High Resolution Local Structure-Constrained Image Upsampling

Yang Zhao, Ronggang Wang, *Member, IEEE*, Wenmin Wang, and Wen Gao, *Fellow, IEEE*

***Abstract*— With the development of ultra-high-resolution display devices, the visual perception of fine texture details is becoming more and more important. A method of high-quality image upsampling with a low cost is greatly needed. In this paper, we propose a fast and efficient image upsampling method that makes use of high-resolution local structure constraints. The average local difference is used to divide a bicubic-interpolated image into a sharp edge area and a texture area, and these two areas are reconstructed separately with specific constraints. For reconstruction of the sharp edge area, a high-resolution gradient map is estimated as an extra constraint for the recovery of sharp and natural edges; for the reconstruction of the texture area, a high-resolution local texture structure map is estimated as an extra constraint to recover fine texture details. These two reconstructed areas are then combined to obtain the final high-resolution image. The experimental results demonstrated that the proposed method recovered finer pixel-level texture details and obtained top-level objective performance with a low time cost compared with state-of-the-art methods.**

***Index Terms*— Image upsampling, image upscaling, super-resolution, gradient morphology, local binary pattern.**

I. Introduction

IMAGE upsampling, also known as single-image super-resolution or image upscaling, is a fundamental technique for various applications in machine vision and image processing, such as digital photographs, image editing, high-definition television and ultra-high-definition television, medical image processing, and object recognition. Recently, face super-resolution has also drawn significant attention [48], [49]. The goal of image upsampling is to recover a high-resolution image (HRI) from a low-resolution image (LRI). The means by which to reconstruct high-quality HRIs at a low cost are still elusive.

One classic method of image upsampling is the use of interpolation techniques such as nearest neighbor, bilinear interpolation, bicubic interpolation, and splines [1], [2]. Unfortunately, these methods often produce some unnatural artifacts such as blurring, ringing, and jagged edges. It is recognized that edges are more obvious than random textures for image perception [5]; thus many methods have been proposed to suppress the unnatural artifacts of interpolated edges. Some methods refined the edges by enforcing edge knowledge, such as smooth edge [6] and geometric regularity [7]. Some other methods, such as the triangle interpolation [11], parallelogram interpolation [12], and local segment adaptive interpolation [13], [14] methods, alter the local interpolation grids. Some edge-enhancing algorithms have been proposed to sharpen upscaled edges, including contrast enhancement [15], graph-based interpolation [34], and iterative edge refining interpolation methods [16]. These interpolation-based methods are capable of producing sharp edges but cannot recover high-frequency details.

The example-based methods aim to learn the high-frequency information from the low-resolution (LR)/ high-resolution (HR) example pairs in the dictionary. This type of method was first proposed in [3] and was further developed in [17]–[23] and [35]–[39]. In [19], Yang *et al.* proposed an effective learning method based on sparse representation and achieved impressive results. Many sparse representation based methods were then proposed, such as bilevel dictionary learning [20], beta process joint dictionary learning [38], a statistical prediction model-based method [39], nonlocally centralized sparse representation [35], and a subdictionary-based method [21]. In [24], [25], and [45], local self-example-based methods were introduced by consideration of the similarities of local patches in different scales of the same image. The computational cost of example-based methods is very high as a result of patch-by-patch computing. In [37], Timofte *et al.* proposed an accelerated method in which the LR patch was replaced with the most similar dictionary atom and the corresponding embedding matrix was precomputed. Recently, Zhang *et al.* [47] proposed an efficient super-resolution method by means of multiple linear mapping. These example-based methods can recover some high-frequency components and are effective for areas with a stable structure. However, as noted by Cohen *et al.* [4], these methods still fail to reproduce fine pixel-level high-frequency texture details.

The reconstruction-based approaches impose the similarity constraint between the downsampling of the target HRI and the original LRI. Early multiframe reconstruction-based methods [40] fused multiple LRIs of the same scene to recover an HRI. However, the multiple frames were difficult to align and tended to produce new artifacts. Recently, many

Manuscript received December 4, 2014; revised April 2, 2015, June 2, 2015, and July 3, 2015; accepted July 9, 2015. Date of publication July 14, 2015; date of current version August 18, 2015. This work was supported in part by the National Science Foundation of China under Grant 61370115 and Grant 61402018, in part by the China Post-Doctoral Science Foundation under Grant 2014M550016, and in part by the Shenzhen Peacock Plan. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dacheng Tao.

The authors are with the School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen 518055, China (e-mail: zyknight@mail.ustc.edu.cn; rgwang@pkusz.edu.cn; wangwm@pkusz.edu.cn; wgao@pku.edu.cn).

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

Digital Object Identifier 10.1109/TIP.2015.2456416

methods based on single-image reconstruction have been proposed with the use of different image models or extra constraints. For example, some edge-directed constraints have been introduced in recent reconstruction-based methods, such as the gradient profile prior [8], [9], the estimated gradient map [10], [41], and the displacement field [42]. In [46], nonlocal means and steering kernel regression were adopted as reconstruction constraints. Some de-blurring-based and de-convolution-based methods have also been proposed [31]–[33]. Most reconstruction-based methods have focused on edge constraints and have also failed to recover fine pixel-level textural details.

Motivated by the methods based on edge-directed reconstruction [9], [10], we propose a simple sharpness-preserving reconstruction process to recover the sharp edges. In addition, we first propose the reconstruction of fine texture details by means of the estimated local texture structure constraint.

In this paper, we propose a fast and efficient image upsampling method to recover both sharp edges and fine texture details by means of HR local structure constraints. We make use of the average local difference (ALD) to segment the bicubic-interpolated LRI into a sharp edge area and a texture area, and these two areas are reconstructed separately with specific constraints. For the sharp edge area, an HR gradient map is estimated as an extra constraint to recover the sharp and natural edges; for the texture area, an HR local texture structure map is estimated as an extra constraint to reconstruct the fine texture details. Finally, the reconstructed sharp edge area and texture area are combined to obtain the HRI. The experimental results demonstrate the advantages of the proposed method over state-of-the-art methods.

The main technical contributions of our algorithm are summarized as follows.

1) For reconstruction of the sharp edge area, we propose a simple morphology-based method to estimate an HR gradient map; this estimated HR gradient map is then used as an edge constraint to suppress the blurry and jagged artifacts caused by the traditional interpolation methods. Compared with other gradient constraint based methods proposed in [9] and [10], our morphology-based method has two merits. First, our method does not require that the gradient profile prior be learned from a training dataset as in [9] or calculation of the gradient direction as in [10], so our method is much faster. Second, our method is only applied to the sharp edge area to prevent an over-sharpening artifact.
2) For reconstruction of the texture area, we first propose to estimate the HR local texture structure map by adding high-frequency components to the local binary pattern (LBP) map of the interpolated image and using the estimated HR local texture structure map as an extra constraint to reconstruct the texture details. Our method obtains finer pixel-level texture details than example-based methods.

The rest of this paper is organized as follows. The texture blurring origins of the traditional methods are analyzed in Section II. Section III presents the proposed method in detail. Section IV provides the experimental results to demonstrate the effectiveness of our method, and Section V concludes the paper.

II. Texture Blurring Analysis of Traditional Methods

The local texture is described by the particularity of individual pixels and the regularity of the local region. The pixel values of HR textures are abundant and various. Interpolation methods tend to retain the regularity of the local region, whereas they ignore the recovery of the particularity of individual pixels, which causes texture blurring. Many upsampling methods have recently been proposed, but they still cannot recover fine details at the pixel level.

The back projection (BP) method attempts to refine texture details by means of the following global similarity constraint,

$$\boldsymbol{X} = \underset{\boldsymbol{X}}{argmin} \parallel DH\boldsymbol{X} - \boldsymbol{Y} \parallel_2^2 \tag{1}$$

where $\boldsymbol{X}$ denotes the reconstructed HRI, $\boldsymbol{Y}$ is the LRI, D denotes a downsampling operator, and H is a blurring filter. This optimization problem can be solved by means of gradient descent. In the BP method, the residues between $DH\boldsymbol{X}$ and $\boldsymbol{Y}$ are iteratively added to the upscaled image to increase the high-frequency details. However the LR residues of $(DH\boldsymbol{X} - \boldsymbol{Y})$ must be upscaled before they can be added to the HRI. This upscaling process blurs the residues; thus, the BP-like method cannot recover fine texture details at the pixel level.

Many example-based methods have been proposed that make use of pretrained HR patches to recover image details. These methods can recover some details in areas with a stable structure, but they also tried to reconstruct HR random texture with several fixed prelearned atoms. However, unlike the structural textures or sharp edges with a stable local structure, random texture contains various local gray value distributions and abundant pixel-level details. The atoms in the pretrained dictionary often have stable local structures, such as edge-like patches. Thus, it is difficult to recover the pixel-level random texture details by combining these atoms that lack high-frequency information. Consequently, these example-based methods can recover the details of sharp edges and structural textures but still cannot obtain fine pixel-level random texture details [4]. The term "texture details" as used in this paper mainly denotes the random texture that contains abundant high-frequency pixel-level details. Moreover, these methods attempt to use the relationship between LR patches and LR examples to represent the relationship between HR patches and HR examples. However, it is difficult to set up the bijection between the LR texture and the HR texture, thus this assumption is not sufficiently robust to recover fine texture details.

Some other similarity constraint-based methods, such as the local self-similarity (LSS) based method, have been proposed [25]. In the LSS-based method, a similar patch in the downsampled image of LRI $\boldsymbol{Y}$ $(DH\boldsymbol{Y})$ is used to estimate the HR residues. The merit of LSS is that extra HR examples are not required and the estimated residues do not require upscaling. However, the additional downsampling of

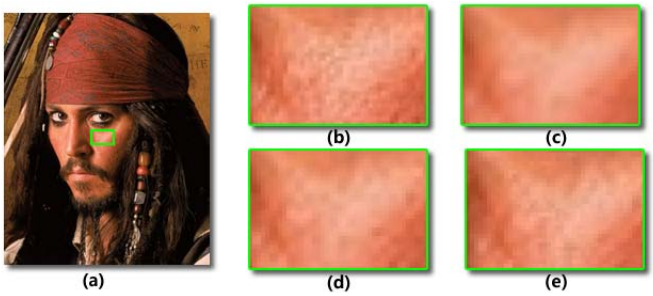

Fig. 1. The close-ups of upsampled facial texture with different methods. (a) the green square shows the selected facial texture, (b) the original HR texture, (c) upsampled texture with bicubic interpolation, (d) upsampled texture with BP method, (e) upsampled texture with our method.

