CrossMark

Optimal filter based on scale-invariance generation of natural images

Feng Jiang · Bo-Wei Chen · Seungmin Rho · Wen Ji · Liqiang Pan · Hongwei Guo · Debin Zhao

Abstract One of the most striking properties of natural image statistics is the scale invariance. Some earlier studies have assumed that the kurtosis of marginal band pass filter response to be constant throughout scales for a natural image. In our study, this assumption is loosened by adaptively estimating an optimal filter computation whose response distributions through scales have the least Kullback–Leibler divergence. The adaptive filter and its responses characterize the scale-invariance property more accurately and effectively and are further utilized to model the statistics scale-invariance prior in this paper. Extensive experiments on image super-resolution and de-noising manifest that the explored natural images scale-invariance prior model achieves significant performance improvements over the current state-of-the-art schemes.

Keywords Image restoration · Statistical modeling · Scale invariance

F. Jiang · L. Pan · H. Guo · D. Zhao
Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
e-mail: fjiang@hit.edu.cn

D. Zhao
e-mail: dbzhao@hit.edu.cn

B.-W. Chen (✉)
Department of Electrical Engineering, National Cheng Kung University, Tainan City, Taiwan
e-mail: dennisbwc@gmail.com

S. Rho
Department of Multimedia, Sungkyul University, Anyang City, Korea
e-mail: korea.smrho@gmail.com

W. Ji
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
e-mail: jiwen@ict.ac.cn

1 Introduction

One core task for computer vision is to interpret, model, and learn from the highly specialized image data [1,2]. It is also plausible that our brains have learned an approximation to this probabilistic model, allowing for efficient representations and effective solutions of many difficult computational problems [3,4]. The processing of vision information in the human brain depends not only on the characters of the vision information, but also on the prior knowledge, i.e. the general regulations of the natural images generalized and memorized by the human brain. Natural images adding prior knowledge gives us perception. Aspects of vision ranging from the responses of individual neurons to gestalt perceptual rules would be seen not as artifacts of the brain's circuitry, but rather as matched to the statistical structure of the physical world [5,6]. In recent years, how to capture the prior of natural images and provide precise mathematical model of the structure in the visual world have become a prevalent and critical problem [7,8]. Such models have pivotal role in many low-vision tasks which would provide a rigorous basis for practical algorithms in image coding, processing, and recognition [9,10].

In this view, aspects of vision ranging from the responses of individual neurons to gestalt perceptual rules would be seen not as artifacts of the brain's circuitry, but rather as matched to the statistical structure of the physical world [5,11–13]. One of the most notable properties of natural image statistics is scale invariance; it is exhibited as:

$$Q[\varphi(x)] = Q\left[\alpha^v \phi(\alpha x)\right], \tag{1}$$

where $Q[\phi(\alpha x)]$ is any ensemble statistic of $\phi(x)$ of scale α and v is a universal exponent [14,15]. In particular, Field observed that the spatial patterns of image intensity from natural environments have power spectra that approximate $S_I \propto 1/k^2$, which is consistent with the hypothesis of scale-invariance and simple dimensional analysis, and he suggested a direct connection to the distribution of receptive field parameters across neurons in the visual cortex [16]. The intuition that scale invariance is a strong constraint on the form of the probability distribution comes from statistical mechanics. Scale invariance is researched at many aspects. Natural images exhibit a highly non-Gaussian and heavy-tailed distribution when derivative-like filters are applied to them [17] and scale-invariance states that natural images exhibit similar heavy-tailed distributions at different scales in one image [18].

Some earlier studies have assumed the kurtosis of marginal band pass filter response distributions to be constant throughout scales. Lam's paper [19] shows that the kurtosis is constant at different scales for DCT filters' marginal distributions. In Bethge's work [20], it is reported that the kurtosis value is lower for high-frequency filters than for lower ones. Also, the scaling of power spectrum, local histograms, and pixel information can exhibit that natural scenes possess scale invariance [15].

Intuitively, a natural image always contains the same contents of different scales and dually the same contents of the same scale exist throughout scales of the image. This paper concentrates on the essence and modeling of natural image scale-invariance character. The main contributions are manifested at two aspects:

- Earlier studies have reported the kurtosis of marginal band pass filter response distributions to be constant throughout scales, which is a strong assumption. We loosened this assumption by estimating an optimal derivative-like filter, whose response distribution throughout scales best match each other compared with the general derivate-like filters. Such a filter satisfies the inner essence of the scale-invariance character of the natural image.
- Utilizing the scale-invariance optimal filter and its statics, we develop a novel scale-invariance prior model, which is applied to image super-resolution, where the response of the optimal filter throughout scales remaining the same is exploited as an effective constraint.

The remainder of the paper is organized as follows. Related works are overviewed in Sect. 2. Section 3 elaborates the filter response analysis and the estimation of optimal filter for scale invariance and the image super-resolution based on the optimal filter. Extensive experimental results are reported in Sect. 4.

2 Related work

Computer vision can be considered a highly specialized data collection and data analysis problem. We need to understand the special properties of image data to construct statistical models for representing the wide variety of image patterns. One special property of vision that distinguishes itself from other sensory data such as speech data is that the distance or scale plays a profound role in image data [14]. More specifically, visual objects and patterns can appear at a wide range of distances or scales, and the same visual pattern appearing at different distances or scales produces different image data with different statistical properties, and thus entails different regimes of statistical models.

2.1 Scale-invariance character of natural images statistics

One of the most striking properties of natural image statistics is their scale invariance [21]. The most notable scale-invariant property is the power-law spectrum [22]. This property is very robust and holds across different images and scenes. Various other properties of natural images have been shown to be scale invariant as described in [21, 23]. Natural images, in addition to having scale-invariant statistics, are also extremely non-Gaussian. The distributions of the different Fourier coefficients, for example, have very large peaks, heavy tails, and are highly kurtotic. These distributions can be generally well fitted with a generalized Gaussian distribution, which captures this distinctive shape [20]. The kurtosis of a generalized Gaussian distribution is directly dependent on its shape parameter α. Assuming x is generalized Gaussian distributed such that $x \sim GG(\mu, \sigma^2, \alpha)$ where μ is the mean, σ^2 the variance, and α is the shape parameter, the kurtosis of x is:

$$\kappa_x(\alpha) = \frac{\Gamma(\frac{1}{\alpha})\Gamma(\frac{5}{\alpha})}{\Gamma(\frac{3}{\alpha})^2}, \tag{2}$$

where Γ is the standard gamma function [24]. As can be seen, the kurtosis is inversely related to the shape parameter α. For natural images, α is usually rather small, having values of between 0.5 and 1 [22]. In light of this, one would expect to see some sort of scale invariance in the kurtosis of marginal coefficient distributions; specifically, a reasonable assumption would be that the kurtosis should be constant throughout scales. This, however, is not always the case. There are inconclusive evidence on whether the kurtosis values change with the scale of the measured filter response distribution. In [19] it has been reported that the kurtosis is constant throughout scales for DCT filters marginal distributions, whereas in [20] it has been reported that the kurtosis changes with scale. Specifically, in [20] it is reported that for higher frequencies, the kurtosis values are lower than for the low frequencies.

