Manifold-Regularization Super-Resolution Image Reconstruction

Xian-Hua Zeng[12*], Su-Li Hou[12]

[1] College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[2] Chongqing Key Lab of Computational Intelligence, Chongqing 400065, China
  zengxh@cqupt.edu.cn, 1271075315@qq.com

Abstract. The conventional sparse coding-based super-resolution image reconstruction methods cannot preserve the image local smoothing structures, i.e., has not considered the smoothness between patches in the reconstruction process. In this paper, we introduce the manifold regularization to constrain the image patches lay on the intrinsic smoothing manifold, and incorporate the image nonlocal self-similarity into sparse representation model to improve the accuracy of the sparse coefficients, thus more accurate high-resolution patches can be obtained and used to reconstruct the better high-resolution image. Accordingly a novel Manifold-Regularization Super-Resolution Image Reconstruction algorithm (MSIR algorithm) is proposed. Experimental results on benchmark medical images and the publicly available medical image sets (100 images) demonstrate the effectiveness of MSIR algorithm outperform the four classical super-resolution reconstruction methods.

Keywords: medical imaging, nonlocal self-similarity, sparse representation, super-resolution, manifold smoothing

1 Introduction

Image super-resolution has wide applications in video surveillance, medical imaging, remote sensing, digital photography and computer vision. Generally speaking, the target of image super-resolution methods is to recover a high-resolution image from one or a sequence of low-resolution images. In the field of computer vision, analysis and research have been focused on super-resolution image reconstruction techniques, with which more detailed image texture and relevant information could be obtained without using the expensive physical instrument. For example, when conducting computer tomography scan in medical, an improvement in the spatial resolution is expected with constant radiation dose. Thus, the width of the focal point and collimator aperture of the detector could be reduced. However, on the one hand, if the focal point is small, radiation dose received will be reduced, which will affects the density resolution, on the other hand, the detector count rate will decrease, which will have to extend the acquisition time to ensure the count rate. If the probe unit is reduced, it will be not only expensive but also reduce the detector count rate. So in order to improve the count rate of the detector, it must extend the acquisition time or increase focal point width [1]. In this case, improving the image resolution by the way of the software has become a research focus.

Spatial resolution and density resolution are two key indicators of medical images. Density resolution represents the minimum density difference which can be displayed in the image. Density resolution of the medical image was constrained by the noise and the size of the display, the less noise and bigger display, the density resolution is better. In the medical equipment, spatial resolution is referred to as geometric resolution or high contrast resolution, which refers to the ability to identify the fine structure in the case of high contrast. In the evaluation of the quality of medical images, the spatial resolution is the first con-

---

* Corresponding Author

sider content. So in order to improve the resolution of image, the spatial resolution is the main content to describe. That is enhancing spatial resolution is an alternative solution to improve the resolution, i.e. it can be used to discriminate the smallest possible details which can be seen, and so on. This issue has attracted researchers with high interest for medical applications, such as literature [2-7] for medical images. But how to strengthen the spatial resolution meanwhile effectively reducing noise is still a challenging problem in medical imaging.

The current image super-resolution reconstruction methods fall into three main categories: interpolation-based methods [6, 9, 10, 17], reconstruction-based methods [26] and learning-based methods [8, 11, 12-16]. The conventional interpolation-based super-resolution reconstruction methods [9] improved image resolution, but noise and blur will be introduced. The reconstruction-based super-resolution methods [26] need to develop and seek new motion model to accurately estimate motion. The learning-based methods, the hotspot in the super-resolution, learn the relationship between high and low resolution image from the training samples and extract the information needed through machine learning methods. As a priori knowledge, this relationship is able to guide the super-resolution reconstruction and satisfactory results have been achieved. For example, Yang et al. introduced the sparse coding to super-resolution reconstruction [16] for the first time, and assumed that the low-resolution image can be linearly expressed by the training sample set. With linear programming methods, the high/low dictionary pairs were trained and sparse coefficients of the low-resolution image patch were obtained under the complete dictionary. Then, the high-resolution patch was expressed with the dictionary and the sparse coefficients. However, artifacts would appear in the image with this method, such as ringings, jags and zippers in the region of edges. The main reason is that the sparse coding does not consider the image local smoothing structures and that the coefficients rely on the over complete dictionary. Being aware of the similar shortcomings, Dong et al. proposed a method named sparse representation based on image interpolation with nonlocal autoregressive modeling [17]. Nonlocal redundancy was employed to adjust the sparse representation model minimization to reconstruct the low-resolution image. Such kind of nonlocal similarity has been successfully used in the image processing [18-25].

Considering the facts that (1) sparse coefficients highly depend on the over-complete dictionary pairs, which possibly generate over-sampling image or under-sampling image; (2) the image local smoothing structures, produce artifacts, such as the ringings and jags, are not considered in the process of image reconstruction; (3) the medical images have a rich amount of nonlocal redundancy, and nonlocal similar patches to a given patch could provide nonlocal constraint to the local smoothing structure [18-25]; (4) using image intrinsic manifold constrains image structure such as used in literature [27-29], which can make the reconstructed image natural and smooth. So we introduce the manifold regularization to constrain image patch lay on the intrinsic smooth manifold, and make full use of the non-local similarity patch as priori knowledge, and get the best sparse coefficient by weights averaging model, which make the coefficient more accurate. Moreover we also redefine the concept of similarity between two patches. Experimental results demonstrate that the proposed method outperforms other state-of-the-art super-resolution methods in terms of Peak Signal to Noise Ratio (PSNR) as well as perceptual quality metrics Structural Similarity Index Measurement System (SSIM). The benefits are: (1) searching k similar patches of the given patch, and using the non-local similar patches to adjust the sparse coefficients, which can make the sparse coefficients more close to the real sparse coefficients; (2) introducing the manifold regularization to constrain the image lay on the intrinsic manifold structure, which can make the reconstructed image keeping smoothing structures; (3) to improve the quality of reconstructed image, a new criterion which is alterable (expanding several pixel around the object pixels), different from the literature [12] and [17] using the fixed window to measure the similarity between two patches, is adopted to measure the similarity between two patches, the detailed information can be found in subsection 3.1.

In the paper, a new method named manifold-regularization image super-resolution reconstruction algorithm is proposed to make the reconstructed image appear more natural. Manifold regularization constrain image patch is utilized to lay on the intrinsic smooth manifold, so that the high-resolution image structure can be estimated effectively and the image local smoothing structures can be better constructed with the nonlocal self-similarity. The reconstructed image has better texture and visual, greatly improving the accuracy of diagnosis.

The rest of this paper is organized as follows. Section 2 presents some related works. Section 3 describes the proposed method. Section 4 presents extensive experimental results and Section 5 concludes the paper.

2 Related Works

Learning-based methods is the hot spot in the field of super-resolution reconstruction, in which neighbor embedding [12,15], example-based method [11] are commonly used. However, most learning-based super-resolution algorithm lack of effective learning method to express the a priori knowledge, and finding out the matching information from the priori knowledge is difficult. To solve this problem, Yang et al. introduced sparse coding to super-resolution reconstruction [16] for the first time, significantly reducing the computational cost substantially and improves the reconstructed image quality.

The target of sparse coding is to represent an input vector approximately as a weighted linear combination of a small number of basis vectors called basis atoms by using the sparse combination coefficients. Assuming that the matrix $D \in R^{n \times K} (n \prec K)$ is an over-complete dictionary, in which each column vector is a n-dimension atom, the vector $x \in R^n$ is a signal which can be linear represented by the over-complete dictionary D and the coefficient $\alpha = [\alpha_1, ..., \alpha_K]^T \in R^K$, that is

$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \|D\alpha - x\|_2^2 \leq \varepsilon \tag{1}$$

Reconstruction constraint assume that the low-resolution signal is defined as the follows [16]

$$y = SHx = SHD\alpha , \tag{2}$$

where H represents a blurring filter, and S the down-sampling operator.

The Eq.(1) is an underdetermined equation, as the dictionary is over-complete. Correspondingly, the Eq.(2) cannot be determined either. Donoho proved that the Eq.(1) can find the answer only when the coefficient is sparse enough. That is to say that the high-resolution image patch x can be represented by the sparse linear combination of the high-resolution dictionary D_h, that is

$$x \approx D_h \alpha \quad s.t. \quad \alpha \in R^K, \|\alpha\|_0 \leq K \tag{3}$$

where α is the sparse coefficient.