LRI $\boldsymbol{Y}$ ($\boldsymbol{DHY}$) loses high-frequency information in LRI $\boldsymbol{Y}$ and causes texture blurring.

Fig.1 illustrates some close-ups of facial textures upsampled with different methods. The bicubic interpolation blurred the high-frequency details in the original HRI. The BP method recovered some texture details with the global similarity constraint, but it cannot recover the pixel-level details, as mentioned previously. Our proposed method obtains finer and clearer texture details with the HR texture structure constraint.

III. The Proposed Method

The natures of the sharp edge and the texture are quite different. The sharp edge has a stable local structure, whereas the texture has a high degree of randomness and abundant high-frequency information. The goals of upsampling the sharp edge area and the texture area are also different. For the sharp edge area, the objective is to obtain visually sharp edges without unnatural artifacts; for the texture area, the objective is to recover fine pixel-level details. Thus, we divide an image into a sharp edge area and a texture area and propose different methods with which to upscale these two respective areas.

We first upsample the LRI with bicubic interpolation. The sharp edge area of the upsampled image is then extracted, and the remaining area is designated as the texture area. The accurate location of edges is difficult. Fortunately, in this image upsampling scenario, rough extraction of the area containing the sharp edges is sufficient. It is easy to understand that local acute gradient transformations, such as object boundaries and sharp lines, often occur in the sharp edge area. Hence, we apply the ALD as the local sharp edge detector, which is calculated as

$$ALD_{P,R} = \frac{1}{P}\sum_{p=0}^{P-1}|g_p - g_c| \tag{2}$$

where g_c represents the gray value of the center pixel and g_p $(p = 0, 1, 2, \ldots, P-1)$ denotes the gray value of the neighboring pixel within a circle of radius R, and P is the total number of neighboring pixels.

After the ALD of each pixel is obtained, we set a threshold to determine the sharp edge area. The sharp edge area is then upscaled with the proposed HR gradient constrained method, and the remaining texture area is upscaled with the proposed HR local texture structure constrained method. Finally, the two upscaled areas are combined to obtain the final HRI, as illustrated in Fig.2.

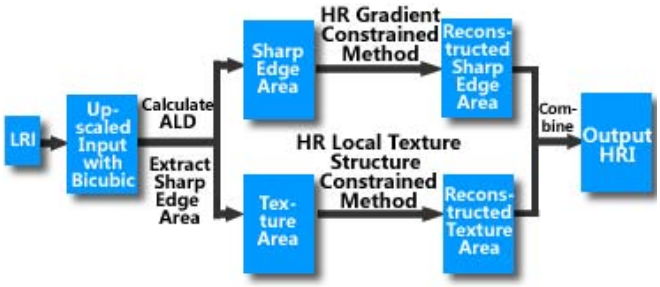

Fig. 2. Architecture of the proposed method.

A. HR Gradient Constrained Method for Sharp Edge Area

Motivated by the previous gradient sharpening methods [8]–[10], we propose an edge-directed upsampling method based on the HR gradient constraint. The main idea is to estimate the HR gradient map by sharpening the gradient map of a bicubic-interpolated image. The estimated HR gradient map is then used as a gradient constraint to recover the sharp edges.

The gradient map is composed of the gradient magnitude of each pixel. For image I, the gradient magnitude at pixel $I(x, y)$ is calculated by

$G(x, y) = \| \nabla I \|_2$, where $\nabla I = (\partial_x I, \partial_y I)$,
$\partial_x I = I(x+1, y) - I(x-1, y)$,
$\partial_y I = I(x, y+1) - I(x, y-1)$.

We estimate the HR gradient map by means of simple morphology processing,

$$\hat{\boldsymbol{G}} = \alpha E(\boldsymbol{G}_b) \tag{3}$$

where $\boldsymbol{G}_b$ denotes the gradient map of a bicubic-upsampled image, E is a morphologic eroding operator, and α denotes a gain factor ($\alpha > 1$) to enlarge the range of the gradient magnitude. In our experiments, the value of α is set as 2 for all test images.

Note that the global erosion process tends to narrow lines and thus some tiny lines may be eliminated. To retain tiny sharp lines, we first upsample the gradient map by a factor of 2 before the morphology process and then downsample it back to the original resolution. Therefore, Eqn. (3) is changed to the following equation:

$$\hat{\boldsymbol{G}} = \alpha DHE(H^T U \boldsymbol{G}_b) \tag{4}$$

where D and U represent the downsampling operator and upsampling operator, respectively, H is a blurring filter, E, $\boldsymbol{G}_b$ and α are defined as in Eqn (3). The jaggy artifact can also be effectively suppressed by applying the above upsampling and downsampling operations.

The erosion E is calculated by

$$E(I)_{withB} = \min_{(s,t)\in B}\{I(x+s, y+t)\} \tag{5}$$

where I denotes an image matrix, B is a structuring element, and (s, t) are the coordinate offsets within B. In this paper,

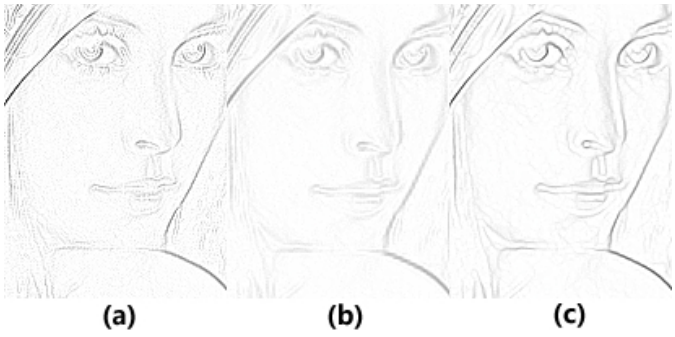

Fig. 3. The gradient maps of different images. (a) gradient map of original HRI, (b) gradient map of bicubic upsampled image, (c) estimated gradient map with our method.

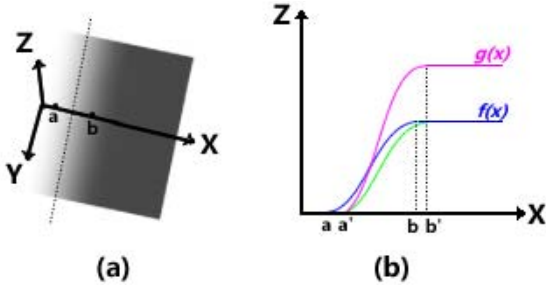

Fig. 4. Illustration of sharpening gradient magnitude. (a) close-up of an upsampled boundary between gray foreground and white background. (Note that X-axis and Y-axis are on the image plane, and Z-axis denotes the gradient magnitude) (b) gradient magnitude curves along the X-axis. (The blue curve is the original gradient magnitude; the green curve denotes the gradient magnitude after the erosion; the red curve represents the final result multiplied by gain factor α).

B is set as a 3×3 local area. The erosion is the minimum of all pixel values in the region of B.

From the different gradient maps of the "lenna" image shown in Fig.3, we can see that our method can estimate sharp and alias-free edges in the gradient map.

In the following paragraphs, we use a simple sharp edge as an example to demonstrate that the proposed morphology-based method can sharpen the gradient magnitude (see Fig.4). Let $f(x)$ be a continuous and monotonic function defined in the interval of (a, b), and let $g(x)$ be the processed function of $f(x)$ by our method. The acutance of $f(x)$ is denoted by

$$\omega_{(a,b)}(f(x)) = \frac{|f(b) - f(a)|}{|b - a|} \tag{6}$$

The larger the $\omega_{(a,b)}(f(x))$, the sharper the function $f(x)$.

First, the size of structuring element B is fixed, thus $|b' - a'|$ equals $|b - a|$.

Second, the gain factor $\alpha > 1$, thereby, the $g(b') = \alpha f(b') > f(b')$. According to $f(b) = f(b')$ and $f(a) = g(a) = 0$, we obtain $|g(b') - g(a')| = |g(b')| > |f(b)| = |f(b) - f(a)|$.

Thereby, $\omega_{(a',b')}(g(x)) > \omega_{(a,b)}(f(x))$, that is, the function $g(x)$ obtained by our proposed method is sharper than $f(x)$.

Fig.5 illustrates the principles of three different methods of gradient magnitude sharpening. Gradient profile prior (GPP) method [9] makes use of a trained gradient prior to sharpening the gradient curve, and Wang *et al.* [10] proposed sharpening the curve by coordinate shifting. Compared with the two methods above, our global morphology-based process has two merits. First, our method does not require that the gradient profile prior be learned from a training dataset or the calculation of the gradient direction, so our method is much faster. Second, the parameter and structuring element B are fixed for all images, so our method is less sensitive to parameter values.

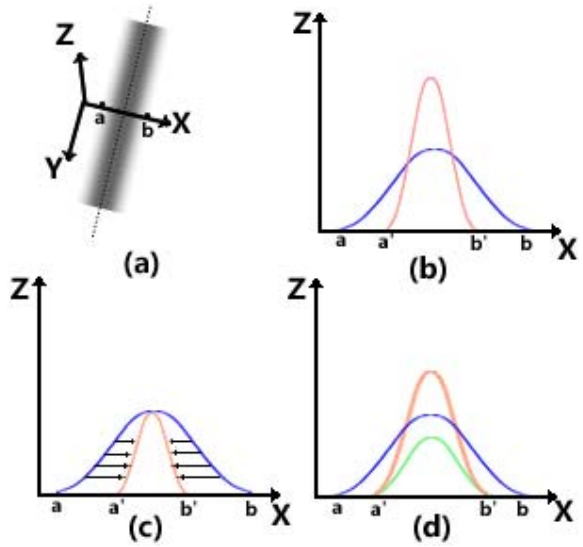

Fig. 5. Comparison of different gradient magnitude sharpening methods. (a) close-up of an upsampled line, (b) gradient magnitude sharpened by GPP [9], (c) gradient magnitude sharpened by wangs method [10], (d) gradient magnitude sharpened by our method. (The blue curves in (b), (c) and (d) are the original gradient magnitudes along the X-axis; the green curve in (d) is the gradient magnitude after erosion; the red curve in (d) denotes the final result multiplied by gain factor α).