With the exploration of scale-invariance character, more and more applications have emerged. In Hou's algorithm [25], multi-resolution sparse prior was proposed; it is computed throughout scales, and then a minimization problem with the prior is solved by a strategy of multi-stage convex relaxation. Zoran's paper [14] proposed a noise estimation method according to the assumption that different scale filter distributions' kurtosis values should not change with scale and that added noise can cause these values to change. Gluckman [26] showed a representation which separates scale variant and invariant information. To solve the problem of contour completion and extraction, Ren [27] proposed a multi-scale Bayesian approach.

Many proposed scale-invariance prior models base on the assumption that the kurtosis values of the filter distributions remain the same. In fact, it is an exaggerated assumption as analyzed in this paper. In our proposed method, we loosened this assumption by estimating an optimal derivative-like filter adaptively, whose response distribution of different scales is expected to be the same. With this loosening assumption, our proposed method is more effective and robust in the scale-invariance prior expression.

2.2 Image super-resolution

Image super-resolution, which is the art of rescaling a low-resolution (LR) image to a high-resolution (HR) version, has become a very active area of research in image processing. The interest in image super-resolution is born not only in the great practical importance of enhancing resolution of images, such as in the fields of digital photography, computer vision, computer graphics, medical imaging, satellite remote sensing, and consumer electronics, but also the important theoretical value of using image super-resolution to understand the validity of different image models in inverse problems. In the last several years, there has been a great deal of work on image super-resolution. In general, image super-resolution techniques can be categorized into three families: interpolation-based methods [28–31], reconstruction-based methods [32,33], and learning-based methods [34,35].

Considering the underlying image models during interpolation, most of image super resolution algorithms can be categorized as global or local ones. A global algorithm

trains the interpolation model using the whole image sample set, while a local algorithm aims to train the model by using only useful local information. The representative global methods are those based on classical data-invariant linear filters, such as bilinear and bicubic [36]. These methods have a relatively low complexity, but suffer from the inability to adapt to varying pixel structures, which result in blurred edges and annoying artifacts. The local algorithms usually result in better empirical results, since it is hard to find a unified model with a good predictability for the whole image sample set. In the literature, some local learning methods have been proposed with great success. Li and Orchard [33] propose adapting the interpolation based on the geometric duality between the LR and the HR covariance. Zhang and Wu [37] propose partitioning the local neighborhood of each missing sample into two oriented subsets in orthogonal directions, and then fuse the directional interpolation results by minimum mean square-error estimation. Takeda et al. propose using kernel regression as an effective tool for interpolation in image processing [34]. Recently, Zhang and Wu [38] propose the named SAI algorithm, which learns and adapts varying scene structures using a locally linear regression model and interpolates the missing pixels in a group by a soft-decision method.

From a machine learning perspective, the available LR image pixels can be regarded as labeled samples and the missing HR image pixels as unlabeled ones. What image super-resolution does is to learn latent models of the image sample set in a supervised manner. In all local methods mentioned above, interpolation models are learned only according to the labeled samples in a local neighborhood and then mapped to the missing HR samples to perform inference. During this procedure, unlabeled samples are left out and the information hidden are not sufficiently explored. Image super-resolution is an ill-posed problem.

To cope with the ill-posed nature of inverse problems, one type of scheme in literature employs image prior knowledge for regularizing the solution to the following minimization problem:

$$\arg\min_u \frac{1}{2} \|\boldsymbol{H}\boldsymbol{u} - \boldsymbol{y}\|_2^2 + \lambda \cdot \Psi(\boldsymbol{u}), \tag{3}$$

where $\frac{1}{2}\|\boldsymbol{H}\boldsymbol{u} - \boldsymbol{y}\|_2^2$ is the ℓ_2 data-fidelity term, $\Psi(\boldsymbol{u})$ denotes image prior named the regularization term, and λ is the regularization parameter. In fact, the above regularization-based framework (2) can be strictly derived from Bayesian inference with prior knowledge as some image prior possibility models. Many optimization approaches for regularization-based image inverse problems have been developed. The key lies in employing effective prior knowledge for regularizing to cope with the ill-posed nature.

3 Multi-scale image filter response analysis

From the viewpoint of natural image prior modeling, natural images have an intrinsic statistics property of their gradient magnitudes: the heavy-tailed distribution. It generates sparse regularization in the objective function and, therefore, makes it non-
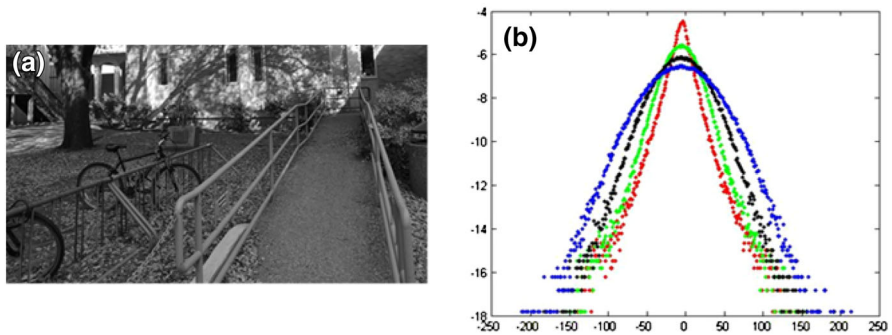
Fig. 1 **a** Original image; **b** different additive Gaussian filtering statistical noise response in case of distribution and performance (color figure online)
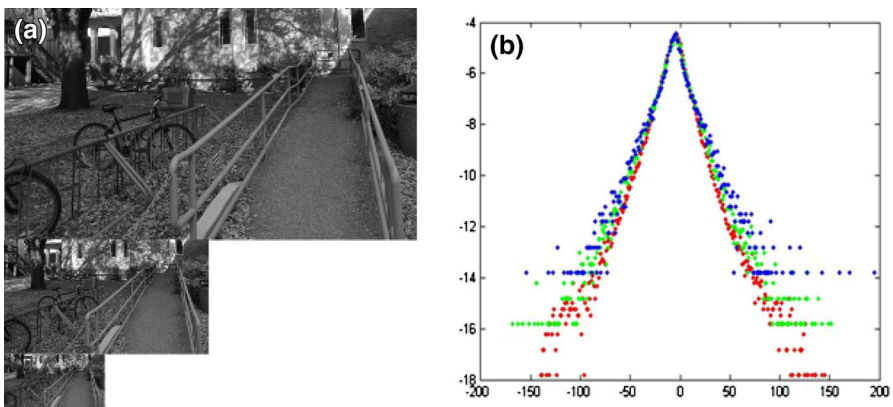

