing artifacts, especially along the image edges. But our result (f) is closer to the ground truth by enforcing the reconstruction constraint. See text for details. (a) Input. (b) Bicubic. (c) Sharpened bicubic. (d) Back-projection. (e) Gradient reconstruction. (f) Our result. (g) Ground truth.

Given the LR image I_l, in order to reconstruct the HR image I_h, we minimize the following energy function by enforcing the constraints in both image domain and gradient domain

$$E(I_h \mid I_l, \tilde{\nabla I_h}) = E_i(I_h \mid I_l) + \beta E_g(\nabla I_h \mid \tilde{\nabla I_h}) \quad (9)$$

where $\tilde{\nabla I_h}$ is the estimated HR gradient field, $E_i(I_h|I_l)$ is the reconstruction constraint in the image domain and $E_g(\nabla I_h|\tilde{\nabla I_h})$ is the gradient constraint in the gradient domain.

The LR image is generally modeled as the blurring and then down-sampling of the HR image, and the blur kernel is determined by the point spread function of the camera sensor through analyzing the image generation process in camera system [28]. Then the reconstruction constraint measures the difference between the LR image I_l and the smoothed and down-sampled version of HR image I_h, i.e.,

$$E_i(I_h \mid I_l) = |(I_h * G) \downarrow - I_l|^2 \quad (10)$$

where G is a spatial filter, $*$ is the convolution operator, and $\downarrow$ is the down-sampling operation. In real applications, the convolution kernel G is commonly approximated by the Gaussian function [6], [5], [4], [7]. The kernel standard deviations are empirically set to 0.8, 1.2 and 1.6 for the up-sampling factors of 2, 3, and 4 [7], which increases with respect to the larger up-sampling factor.

The gradient constraint requires that the gradient field of the recovered HR image should be close to the transformed HR gradient field $\tilde{\nabla I_h}$

$$E_g(\nabla I_h \mid \tilde{\nabla I_h}) = |\nabla I_h - \tilde{\nabla I_h}|^2 \quad (11)$$

where ∇I_h is the gradient of I_h. Using this constraint, we encourage the gradient profiles in I_h to have the desired statistics which are learned from natural images.

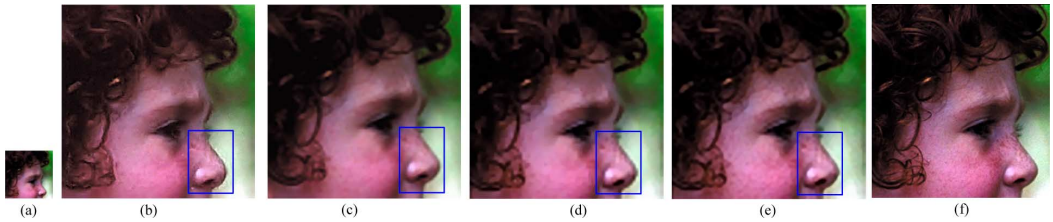

Fig. 13. Super-resolution comparison (4X) of learning-based method [33], alpha channel super-resolution [7], gradient statistical prior using Gaussian mixture model and our approach. Both large scale edges and small scale details (on the face) are recovered in our result. Please refer to the electronic version for better comparison. (a) Input. (b) Learning-based. (c) Alpha channel super resolution. (d) Gradient statistical prior. (e) Our result. (f) Ground truth.

Fig. 14. Super-resolution results with up-sampling factor of 4 for images with highly textured regions. The left images of (a) and (b) are the nearest neighbor interpolated input images, and the right images of (a) and (b) are the results of our method.

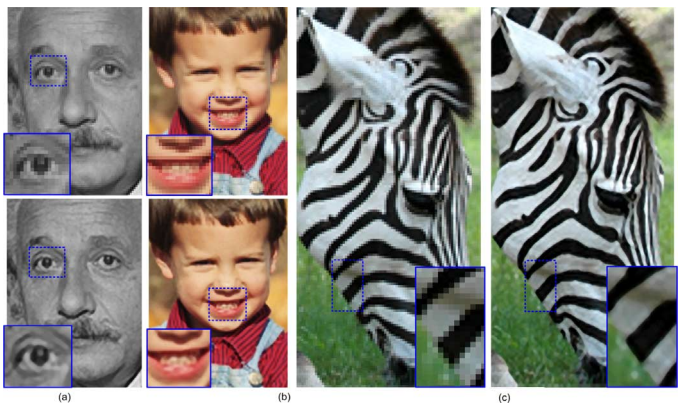

Fig. 15. Super-resolution results with up-sampling factors of 8 and 16. (a) 8X. (b) 8X. (c) 16X.