After the HR gradient map $\hat{\boldsymbol{G}}$ is estimated, it is used as an additional constraint to recover sharp edges. Consequently, the HRI is reconstructed by minimizing the following objective energy function:

$$\boldsymbol{X} = \underset{\boldsymbol{X}}{argmin} \|DH\boldsymbol{X} - \boldsymbol{Y}\|_2^2 + \lambda \|\boldsymbol{G}_X - \hat{\boldsymbol{G}}\|_2^2 \tag{7}$$

where $\boldsymbol{X}$ denotes the reconstructed HRI, $\boldsymbol{Y}$ denotes the given LRI, $\boldsymbol{G}_X$ is the gradient map of $\boldsymbol{X}$, and λ is a parameter to balance the similarity between $DH\boldsymbol{X}$ and $\boldsymbol{Y}$ and the similarity between $\boldsymbol{G}_X$ and $\hat{\boldsymbol{G}}$. This optimization problem can be solved by gradient descent. Eqn. (7) can then be estimated by updating $\boldsymbol{X}$ using an iterative formula as

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \lambda_1 H^T U(\boldsymbol{Y} - DH\boldsymbol{X}_t) + \lambda_2(\hat{\boldsymbol{G}}^2 - \boldsymbol{G}_X^2) \tag{8}$$

where $\boldsymbol{X}_t$ is the estimated HRI after the t-th iteration, λ_1 and λ_2 are iterative constants, and U, D, and H are defined as in Eqn. (4).

B. HR Local Structure Constrained Method for Texture Area

It is difficult to estimate the HR texture from the LR texture. We thus propose estimation of the HR local texture structure instead of direct estimation of the HR texture. The estimated

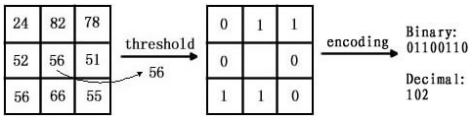

Fig. 6. The illustration of LBP process.

Fig. 7. Illustration of LBP maps. The left column shows some HRIs; the middle column shows the LBP maps of the HRIs; the right column shows the LBP maps of the bicubic upsampled images.

HR texture structure is then used as an extra constraint to recover the fine texture.

We adopt the LBP [26] as the local texture structure descriptor. The LBP is one of the most popular local texture descriptors because of its simplicity and efficiency [27], [28]. In the LBP, the values of the neighboring pixels are converted into binary values by comparing them with the central pixel; these binary values are then combined to characterize the local pattern, as illustrated in Fig.6. Usually, the LBP coding strategy can be described as follows [26],

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{9}$$

where g_p, g_c, R, and P have been defined in Eqn. (2).

In the local texture analysis theory, the local texture can be charaterized by two orthogonal components: the local structure (e.g., the LBP) and the magnitude of the pixel value difference (e.g., the ALD computed by Eqn. (2)). The LBP maps of different images are shown in Fig. 7. We find that the LBP maps of natural HRIs often have abundant high-frequency details, even in smooth areas. However, the LBP maps of the interpolated images lose much spatial frequency information. The loss of the high-frequency information in the local texture structure leads to a blurry texture. To increase the texture details, we propose estimation of the HR LBP map by the addition of high-frequency information into the LBP map of a bicubic upsampled image. The HR LBP map $\hat{\boldsymbol{T}}$ is calculated by

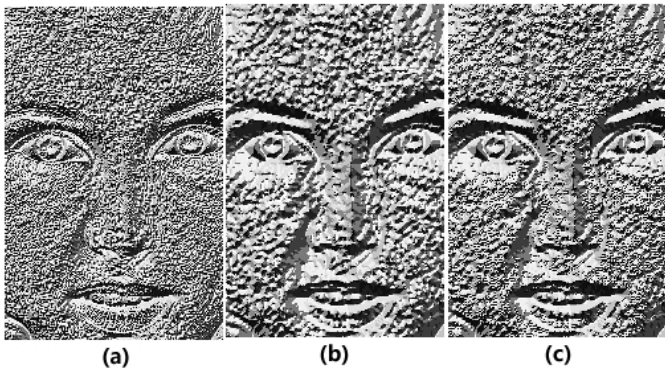

Fig. 8. The LBP maps of different images. (a) LBP map of original HRI, (b) LBP map of bicubic upsampled image, (c) estimated LBP map with our method.

$$\hat{\boldsymbol{T}} = T\boldsymbol{X}^* \tag{10}$$

where T is the LBP encoding operator, and $T\boldsymbol{X}^*$ denotes the LBP map of $\boldsymbol{X}^*$. $\boldsymbol{X}^*$ can be estimated using the result of the following iteration:

$$\begin{aligned} T\boldsymbol{X}_{t+1} = {} & T\boldsymbol{X}_t + \lambda_1 H^T U(T\boldsymbol{Y} - TDH\boldsymbol{X}_t) \\ & + \lambda_2[T(\boldsymbol{X}_t + \gamma(\boldsymbol{X}_t)) - T\boldsymbol{X}_t] \end{aligned} \tag{11}$$

where $\boldsymbol{X}_t$ is the estimated HRI after the t-th iteration, $\boldsymbol{Y}$ is the LRI, λ_1 and λ_2 are iterative constants, $\gamma(\boldsymbol{X}_t)$ denotes the addition of high-frequency components into $\boldsymbol{X}_t$, and U, D, and H are defined as in Eqn. (4). The term $[T(\boldsymbol{X}_t + \gamma(\boldsymbol{X}_t)) - T\boldsymbol{X}_t]$ changes the LBP map by the addition of high-frequency components into $\boldsymbol{X}_t$. Another term $(T\boldsymbol{Y} - TDH\boldsymbol{X}_t)$ is a similarity constraint to maintain the consistency of the new LBP map with the LBP map of LRI $\boldsymbol{Y}$. The initial $\boldsymbol{X}_0$ of the iteration is set as the bicubic upsampled image of LRI.

Here, we simply set $\gamma(\boldsymbol{X})$ as $\lambda' T\boldsymbol{X}$, λ' is a scale number, so that this high-frequency component is consistent with the LBP structure. For example, the LBP map has abundant high-frequency details on the texture area but few details on the sharp edge area. The $\lambda' T\boldsymbol{X}$ term can then add considerable high-frequency information into the texture area without destroying the local structure of the sharp edge area. Note that the parameter λ' is a small number to restrict the magnitude of the high-frequency components. The value of $T\boldsymbol{X}$ ranges from 0 to 255, and the λ' is set as 0.005 to 0.01 in our method. Hence, the LBP map $T\boldsymbol{X}_t$ is changed step-by-step by the addition of this small high-frequency component into $\boldsymbol{X}_t$ in each iteration.

As shown in Fig.8, our estimated LBP map contains more high-frequency texture structure information than the LBP map of the bicubic upsampled image. In addition, the additional high-frequency components do not change the local structure of the sharp edge area.

After the estimated HR LBP map $\hat{\boldsymbol{T}}$ is obtained, it is used as an additional constraint. The HRI texture can then be

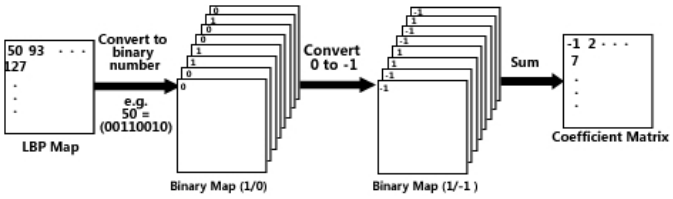

Fig. 9. Converting the LBP map to a coefficient matrix.

reconstructed by minimizing the following objective energy function:

$$\boldsymbol{X} = \underset{\boldsymbol{X}}{argmin} \|DH\boldsymbol{X} - \boldsymbol{Y}\|_2^2 + \lambda \|T\boldsymbol{X} - \hat{\boldsymbol{T}}\|_2^2 \tag{12}$$

where $T\boldsymbol{X}$ is the LBP map of $\boldsymbol{X}$. This optimization problem is difficult to solve because T is the LBP encoding operator rather than a function related to the pixel value. Instead of solving Eqn. (12) directly, we make use of the estimated HR LBP map and ALD calculated in Eqn. (2) to reconstruct the HRI. Although many image reconstruction algorithms were available to utilize the local descriptors [29], [30], these algorithms are usually based on dictionary learning and are very time consuming. Because the LRI is known in the upsampling scenario, we propose a simple method of reconstruction by making use of the LRI. In our method, we first transform the HR LBP map into a coefficient matrix, which is then used to reconstruct the HRI. As illustrated in Fig.9, each number in the LBP map is first converted to an 8-bit binary number, and the LBP map is split into eight binary maps. The number 0 is then converted to -1 in each binary map. Finally, the eight maps are combined to form the coefficient matrix.