The super-resolution problem via sparse representation is often performed in two phases: training phase and Sparse Representation phase (SR phase).

· Training phase: Given the sampled training image patch pairs $P = \{X^h, Y^l\}$ where $X^h = \{x_1, x_2, ..., x_n\}$ are the set of sampled high-resolution image patches and $Y^l = \{y_1, y_2, ... y_n\}$ are the corresponding low-resolution image patches, the goal of this phase is to learn dictionaries for high-resolution and low-resolution image patches, so that the sparse representation of the high-resolution patch is the same as the sparse representation of the corresponding low-resolution patch. D_l and D_h are determined by minimizing the following problem [16]

$$\min_{\{D_h, D_l, Z\}} \frac{1}{N} \|X^h - D_h Z\|_2^2 + \frac{1}{M} \|Y^l - D_l Z\|_2^2 + \kappa(\frac{1}{N} + \frac{1}{M})\|Z\|_1 \tag{4}$$

where N and M are the dimensions of the high-resolution and low-resolution image patches in vector formation, Z denotes the co-codes of all high-low image patches. Finally, the same learning strategy is used in the single dictionary case for training the two dictionaries for SR purpose.

· SR phase: To estimate the high-resolution image from the low-resolution image Y, the sparse representation of the patch y is determined by solving the sparse-coding problem:

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \|FD_l \alpha - Fy\|_2^2 \leq \varepsilon_1$$
$$\|RD_h \alpha - \varpi\|_2^2 \leq \varepsilon_2 \tag{5}$$

where F is a feature extraction operator, and the matrix R extracts the region of overlap between the current target patch and previously reconstructed high-resolution image, and ϖ contains the values of the previously reconstructed high-resolution image on the overlap [16]. Given the optimal

solution α_i to Eq.(5), the high-resolution patch x_i can be reconstructed as $x_i = D_h \alpha_i$.

In order to improve the quality of reconstruction, it is noted that Eq. (5) does not require exact equality between the low-resolution patch y and the reconstructed patch $D_l \alpha$. As a result, it may not satisfy the reconstruction constraint (2) exactly. This discrepancy is eliminated through projection X_0 onto the solution space of $SHX = Y$, that is

$$X^* = \arg \min_X \left\| SHX - Y \right\| + c \left\| X - X_0 \right\|_2^2 \tag{6}$$

The solution to this optimization problem can be efficiently computed using gradient descent. The update equation for this iterative method is

$$X_{t+1} = X_t + \upsilon [H^T S^T (Y - SHX_t) + c(X - X_0)] \tag{7}$$

where X_t is the estimate of the high-resolution image after the t th iteration, υ is the gradient descent. The result X^* is taken from the previously mentioned optimization as the final estimate of the high-resolution image. This image is as close as possible to the initial SR X_0 given by the sparse, while respecting the reconstruction constraint [16].

3 The Proposed Method

The conventional sparse representation for super-resolution method cannot preserve the image local smoothing structures. It is assumed that the low-resolution image is directly down-sampled from the original high-resolution image, but only use a one-pass algorithm which constrain the current reconstructed patches to closely agree with the previously computed adjacent high-resolution patches to enforce the compatibility between adjacent patches (refer to Eq.(5)). It really improves the compatibility between adjacent patches, but only this and the trained dictionary cannot impose structural constraint on the missing pixels or well preserve the local smoothing structures. Fortunately, the manifold can constrain the image patches local smoothing structures [27-29] and the nonlocal self-similarity is effective in preserving the local structure [18-21]. So the manifold regularization is introduced to constrain the image patches lay on the intrinsic smoothing manifold, and to incorporate the image nonlocal self-similarity into sparse representation model to improve the accuracy of the sparse coefficients (refer to subsection 3.2). In addition, in order to get a reasonable similarity measure between two patches, we adopt a new criterion (refer to subsection 3.1) to obtain the similar patches, which is different from literature [12] and [17].

The main flow chart of our algorithm is shown in Fig. 1 where LR and HR respectively denote the abbreviations of Low-Resolution and High-Resolution.

As is shown in Fig.1, the t image patches are firstly obtained by partitioning the up-sampled LR image into the overlapping patches. Then, a HR image patch with corresponding to each image patch is recovered by the MSIR algorithm. All image patches can be handled to be HR image patches. Finally, each pixel of the HR image is acquired by computing the mean value of the accumulated recovered t patches in the overlapping position. Shown in the dashed box are the details about the HR image patch reconstruction by the MSIR algorithm. The k similar patches obtained by a new criterion of the given patch is firstly used to approximately represent the given patch, denoted by p_1, which can constrain the image patch local smoothness, and the weight of patches can be obtained, denoted by ω_{ij}. Then, the image nonlocal self-similarity is incorporated into sparse representation model to obtain the amended sparse coefficient, using the patch weights and corresponding sparse coefficients to mend the sparse coefficients. In which the LR dictionary and the HR dictionary are used to compute the sparse coefficient and the recovered patch p_2, and the sparse coefficient is obtained by solving the approximate equation (refer to Eq.(15)). Finally, the HR image patch can be obtained through the fusion patch equation (refer to Eq.(16)).
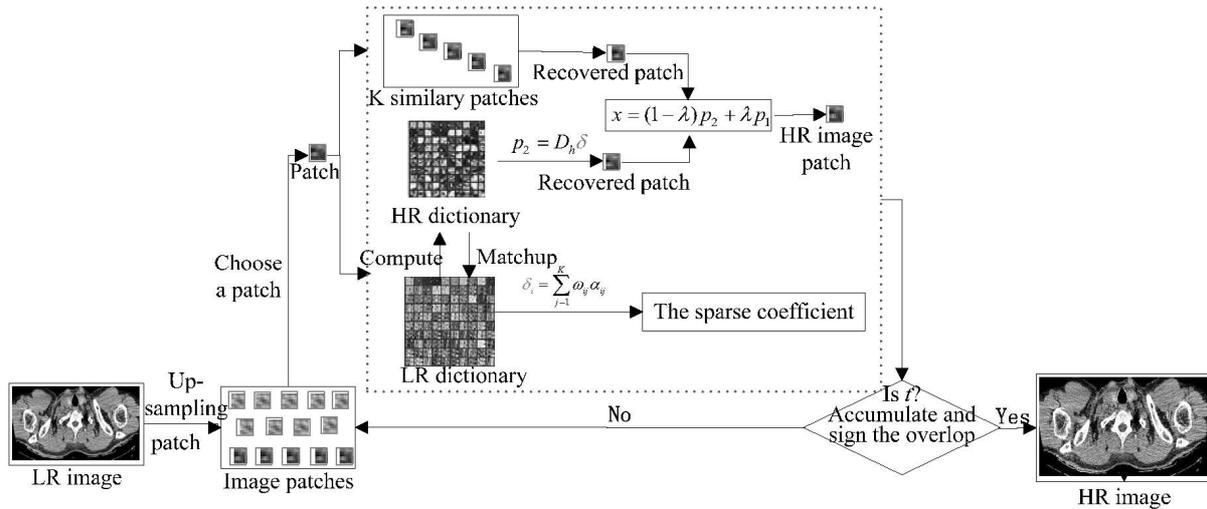
Fig. 1. Overview of the super-resolution reconstruction on the obtained high-low resolution dictionary pair

3.1 The Similarity Measure between Image Patches

Medical images have a rich amount of nonlocal redundancy (e.g., repetitive image structures across the image). In the image processing, the similarity is decided by the space structure, that is to say, the evaluation object is often the window region surrounding the pixels. Fig. 2 illustrates a more effective method for the nonlocal similarity. For pixels *m* and *n*, we can get the square window of *b1* and *b2*. Computing the similarity between *b1* and *b2* decides the similarity between pixels *m* and *n*. i.e. literature [12] and [17], which can find the similar patches but cannot ensure that the similar patches are compatible with the object patch especially in the region of edges. In this case, we can appropriately expand several pixels around the object pixels to search the similar patches, and later resize them to get the similar patches, which not only can gain the similar patches of the object patch, but also can ensure the similar patches are compatible with the object patch.
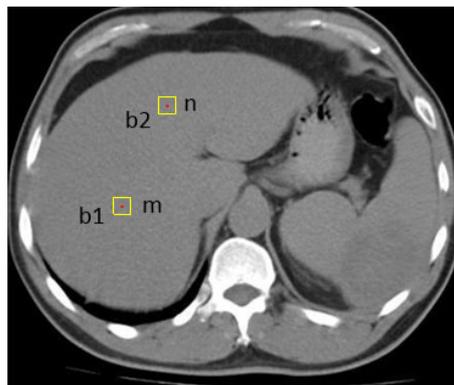