The energy function of (9) is a quadratic function with respect to I_h, therefore it is convex and the global minimum can be obtained by gradient descent flow

$$\frac{\partial I_h}{\partial t} = -\frac{\partial E(I_h|I_l, \tilde{\nabla I_h})}{\partial I_h}$$

where $(\partial E(I_h|I_l, \tilde{\nabla I_h}))/(\partial I_h) = ((I_h * G) \downarrow -I_l) \uparrow *G - \beta \cdot (\mathrm{div}(\nabla I_h) - \mathrm{div}(\tilde{\nabla I_h}))$, and $\mathrm{div}(\nabla I_h) = (\partial^2 I_h/\partial x^2) + (\partial^2 I_h/\partial y^2)$. The divergence operator (i.e., div) can be implemented using standard finite difference scheme.

In our implementation, we use the following iterative schemes to optimize (9), i.e., $I_h^{t+1} = I_h^t - \tau \cdot (\partial E(I_h)/\partial I_h)$. We set the step size τ to 0.2 and use the up-sampled image I_l^u as the initial value of I_h. In experiments, one hundred iterations are enough to produce sharp and clear high-resolution results.

Fig. 16. More results on images from Berkeley segmentation database. The images are down-sampled and then up-sampled in factor of 3. The quality of results and computational time are listed in Table I. (a) Lady (3X). (b) Starfish (3X).

Parameter β balances the image domain constraint and gradient domain constraint. Larger value of β places larger importance on the gradient domain constraint, which helps to produce sharp edges with little artifacts. On the other hand, smaller value of β places larger importance on the image domain constraint, which produces better image color and contrast, however, may introduce ringing or jaggy artifacts along edges. We set $\beta = 0.05$ by experimental justification for better balance between the artifacts removal and image color and contrast recovery. Fig. 9 presents an example to show the effect of β on the super-resolution result, and $\beta = 0.05$ produces best result.

Fig. 10 gives a real example of our method. Fig. 10(b) are back-projection [27] result using the reconstruction constraint only and its gradient field. Notice the ringing artifacts in both image and gradient field. The bottom image in Fig. 10(c) is our transformed gradient field. It is much closer to the ground truth gradient field shown in Fig. 10(d). The top image in Fig. 10(c) is our final reconstructed HR image using both image and gradient constraints. The ringing artifacts are substantially suppressed by the gradient constraint.

Fig. 11 also shows an example on a synthetic image. Our approach can reconstruct a very sharp HR image guided by the transformed gradient field.

B. More Examples and Evaluation

We test our approach on a variety of images. For the color images, we transform the image from RGB color space to YUV color space, and only perform image super-resolution on the luminance channel. The color channels are up-sampled using the bicubic interpolation.

In Fig. 12, we compare our approach with bicubic interpolation, sharpened bicubic interpolation, back-projection [27], and reconstruction from the transformed gradient field by solving

TABLE I
SUPER-RESOLUTION QUALITY MEASUREMENT AND CPU TIME. THE CPU TIME IN BRACKETS ARE FOR GRADIENT FIELD TRANSFORMATION AND ENERGY OPTIMIZATION RESPECTIVELY

test images	bicubic			back-projection			our method			CPU time (in seconds)
	RMS	ERMS	SSIQ	RMS	ERMS	SSIQ	RMS	ERMS	SSIQ	
Monarch	16.4	26.0	0.883	13.6	21.3	0.890	13.2	20.9	0.902	1.018 (0.809, 0.192)
Lena	8.8	11.5	0.748	8.2	10.8	0.771	7.8	10.1	0.785	0.873 (0.640, 0.169)
Head	8.7	10.9	0.742	8.6	10.6	0.744	8.4	10.3	0.749	0.498 (0.411, 0.085)
Lady	11.3	14.9	0.856	9.7	12.6	0.868	9.5	12.3	0.891	1.052 (0.862, 0.168)
Starfish	12.6	13.3	0.756	11.7	12.2	0.778	11.5	12.0	0.790	0.996 (0.829, 0.164)

poisson equation. The result of bicubic interpolation is over-smooth. The sharpened bicubic interpolation and back-projection introduce ringing or jaggy artifacts, especially along salient edges. The result of reconstruction from the transformed gradient field is sharp and with rare artifacts, but the color is not close to the ground truth HR image. By combing gradient constraint and reconstruction constraint, our final result is the best.