Accordingly, the HRI $\boldsymbol{X}$ is reconstructed by utilizing the coefficient matrix $\hat{\boldsymbol{C}}$ and ALD matrix $\hat{\boldsymbol{D}}$,

$$\boldsymbol{X} = H^T U\boldsymbol{Y} + \lambda \hat{\boldsymbol{C}} \bigotimes \hat{\boldsymbol{D}} \tag{13}$$

where ($\bigotimes$) represents the element-wise multiplication of two matrices, and λ is a constant. The $\hat{\boldsymbol{C}}$ can increase the variety of local differences, and $\hat{\boldsymbol{D}}$ restrains the magnitude of these local differences to prevent noise or unreasonable pixel values. Eqn. (13) adjusts the local distribution and enhances the high frequency spatial details according to the HR LBP map $\hat{\boldsymbol{T}}$ by the addition of the $\hat{\boldsymbol{C}} \bigotimes \hat{\boldsymbol{D}}$ term. Because $\hat{\boldsymbol{T}}$ is an estimated HR LBP map in the proposed method, we then utilize the following global similarity constraint to ensure the reconstructed result is still consistent with the known LRI,

$$\boldsymbol{X} = \underset{\boldsymbol{X}}{argmin} \|DH\boldsymbol{X} - \boldsymbol{Y}\|_2^2 \tag{14}$$

The reconstruction of the HR LBP map can be regarded as iteratively adding high-frequency "noise" under the texture structure constraint. Visually fine texture details can be recovered with the use of our estimated HR LBP map. At the meantime, the LBP only reflects the local structure and does not directly relate to the pixel value; thus, the added high-frequency information in the LBP map will not generate noise in the image. In other words, the reconstructed HR LBP map only increases the spatial frequency of the local structure rather than the magnitude of the pixel value. The ALD

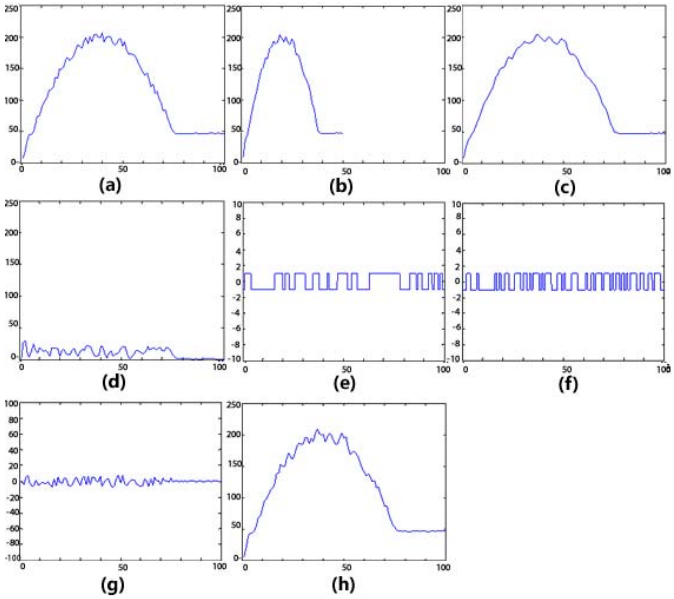

Fig. 10. Illustration of our method in 1D case. (a) original curve, (b) downsampled curve, (c) interpolation reconstructed curve, (d) the ALD value of reconstructed curve in (c), (e) the 1D LBP map of reconstructed curve in (c), (f) the estimated 1D LBP map after adding high frequency components, (g) high frequency details combined by ALD in (d) and estimated 1D LBP map in (f), (h) our reconstructed curve by using the curve in (c) and high frequency details in (g).

describes the magnitude of the local pixel value difference, which is used to control the contribution of the LBP map to the HRI. For example, although the reconstructed LBP map has abundant details in the flat texture area, the ALD of the flat area is small; therefore, noise will not be involved in the flat area. Outside the flat texture area, the ALD value is larger, and thus the estimated LBP map can recover more abundant details.

We use a 1D case to illustrate the way by which our method increases the high-frequency details. As shown in Fig.10, the original curve consists of a series of discrete points in the range of (0, 100). This curve is first downsampled into the range of (0, 50) and then upsampled to (0, 100). A large amount of high-frequency information is lost during downsampling and interpolation-based upsampling (see Fig.10(c)). We estimated the HR 1D LBP map (Fig.10(f)) with the addition of some high-frequency components to the 1D LBP map of the upsampled curve (Fig.10(e)). The similarity constraint in Eqn. (11) means that the downsampling of the estimated HR 1D LBP map should be consistent with the original 1D LBP map of the upsampled curve. The high-frequency details can then be obtained by combining the ALD in Fig.10(d) and the estimated HR 1D LBP map. The HR 1D LBP map decides the sign symbol, and the ALD value restricts the magnitude. Finally, the curve can be reconstructed by adding the high-frequency details to the upsampled curve, as in Fig.10(h). This reconstructed curve has better visual quality than the interpolated curve.

Finally, we analyze the limitations of our HR local texture structure constrained method. First, our method can increase the details in the texture area but cannot recover the

details in the sharp edge area. Because the LBP map has few high-frequency details in the sharp edge area, we did not apply this method in the sharp edge area. Second, the $\gamma(\boldsymbol{X})$ is simply set as $\lambda' T\boldsymbol{X}$ in our method; therefore, the estimated result is partially dependent on the the high-frequency information contained in the original LBP map ($T\boldsymbol{X}_0$). When the texture details are seriously damaged or the LBP map of the original LRI contains scarce high-frequency information, the HR LBP map is difficult to estimate, and the performance of our method is limited. Third, Eqn. (11) is not convergent because a high-frequency component is added in each iteration. Hence, the number of iterations of Eqn. (11) is empirical. In our experiments, we set the iteration number at 50 for all images.

C. Analysis of Computation Time

Given an image of size N and a number of iterations T, the complexity of our HR gradient constrained method is *O(TN)*. The complexity of the LBP encoding process ($R = 1$, $P = 8$) with a fast algorithm is *O(PN)*. Thus the complexity of the HR local texture structure constrained method is *O(TPN)*. The overall complexity of our method is *O(N)* because T and P are constants. The overall memory complexity of our constrained iterative process is also *O(N)*. *O(N)* is used to hold the updated image matrices and residue matrix in each iteration.

The computation time of our method depends on the image size and the number of iterations. In this paper, all experiments are performed with MATLAB on an Intel Core i5-3317U laptop PC. For upsampling of the sharp edge area, the erosion process runs once to obtain the HR gradient map. In each iteration, we perform gradient filtering twice (along x and y directions). For example, the erosion process takes about 80 ms for the 512×512 image "lenna", and each gradient filtering process takes about 5 ms. Each iteration is run 50 times, and the total upsampling process for the edge area for "lenna" takes about 580 ms. For the texture area upsampling, the most computational step is the calculation of the LBP map in each iteration. For example, calculation of the LBP map in each iteration takes 80 to 100 ms for "lenna", and the total time for the texture area upsampling process is about 5 seconds. After taking into account other modules, such as image combing, it takes about 6 seconds to generate the final upsampled "lenna". Note that the proposed algorithm has much room for acceleration, such as implemention of these iterations with C instead of Matlab, or the use of CUDA implementation [16].

D. Summarization of the Proposed Algorithm

The proposed algorithm for HR local structure constrained image upsampling is summarized in **Algorithm 1**.

Algorithm 1 HR Local Structure Constrained Image Upsampling

1: **Input**: LR image $\boldsymbol{Y}$; upsampling factor s; sharp edge area threshold t;
2: **Initialization**: upsample $\boldsymbol{Y}$ with bicubic interpolation by a factor of s and obtain initial HR $\boldsymbol{X_0}$;
3: Calculate the ALD map $\hat{\boldsymbol{D}}$ of $\boldsymbol{X_0}$ by Eqn. (1); divide image into two areas, sharp edge area and texture area using $\hat{\boldsymbol{D}}$ with threshold t;
4: **Upsample sharp edge area with gradient constraint:**
   a) calculate gradient map G_b of X_0;
   b) estimate HR gradient map $\hat{\boldsymbol{G}}$ by Eqn. (4);
   c) use gradient descent to reconstruct HR sharp edge area $\boldsymbol{X}_e^*$, which satisfies the reconstruction constraint Eqn. (7);
5: **Upsample texture area with local texture constraint:**
   a) calculate the LBP map $T\boldsymbol{X_0}$, T denotes the LBP encoding process by Eqn. (9);
   b) estimate HR LBP map $\hat{\boldsymbol{T}}$ by using the Eqn. (11);
   c) convert the $\hat{\boldsymbol{T}}$ into a coefficient matrix $\hat{\boldsymbol{C}}$;
   d) reconstruct HR texture area $\boldsymbol{X}_t^*$ by means of Eqn. (13) and reconstruction constraint Eqn. (14);
6: Combine the HR sharp edge area $\boldsymbol{X}_e^*$ and the HR texture area $\boldsymbol{X}_t^*$ to get the HR image $\boldsymbol{X}^*$;
7: **Output**: the final HR image $\boldsymbol{X}^*$;

IV. Experimental Results

We test the proposed method on a set of natural images downloaded from the web and from the Berkeley image dataset [44].[1] Color images are first converted from RGB to YUV. The proposed method is then applied only on the Y (intensity) component, and bicubic interpolation is used for the U and V components. In our experiments, the input LRIs are obtained by downsampling the original HRIs with bicubic interpolation, and the LRIs are then upsampled to their original size, respectively, with different methods for the performance comparison. The upsampling factor in our experiment is set as 2 and 3. In this paper, our method is compared with bicubic interpolation, the ICBI method [16], Shan's deconvolution-based method [31], Wang's edge-sharpening method [10], and several example-based methods including ScSR [19], ASDS [21], Zeyde's method [36], BP-JDL [38], and ANR [37]. In our method, the threshold to extract the sharp edge area is set to 1.5 times the mean value of the ALD of the entire image. The iterative constants (λ_1, λ_2) in Eqn. (8) and Eqn. (11) are set to (0.2, 0.004) and (0.1, 0.01) respectively, and the iterative constant λ in Eqn. (13) is set to 1.