Fig. 2. Illustration of the nonlocal similarity

In Fig. 3, assuming that the low-resolution image patch (the yellow square) size is $\sqrt{s} \times \sqrt{s}$, and putting it as the center to form a matching patch (the blue square) with the size of $(\sqrt{s}+u) \times (\sqrt{s}+u)$, the similar patches of the matching patch is searched within a large range (the red line square). For simplicity, the Euclidean distance is utilized to decide the similarity between matching patches. Then the similar patches of the low-resolution patch (the yellow square) can be acquired by extracting $\sqrt{s} \times \sqrt{s}$ pixels from the center of the similar patches. The Euclidean distance is computed as follows:

$$d_i^j = \left\| y_i - y_{ij} \right\|_2^2 \tag{8}$$

where y_i is the low-resolution patch, y_{ij} is the *j*-th neighbor patch of the patch y_i.
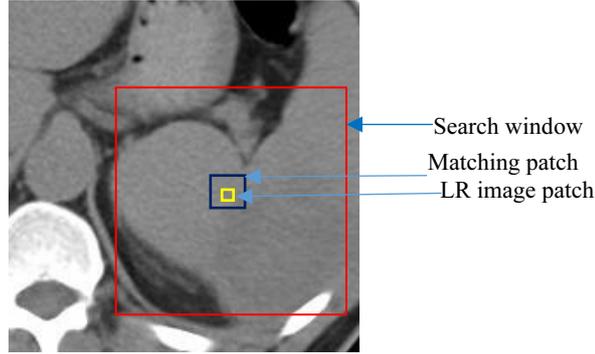

Fig. 3. The searching of the nonlocal similarity

3.2 Manifold-regularization Super-resolution Algorithm between Neighborhood Patches

Presented in this subsection is the manifold-regularization super-resolution reconstruction algorithm, including two main phases: training phase and super-resolution reconstruction phase.

Training phase. From the training images, we can extract a high-resolution image patch set X^h and be vectorized, denoted as $X^h = \{x_1, x_2, ..., x_n\}$. Then, for each x_i of X^h, the down-sampling processing is apply to produce a corresponding low-resolution image patch set $Y^l = \{y_1, y_2, ..., y_n\}$, that is

$$Y^l = SHX^h \tag{9}$$

We consider X^h as high-resolution patches and Y^l as the corresponding low-resolution patches, and H is a blurring filter, and S is the down-sampling operator. Consequently, we obtain a database of high-resolution/low-resolution patch pairs, in which we use the first- and second-order gradient as the feature for the low-resolution patch due to their simplicity and effectiveness. They are normalized into the dictionary pair with corresponding to the high-resolution and low-resolution, i.e.

$$(D_h, D_l) = (\frac{X^h}{\|X^h\|}, \frac{Y^l}{\|Y^l\|}) \tag{10}$$

Super-resolution reconstruction phase. In sparse representation of traditional super-resolution image construction algorithm, the sparse coefficients are obtained by solving the optimization problem Eq.(5), it considered the compatibility between adjacent patches, but it still cannot impose structural constraint on the missing pixels and well preserve the image local smoothing structures. The main reason is that it doesn't consider itself relevant information, and use the trained dictionary to linear represent image patches, which may cause reconstructed image over-sampling or under-sampling. In medical image, fortunately, the nonlocal similar patches to a given patch could provide nonlocal constraint to the local smoothing structure. So in this paper, we make full use the *k* similar patches to adjust the sparse coefficients, and use the manifold regularization to constrain the image local smoothing structures.

In the reconstruction process, we use the bi-cubic up-sampling the low-resolution image Y, and partition it into the overlapping patches, denoted by y_i. The patch y_i may have many nonlocal neighbors y_{ij} which are similar to it but spatially far from it in the image. The similarity weight between the image patch y_i and its *j*-th nonlocal neighbor y_{ij} is generally computed by the following heat-kernel function:

$$\omega_{ij} = \exp(-\|y_i - y_{ij}\|_2^2 / h) / W \tag{11}$$

where h is a predetermined scalar and W is a normalization factor.

The manifold smoothing between neighborhood image patches can be guaranteed by minimizing the weighted L$_2$-norm divergence, that is:

$$\min_{y_i} \sum_j \omega_{ij} \left\| y_i - y_{ij} \right\|_2 \tag{12}$$

According to the normalized weight ω_{ij}, obviously, these nonlocal similar patches y_{ij} can be used to approximate y_i, i.e.

$$y_i \approx \sum_{j=1}^K \omega_{ij} y_{ij} \tag{13}$$

Moreover, co-sparsity characteristics should be used as reconstructing the high-resolution image patch. The weights of the similar patches are used to average the sparse coefficients, and constrain the reconstructed image patch close to the high-resolution patch. This is different from the constraint of the sparse coding method [16] which guarantees the compatibility between adjacent patches with a consistent overlap portion. So the sharing sparse coefficient δ_i is computed by the following objective function:

$$\min \left\| \delta_i \right\|_1$$

$$s.t. \begin{cases} \left\| D_l \delta_i - \sum_{j=1}^K \omega_{ij} D_l \alpha_{ij} \right\|_2 \leq \varepsilon_1 & (a)) \\ \left\| x_i - D_h \delta_i \right\|_2 \leq \varepsilon_2 & (b)) \end{cases} \tag{14}$$

where D_l is the low-resolution dictionary, α_{ij} is the sparse coefficients which can be obtained by solving the optimization problem Eq.(5), and D_h is the high-resolution dictionary.

Get the optimal solution to Eq.(14), the high-resolution patch can be denote as $p_2 = D_h \delta_i$, where p_2 is the high-resolution patch by sparse representation. In this paper, for simplicity, we only use the constraint (a) of the Eq.(14) to solve the coefficient that is we directly use the coefficients averaged by weights to approximately represent the sparse coefficient of the patch, denote as

$$\delta_i \approx \sum_{j=1}^K \omega_{ij} \alpha_{ij} \tag{15}$$

So we can get p_2 by sparse representation. Then we can obtain the high-resolution patch x_i by integrating the patch p_2 and the patch reconstructed by manifold smoothing. The reconstructed high-resolution patch x_i is determined by the following formula:

$$x_i = (1-\lambda)p_2 + \lambda \sum_{j=1}^K \omega_{ij} y_{ij} \tag{16}$$

where λ is a fidelity parameter.

The high-resolution patch obtained is located correspondingly in the image, the overlap is marked until all the patches finish. Then the mean of the pixel in the overlap area is computed. If the super-resolution image patches are filled into the corresponding region of a zero matrix with the dimensions of the required super-resolution image, noted as X^r, then the initial reconstruction image X_0 is computed as

$$X^r = \sum_{i=1}^t B_i x_i \tag{17}$$

$$X_0 = X^r ./ sign \tag{18}$$

where B_i is an operator which keeps the patch in its position in the whole super-resolution image region. x_i is the i-th high-resolution patch, $./$ is element-wise division, X_0 is the high-resolution image, *sign* is a matrix which content is the time number of overlap pixel in a certain pixel.

But, we notice that the high-resolution image X_0 probable does not satisfy the constraint (b) of the

Eq.(14), so we assume that all the sparse coefficients satisfy the constraint (b), the reconstructed image will satisfy

$$\left\|X - X_0\right\|_2 \leq \varepsilon_2 \qquad \textbf{(19)}$$

where X is the high-resolution image, X_0 is the reconstructed image. In order to solve this equation, we project X_0 onto the solution space of $SHX = Y$, one hand it can reduce the influence of the noise, on the other hand it can ensure the equality between the low-resolution patch y and its reconstruction $D_l\delta$

$$X^* = \arg\min_X \left\|Y - SHX\right\|_2^2 + c\left\|X - X_0\right\|_2^2 \qquad \textbf{(20)}$$

We can solve this problem using gradient descent. The update equation for this iterative method is

$$X_{t+1} = X_t + v[H^T S^T (Y - SHX_t) + c(X - X_0)] \qquad \textbf{(21)}$$

where X_t is the reconstructed high-resolution image after the t-th iteration, v is the step size of the gradient descent [16]. Then we can get the final high-resolution image.