Fig. 2 Multi-scale image filter response. **a** Images at different scales and **b** derivative-like statistics of the three images at different scales (color figure online)

convex. In observations, similar highly non-Gaussian and heavy-tailed distributions are obtained when applying derivative-like filters throughout scales of a natural image. For image in a certain scale, the image quality will affect the final statistical response distribution. We conducted a simple experiment. For the image in Fig. 2, Gaussian noise $\sigma = 15, 25, 35$ are added to the image, respectively.

In Fig. 1, red dots indicate the response distribution of the original image, green black and blue dots indicate the response distribution of the original image on adding white Gaussian noise $\sigma = 15, 25, 35$, respectively. It can be seen that with the increase of the noise, the response distribution drifts farther away from that of the original image.

The response distribution of derivate-like filters provides crucial cues about the quality of the image. The relationships between response distributions of different scales further demonstrate the essences. As shown in Fig. 1, a 3×3 derivative-like filter is applied to an image at different scales and the response distributions are illustrated with different colors in Fig. 2.

In Fig. 2, the response distributions of original higher-resolution image, intermediate-scale image and lower-resolution scale are indicated with red, green and blue dots, respectively. From the responses it can be seen that distributions of different

scale responses are very close. Although this property provides exciting cues for image prior modeling and pursuing, it is not precise enough for most of the image inverse problems. Two inspirations can be drawn from the observation. For a natural image, response to different scales are not independent, but they have similar kurtosis. In fact, this is a week constraint. The response distributions are not the same precisely, which is more effective to constrain in the image inverse problem. For the different derivative filters, if the response distribution of a filter is more similar to those at different scales, the filter can better keep scale-invariant features. An optimal filter is expected to get exactly the same response in all scales and such a filter best meets the scale-invariant feature. If such optimalilty is obtained, whose corresponding response distributions are the same throughout the scales, it will be able to provide more information for image inverse problem.

3.1 Optimal filter estimation based on response specification

Natural image scale invariance demonstrates that different scale images exhibit similar heavy-tailed filter response distribution when derivative-like filters are applied to them. If we can find the filter whose responses to the multi-scale images have the least difference, such a filter will inflect the inner essence of the scale-invariance character of the natural image. In this paper, we estimate the optimal filter from the multi-scale of the images and adopt a recursive procedure based on response histogram specification. In each iteration, the response specification is utilized until the algorithm converges (Fig. 3).

More specifically, consider two different scales of the same image I_o and I_l, where I_o is large scale with a size of $M \times N$ and I_l is small scale with a scale $\frac{M}{2} \times \frac{N}{2}$. Assuming the size of the image block to be b_s, divide I_o and I_l into overlapping blocks of size b_s and reorganize each block into one row of S_o and S_l, which can be expressed as

$$\begin{aligned} S_o &= \varphi(I_o) \\ S_l &= \varphi(I_l) \end{aligned}. \tag{4}$$

For each pixel in I_o and I_l, its b_s neighborhoods are reorganized into one row in S_o and S_l. Accordingly, $S1$ and $S2$ contain $M \times N$ and $\frac{M}{2} \times \frac{N}{2}$ rows, respectively. Denote the filter as f; then the filter response can be represented in the matrix
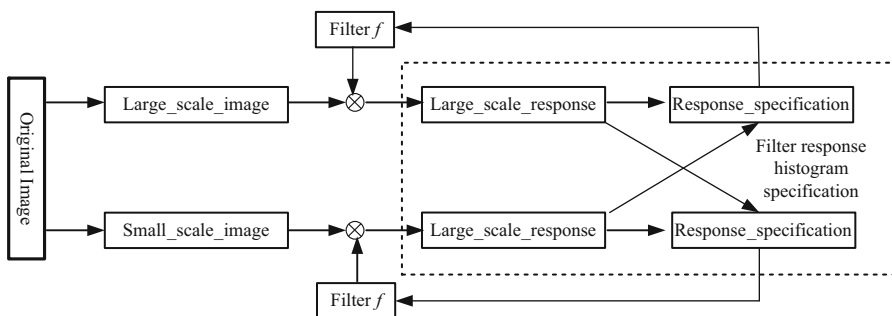

Fig. 3 Optimal filter estimation based on response histogram regulation

form. Define that R_o and R_l are the filter responses, corresponding to the large-scale response and small-scale response, respectively. $R_o = \begin{bmatrix} R_{o\otimes 1}, & R_{o\otimes 2}, & \dots, & R_{o\otimes MN} \end{bmatrix}^T$ and $R_l = \begin{bmatrix} R_{o\otimes 1}, & R_{o\otimes 2}, & \dots, & R_{o\otimes \frac{MN}{4}} \end{bmatrix}^T$, which can be simplified into the following two expressions:

$$\begin{aligned} R_o &= S_o \times f \\ R_l &= S_l \times f \end{aligned} \qquad (5)$$

Histogram specification is adopted to make a histogram of the filter response in a certain scale closer to the reference histogram distribution. Given two responses R_o and R_l, the filter response specification processing provides the estimation of the expected filter response to one scale image, given that of the other scale image, which can be expressed as Eq. 6. The detailed procedure is provided in Algorithm 1:

$$R'_o = \text{hist\_spe } (R_o, R_l). \qquad (6)$$

Algorithm 1: the response specification.

Input: filter responses of two different scales of the same nature image, R_1 and R_2.

Initialization: $R'_o = R_o \times 100$ and $R'_l = R_l \times 100$.