Fig.11 shows some upsampled edges of the "lenna" image with different methods. By comparing the close-ups of these edges, we can make the following observations. First, bicubic interpolation produces blurry edges. Although the ICBI method can enhance lines, the recovered edges are still blurry. Second, Shan's method can sharpen the edges, but it also generates some unnatural artifacts. Third, Wang's method, ScSR, and ASDS can recover sharp edges. Finally, our method produces sharper edges than the other methods. Fig.12 compares the upsampled results of the "zebra" image with those of various other methods. By comparing the close-ups of the stripe and meadow, our method produces the sharpest

[1]The results and demo video are available at http://zycv.890m.com/zyProjectPages/proj1.html

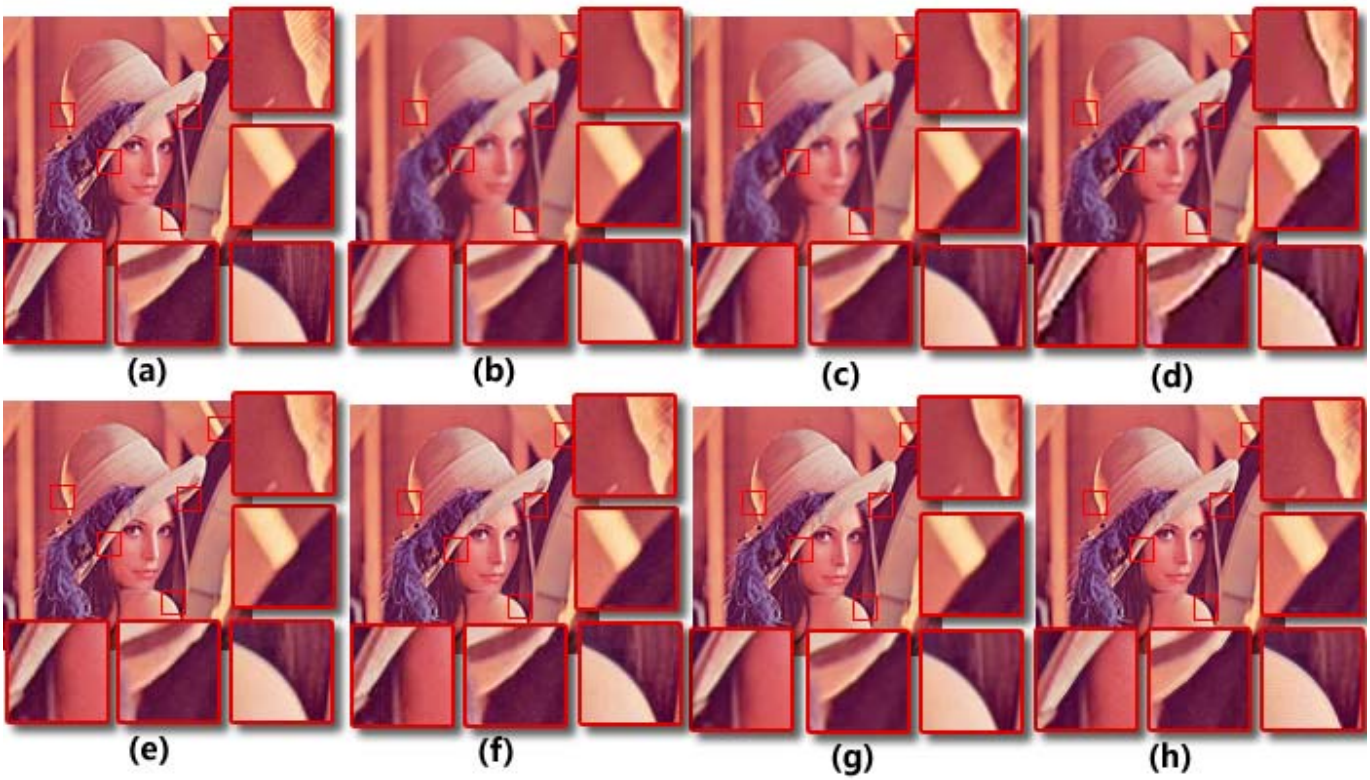

Fig. 11. Upsampled edges of the "lenna" image (2X), (a) the original HRI, (b) with the bicubic interpolation, (c) with ICBI [16], (d) with Shan's method [31], (e) with Wang's method [10], (f) with ScSR [19], (g) with ASDS [21], (h) with our method.

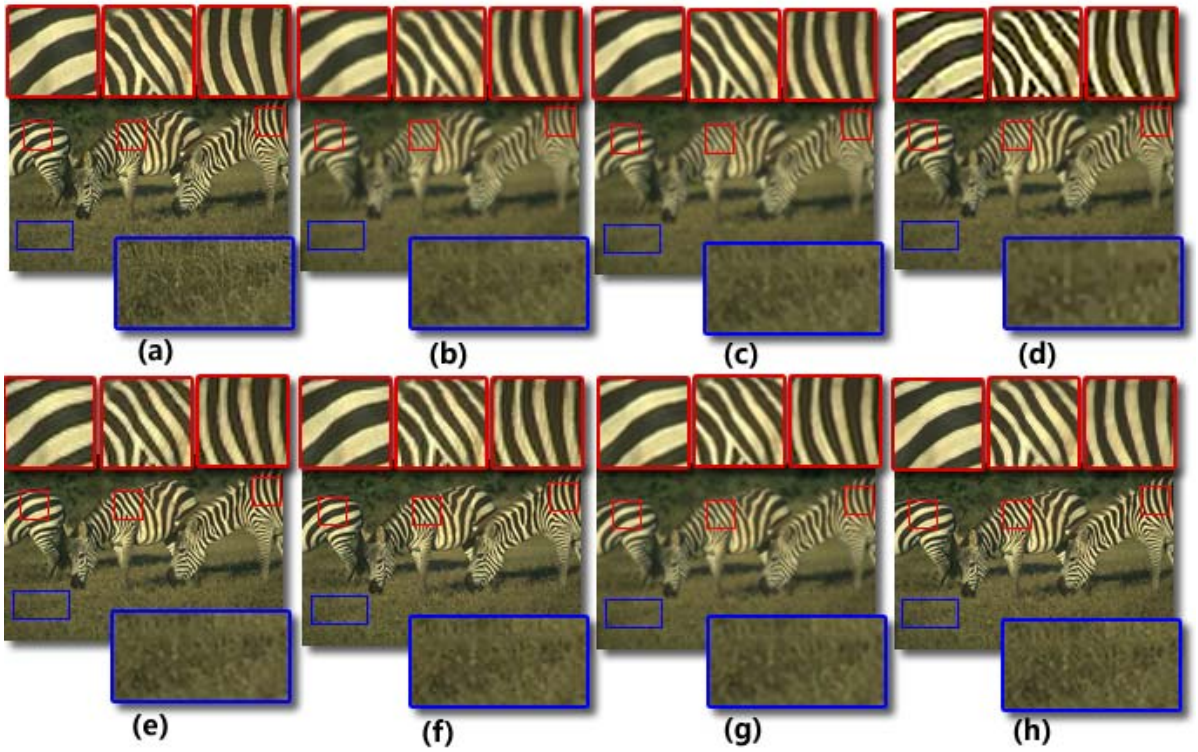

Fig. 12. Upsampled results of the "zebra" image (2X), (a) the original HRI, (b) with the bicubic interpolation, (c) with ICBI [16], (d) with Shan's method [31], (e) with Wang's method [10], (f) with ScSR [19], (g) with ASDS [21], (h) with our method.

edges and the clearest texture details in the meadow area. Our method is less sensitive to parameters than Wang's HR gradient-constrained method. Although the two learning-based methods of ScSR and ASDS can also recover natural and sharp edges, the edges obtained with the proposed method are still sharper and clearer.

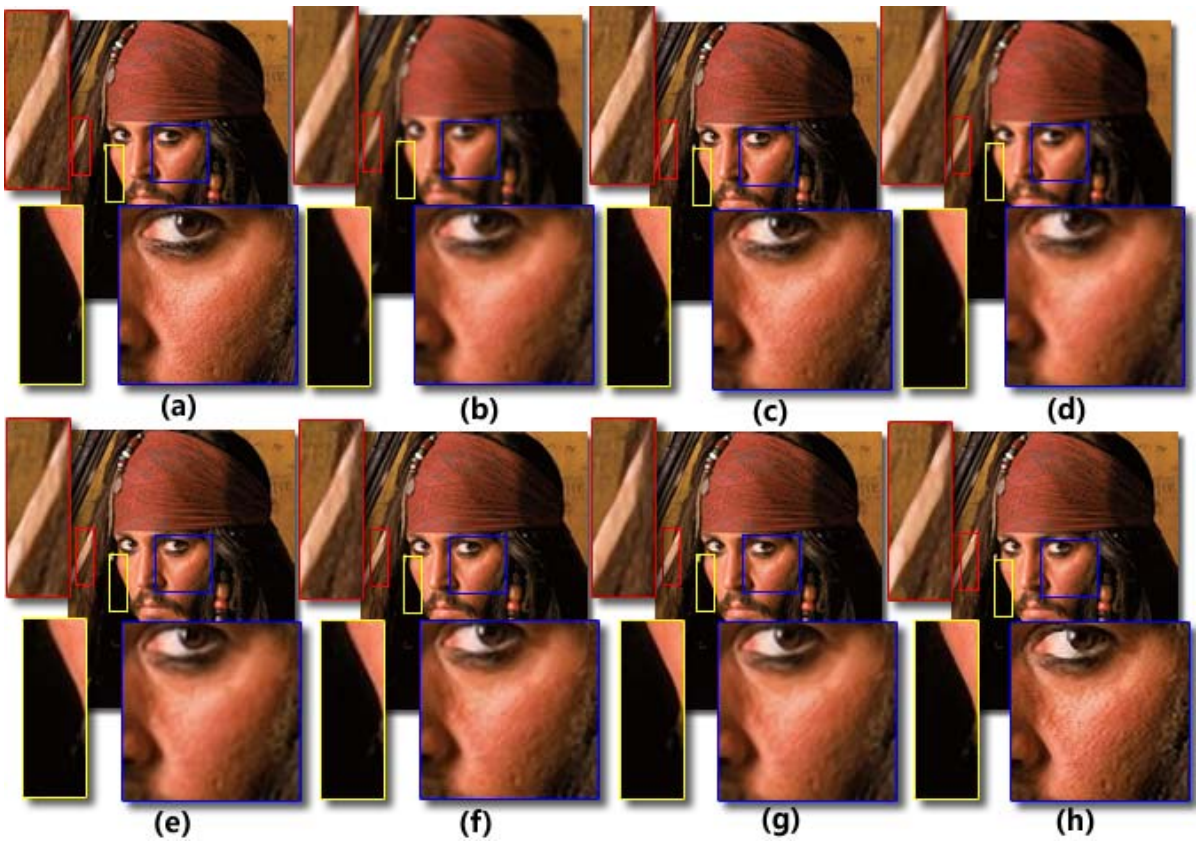

Fig. 13. Upsampled results of the "pirate" image (2X), (a) the original HRI, (b) with the bicubic interpolation, (c) with ScSR [19], (d) with Zeyde's method [36], (e) with ASDS [21], (f) with ANR [37], (g) with BP-JDL [38], (h) with our method.