3.3 Summary of the Algorithm

The manifold regularization is introduced to constrain the image patch local smoothing structures and incorporate nonlocal similarity into sparse representation to improve the accuracy of the coefficient. The algorithm is to search k similar patches of the given patch, and to calculate the weighted average of the sparse coefficients. The coefficients can depend not only on the sample patches in the given database but also on the similarity between the input patches. And adding the manifold regularization to constrain the image patches lay on the intrinsic smoothing manifold, which can insure the reconstructed patches having local smoothing structures. Meanwhile, in order to ensure the similar patches are close to the object patch especially in the region of edges, a new similarity criteria, different from literature [12] and [17], is introduced. The main procedures of the proposed method i.e. Manifold-Regularization Super-Resolution Image Reconstruction (MSIR algorithm) are summarized in Algorithm 1.

Algorithm 1: Manifold-Regularization Super-Resolution Image Reconstruction

INPUT:
- Input the trained dictionaries D_h and D_l by Eq.(9) and Eq.(10).
- Use bi-cubic up-sample the low-resolution image Y.

OUTPUT:
- HR image X^*.

BEGIN
(1) Put the up-sampled image into patches y_i, size is $\sqrt{s} \times \sqrt{s}$, overlap l pixels between adjacent patches.
(2) For each patch y_i of the image $i = 1, 2, 3, ...t$, perform the following step iteratively:
   (a) Search the similar patch $y_{ij}, j = 1, 2, 3, ...k$ of y_i using the new similar measure between patches, and compute the mean pixel value $\bar{x}$ of patch y_i.
   (b) Compute the weight coefficient ω_{ij} for the similarity patch y_{ij} according to Eq.(11).
   (c) Solve the optimization problem defined in Eq.(5).
   (d) Optimization coefficient δ_i according to Eq.(15).
   (e) Obtain the high-resolution patch x_i by Eq.(16). Put the patch $x_i + \bar{x}$ into a high-resolution image patch.
(3) **End**
(4) Get the high-resolution image X_0, i.e. put the high-resolution patch x together and compute the average value on the overlapping pixel of all high-resolution image patches using Eq.(17), (18).

---

(5) Using the gradient descent to get the final high-resolution image which satisfies the reconstruction constraint (b) of Eq.(14), according to Eq.(20) and Eq.(21) with T times iterations.

END

3.4 Analysis of our MSIR Algorithm

Computational complexity. As outlined in Algorithm 1, major cost of the proposed algorithm is in the implementation of Step 2, for t iterations. For each iteration, the computational complexity of Step 2(a) for searching the k similar patches. First, compute the similarity among the patch is $O(Ms)$, M is the total number of the patch in

the searching window. Then, obtain k patches which values are the top k from the similar patch, this step yields $O(M \log M)$ complexity. Step 2(b) has $O(k)$ complexity. In Step 2(c), assume that the complexity of obtaining the sparse coefficient is $O(T_s)$ in the sparse coding super-resolution [15], so this step has $O(kT_s)$ complexity. Step 2(d) has $O(k^2)$ complexity. In Step 2(e), assume that the complexity of reconstructing a patch is $O(T_c)$ in the sparse coding super-resolution, this step has $O(T_c + k^2)$ complexity. As a result the computational complexity of Step 2 is $O(Mts + Mt \log M + kt + ktT_s + 2tk^2 + tT_c)$. As the computational complexity of Algorithm 1 mainly depends on Step 2, this result is also the computational complexity of this algorithm.

Note that the computational complexity of the proposed method strongly depends on the parameter k. In order to verify this relationship experimentally, simulations have been performed with magnification sc=3 on three images of sizes 50×50, 100×100 and 200×200 for several values of k ranging from 1 to 10. We used a database comprising 1024 low-resolution and high-resolution patch pairs. Simulation was performed on a desktop computer with a dual core processor and 2GB RAM. Fig. 4 shows the overall computational time for various size of low-resolution images and values of k. We can see that these curves with the value of k increases, the slope becomes larger.
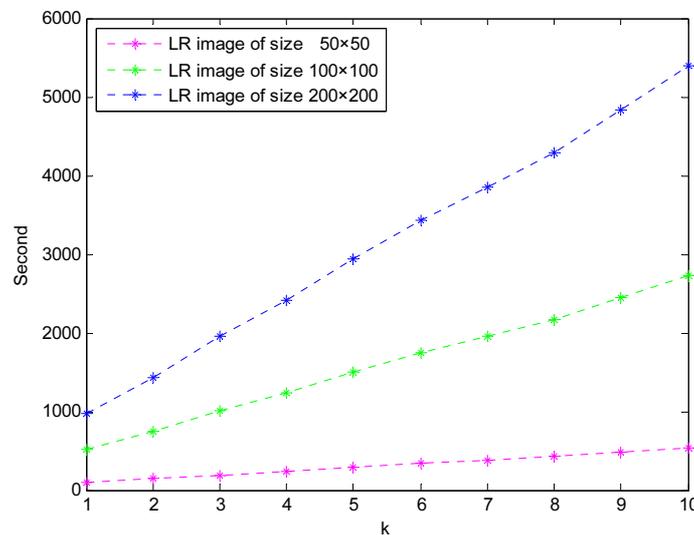

Fig. 4. The increasing of computational time via the increasing of k in our MSIR algorithm

In order to compare the computational speed of the proposed method with the NARM-SRM method [17] and the ScSR method [16], we have tested three methods on the high-resolution heart image of size 510×510. All tests have been done using Matlab R2012b on a dual core processor and 2GB RAM desktop. For the NARM-SRM, it takes about 1194.08 seconds. For the ScSR method, the reconstruction phase takes about 860.34 seconds. For our method, with the number of similar patch k set to 4 and λ to 0.35, the overall time is about 1432.31 seconds. The operation time required in SR phase of the proposed method is higher than that of the ScSR method. This is because the number of the similar patch k and our implementation has not been optimized yet. So, how to reduce the computational complexity is the orient

of the efforts in the future.

Empirical estimation on MSIR algorithm parameters. In this subsection, we investigate the effects of some parameters on the performance, namely the parameter k and the regularization parameter λ in Eq.(16). Here, we take the part of chest image for example to study how the parameters influence the reconstructed results in Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measurement System (SSIM). The PSNR and SSIM are a common criterion in image processing for recovery, and PSNR measure the intensity difference between two images and SSIM is evaluation the similarity between two images, but the SSIM can better express the structure similarity between the recovered image and the reference one, the detailed information can be found in the section 4.2. Fig. 5 shows the change of the PSNR and SSIM versus several values of k in [1 2 3 4 5 6 7 8 9 10] and λ in [0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50] for various value of k and λ.
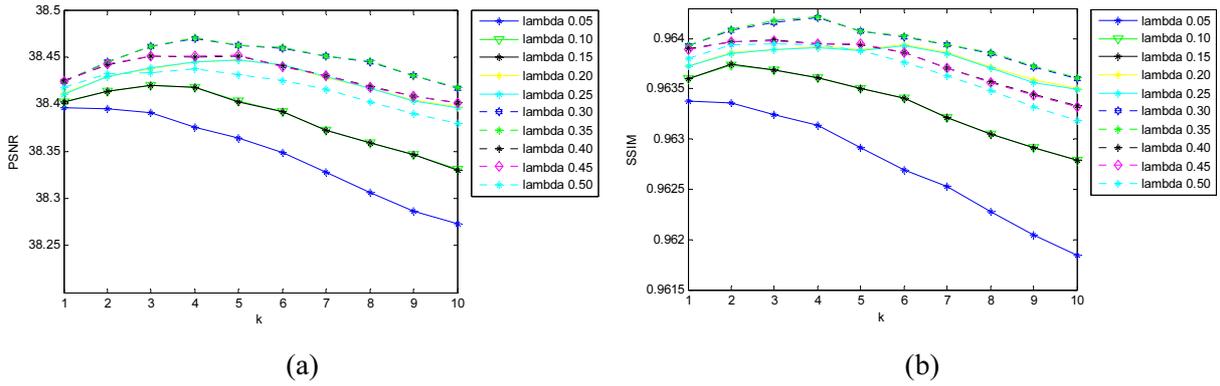

(a)   (b)

Fig. 5. The different value of PSNR and SSIM via the increasing of k and λ, (a) PSNR (b) SSIM

From Fig. 5, we can see that for different k and λ, the PSNR value is different. The λ is constant, the PSNR value increase smoothly and then decrease within the range of k from 1 to 10, and when k reach a certain value, the slope become larger. When k is constant, the PSNR value increase within the range 0.05 to 0.35, but the span of increase is different, when the parameter λ is 0.35, the polyline reached a peak, and larger than 0.35, the PSNR value decrease gradually. The trend of SSIM is pretty much the same as the PSNR, so here we don't analysis it. That is why in our experiments, the parameter k is set to 4 and λ is set to 0.35.