Fig. 13 shows the comparison of our approach with learning-based method [31], alpha channel super-resolution [6] and gradient statistical prior which is modeled as the marginal distribution of gradients using Gaussian mixture model [10]. The minus logarithm of the gradient statistical prior is applied to substitute the gradient constraint in (9) to produce high-resolution result. The result of learning-based method is sharp in appearance. However, high frequency artifacts are also introduced from the training samples, for example the artifacts around the nose. The salient edges in alpha channel super-resolution result are sharp, but the small scale edges, for example flecks on the face, are not well recovered. That's because it's hard to estimate alpha channel value for the edges with weak contrast and large blur. The gradient statistical prior tends to enhance the high contrast edges and smooth out the lower contrast details. The weight of prior term is carefully selected to better preserve details while enhancing salient edges in Fig. 13(d), and the edges around the nose still have ringing and jaggy artifacts. Compared with these results, our approach can recover both sharp edges and small scale details, and introduce minimal additional artifacts.

Fig. 14 shows two super-resolution results of textured images with up-sampling factor of 4. It is shown that the textures are well recovered in the HR images with rare artifacts. Fig. 15 shows three examples with up-sampling factor of 8 and one example with up-sampling factor of 16, in which the HR results are produced by repeatedly running our super-resolution algorithm with up-sampling factor of 2. All of the results show that our method can reliably recover the image details and produce sharp edges with minimal additional artifacts.

We also compute RMS, ERMS [5], and SSIM [47] to qualitatively measure the super-resolution results. These measurements on examples of Monarch (Fig. 10), Lena (Fig. 12), Head (Fig. 13), Lady [Fig. 16(a)] and Starfish [Fig. 16(b)] are listed in Table I. Our model outperforms the bicubic and back-projection with lower reconstruction error and higher visual similarity.

Table I also presents the CPU time (in 2.7 Ghz platform) for Monarch (LR is 133×141 and HR is 399×423), Lena (LR is 125×125 and HR is 375×375), Head (LR is 70×70 and HR is 280×280), Lady and Starfish (LR is 107×160 and HR is 321×480). The computation time of the gradient field transformation for these examples are 0.809 s, 0.640 s, 0.411 s, 0.862 s, 0.829 s and the total time of super-resolution are 1.018 s, 0.873 s, 0.498 s, 1.052 s, 0.996 s. The computation complexity is linearly dependent on the number of pixels in the HR image.

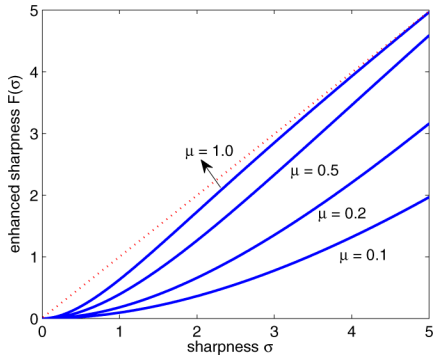

Fig. 17. Sharpness enhancement function defined in (12) with different μ.

V. GRADIENT PROFILE PRIOR FOR IMAGE ENHANCEMENT

Images are commonly degraded by blurring due to a low-pass filtering or out-of-focus caused by the optics of cameras or some image filtering procedures. In this section, we will show how to apply the gradient profile prior to enhance the image sharpness.

The basic idea of sharpness enhancement is to sharpen the gradient profiles of blurry image by gradient field transformation, then recover the sharpened image by solving poisson equation over the sharpened gradient field. Compared with super-resolution, the major difficulty in sharpness enhancement is that no prior knowledge of the sharpness relationship between the blurry image and the enhanced image can be learned due to the diversity of blurry images. We will design a parametric sharpness enhancement function and an automatic sharpness transfer to enhance the blurry image.

A. Sharpness Enhancement Function

We first define sharpness enhancement function F, which maps the sharpness of gradient profile from the blurry image to the enhanced image. A sharpness enhancement function should be defined to adaptively and smoothly enhance the edges with

different sharpness in the blurry image, so F should be continuous and $F(\sigma) < \sigma$, therefore we define

$$F(\sigma) = [1 - \exp(-\mu\sigma)]\sigma \tag{12}$$

where μ is a parameter to control the shape of this function. Obviously, $F(\sigma) < \sigma$, which means that the gradient profiles will be sharpened in the enhanced image. Fig. 17 shows the shape of this function with respect to different μ, and they have the similar shapes as the sharpness dependency learned in Section II-D. The smaller value of μ implies that the edge sharpness is more enhanced. Using this definition, the gradient transform ratio is defined as $r_s(x, x_0) = (g(x; F(\sigma(x)), \lambda^*))/(g(x; \sigma(x), \lambda^*))$, where x_0 and $\sigma(x)$ is the edge pixel and sharpness of the gradient profile passing through x. λ^* is the learned optimal shape parameter in Section II-C. After gradient field transformation guided by the transform ratio, the enhanced images are recovered by solving poisson equation over the transformed gradient field.