Fig.13 illustrates the upsampled results of the "pirate" image with the recent state-of-the-art example-based methods and our proposed method. We first compare the close-ups of the sharp edge area. All of these example-based methods can recover sharp edges. The BP-JDL method and our proposed method produce sharper edges than the other methods. Note that our proposed method is much faster than the BP-JDL method. We then compare the close-ups of the texture area. Our proposed HR local texture constrained method can reproduce finer texture details than these state-of-the-art example-based methods.

Fig.14 and 15 illustrate the upsampled local textures with these example-based methods and our proposed method. By comparing the further close-ups of a local texture patch, we find that the HR textures have fine pixel-level details, even in the flat background area. The example-based methods can reconstruct the sharp edges and fine texture, but they still fail to obtain pixel-level texture details. By making use of the HR texture structure constraint, our method can recover better high-frequency details than the other methods. The LBP maps of the texture area are shown in the red squares in Fig.14 and 15. These LBP maps can be used to visualize the local texture structures of the upsampled results. By comparing the LBP maps of different methods, we obtain the following findings. The LBP maps of the original HRIs have abundant pixel-level high-frequency details. However, the LBP maps of the upsampled images with bicubic interpolation or example-based methods still lose a great deal of spatial frequency information. Our proposed method enhances the spatial frequency information of the local structure, and therefore recovers finer pixel-level texture details.

Table I lists the image reconstruction performance (PSNR, SSIM index [44], and running time) of the various methods with an upsampling factor of 2. By comparing the PSNR values, we find that our proposed method, the BP-JDL method, and the ANR method show better performance than the other methods. Although our method aims to refine the visual quality, the average PSNR value of our method is still comparable to or slightly better than these state-of-the-art example-based methods. As mentioned earlier, the performance of the proposed HR local texture constrained method is partially dependent on the high-frequency information contained in the original LBP map. When the original LRI contains scarce pixel-level high-frequency information, it is difficult to reproduce abundant pixel-level high-frequency details with the estimated HR LBP map. Our method thus performs better on images with abundant high-frequency texture details but slightly worse on images with less high-frequency information. However, our method still achieves the best performance on average because most of the natural images contain various high-frequency details. By using the SSIM index [43] to measure the structural similarity of the upsampled images with the original HRI, our method performs better than the others in almost all cases.

The running times of these methods are also listed in Table I. The ANR method is the fastest, and our method is

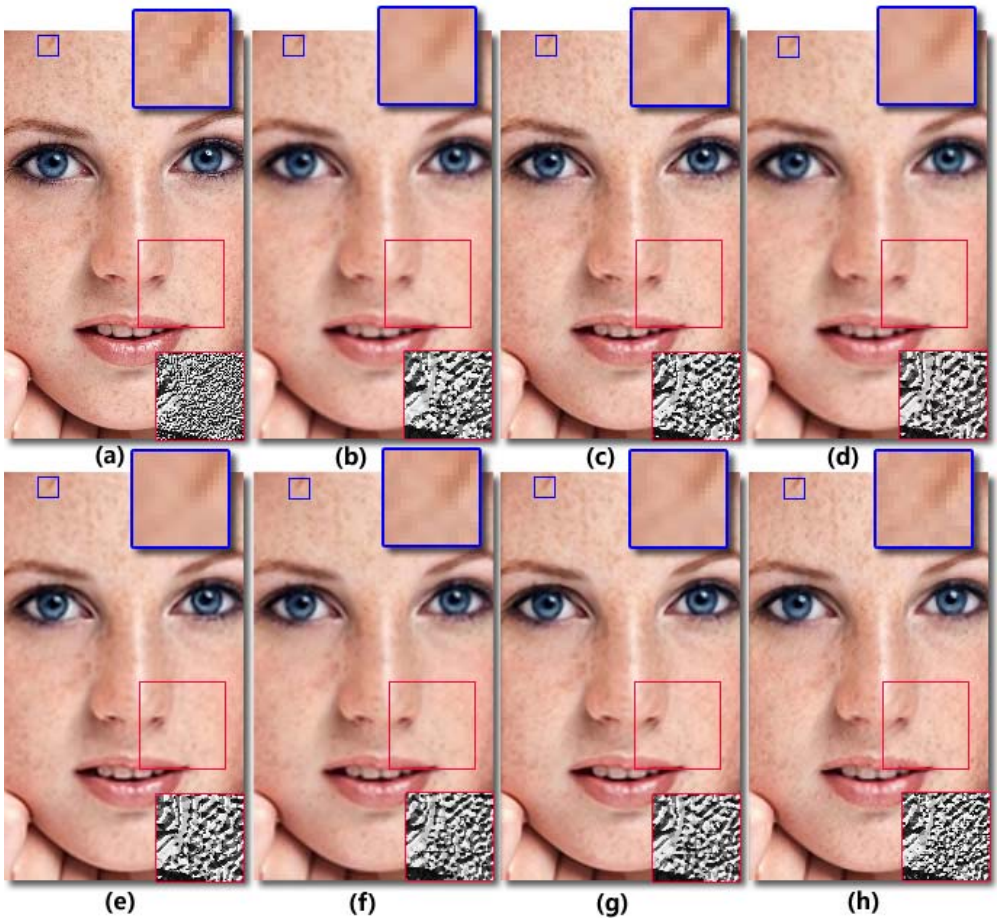

Fig. 14. Upsampled results of the "face" image (2X), (a) the original HRI, (b) with the bicubic interpolation, (c) with ScSR [19], (d) with Zeyde's method [36], (e) with ASDS [21], (f) with ANR [37], (g) with BP-JDL [38], (h) with our method.

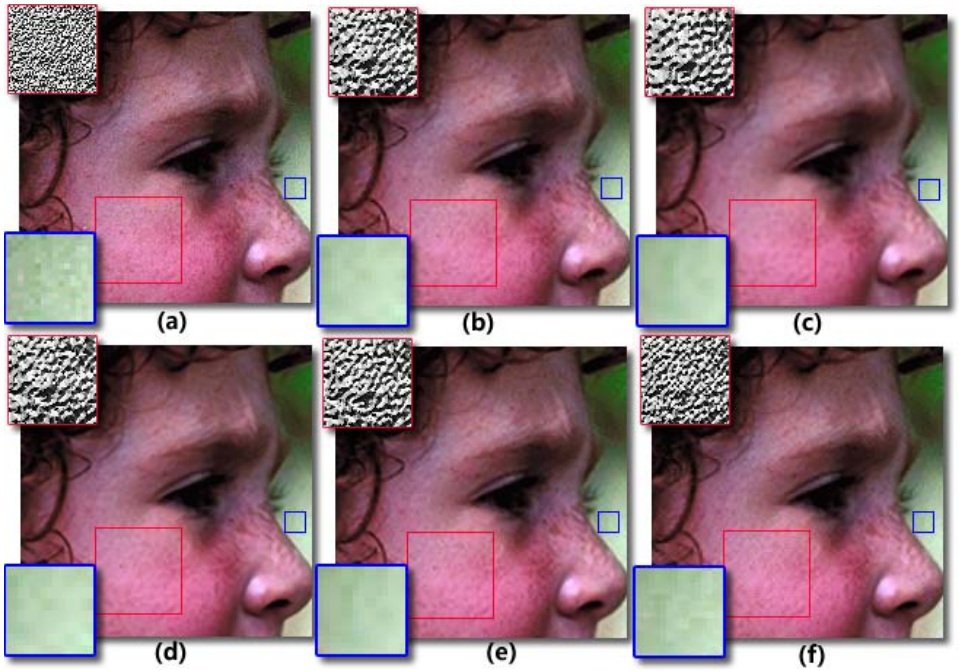

Fig. 15. Upsampled results of the "girl" image (2X), (a) the original HRI, (b) with ScSR [19], (c) with ASDS [21], (d) with ANR [37], (e) with BP-JDL [38], (f) with our method.

the second-fastest. The ANR method saves running time by making use of the precomputed projection matrix of the nearest atom as an approximation to avoid calculating the projection matrix of each input. However, the accuracy of the approximation mainly depends on the similarity between the LR input and its nearest dictionary atom. This inexact

TABLE I

EXPERIMENTAL RESULTS OF DIFFERENT METHODS (UPSAMPLING FACTOR IS 2)