Step1 Obtain the histograms of $\lfloor R_o \times 100 \rfloor$ and $\lfloor R_l \times 100 \rfloor$ as h_o and h_l;

Step2 find its cumulative $H_o[j] = \sum_{i=0}^{j} h_o[i]$ and $H_1[j] = \sum_{i=0}^{j} h_l[i]$;

Step3 Relate the two mapping above to build a lookup table for the over all mapping. Specifically, for each input level i in $\lfloor R_o \times 100 \rfloor$, find an output level j, so that $H_1[j]$ best matches $H_o[i]$:

$$|H_o[i] - H_1[j]| = min_k |H_o[i] - H_l[k]|$$

Step5 setup a lookup entry in R'_o and update R'_o

$$lookup[i] = j$$

Step6 $R'_o = \frac{R'_o}{100}$;

Output: The specified response R'_o.

The Algorithm 1 aims at making one scale filter response as reference and estimating the filter response of the other scales to meet the gap of the inconformity between the response of different scales. The response histogram specification provides an effective basis of the optimal filter estimation. In the procedure of the estimation of optimal filter, firstly the initialized filter is utilized to obtain the response of different scales and a recursive one is adopted. In each interaction, the response of one scale is fixed and utilized to estimate the responses of the other scales. One the responses of the other scales is obtained, and the filter is updated. With the recursive process and the updating

filter, the response distributions throughout the scales became identical gradually. The finally estimated optimal filter makes the response distributions of different scales have the least distribution distance and the whole procedure is provided in Algorithm 2.

Algorithm 2: estimation of the optimal filter f_{opt}.

Input: two different scales of the same nature image, I_o and I_l.

Initialization: $\lambda \in (0, 1)$, randomly selected derivative-like filter f_0.

Step1 Set the initial values $f_{opt} = f_0$ for the filter.

Repeat

Step2 obtain R_o and R_l, the filter f_{opt} responses to I_o and I_l, by computing Eq. (5);

Step3 perform the response histogram specification, in which R_2 denoted as reference:

$$R_o^{'} = histfml(R_o, R_l);$$

Step4 obtain filter $f_1 = S_1^{\dagger} \times R_1^{'}$, where $S_1^{\dagger}$ is the pseudo-inverse of S_1;

Step5 update R_o and R_l as the filter f_1 responses to I_o and I_l;

Step6 perform the response histogram specification, in which R_l denoted as reference:

$$R_l^{'} = histfml(R_l, R_o);$$

Step7 obtain filter $f^{'} = S_2^{\dagger} \times R_2^{'}$, where $S_2^{\dagger}$ is the pseudo-inverse of S_2;

Step8 update $f_{opt} = f_{opt} + \lambda \cdot \left(f^{'} - f_{opt} \right)$

Until stopping criterion is satisfied

Output: The obtained optimal scale invariance filter f_{opt}.

3.2 Image super-resolution based on the optimal filter

The natural image scale-invariance prior is adopted in this paper as a constraint in the image super-resolution task. By incorporating the proposed prior modeling (18) into the regularization-based framework (2), a new formulation for image restoration can be expressed as follows:

$$\text{argmin}_u \frac{1}{2} \| \boldsymbol{H}\boldsymbol{u} - \boldsymbol{y} \|_2^2 + \tau \cdot \left\| \text{Resp}(u) - \text{Resp}(u') \right\|_2^2, \tag{7}$$

where τ is the control parameter. Note that the first term of Eq. 7 actually represents the observation constraint and the second term represents the image prior scale-invariance
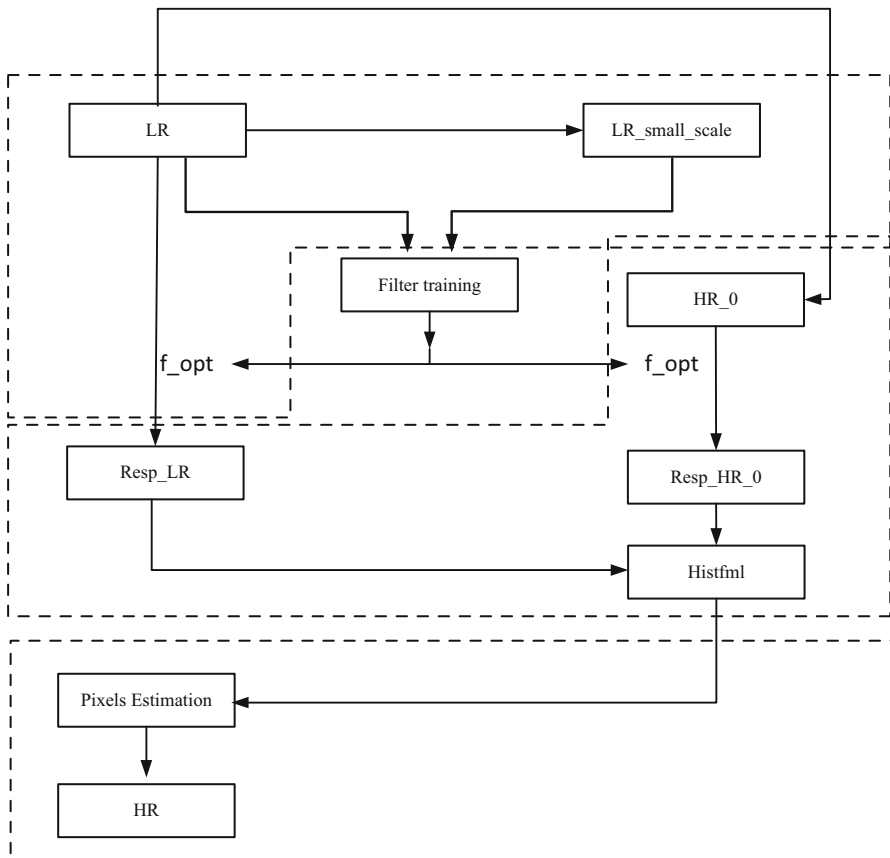
Fig. 4 Framework of image super-resolution based on optimal filter

constraints, respectively. Note that the second regulation term cannot compute the derivation directly, and the proposed image super-resolution based on the optimal filter is shown in Fig. 4. Firstly, the optimal filter is obtained based on the original low-resolution images. Then the filter response to the high-resolution images is estimated based on the response of low-resolution images; finally, the pixels of the estimated high-resolution image are refined.