Fig. 18 shows the results of image enhancement using (12) with different parameter μ. Please refer to the figure for details. Obviously, the enhancement results in Fig. 18(b), (c) are much sharper than the observed image. The parameter μ is reasonably set between 0.1 and 0.5 in applications. Fig. 19 illustrates an example for enhancement of real captured image due to out-of-focus, and the blurry image is successfully enhanced.

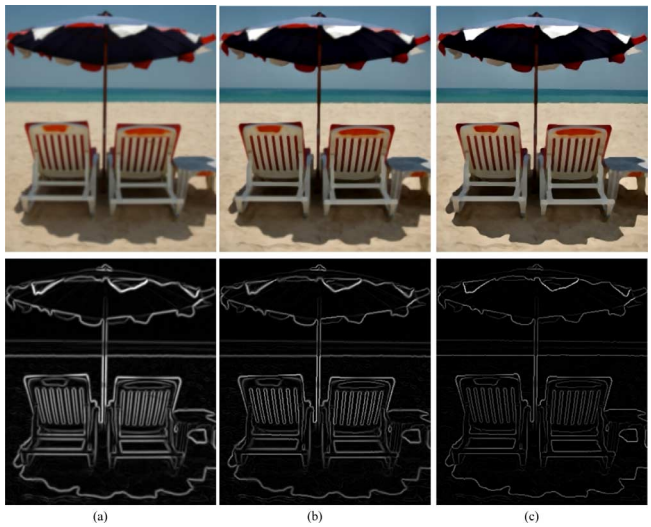

Fig. 18. Image enhancement using sharpness enhancement function. (a) The blurry image and its gradient magnitude map. (b) Enhanced gradient field (bottom) with $\mu = 0.5$, and the reconstructed image (top). (c) Enhanced gradient fields (bottom) with $\mu = 0.1$ and the reconstructed images (top). Enhanced image in (c) is sharper than (b), and the enhanced image in (b) is sharper than the blurry image in (a).

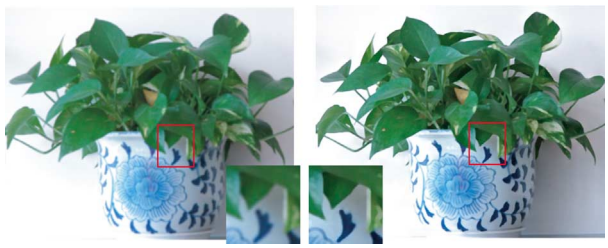

Fig. 19. Example of image enhancement for a real captured blurry image due to out-of-focus. (a) Blurry image. (b) Enhanced image. μ is set to 0.2.

B. Sharpness Transfer

Inspired by the idea of image statistics transfer, we define an automatic method to enhance image sharpness. The idea is to transfer the gradient profile sharpness distribution of a source image (or natural images) to the target blurry image I_t. To implement this idea, we first compute the accumulative sharpness distributions of the source image (or natural images) and the blurry image respectively, which are denoted as H_s and H_t. Then the transferred sharpness σ' for a gradient profile with sharpness σ in I_t is estimated as $\sigma' = H_s^{-1}(H_t(\sigma))$, where H_s^{-1} is the inverse of H_s. Using this automatic sharpness mapping function, we perform image enhancement by gradient field transformation and poisson reconstruction.

Fig. 20 gives one example of image sharpness transfer. Fig. 20(e), (f) show the sharpness distributions of source image and target image. Then the sharpness distribution of the source image is transferred to the target blurry image by transfer mapping in Fig. 20(g), and the enhanced image in Fig. 20(c) is sharp and clear. We also apply the sharpness distribution of a collection of natural images shown in Fig. 3 to the target image. Obviously, the enhanced images are sharp and rarely has artifacts along structures.