4　Performance Evaluation

In this section, we test multiple images to verify effectiveness of the proposed method, and the proposed method is compared with the Bicubic interpolation (Bb) [9], the Neighbor Embedding-based SR method (NE) of Chang et al.[12], the Sparse coding-based SR method (ScSR) [16], and the Sparse representation based image interpolation with nonlocal autoregressive modeling (NARM-SRM) [17] in terms of both objective and subjective measures.

4.1　Experimental Configuration

The experimental simulation are performed on seven images used as Original test HR images and shown in Fig. 6: CT of chest, CT of heart, MRI of knee, CT of chest, CT of chest, CT of thorax, and CT of thorax. The training databases of all the methods are CT and MRI medical images which are public and can be downloaded from the internet. The high/low dictionaries include 1024 patch pairs extracted randomly from the training databases. The parameter λ in Eq.(16) and the number of similarity patch k are experimentally set to 4 and 0.35, respectively. The size of the up-sampling image patch is 5×5, the size of searching window is 15×15 and matching patch is 9×9.
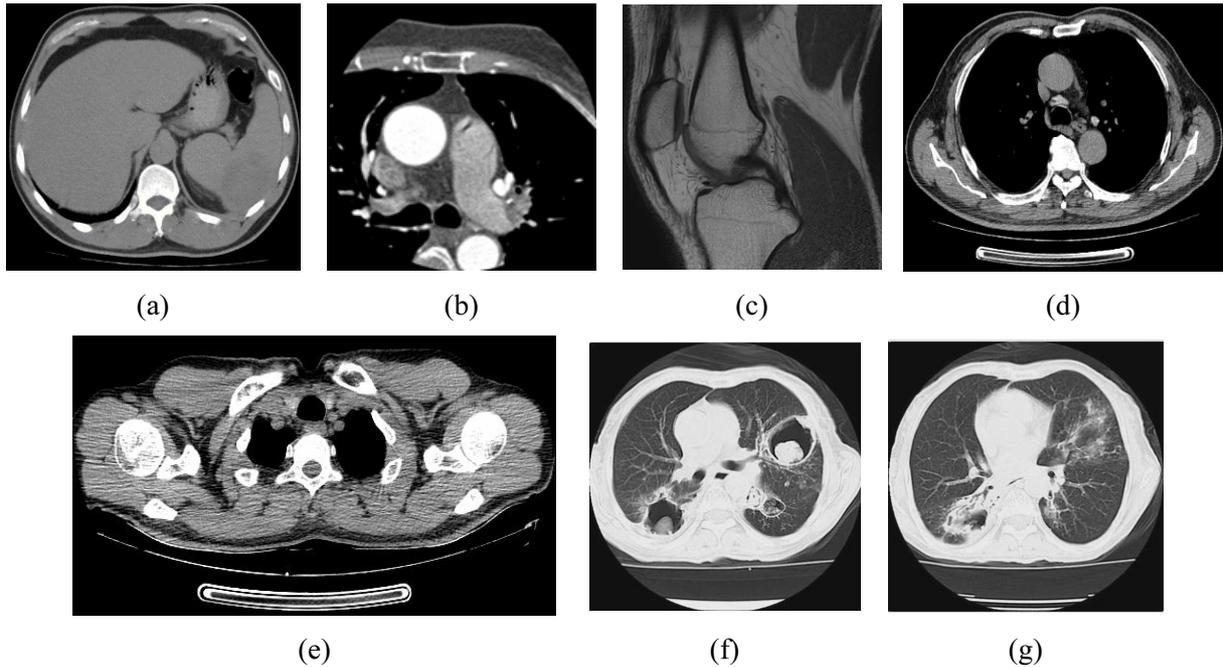
Fig. 6. Original test HR images (a) CT image of chest, (b) CT image of heart, (c) MRI image of knee, (d) CT image of chest, (e) CT image of chest, (f) CT image of thorax, (g) CT image of thorax

4.2 Experimental Results

To prove the effectiveness of proposed method and evaluate the objective quality of the reconstructed images, we use two quality metrics, namely Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measurement system (SSIM) [30] to estimate the quality of the recovered image. We all know that the PSNR measures the intensity difference between two images, but it fails to describe the subjective quality of the image, the SSIM is evaluation the similarity between two images. Compared with PSNR, SSIM is one of the most frequently used metrics in the image processing, and it can better express the structure similarity between the recovered image and the reference one, and both of the PSNR and SSIM bigger are better according to the definition in literature [30].

Results on benchmark image. To better verify the proposed method, firstly, we stimulate on single image and describe the result. The parameter and criterion are mentioned in the above section. The best results of the methods are reported in Table 1 and the best value obtained is in bold number. We can observe that, the quantitative evaluations of the proposed method are superior to those of the Bicubic interpolation and the NARM-SRM, and slightly better compared with the NE-based method and the ScSR method.

Table 1. PSNR and SSIM comparison for $3\times$ magnification

Image	PSNR(dB)					SSIM				
	Bb	NE	ScSR	NARM[17]	**Ours**	Bb	NE	ScSR	NARM[17]	**Ours**
(a)	24.92	36.63	36.88	28.98	**37.09**	0.79	0.90	0.91	0.87	**0.92**
(b)	26.38	38.40	38.64	30.05	**39.04**	0.85	0.93	0.94	0.89	**0.95**
(c)	29.07	36.65	36.80	31.62	**36.89**	0.77	0.85	0.85	0.84	**0.86**
(d)	19.97	32.60	32.81	21.32	**33.11**	0.69	0.79	0.81	0.80	**0.82**
(e)	17.99	30.69	30.77	17.72	**30.85**	0.53	0.64	0.65	0.62	**0.67**
(f)	20.63	32.56	32.79	23.25	**32.93**	0.66	0.77	0.79	0.79	**0.82**
(g)	19.78	32.14	32.34	21.46	**32.42**	0.64	0.75	0.76	0.75	**0.78**

In Fig.7, we compare our method with several more state-of-the art methods on the CT image, including Bicubic [9], NE [12], ScSR [16], and the recently proposed method based image interpolation with nonlocal autoregressive modeling (NARM-SRM) [17]. The result from bi-cubic has many jogged effects along the edges and lost local texture. NE and ScSR generate sharp edges, but also lost local texture, especially in the salient edges. The NARM-SRM method gives a clear image, which has clear edges, but