3.2.1 Estimation of the most likely response for high-resolution image

For an input images LR, a low-resolution image LR_small_scale is obtained by first blurring with a blur kernel and then downsampled by a scaling factor. The low-resolution images LR and image LR_small_scale is then used to estimate the optimal filter f_{opt} for the current image with histogram specification. By applying bicubic interpolation to low-resolution images LR we obtain an initialization of high-resolution image HR_0 and then compute the response distributions of HR_0 and LR images with the optimal filter as follows:

$$\varphi(\text{LR}) \times f_{\text{opt}} = \text{Resp\_LR}$$
$$\varphi(\text{HR\_0}) \times f_{\text{opt}} = \text{Resp\_HR\_0}'$$

(8)

where Resp_LR and Resp_HR_0 denote optimal filter response distribution of LR images and HR_0 image. Based on the scale invariance of the natural image, for the optimal filter, the response distribution at different scales should be very close. Accordingly, the optimal filter response distribution in the low-scale image is used to predict that in the high-resolution image:

$$\text{Resp\_HR}' = \text{histfml}(\text{Resp\_HR\_0}, \text{Resp\_LR}).$$

(9)

In the formula 9, Resp_HR$'$ represents the most likely response of the high-resolution image.

3.2.2 Pixel estimation

Nonlocal self-similarity is another significant property that characterizes the repetitiveness of the textures or structures embodied by natural images within nonlocal area, which can be used for retaining the sharpness and edges effactually to maintain image nonlocal consistency.

In this study, we mathematically characterize the nonlocal self-similarity for natural images by means of the distributions of the transform coefficients, which are achieved by transforming the three-dimensional group generated by stacking similar image patches. More specifically, as illustrated in Fig. 3, firstly, we divide the image u into many overlapped blocks, and for each block in red denoted by u^i, we find a fixed number of blocks that are most similar to it within the blue search window. For simplicity, the criterion for calculating similarity between different blocks is Euclidean distance. Then, these nine blocks are stacked into a three-dimensional array, which we call a group (Fig. 5).

Having obtained D_v, the similar image block group of the reference block V_HR_0, SVD decomposition is utilized to obtain the most likely direction of the reference block.
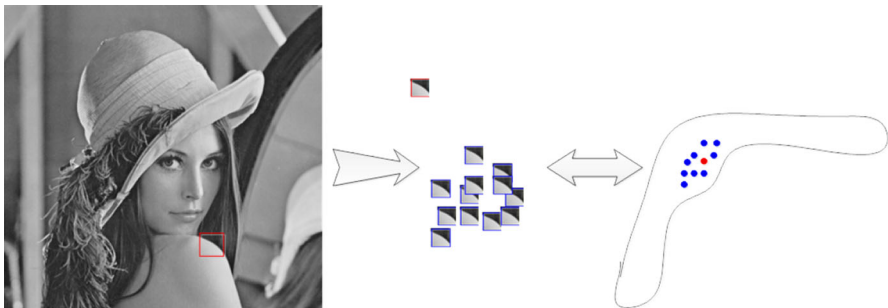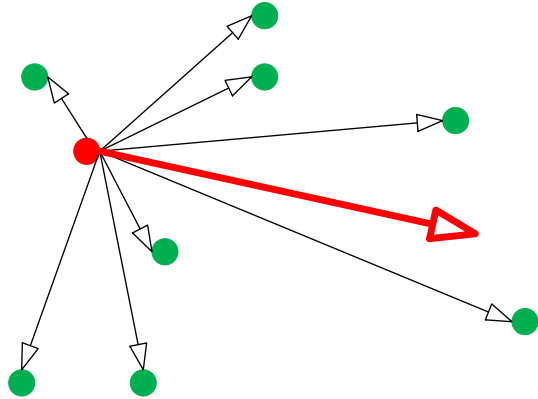

Fig. 5 Nonlocal self-similar blocks in the natural image carve out different manifolds

Fig. 6 Find the most likely
direction of change (color figure
online)

$$[S_v, S_\lambda] = \text{SVD}(D_v)$$
$$D_v = \{v_{c,1} - v_c, v_{c,2} - v_c, \ldots, v_{c,n} - v_c\},$$ (10)

where $V_c = \sum_{i=1}^{n} V_{c,i}$, S_v and S_λ are the obtained eigenvectors and eigenvalues.

In Fig. 6, the red block points indicate the current image patch and the green points indicate the n similar blocks in the bicubiced image. Based on the reference point of the block to the point of similar blocks, the thick red line shows the approximate trend of the reference block. The reference block direction vector can be approximated as:

$$S_v = \{k_1, k_2, \ldots, k_n\}$$
$$S_\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$$
$$d = \sqrt{\lambda_1}k_1 + \sqrt{\lambda_2}k_2 + \cdots + \sqrt{\lambda_n}k_n,$$ (11)

where d indicates the reference block μ and f_{opt} describe the change along the reference block direction and the optimal filter respectively, the following equation can be obtained:

$$\mu \cdot d \cdot f_{\text{opt}} = \text{dis}(\text{Resp\_HR}(\omega), \text{Resp\_HR}'(\omega)).$$ (12)

In Eq. 12, $\text{Resp\_HR}(w)$ and $\text{Resp\_HR}'(w)$ are the response values of the optimal filter to the current patch in the bicubiced image and the estimate response according to the original image.

In Fig. 7, V_HR_0 represents the block image, V_HR is the block vector to estimate ultimately, and V_c is a gap between the two vectors, which can be expressed as

$$V\_c = \mu \cdot d.$$ (13)

So, the estimated high-resolution vector can be solved as a block:

$$V\_\text{HR} = V\_\text{HR}\_0 + V\_c$$
$$V\_c = \mu \cdot d$$
$$\mu = \frac{\text{dis}(\text{Resp\_HR}(\omega), \text{Resp\_HR}'(\omega))}{d \cdot f\_\text{opt}}.$$ (14)
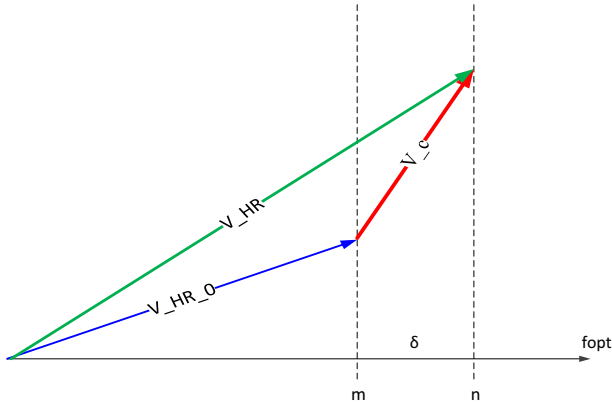
Fig. 7 The image block update

These obtained image blocks are then combined to get the final high-resolution image. The strict mathematical description on the proposed image super-resolution method is provided in Algorithm 3.