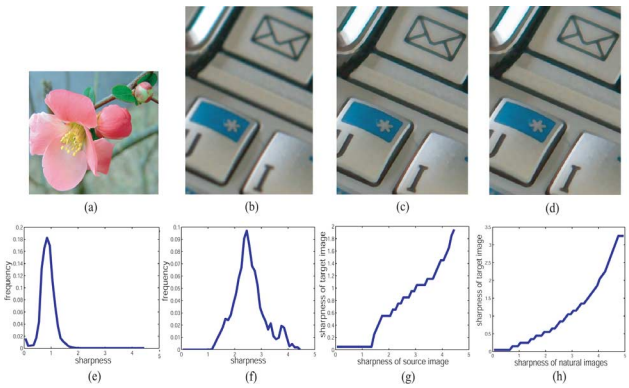

Fig. 20. Image sharpness transfer. (a) Source image. (b) Target image. (c) Result after transferring the sharpness distribution from (a) to (b). (d) Result after transferring the sharpness distribution of natural images in Fig. 3 to (b). (e) Sharpness distribution of (a). (f) Sharpness distribution of (b). (g) Sharpness transfer mapping from (a) to (b). (h) Sharpness transfer mapping from natural images in Fig. 3 to (b).

C. Comparison With the Other Edge Enhancement Methods

In Fig. 21, we also compare the proposed enhancement algorithm with unsharp masking, Osher-Rudin's shock filter [40] and Kramer's improved shock filter [41], [39], for comparison. For the unsharp masking algorithm, the bandwidth of Gaussian is set to 2.0, and the fraction of difference is set to 1.0. The unsharp masking well recovers the contrast of image, however the enhanced image has halo and ringing artifacts and the edges are not sharp enough in appearance. Both of Osher-Rudin's shock filter and Kramer's improved shock filter produce very sharp edges. However they tend to sharpen all the detected edges to be step edges, and the sharpened edges are not regular and with halo or jaggy artifacts, e.g., the shadow or small details in the

Fig. 21. Comparison of image enhancement algorithms (Please zoom-in for better comparison). (a) Blurry image. (b) Our result with $\mu = 0.2$. (c) Result of Osher-Rudin's shock filter. (d) Result of Kramer's improved shock filter. (e) Result of unsharp masking (Gaussian bandwidth and fraction of difference are set to 2.0 and 1.0).

boy's face and the rails appeared in the background layer. Compared to these work, our enhancement algorithm adaptively enhances the sharpness of edges controlled by the sharpness enhancement function (μ is set 0.2 in this example), and it produces sharp image with rare halo or jaggy artifacts.

VI. Conclusion and Discussion

In this paper, we have established a novel generic natural image prior—gradient profile prior. Using this prior, gradient field transformation was proposed to transform the image gradient field guided by the prior knowledge of natural image gradient profiles. The transformed gradient field was applied to constrain the gradient field of the HR image and the enhanced image in image super-resolution and sharpness enhancement. The gradient domain constraint helps to sharpen the details and suppress ringing or jaggy artifacts along edges. Encouraging results have been obtained on a variety of natural images or synthetic images in image super-resolution and enhancement.

As a generic prior of natural image, the gradient profile prior is learned from a large collection of natural images, which reflects the generic regularities of natural images. For a specific image, though its gradient profiles may not fit the prior perfectly, we can always produce sharp and clear high-resolution or enhanced images by gradient field transformation and forcing the sharpness distribution of the enhanced gradient profiles to fit the learned statistical relationship or distribution in natural image. This can be verified by all the examples in this paper, which are not intentionally selected to fit the prior.

A. Comparison With Fattal's Edge Statistics

Fattal [34] proposed an effective edge statistics, which is called edge-frame continuity modulus (EFCM), for image up-sampling and achieved state-of-the-art results. Compared with this edge statistics, the gradient profile prior has the following distinct differences or advantages. Firstly, the gradient profile model parametrically describes the gradient field of natural image by only two parameters of shape and sharpness. As the shape parameter is found to be stable across resolutions, then the gradient profile prior is simply imposed on the HR image based on the sharpness dependency between HR and LR images. Second, both of the gradient profile prior and Fattal's edge statistics achieve convincingly excellent super-resolution results (see Fig. 22 for example), however, our method is significantly faster due to its simplicity. It was reported in [34] that