introduces undesired smoothing that make the image unnatural. We also give the PSNR and SSIM for all the methods in the followed parentheses in the caption. Again, our method achieves the higher PSNR and SSIM among these methods.
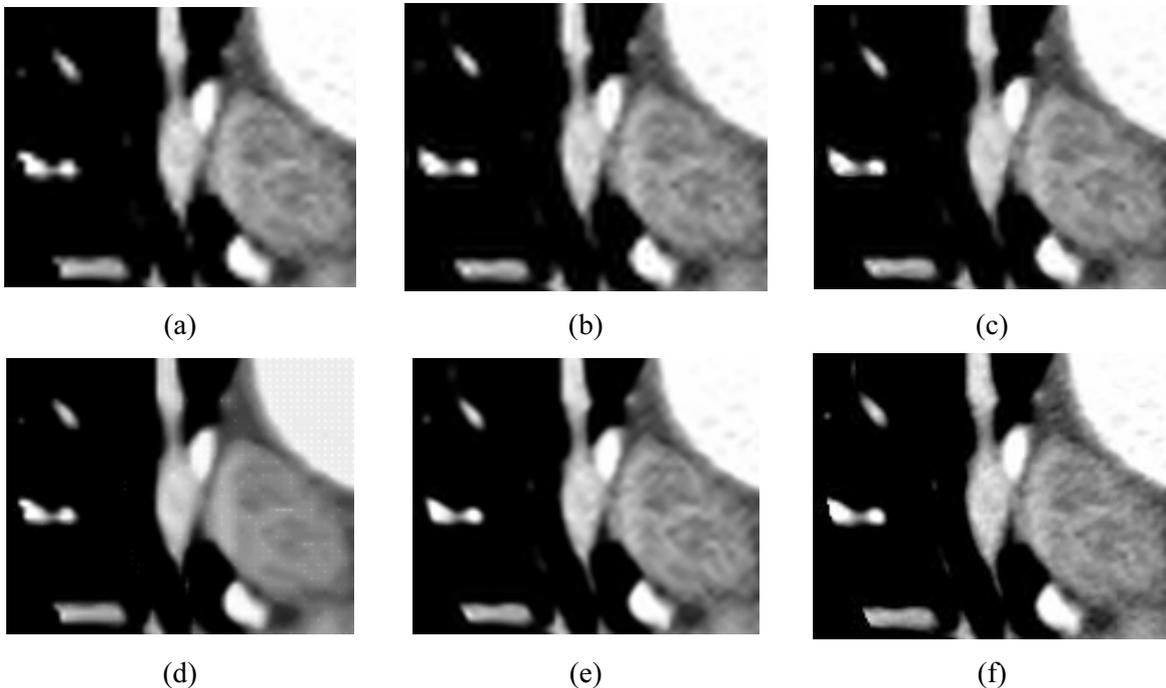

Fig. 7. The super-resolution results of a desired region of interest of the CT image of heart [Fig. 6(b)] with magnification factor of 3. (a) Result of the bicubic interpolation (PSNR=26.38dB, SSIM=0.85). (b) Result of the NE method [12] with the number of neighbors k=8 (PSNR=38.40dB, SSIM=0.93). (c) Result of the ScSR method [16] (PSNR=38.64dB, SSIM=0.94). (d) Result of the NARM-SRM method [17] (PSNR=30.05dB, SSIM=0.89). (e) Result of our method (PSNR=39.04dB, SSIM=0.95). (f) Original test image.

Fig. 8 and Fig. 9 compare the outputs of our method with the based image interpolation with nonlocal autoregressive modeling (NARM) [17]. The NARM-SRM method is similar to ours in the sense that both methods are sparse representation problem, it cannot be denied that NARM-SRM image interpolation method can effectively reconstruct the edge structures and suppress the jaggy artifacts, achieving the best image interpolation results, but it introduces undesired smoothing, which make the image unnatural, and in the salient edge or line it losts local texture. We also can see the PSNR and SSIM in the followed parentheses in the caption, our method achieve the best result so far in terms of PSNR as well as perceptual quality metrics such as SSIM.

In Fig. 10, we compare our method with ScSR [16] on an image of the chest used in the literature [8]. From these figures, we can see that our method retains much more structures especially in the salient edge and has the smoothed structure in the edge. This is because the ScSR method does not consider the image local smoothing structures and the nonlocal similarity. Simultaneously, we also can see the PSNR and SSIM in the followed parentheses in the caption, the PSNR and SSIM value produced by our method are higher than other methods.

From the experiment, we can see that whether the subjective or objective aspect our method outperform other state-of-art super-resolution. But the proposed method also has several areas for improvement, such as the parameter λ and k, the values we used in simulation are the experience value which is not always apt for all images, so what we record in the previous section is not the best and constant. Fig. 11, shows the scope of PSNR and SSIM when using different parameter, k in [1 3 5 7 9] and λ in [0.1 0.2 0.3 0.4 0.5], in different images which are the sequence of the heart, we can see that for a given image different parameter have several values and there are a span between the minimum and maximum value. So how to find an efficient and appropriate parameter for the given image is the direction of our efforts.
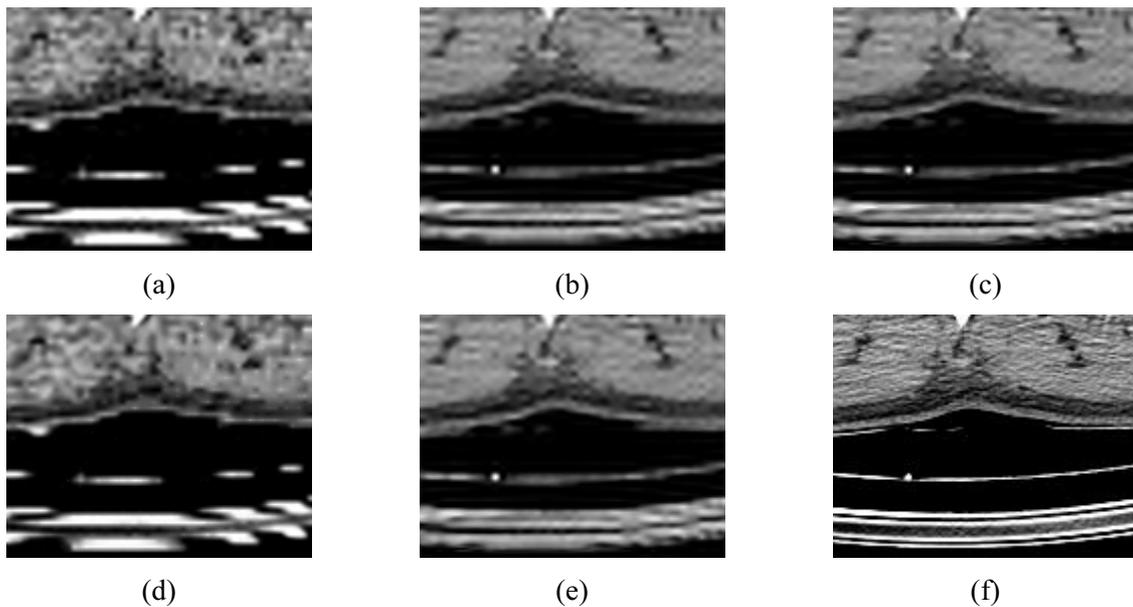
Fig. 8. The super-resolution results of a desired region of interest of the CT image of chest [Fig. 6(e)] with magnification factor of 3. (a) Result of the bicubic interpolation (PSNR=17.99dB, SSIM=0.53). (b) Result of the NE method [12] with the number of neighbors k=8 (PSNR=30.69dB, SSIM=0.64). (c) Result of the ScSR method [16] (PSNR=30.77dB, SSIM=0.65). (d) Result of the NARM-SRM method [17] (PSNR=17.72dB, SSIM=0.62). (e) Result of our method (PSNR=30.85dB, SSIM=0.67). (f) Original test image.

Results on multiple image. The following is the results simulating on the medical image set, the test image main include 100 medical images (50 chests and 50 hearts), and the medical image set is the public database and can be downloaded from the internet.

From the Fig.12, we can see that the PSNR and SSIM of our method are higher than other four methods, no matter the minimum value or maximum value. From the recorded data, we know that it compared with the bicubic interpolation (Bb) [9] PSNR improvement on average 10.10dB, SSIM average upgrade 0.16, compared with the Neighbor Embedding-based SR method (NE) [12] PSNR improvement on average 0.87dB, SSIM average upgrade 0.05, compared with the Sparse coding-based SR method (ScSR) [16], PSNR improvement on average 0.70dB, SSIM average upgrade 0.04, compared with the Sparse representation based image interpolation with nonlocal autoregressive modeling (NARM-SRM) [17] PSNR improvement on average 6.98dB, SSIM average upgrade 0.07. Those data prove that our idea is feasible and effective on the medical image.