	lenna	*girl*	*pirate*	*face*	*zebra*	*15004*	*22093*	*23025*	*24004*	*253027*	*311081*	*56028*	*223061*	*Average*
Bicubic	35.42	34.58	33.29	34.64	27.26	27.44	26.73	31.29	25.60	25.27	24.86	26.50	26.53	29.18
	0.991	0.862	0.984	0.903	0.875	0.886	0.851	0.892	0.808	0.829	0.742	0.794	0.791	0.862
	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ICBI	31.97	33.03	30.57	32.47	25.23	25.96	25.26	29.81	24.32	24.04	23.48	25.08	25.44	27.43
	0.966	0.832	0.953	0.882	0.836	0.856	0.814	0.861	0.773	0.795	0.685	0.752	0.754	0.827
	89.1	16.6	70.5	35.9	82.3	74.2	83.8	55.2	75.4	86.4	72.7	76.7	60.4	61.7
Shan's	31.20	32.15	29.18	31.28	24.34	24.05	24.29	28.41	23.51	22.68	22.36	23.81	24.28	26.27
	0.941	0.779	0.887	0.826	0.758	0.742	0.723	0.772	0.688	0.680	0.581	0.611	0.631	0.740
	15.8	2.8	7.2	6.5	5.9	5.9	6.0	6.1	6.0	6.3	6.4	6.2	6.3	6.7
ScSR	36.75	35.37	34.40	35.36	28.49	28.46	27.40	31.90	26.06	26.07	25.35	27.33	27.13	30.00
	0.994	0.875	0.990	0.913	0.899	0.913	0.872	0.908	0.834	0.853	0.778	0.830	0.828	0.883
	480.9	141.3	329.4	234.5	279.6	286.4	283.9	325.7	307.3	317.3	299.2	299.9	351.2	302.8
ASDS	36.18	34.52	33.81	34.66	28.53	28.40	27.42	31.95	26.24	26.31	25.30	27.39	26.83	29.81
	0.926	0.846	0.888	0.892	0.881	0.902	0.857	0.887	0.816	0.837	0.771	0.811	0.806	0.855
	819.9	187.4	554.6	301.8	574.5	615.4	672.0	506.4	612.2	605.1	692.9	641.0	614.5	569.0
Zeyde's	35.37	34.82	32.89	34.20	27.18	27.23	26.61	31.34	25.47	25.08	24.69	26.38	26.33	29.05
	0.991	0.862	0.880	0.890	0.874	0.876	0.841	0.878	0.796	0.815	0.731	0.781	0.778	0.845
	13.8	4.4	10.5	6.6	10.6	10.7	10.6	10.8	10.8	10.6	10.8	10.8	10.6	10.1
BPJDL	37.13	35.44	34.63	35.43	29.12	**29.03**	27.47	32.02	26.17	26.43	25.53	**27.61**	**27.31**	30.25
	0.991	0.788	0.990	0.888	0.900	**0.917**	0.867	0.900	**0.841**	0.856	0.758	0.835	0.833	0.874
	995.8	294.7	1267.1	906.1	589.6	1097.3	1109.7	973.6	711.9	599.7	596.5	818.1	993.6	842.6
ANR	37.26	35.66	34.16	34.92	**29.13**	28.82	27.52	**32.18**	26.28	**26.54**	25.57	27.41	27.20	30.20
	0.997	0.883	0.906	0.903	**0.912**	0.914	0.874	0.900	0.835	0.859	0.793	0.828	0.831	0.880
	2.5	**0.9**	**2.2**	**1.3**	**1.6**	**1.6**	**1.7**	**1.7**	**1.8**	**1.6**	**1.7**	**1.8**	**1.7**	**1.7**
Ours	**37.28**	**35.83**	**34.84**	**35.64**	28.79	28.56	**27.59**	32.04	**26.30**	26.46	**25.70**	27.43	27.22	**30.28**
	0.997	**0.884**	**0.996**	**0.916**	0.906	**0.917**	**0.880**	**0.912**	**0.841**	**0.871**	**0.805**	**0.837**	**0.835**	**0.892**
	6.3	2.6	5.7	3.6	4.9	5.1	5.0	5.0	5.1	4.9	5.0	5.2	4.9	4.8

Three results are reported in each cell: Top - PSNR (dB), Middle - SSIM index, Bottom - running time (s). Note that the best result for each image is highlighted and the second best result for each image is underlined.

TABLE II

EXPERIMENTAL RESULTS OF DIFFERENT METHODS (UPSAMPLING FACTOR IS 3)

	lenna	*girl*	*pirate*	*face*	*zebra*	*15004*	*22093*	*23025*	*24004*	*253027*	*311081*	*56028*	*223061*	*Average*
Bicubic	32.12	32.89	30.37	32.54	24.14	24.48	24.45	29.00	23.69	22.62	22.60	24.26	24.75	26.76
	0.957	0.800	0.924	0.859	0.766	0.772	0.751	0.814	0.706	0.710	0.556	0.655	0.670	0.764
	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ICBI	27.45	29.85	26.56	28.83	21.60	22.30	22.53	26.78	22.05	21.27	20.89	22.01	23.26	24.26
	0.883	0.740	0.847	0.813	0.670	0.689	0.676	0.746	0.646	0.641	0.449	0.565	0.594	0.689
	139.1	28.0	102.4	54.3	122.7	98.9	114.3	81.6	106.8	125.2	90.3	109.4	78.5	96.2
Shan's	28.16	30.26	26.74	28.96	22.07	22.14	22.60	26.64	22.16	21.32	20.93	22.12	23.20	24.41
	0.892	0.738	0.824	0.792	0.675	0.651	0.650	0.710	0.630	0.609	0.421	0.532	0.559	0.668
	12.9	3.9	18.1	11.3	15.9	16.1	15.9	16.0	16.0	16.1	16.2	15.9	16.0	14.6
ScSR	32.97	33.10	31.47	32.81	24.98	25.66	25.41	30.16	23.96	23.26	23.28	25.09	25.38	27.50
	0.962	0.811	0.937	0.879	0.794	0.827	0.794	0.838	0.745	0.742	0.631	0.696	0.722	0.798
	490.2	146.9	344.2	227.5	275.8	294.7	276.9	337.7	313.7	325.3	289.2	294.1	367.9	306.4
ASDS	33.17	33.56	31.65	33.61	**25.20**	**25.89**	**25.51**	30.22	**24.26**	23.28	23.14	25.22	25.28	27.69
	0.903	0.818	0.892	0.873	0.809	0.831	0.791	0.842	0.742	0.750	0.625	0.709	0.718	0.792
	791.6	194.1	516.3	307.4	548.3	596.5	534.3	477.1	579.2	527.1	627.4	562.0	543.3	523.4
Zeyde's	32.08	32.89	30.45	32.81	24.05	24.48	24.51	29.32	23.67	22.54	22.50	24.19	24.52	26.77
	0.957	0.800	0.815	0.882	0.763	0.766	0.747	0.813	0.703	0.703	0.549	0.652	0.656	0.754
	10.9	3.9	8.2	4.9	6.9	6.8	6.9	6.9	7.0	6.8	6.8	6.9	6.9	6.9
BPJDL	33.28	33.44	31.15	33.22	25.03	25.21	24.96	29.55	24.14	23.21	22.99	24.90	25.15	27.40
	0.967	0.705	0.945	0.836	0.791	0.805	0.761	0.815	0.732	0.730	0.554	0.687	0.692	0.771
	1195.2	404.7	684.6	462.9	828.2	590.2	598.4	601.3	599.7	609.1	603.1	595.0	608.5	644.6
ANR	33.58	33.62	31.47	**33.74**	25.12	25.46	25.16	30.05	24.16	23.16	23.00	24.98	25.00	27.58
	0.971	0.822	0.843	**0.897**	0.806	0.814	0.781	0.840	0.737	0.744	0.610	0.700	0.706	0.790
	1.8	**0.6**	**1.4**	**0.9**	**1.2**	**1.2**	**1.3**	**1.1**	**1.2**	**1.1**	**1.1**	**1.2**	**1.3**	**1.3**
Ours	33.29	**33.67**	**31.73**	33.64	25.12	25.78	25.44	**30.25**	24.21	**23.38**	**23.62**	**25.32**	**25.47**	**27.76**
	0.973	**0.823**	**0.946**	0.885	**0.817**	**0.837**	**0.799**	**0.851**	**0.750**	**0.761**	**0.647**	**0.716**	**0.727**	**0.810**
	8.9	2.7	6.1	3.9	5.3	5.4	5.4	5.4	5.3	5.3	5.3	5.4	5.4	5.4

Three results are reported in each cell: Top - PSNR (dB), Middle - SSIM index, Bottom - running time (s). Note that the best result for each image is highlighted and the second best result for each image is underlined.

estimation may lead to the loss of some high-frequency details in the texture area, whereas our method can generate finer pixel-level texture details than the ANR method by means of the estimated HR LBP maps as an extra constraint. In addition, most of the time required for our method is consumed in the calculation of LBP maps. Therefore, our method is slightly slower than the ANR method, but it can recover better texture details and is still cost-effective.

We also test the proposed method for 3X magnification, and the experimental results for different methods are listed in Table II. The following observations can be made. First, the performance of these methods decreases when the upsampling factor increases. Second, although the performance of our method also degrades, it still obtains the highest average PSNR value and SSIM index among the methods. Third, our method is faster than other methods except the ANR method.

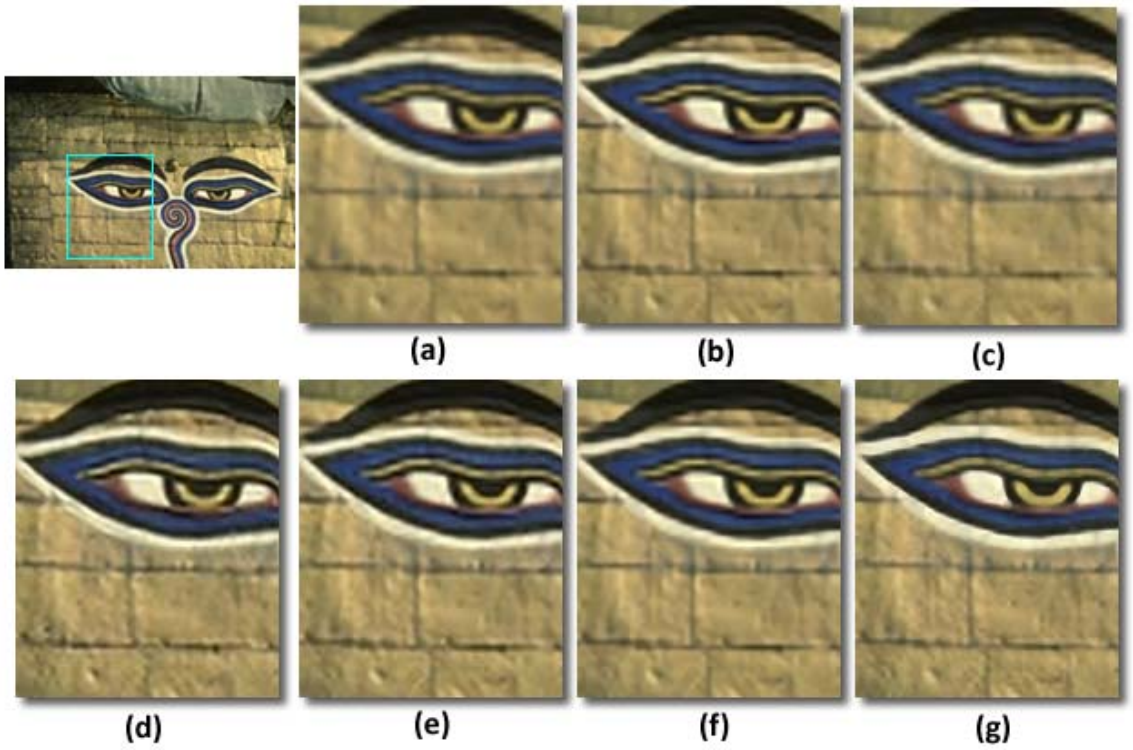

Fig. 16. Upsampled results of the "BDS_56028" image (3X), (a) with the bicubic interpolation, (b) with ScSR [19], (c) with Zeyde's method [36], (d) with ASDS [21], (e) with ANR [37], (f) with BP-JDL [38], (g) with our method.