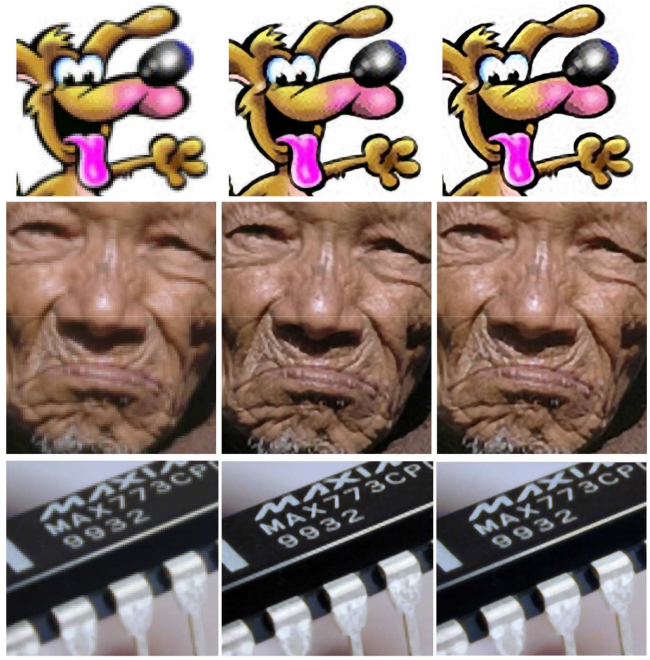

Fig. 22. Comparison of super-resolution results with Fattal's method [34]. (a) Input images. (b), (c) Results of Fattal's method and our method, respectively.

2 s (2.1 GHz CPU) were taken to up-sample an image of 128^2 pixels with factor of 2. However, our method only takes 0.395 s (2.7 GHz CPU) in the same case. Thirdly, as the gradient profile prior is a generic image prior for natural image, it can also be applied to generic image enhancement task by directly enhancing its gradient field.

B. Noisy Image Super-Resolution and Enhancement

For the input image with heavy noise, the noise might also produce local maximums in the gradient domain, which will introduce the noisy gradient profiles. After gradient field transformation, the noise might also be enhanced. See Fig. 23(a) for example, in which noise with standard deviation 10 are added. One possible solution is to denoise the input image first, then add the (up-sampled) noise back after the super-resolution or enhancement. For example, in Fig. 23(d), the noisy LR image is firstly denoised by the method in [48], then the denoised image is up-sampled by the proposed method, and the noise layer is up-sampled by bilinear interpolation. After adding back the up-sampled noise, the super-resolution result is more pleasant to our eye due to the removal of noisy gradient profiles.

C. Comparison Over the LR Image Generated With and Without Smoothing on the HR Image

In the super-resolution methods, the LR image is generally assumed to be generated by prefiltering (i.e., smoothing) and then down-sampling of the HR image, which approximate the image generation process of camera system. On the other hand, in the literatures of image interpolation, the LR image is also assumed to be generated by directly down-sampling of the HR image without prefiltering of the HR image. In this case, the LR signal is commonly aliased along the image

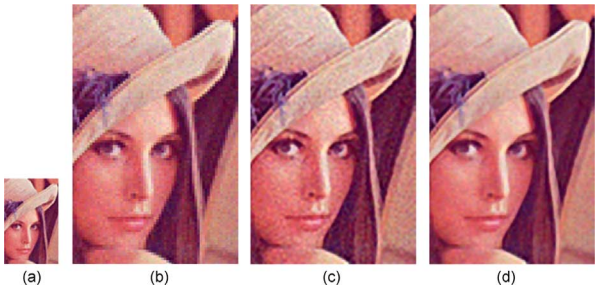

Fig. 23. Super-resolution on a noisy image (3X), standard deviation of the added noises in LR image is 10. (a) Noisy LR image. (b) HR image using nearest neighbor interpolation. (c) HR image using the proposed image on the noisy LR image. (d) LR image is firstly denoised by nonlocal denoising method [48], then the denoised image is up-sampled by the proposed method, and the noises are up-sampled by bilinear interpolation.

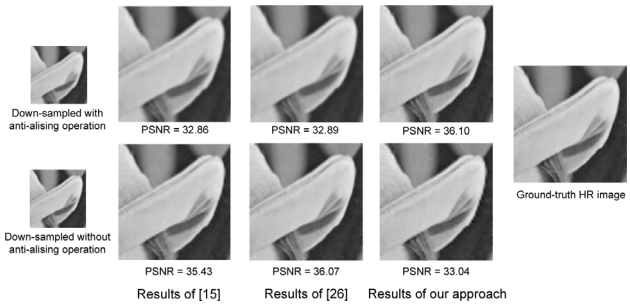

Fig. 24. Comparisons on the LR images generated with and without smoothing (i.e., anti-aliasing) operation on the HR image. We compare our method with the interpolation methods in [15], [26], which assume that the LR image is directly generated by subsampling of HR image without the smoothing operation.

edges, therefore, the methods in [15], [26] attempts to remove the aliasing artifacts in the HR image. To fairly compare the interpolation-based methods with our proposed method, we quantitatively compare them on the LR images down-sampled with and without prefiltering operations in the first and second row of Fig. 24 respectively. We can observe that, for the LR image generated with smoothing operation, the interpolation based methods in [15], [26] produce blurry HR results, while our method produces sharp HR image with much more details and highest PSNR value of 36.10. However, for the LR image generated without smoothing operation, the methods in [15], [26] produce more reasonable HR images with higher PSNRs than our method. That is understandable because these interpolation methods and our method are designed for the LR images generated with different assumptions. In the real applications, we should choose the appropriate methods based on how the LR image is generated.