Algorithm 3: Image super-resolution based on the optimal filter.

Input: low resolution images I_0

Initialization: Down sample the original I_0 and obtain a low resolution I_l

Step1 compute the optimal filter f_{opt} with I_0 and I_l;

Step2 Applying bicubic interpolation to I_0 and obtain I_1';

Step3 compute R_1' and R_0, the filter response of I_1' and I_0;

Step4 $R_0 = hist\_spe(R_1', R_0)$;

Step6 For each patch V_HR_0 in I_1'

Group similar image patches and perform SVD and obtain d;

Obtain $\mu = \frac{dis(Resp_{HR(w)}, Resp\_HR'(w))}{d.f\_opt}$;

Update the image patch value $V_{HR} = V_{HR\_0} + V\_c$;

End for

Output: Final super-resolution image.

4 Experimental results

In this section, extensive experimental results are presented to evaluate the performance of the proposed algorithm, which is compared with the state-of-the-art methods. We

Fig. 8 Test images: bird, building, butterfly in Berkeley Segmentation Dataset, whose size is 481×321 or 321×481

evaluate the capability of optimal filter describing the scale-invariance prior with Berkeley Segmentation Dataset [39]. We apply our algorithm to the application of image super-resolution and image restoration.

4.1 Scale-invariance performance with the optimal filter

To evaluate the capability of the optimal filter satisfying the scale invariance, we select ten grayscale images from the Berkeley Segmentation Dataset [39]. The original size of images are 481×321 or 321×481, as shown in figure. The adopted image block size is 3×3, and we use Kullback–Leibler divergence to measure the difference between the responses distribution of different scales. We compared the optimal filter with the one-ring and two-ring derivate filters. The results are reported in the table (Fig. 8).

Here, we will respond to the x-direction differential filter located in the distance and the obtained optimal filter response distribution between different scales compared, as shown in Table 1.

Compared with general derivate filters, the proposed optimal filter obtains the least response distribution distance between different scales.

Table 1 of this method shows that the difference between the proceeds of the optimal filter reduces the filter response at different scales; using the kind of optimal filter to image processing can better meet the scale invariance.

	Images	The proposed optimal filter	1-Ring derivative filter	2-Ring derivative filter
Table 1 The response distribution comparison between the proposed optimal filter and general derivate filters (KL-divergence)	1	0.00290	0.00732	0.00732
	2	0.00703	0.00902	0.00902
	3	0.01105	0.01324	0.01324
	4	0.00284	0.00967	0.00978
	5	0.00200	0.00506	0.00546
	6	0.00996	0.01096	0.01109
	7	0.00144	0.00380	0.00191
	8	0.00588	0.00731	0.00751
	9	0.01615	0.01878	0.01912
	10	0.01294	0.02604	0.02508

4.2 Image super-resolution based on optimal filter

In this section, the experimental results are presented to demonstrate the performance of the proposed algorithm. For thoroughness and fairness of our comparison study, we exploit some widely used images as test ones. Our algorithm is compared with some representative work in the literature. More specifically, five methods are included in our comparative study: (1) bicubic interpolation [36]; (2) image interpolation by adaptive 2-D autoregressive modeling and soft decision estimation denoted as SAI [38]; (3) edge-directed interpolation [33].

In our experiments, the observed low-resolution (LR) image is obtained by first blurring with a blur kernel and then downsampling by a scaling factor, from which the original HR images are reconstructed by the proposed and competing methods. A 6×6 Gaussian filter with standard deviation of 1.5 is used for blurring and then downsampling the blurred image in both horizontal and vertical directions. Since the original HR images are known in the simulation, we can compare the interpolated results with the true images and measure the objective and subjective quality of those interpolated images. In the practical experiments, we initialize the HR image $\times$ using the result of bicubic interpolation (Fig. 9).

Given the fact that the human visual system (HVS) is the ultimate receiver of the restored images, we also show the subjective comparison results. Both the superior subjective and objective qualities on test images convincingly demonstrate the potential of the proposed hierarchical and collaborative sparse coding scheme on image super-resolution. Table 2 shows the quantitative quality comparison with respect to PSNR and SSIM of the five compared methods. It can be observed that for all instances the proposed algorithm consistently works better on PSNR and SSIM than other methods. Compared with interpolation-based methods, such as bicubic, the proposed method can significantly improve the objective quality of generated HR images. Compared with the NEDI and SAI method, the proposed method works better and the average PSNR gain can be improved by 0.14 and 0.13 dB. With respect to SSIM, it can be seen that the proposed HCSR method achieves the highest SSIM scores among all
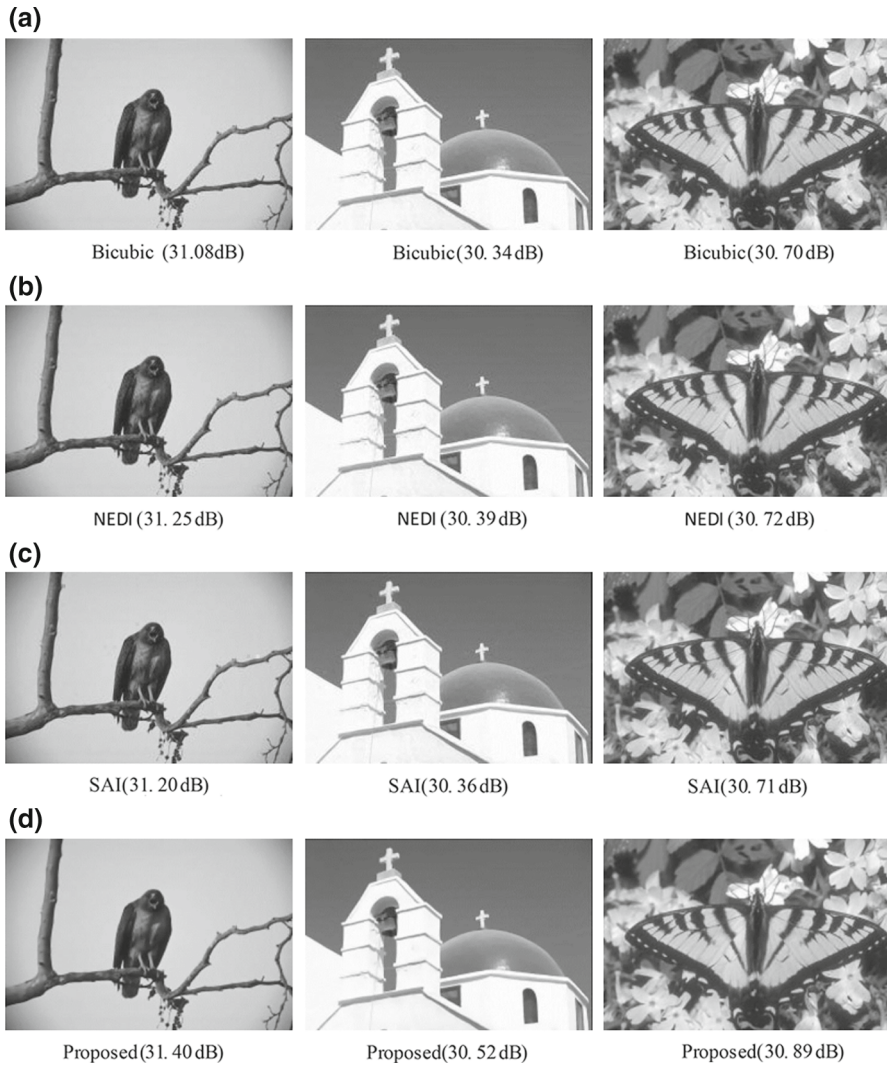
(a)