5 Conclusion

In this paper, a novel method named manifold-regularization super-resolution image reconstruction is proposed. The conventional sparse representation methods can't preserve the image local smoothing structures, since the sparse coefficients obtained by the paired dictionary may not be the most close to the real sparse coefficient, and this method has not considered the smoothness of the patches. The image nonlocal self-similarity is exploited to improve the accuracy of the coefficients, and the manifold regularization to constrain the image patches lay on the intrinsic smoothing manifold. Additionally, in order to ensure the compatible between the similar patches and the object patch especially in the region of edges, a new criterion is adopted to measure the similarity. Extensive experimental results demonstrated that the image local smoothing structures can be effectively preserved with the proposed MSIR method. In the future works, further study is needed to change the rule of the parameters, and find a more reasonable and adaptive setting of them. Moreover, efforts will be made to use the classification to train the dictionary for reducing the complexity of the algorithm.
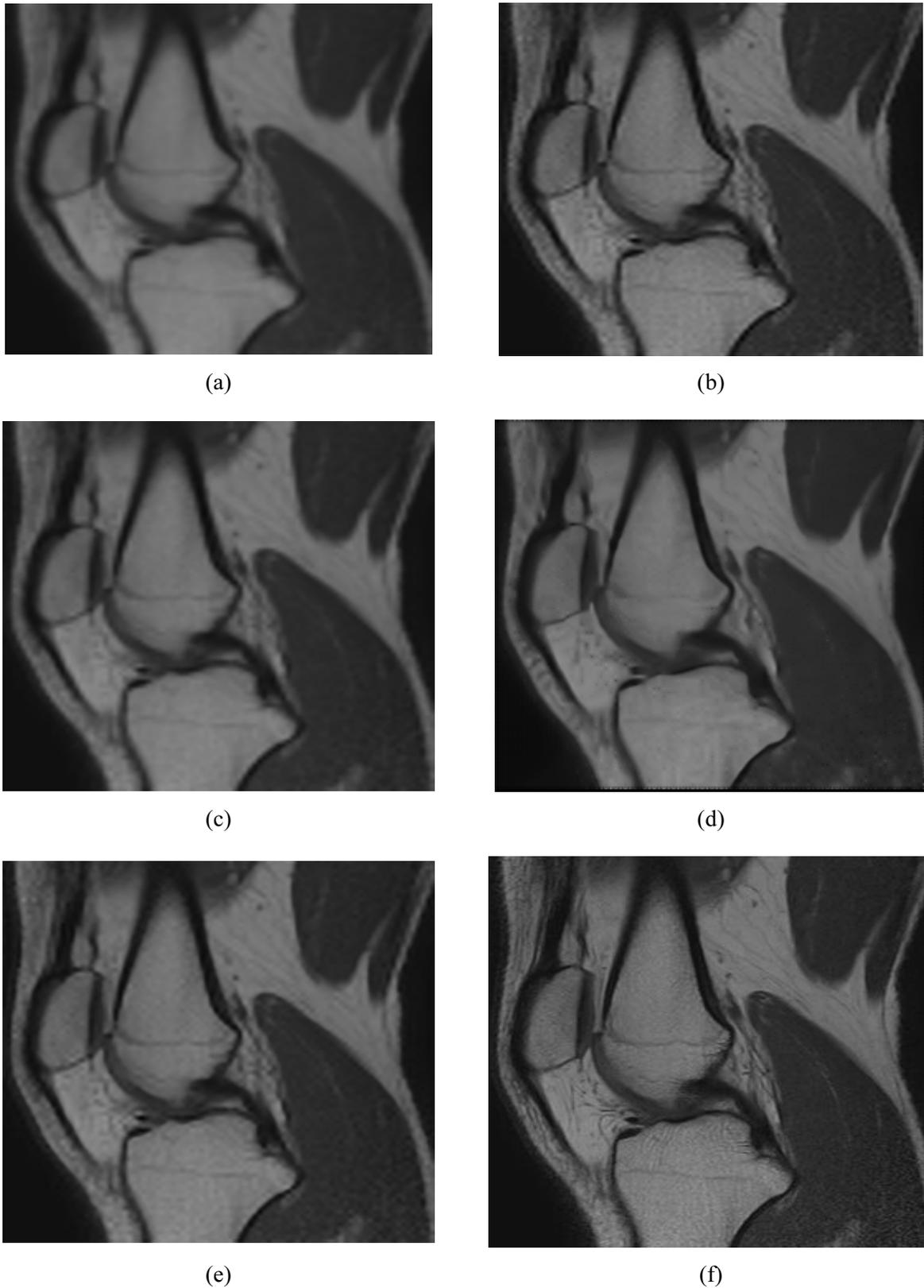
(a)

(b)

(c)

(d)

(e)

(f)

Fig. 9. Results on the MRI image of knee [Fig. 6(c)] with magnification factor of 3. (a) Result of the bicubic interpolation (PSNR=29.07dB, SSIM=0.77). (b) Result of the NE method [12] with the number of neighbors k=8 (PSNR=36.65dB, SSIM=0.85). (c) Result of the ScSR method [16] (PSNR=36.80dB, SSIM=0.85). (d) Result of the NARM-SRM method [17] (PSNR=31.62dB, SSIM=0.84). (e) Result of our method (PSNR=36.89dB, SSIM=0.86).(f) Original test image.
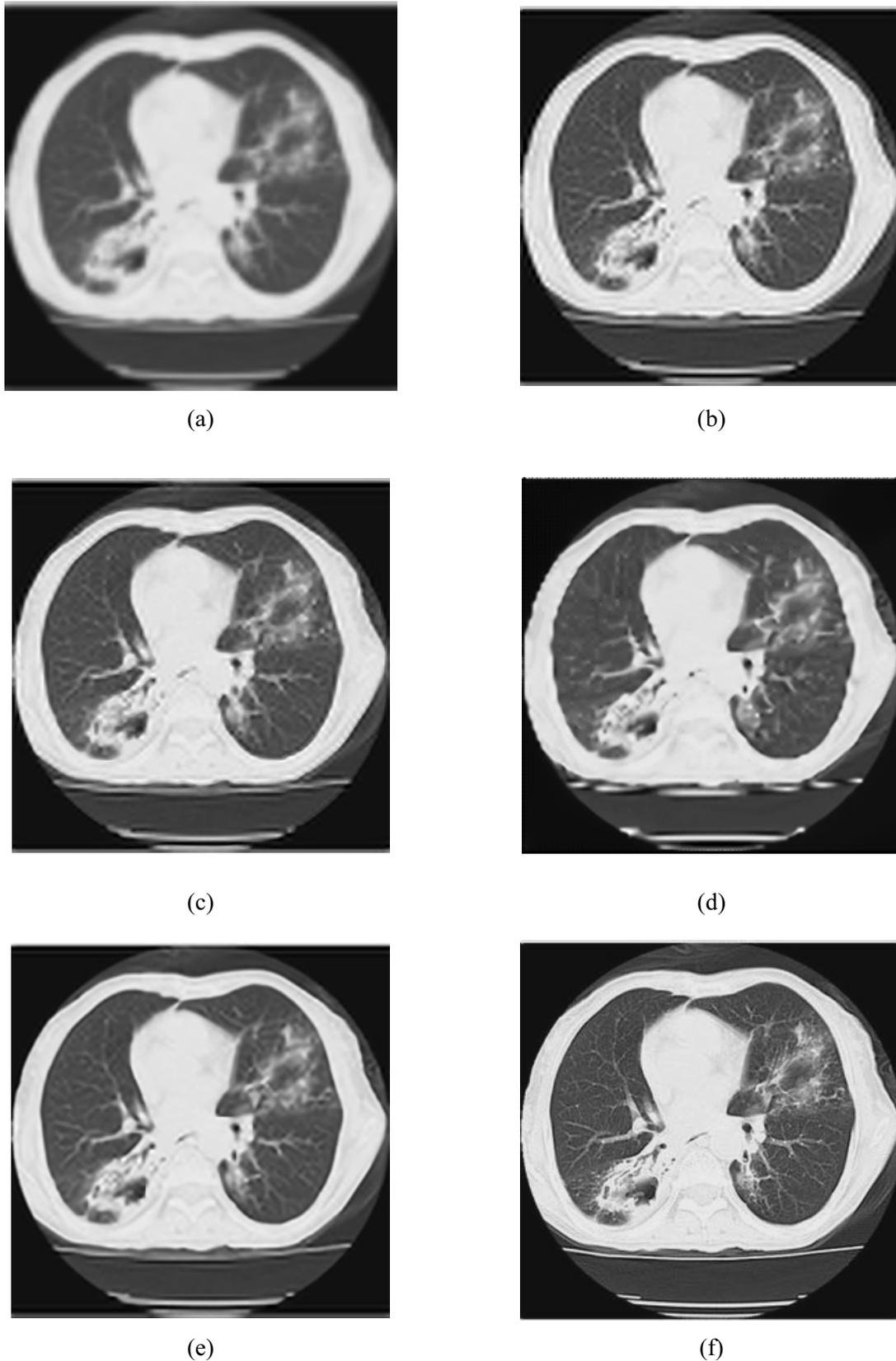
<div align="center">(a)</div>