In the future, we are planning to further speed up the proposed super-resolution and sharpness enhancement algorithm, then extend the proposed method to video super-resolution and enhancement. We are also interested in applying the gradient profile prior to the point spread function estimation [11] and the image deblur.

REFERENCES

[1] S. C. Zhu and D. Mumford, "Learning generic prior models for visual computation," in *Proc. CVPR*, 1997, pp. 463–469.
[2] S. Roth and M. J. Black, "Fields of experts: A framework for learning image priors," in *Proc. CVPR*, 2005, vol. 2, pp. 860–867.
[3] S. Lyu and E. P. Simoncelli, "Statistical modeling of images with fields of gaussian scale mixtures," in *Proc. NIPS*, 2006, pp. 945–952.
[4] 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 Statist. Comput. Theories Vis.*, 2003, pp. 1–24.
[5] J. Sun, N. N. Zheng, H. Tao, and H. Y. Shum, "Image hallucination with primal sketch priors," in *Proc. CVPR*, 2003, vol. 2, pp. 729–736.
[6] S. Y. Dai, M. Han, W. Xu, Y. Wu, and Y. H. Gong, "Soft edge smoothness prior for alpha channel super resolution," in *Proc. CVPR*, 2007, pp. 1–8.
[7] J. Sun, J. Sun, Z. B. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
[8] S. Roth and M. J. Black, "Steerable random fields," in *Proc. ICCV*, 2007, pp. 1–8.
[9] A. Levin, A. Zomet, and Y. Weiss, "Learning how to inpaint from global image statistics," in *Proc. ICCV*, 2003, vol. 1, pp. 305–312.
[10] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, 2006.
[11] N. Joshi, R. Szeliski, and D. J. Kriegman, "Psf estimation using sharp edge prediction," in *Proc. CVPR*, 2008, pp. 1–8.
[12] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
[13] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 22, no. 2, pp. 503–516, 2000.
[14] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, 1992.
[15] S. Mallat and G. Yu, "Super-resolution with sparse mixing estimators," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2889–2900, Nov. 2010.
[16] M. Fadili, J. L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," *Comput. J.*, vol. 52, no. 1, pp. 64–79, 2009.
[17] J. C. Yang, J. Wright, Y. Ma, and T. Huang, "Image super-resolution as sparse representation of raw image patches," in *Proc. CVPR*, 2008, pp. 1–8.
[18] J. G. Huang and D. Mumford, "Statistics of natural images and models," in *Proc. CVPR*, 1999, vol. 1, pp. 540–547.
[19] P. Moulin and L. Juan, "Analysis of multiresolution image denoising schemes using generalized-Gaussian priors," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 909–919, Jun. 1998.
[20] O. David and M. Peyman, "Modeling multiscale differential pixel statistics," *Proc. SPIE -Int. Soc. Opt. Eng.*, pp. 16–24, 2006.
[21] S. C. Zhu and D. Mumford, "Prior learning and gibbs reaction-diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 11, pp. 1236–1250, Nov. 1997.
[22] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. ICCV*, 2009, pp. 349–356.
[23] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Signal Process.*, vol. SP–26, no. 6, pp. 508–517, Jun. 1978.
[24] P. Thevenaz, T. Blu, and M. Unser, "Image Interpolation and Resampling," in *Handbook of Medical Imaging, Processing and Analysis*. New York: Academic, 2000.
[25] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.
[26] X. Zhang and X. Wu, "Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 887–896, Jun. 2008.
[27] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion and transparency," *J. Vis. Commun. Image Represent.*, vol. 4, no. 4, pp. 324–335, 1993.
[28] 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.
[29] Z. C. Lin and H. Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, Jan. 2004.
[30] M. Ben-Ezra, Z. C. Lin, and B. Wilburn, "Penrose pixels: Super-resolution in the detector layout domain," in *Proc. ICCV*, 2007, pp. 1–8.
[31] W. T. Freeman, E. Pasztor, and O. Carmichael, "Learning low-level vision," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.
[32] 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, 2002.