Bicubic (31.08dB)   Bicubic(30. 34 dB)   Bicubic(30. 70 dB)

(b)

NEDI (31. 25 dB)   NEDI (30. 39 dB)   NEDI (30. 72 dB)

(c)

SAI(31. 20 dB)   SAI(30. 36 dB)   SAI(30. 71 dB)

(d)

Proposed(31. 40 dB)   Proposed(30. 52 dB)   Proposed(30. 89 dB)

Fig. 9 Visual quality comparison of the interpolation with different methods

of the competing methods for all test images. This demonstrates that our method can reconstruct the structures of images better.

4.3 Image restoration for noise removal

To test the robustness of the proposed algorithm, we further consider a practical problem, namely, image restoration for mixed Gaussian plus salt-and-pepper noise removal [40], which is an extension of image restoration from partial random samples with Gaussian noise. In the simulations, images will be corrupted by Gaussian noise with

Table 2 Comparison of the proposed optimal filter-based super-resolution and other popular methods

Images	Bicubic		NEDI [33]		SAI [38]		Proposed	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
1	31.08	0.83	31.25	0.86	31.20	0.92	**31.40**	**0.93**
2	30.34	0.75	30.39	0.72	30.36	0.79	**30.52**	**0.80**
3	30.70	0.83	30.72	0.79	30.71	0.85	**30.89**	**0.86**
4	28.51	0.78	28.69	0.81	28.64	0.86	**28.77**	**0.86**
5	28.57	0.69	28.66	0.85	28.71	0.91	**28.80**	**0.91**
6	28.64	0.81	28.60	0.71	28.79	0.75	**28.77**	**0.76**
7	33.81	0.84	33.82	0.80	33.74	0.85	**33.97**	**0.85**
8	27.43	0.71	27.51	0.80	27.56	0.83	**27.66**	**0.84**
9	29.40	0.78	29.38	0.79	29.36	0.79	**29.53**	**0.80**
10	26.27	0.76	26.38	0.78	26.40	0.81	**26.52**	**0.84**
Average	29.47	0.78	29.54	0.79	29.55	0.81	**29.68**	**0.84**

Bold values indicate the best performance compared with the other methods

Table 3 Comparisons of various methods for Gaussian plus salt-and-pepper noise removal

Image	Barbara		House		Boat		Avg.
r	40 %	50 %	40 %	50 %	40 %	50 %	
Noisy	9.36/0.4153	8.39/0.3815	9.46/0.3499	8.50/0.3170	9.42/0.5579	8.46/0.5230	8.93/0.4241
TV [41]	26.18/0.8899	25.40/0.8728	31.10/0.9156	30.36/0.9050	28.53/0.9405	27.66/0.9259	28.21/0.9083
IFASDA [42]	28.59/0.9252	27.45/0.9129	32.26/0.9263	31.69/0.9181	30.28/0.9614	29.50/0.9556	29.96/0.9333
Proposed	**29.77/0.9271**	**28.04/0.9183**	**32.73/0.9280**	**31.77/0.9194**	**30.42/0.9617**	**29.72/0.9593**	**30.41/0.9356**

The left of the slash denotes PSNR (dB) and the right of the slash denotes FSIM
Bold values indicate the best performance compared with the other methods

standard deviation σ and salt-and-pepper noise density level r, where σ is assumed to be known before and r is unknown. Two state-of-the-art algorithms compared with our proposed method are: TV [41] and IFASDA [42]. Experiments are carried out on four benchmark gray images in Fig. 7, where the standard variance σ of Gaussian noise equals 10 and the noise density level varies from 40 to 50 %. To handle this case, we first apply adaptive median filter (AMF [43]) to the noisy image to identify the mask H, that is, change the problem of mixed Gaussian-impulse noise removal into the problem of image restoration from partial random samples with Gaussian noise, and then run the proposed algorithm according to Table 3.

Table 3 presents the PSNR/FSIM results of the three comparative denoising algorithms on all test images for Gaussian plus salt-and-pepper impulse noise removal. Obviously, the proposed method considerably outperforms the other methods in all the cases, with the highest PSNR and FSIM, achieving average PSNR and FSIM improvements over the second best method (i.e., IFASDA), 1.8 dB and 0.01, separately.

5 Conclusion

In this paper, we estimate the optimal filter whose response distributions have the least KL divergence throughout the scales. The adaptively obtained optimal filter reflects the unique scale-invariance character of the current image and, accordingly, a novel scale-invariance natural image prior model is proposed. Extensive experiments on image super-resolution and de-noising manifest that the natural images scale-invariance prior model proposed in this paper achieves significant performance improvements over the current state-of-the-art schemes.