Fig.16 illustrates the upsampled results of the different methods under 3X magnification. The ASDS method produces slightly sharper edges than the other example-based methods, but it amplifies some of the noise in the flat white area at the same time. By comparing the sharp edges and textures of the bricks, we can see that our method recovers sharper edges and finer texture than these state-of-the-art example-based methods.

The performance of our method degrades rapidly when the upsampling factor is increased to 4 or greater. Our method attempts to recover fine texture details by estimating the HR local texture structures. The LRIs for our experiment are obtained by downsampling the HRIs. Therefore, the original LBP maps of these downsampled LRIs contain very few high-frequency details if the downsampling factor is large. In this case, the proposed HR local texture structure constrained method cannot recover pixel-level texture details. The extreme case is when the estimated HR LBP maps contain no extra high-frequency information and the HR local texture constrained method then degrades to the traditional BP method. Therefore, the proposed method is suitable for recovering fine pixel-level texture details with small upsampling factors, such as 2X and 3X magnification, which are sufficient for typical image upsampling scenarios.

V. Conclusion

With the development of ultra-HR display devices, fine texture details are becoming more and more important for visual perception. In this paper, we propose a fast and efficient method for image upsampling by means of HR local structure constraints. We divide the image into a sharp edge area and a texture area and use an HR gradient constraint for reconstruction of the sharp edge area and an HR local texture structure constraint for reconstruction of the texture area. The experimental results demonstrate that our method recovers finer pixel-level high-frequency texture details and obtains top-level objective performance compared with state-of-the-art methods in typical upscaling factors.

Acknowledgment

The authors would like to sincerely thank A. Giachetti, Q. Shan, L. F. Wang, W. S. Dong, J. C. Yang, R. Zeyde, L. He, and R. Timofte for sharing the source codes of the ICBI, Shan's, Wang's, ASDS, ScSR, Zeyde's, BP-JDL, and ANR methods. They also would like to thank Dr. Huifang Sun for useful advice during this work.

References

[1] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[2] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1049–1075, Nov. 1999.

[3] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *Int. J. Comput. Vision*, vol. 40, no. 1, pp. 25–47, Oct. 2000.

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

[5] D. C. Van Essen, C. H. Anderson, and D. J. Felleman, "Information processing in the primate visual system: An integrated systems perspective," *Science*, vol. 255, no. 5043, pp. 419–423, Jan. 1992.

[6] S. Dai, M. Han, W. Xu, Y. Wu, and Y. Gong, "Soft edge smoothness prior for alpha channel super resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[7] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.
[8] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Trans. Graph.*, vol. 26, no. 3, Jul. 2007, Art. ID 95.
[9] J. Sun, Z. Xu, and H. Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
[10] L. Wang, S. Xiang, G. Meng, H. Wu, and C. Pan, "Edge-directed single-image super-resolution via adaptive gradient magnitude self-interpolation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 8, pp. 1289–1299, Aug. 2013.
[11] D. Su and P. Willis, "Image interpolation by pixel-level data-dependent triangulation," *Comput. Graph. Forum*, vol. 23, no. 2, pp. 189–201, Jun. 2004.
[12] Q. Wang and R. K. Ward, "A new orientation-adaptive interpolation method," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 889–900, Apr. 2007.
[13] C. M. Zwart and D. H. Frakes, "Segment adaptive gradient angle interpolation," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 2960–2969, Aug. 2013.
[14] X. Liu, D. Zhao, R. Xiong, S. Ma, W. Gao, and H. Sun, "Image interpolation via regularized local linear regression," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3455–3469, Dec. 2011.
[15] Q. Wang, R. Ward, and J. Zou, "Contrast enhancement for enlarged images based on edge sharpening," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2. Sep. 2005, pp. II-758–II-762.
[16] A. Giachett and N. Asuni, "Real-time artifact-free image upscaling," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2760–2768, Oct. 2011.
[17] J. Sun, N. N. Zheng, H. Tao, and H.-Y. Shum, "Image hallucination with primal sketch priors," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2003, pp. II-729–II-736.
[18] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun./Jul. 2004, pp. I-275–I-282.
[19] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
[20] J. Yang, Z. Wang, Z. Lin, X. Shu, and T. Huang, "Bilevel sparse coding for coupled feature spaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2360–2367.
[21] W. Dong, D. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.
[22] J.-H. Lee, J.-O. Kim, J.-W. Han, K.-S. Choi, and S.-J. Ko, "Edge-oriented two-step interpolation based on training set," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1848–1855, Aug. 2010.
[23] K. K. Chua and Y. H. Tay, "Enhanced image super-resolution technique using convolutional neural network," in *Advances in Visual Informatics*. New York, NY, USA: Springer-Verlag, 2013, pp. 157–164.
[24] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 349–356.
[25] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Trans. Graph.*, vol. 30, no. 2, pp. 12–23, Apr. 2011.
[26] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
[27] Z. H. Guo, D. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.
[28] X. Y. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.
[29] P. Weinzaepfel, H. Jegou, and P. Perez, "Reconstructing an image from its local descriptors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 337–344.
[30] E. d'Angelo, L. Jacques, A. Alahi, and P. Vandergheynst, "From bits to images: Inversion of local binary descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 874–887, May 2014.
[31] Q. Shan, Z. Li, J. Jia, and C. Tang, "Fast image/video upsampling," *ACM Trans. Graph.*, vol. 27, no. 5, Dec. 2008, Art. ID 153.
[32] T. Michaeli and M. Irani, "Nonparametric blind super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 945–952.
[33] N. Efrat, D. Glasner, A. Apartsin, B. Nadler, and A. Levin, "Accurate blur models vs. image priors in single image super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2832–2839.
[34] X. Liu, D. Zhao, J. Zhou, W. Gao, and H. Sun, "Image interpolation via graph-based Bayesian label propagation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1084–1096, Mar. 2014.
[35] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1620–1630, Apr. 2013.
[36] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*. New York, NY, USA: Springer-Verlag, 2010, pp. 711–730.
[37] R. Timofte, V. De Smet, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1920–1927.
[38] L. He, H. Qi, and R. Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 345–352.
[39] T. Peleg and M. Elad, "A statistical prediction model based on sparse representations for single image super-resolution," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2569–2582, Jun. 2014.
[40] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion, and transparency," *J. Vis. Commun. Image Represent.*, vol. 4, no. 4, pp. 324–355, Dec. 1993.
[41] H. Xu, G. Zhai, and X. Yang, "Single image super-resolution with detail enhancement based on local fractal analysis of gradient," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 10, pp. 1740–1754, Oct. 2013.
[42] L. Wang, H. Wu, and C. Pan, "Fast image upsampling via the displacement field," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5123–5135, Dec. 2014.
[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[44] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, Jul. 2001, pp. 416–423.
[45] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with multiscale similarity learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 10, pp. 1648–1659, Oct. 2013.
[46] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with non-local means and steering kernel regression," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4544–4556, Nov. 2012.
[47] K. Zhang, D. Tao, X. Gao, X. Li, and Z. Xiong, "Learning multiple linear mappings for efficient single image super-resolution," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 846–861, Mar. 2015.
[48] Y. Li, C. Cai, G. Qiu, and K.-M. Lam, "Face hallucination based on sparse local-pixel structure," *Pattern Recognit.*, vol. 47, no. 3, pp. 1261–1270, Mar. 2014.
[49] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "A comprehensive survey to face hallucination," *Int. J. Comput. Vision*, vol. 106, no. 1, pp. 9–30, Jan. 2014.

Yang Zhao received the B.E. and Ph.D. degrees from the Department of Automation, University of Science and Technology of China, in 2008 and 2013, respectively.

He has been a Post-Doctoral Fellow with the School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, China, since 2013.

His research interests include pattern recognition and image processing.

Ronggang Wang (M'12) received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences. He was a Research Staff Member with Orange (France Telecom) Labs from 2006 to 2010. He is currently an Associate Professor with the Peking University Shenzhen Graduate School.

He has authored over 50 papers in international journals and conferences, and holds 40 patents. His research interest is on video coding and processing. He has done many technical contributions to ISO/IEC MPEG and China AVS. He led the MPEG Internet Video Coding (IVC) standard and has served as the MPEG IVC AHG Co-Chair since 2012. He is currently the AVS Implementation Sub-Group Co-Chair.

Wenmin Wang received the B.E. degree in computer application from the Harbin University of Science and Technology, in 1983, and the M.E. and Ph.D. degrees in computer architecture from the Harbin Institute of Technology, China, in 1986 and 1989, respectively. Since 1992, he has 18 years of oversea industrial experience in Japan and America. He came back to academia in 2009, and is currently a Professor with the School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, China. His current research interests include video analysis, multimedia retrieval, artificial intelligence, and machine learning.

Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Japan, in 1991.

He was a Professor of Computer Science with the Harbin Institute of Technology from 1991 to 1995, and a Professor with the Institute of Computing Technology, Chinese Academy of Sciences. He is currently a Professor of Computer Science with Peking University Shenzhen Graduate School, China. He has published extensively, including five books and over 600 technical articles in refereed journals and conference proceedings in image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. He served on the Editorial Board of several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, the *EURASIP Journal of Image Communications*, and the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.