[33] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neigbor embedding," in *Proc. CVPR*, 2004, vol. 1, pp. 275–282.
[34] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 95:1–95:8, 2007.
[35] C. Liu, H. Y. Shum, and W. T. Freeman, "Face hallucination: Theory and practice," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 115–134, 2007.
[36] Q. Wang, X. Tang, and H. Y. Shum, "Patch based blind image super resolution," in *Proc. ICCV*, 2005, vol. 1, pp. 709–716.
[37] G. R. A. Polesel and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 505–510, Mar. 1999.
[38] L. I. Rudin, "Images, numerical analysis of singularities and shock filters," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, 1987.
[39] F. Guichard and J. M. Morel, "A note on two classical enhancement filters and their associated pde's," *Int. J. Comput. Vis.*, vol. 52, no. 2/3, pp. 153–160, 2003.
[40] S. Osher and L. I. Rudin, "Feature-oriented image enhancement using shock filters," *SIAM J. Numer. Anal.*, vol. 27, no. 4, pp. 919–940, 1990.
[41] J. G. M. Schavemaker, M. J. T. Reinders, J. J. Gerbrands, and E. Backer, "Image sharpening by morphological filtering," *Pattern Recognit.*, vol. 33, no. 6, pp. 997–1012, 2000.
[42] J. Weickert, "Coherence-enhancing shock filters," *Lecture Notes in Computer Science*, vol. 2781, pp. 1–8, 2003.
[43] M. K. Varanasi and B. Aazhang, "Parametric generalized gaussian density estimation," *J. Acoust. Soc. Amer.*, vol. 86, no. 4, pp. 1404–1415, 1989.
[44] A. L. Yuille and T. Poggio, "Fingerprints theorems for zero crossings," *J. Opt. Soc. Amer. A*, vol. 2, pp. 683–692, 1985.
[45] A. Agrawal, R. Raskar, and R. Chellappa, "What is the range of surfacce reconstructions from a gradient field?," in *Proc. ECCV*, 2006, pp. 578–591.
[46] P. Perez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, 2003.
[47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Quality assessment: From error measurement to structural similarity," *IEEE Trans. Pattern Anal Mach. Intell.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[48] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proc. CVPR*, 2005, vol. 2, pp. 60–65.

Jian Sun (M'11) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, in 2003 and the Ph.D. degree in applied mathematics from Xi'an Jiaotong University, Xi'an, China, in 2009.

He is currently a Faculty Member at the School of Science of Xi'an Jiaotong University. He worked as a Visiting Student at Microsoft Research Asia from November 2005 to March 2008, and a Postdoctoral Research Associate at the EECS of University of Central Florida from August 2009 to March 2010. His current research interests are image denoising, inpainting, segmentation, and super-resolution.

Jian Sun received the B.S. degree, M.S. degree and Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 1997, 2000, and 2003, respectively.

He joined Microsoft Research Asia, Beijing, China, in July, 2003. His research interests include computer vision and computer graphics, especially on stereo matching, interactive computer vision (interactive image/video segmentation, image matting, image completion, image vectorization, and interactive video object tracking), and computational photography (image deblurring).

Zongben Xu received the M.S. and Ph.D. degrees in mathematics from Xi'an Jiaotong University, Xi'an, China, in 1981 and 1987, respectively.

He has been with the Faculty of Science and Institute for Information and System Sciences, Xi'an Jiaotong University since 1982, where he became Associate Professor in 1987 and Full Professor in 1991, and now serves as Director of the Institute for Information and System Sciences. In 1988, he was a Postdoctoral Researcher in the Department of Mathematics, University of Strathclyde, U.K. He now serves as Chief Scientist of one National Basic Research Project of China (973 Project). His current research interests include nonlinear functional analysis, mathematical foundation of information technology, and computational intelligence.

Dr. Xu received the National Natural Science Award of China in 2007, and winner of CSIAM Su Buchin Applied Mathematics Prize in 2008.

Heung-Yeung Shum (S'90–M'90–SM'01–F'06) received the doctorate degree in robotics from the School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Currently, he is a Corporate Vice President at Microsoft and has taken the new role of leading the Core Search Development of Microsoft. He serves on the editorial board of the *International Journal of Computer Vision*, and is a Program Chair of the International Conference of Computer Vision (ICCV) 2007. He has published more than 100 papers in computer vision, computer graphics, pattern recognition, statistical learning, and robotics. He holds more than 50 U.S. patents.