References

1. Mumford D, Gidas B (2001) Stochastic models for generic images. Q Appl Math 59(1):85–111
2. Srivastava A, Grenander U, Liu X (2002) Universal analytical forms for modeling image probabilities. IEEE Pattern Anal Mach Intell 24(9):1200–1214
3. Schellmann M, Gorlatch S, Meiländer D, Kösters T (2011) Parallel medical image reconstruction: from graphics processing units (GPU) to grids. J Supercomput 57(2):151–160
4. Skaruz J (2014) Two-dimensional patterns and images reconstruction with use of cellular automata. J Supercomput 69(1):9–16
5. Atick JJ (1992) Could information theory provide an ecological theory of sensory processing. Network 3:213–251
6. Bialek W (2002) In Physics of biomolecules and cells. In: Flyvbjerg H, Jülicher F, Ormos P, David F (eds) Proceedings of the Les Houches Summer School, Session LXXV, Springer-Verlag, Berlin, pp 485–577. arXiv:physics/0205030
7. Lim H-Y, Kang D-S (2013) Efficient noise reduction in images using directional modified sigma filter. J Supercomput 65(2):580–592
8. Weng C-Y, Zhang YH, Li LC, Wang S-J (2013) Visible watermarking images in high quality of data hiding. J Supercomput 66(2):1033–1048
9. Chen B-W, Chen C-Y, Wang J-F (2013) Smart homecare surveillance system: behavior identification based on state transition support vector machines and sound directivity pattern analysis. IEEE Trans Syst Man Cybern-Part A: Syst 43(6):1279–1289
10. Jiang F, Rho S, Chen B-W, Du X, Zhao D (2014) Face hallucination and recognition in social network services. J Supercomput. doi:10.1007/s11227-014-1257-z
11. Barlow HB (1961) Possible principles underlying the transformation of sensory messages. In: Rosenblith WA (ed) Sensory communication. MIT Press, Cambridge (Mass) London, pp 217–234
12. Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. Annu Rev Neurosci 24(1):1193–1216
13. Bialek W (2002) Thinking about the brain. Physics of bio-molecules and cells. Physique des bio-molécules et des cellules. Springer, Berlin, pp 485–578
14. Zoran D, Weiss Y (2009) Scale invariance and noise in natural images. In: IEEE 12th international conference on computer vision, pp 2209–2216
15. Ruderman DL (1994) The statistics of natural images. Network 5:517–548
16. Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc Am A 4:2379
17. Weiss Y, Freeman WT (2007) What makes a good model of natural images. In: IEEE conference on computer vision and patern recognition (CVPR), pp 1–8
18. Zhang H, Zhang Y, Li H, Huang TS (2012) Generative Bayesian image super-resolution with natural image prior. IEEE Trans Image Process 21(9):4054–4067

19. Lam E, Goodman J (2000) A mathematical analysis of the dct coefficient distributions for images. IEEE Trans Image Process 9(10):1661–1666
20. Bethge M (2006) Factorial coding of natural images: how effective are linear models in removing higher-order dependencies. J Opt Soc Am A Opt Image Sci Vis 23(6):1253–1268
21. Ruderman DL, Bialek W (1993) Statistics of natural images: scaling in the woods. In NIPS, pp 551–558
22. Srivastava A, Lee AB, Simoncelli EP, Zhu S-C (2003) On advances in statistical modeling of natural images. J Math Imaging Vis 18(1):17–33
23. Ruderman DL (1997) Origins of scaling in natural images. Vis Res 37:3385–3398
24. Domínguez-Molina J, González-Farías G, Rodríguez-Dagnino R, Monterrey I A practical procedure to estimate the shape parameter in the generalized Gaussian distribution. Technique report I-01-18_eng. pdf. http://www.cimat.mx/reportes/enlinea/I-01-18_eng.pdf
25. Hou T, Wang S, Qin H (2011) Image deconvolution with multi-stage convex relaxation and its perceptual evaluation. IEEE Trans Image Process 20(12):3383–3392
26. Gluckman J (2006) Scale variant image pyramids. Comput Vis Pattern Recognit 2006 IEEE Computer Society conference on, vol 1. IEEE
27. Ren X, Malik J (2002) A probabilistic multi-scale model for contour completion based on image statistics. Computer vision–ECCV, Springer, Berlin, pp 312–327
28. Freeman WT, Jones TR, Pasztor EC (2002) Example based superresolution. IEEE Comput Graph Appl 22(2):56–65
29. Glasner D, Bagon S, Irani M (2009) Super-resolution from a single image. In: Proceedings of IEEE international conference on computer vision, Oct 2009, pp 349–356
30. Zhang H, Yang J, Zhang Y, Huang T (2010) Non-local kernel regression for image and video restoration. In: Proceedings of European conference on computer vision, 2010, pp 566–579
31. Yang J, Wright J, Huang T, Ma Y (2010) Image super-resolution via sparse representation. IEEE Trans Image Process 19(11):2861–2873
32. Baker S, Kanade T (2002) Limits on super-resolution and how to break them. IEEE Trans Pattern Anal Mach Intell 24(9):1167–1183
33. Li X, Orchard MT (2001) New edge-directed interpolation. IEEE Trans Image Process 10(10):1521–1527
34. Takeda H, Farsiu S, Milanfar P (2007) Kernel regression for image processing and reconstruction. IEEE Trans Image Process 16(2):349–366
35. Tappen MF, Russell BC, Freeman WT (2003) Exploiting the sparse derivative prior for super-resolution and image demosaicing. In: Proceedings of IEEE workshop statistics computer theories vision, Oct 2003, pp 1–24
36. Keys RG (1981) Cubic convolution interpolation for digital image processing. IEEE Trans Acoust Speech Signal Process ASSP–29(6):1153–1160
37. Zhang L, Wu X (2006) Image interpolation via directional filtering and data fusion. IEEE Trans Image Process 15(8):2226–2238
38. Zhang X, Wu X (2008) Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation. Image Process IEEE Trans 17(6):887–896
39. Martin D, Fowlkes C, Tal D, Malik J (2001) A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of ICCV 2:416–423
40. Cai JF, Chan R, Nikolova M (2008) Two-phase methods for deblurring images corrupted by impulse plus Gaussian noise. Inverse Probl Imaging 2(2):187–204
41. Huang Y, Ng M, Wen Y (2009) Fast image restoration methods for impulse and Gaussian noise removal. IEEE Signal Process Lett 16(6):457–460
42. Li Y-R, Shen L, Dai D-Q, Suter BW (2011) Framelet algorithms for de-blurring images corrupted by impulse plus Gaussian noise. IEEE Trans Image Process 20(7):1822–1837
43. Hwang H, Haddad R (1995) Adaptive median filters: new algorithms and results. IEEE Trans Image Process 4(4):499–502