<div align="center">(b)</div>

<div align="center">(c)</div>

<div align="center">(d)</div>

<div align="center">(e)</div>

<div align="center">(f)</div>

Fig. 10. Results on the CT image of thorax [Fig. 6(g)] with magnification factor of 3. (a) Result of the bicubic interpolation (PSNR=19.78dB, SSIM=0.64). (b) Result of the NE method [12] with the number of neighbors *k*=8 (PSNR=32.14dB,SSIM=0.75). (c) Result of the ScSR method [16] (PSNR=32.34dB, SSIM=0.76). (d) Result of the NARM-SRM method [17] PSNR=21.46dB, SSIM=0.75). (e) Result of our method (PSNR=32.42dB, SSIM=0.78). (f) Original test image.
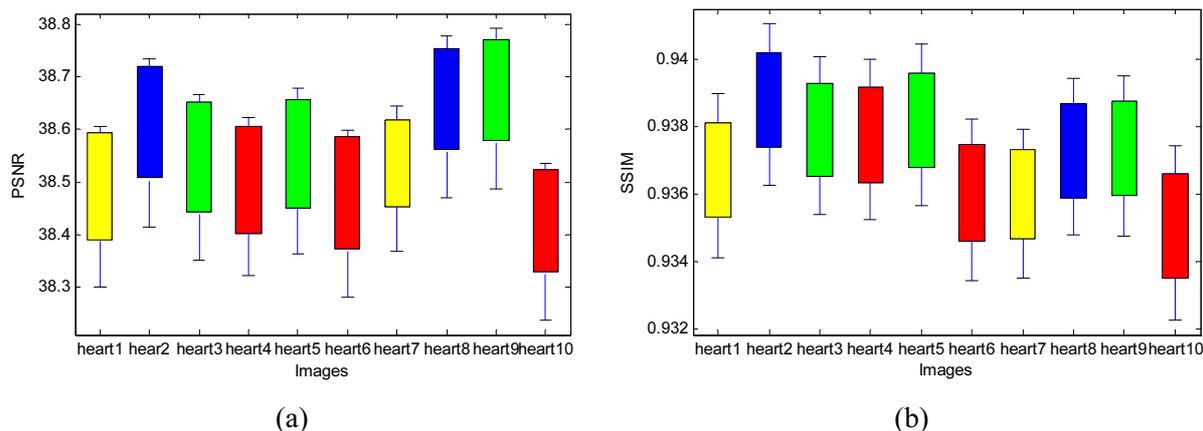
Fig. 11. Objective image quality measures with respect to different CT images of the heart. Each boxplot reflects the minmun, maximun and the median value of the image with different parameters k and λ. (a) PSNR. (b) SSIM.
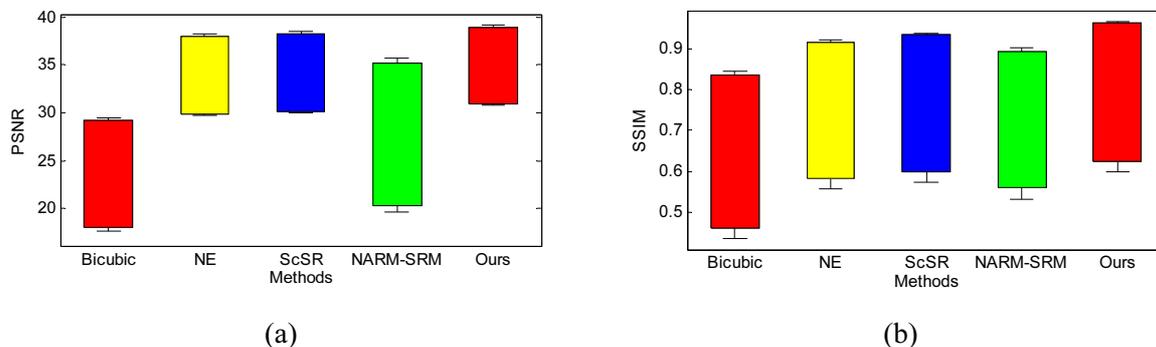

Fig. 12. PSNR and SSIM boxplot illustrate the effectiveness of our method compared with other methods, the Bicubic, NE, ScSR, and NARM-SRM from left to right. (a) PSNR. (b) SSIM.

Acknowledgments

References

[1] T.T. Xin, Study of super-resolution reconstruction for CT images, [dissertation] Nanjing: Southeast University, 2010.

[2] D. Wallach, F. Lamare, J. Rubio, M.J. Ledesma-Carbayo, G. Kontaxakis, A. Santos, P. Marechal, C. Roux, D. Visvikis, Super-resolution in 4D positron emission tomography, in: Proc. of Nuclear Science Symposium Conference Record, 2008.

[3] N. Ben-Eliezer, M. Irani, L. Frydman, Super-resolved spatially encoded single-scan 2D MRI, Magnetic Resonance in Medicine 63(6) (2010)1594-1600.

[4] P. Yan, W. Zhang, B. Turkbey, P.L. Choyke, X. Li, Global structure constrained local shape prior estimation for medical image segmentation, Computer Vision and Image Understanding 117(9)(2013) 1017-1026.

[5] D. Kouame, M. Ploquin, Super-resolution in medical imaging: an illustrative approach through ultrasound, in: Proc. of Biomedical Imaging: From Nano to Macro, 2009..

[6] P. Akhta, F. Azhar, A single image interpolation scheme for enhanced super resolution in bio-medical imaging, in: Proc. of Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference, 2010.

[7] M.D. Robinson, S.J. Chiu, J.A. Izatt, S. Farsiu, New applications of super-resolution in medical imaging, in: P. Milanfar (Ed.), Super-Resolution Imaging, CRC Press, 2010, pp. 384-412.

[8] D.H. Trinh, M. Luong, F. Dibos, J.-M. Rocchisani, C.-D. Pham, T.Q. Nguyen, Novel example-based method for super-resolution and denoising of medical images, IEEE Transactions on Image Processing 23(4)(2014) 1882-1895.

[9] R.G. Keys, Cubic convolution interpolation for digital image processing, IEEE Transactions on Acoustics, Speech and Signal Processing 29(6)(1981) 1153-1160.

[10] W. Chen, Q.-C. Tian, J. Liu, Q.-P. Wang, Nonlocal low-rank matrix completion for image interpolation using edge detection and neural network, Signal, Image and Video Processing 8(4)(2014) 657-663.

[11] W.T. Freeman, T.R. Jones, E C. Pasztor, Example-based super-resolution, IEEE Computer Graphics and Applications 22(2)(2002) 56-65.

[12] H. Chang, D.Y. Yeung, Y.M. Xiong, Super-resolution through neighbor embedding, in: Proc. of Computer Vision and Pattern Recognition, 2004.

[13] C. Damkat, Single image super-resolution using self-examples and texture synthesis, Signal, Image and Video Processing 5(3)(2011) 343-352.

[14] Y. Tang, P. Yan, Y. Yuan, X. Li, Single-image super-resolution via local learning, International Journal of Machine Learning and Cybernetics 2(1)(2011) 15-23.

[15] X. Gao, K. Zhang, D. Tao, X. Li, Image super-resolution with sparse neighbor embedding, IEEE Transactions on Image Processing 21(7)(2012) 3194-3205.

[16] J. Yang, J. Wright, T. Huang, Y. Ma, Image super-resolution via sparse representation, IEEE Transactions on Image Processing 19(11)(2010) 2861-2873.

[17] W. Dong, L. Zhang, R. Lukac, G. Shi, Sparse representation based image interpolation with nonlocal autoregressive modeling, IEEE Transactions on Image Processing 22(4)(2013) 1382-1394.

[18] A. Buades, B. Coll, J. M. Morel, A non-local algorithm for image denoising, in: Proc. of IEEE Computer Society Conference on Computer Vision & Pattern Recognition, 2002.

[19] A. Buades, B. Coll, J.M. Morel, Nonlocal image and movie denoising, International Journal of Computer Vision 76(2)(2008) 123-139.

[20] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Transactions on Image Processing 16(8)(2007) 2080-2095.

[21] L. Zhang, W. Dong, D. Zhang, G. Shi, Two-stage image denoising by principal component analysis with local pixel grouping, Pattern Recognition 43(4)(2010) 1531-1549.

[22] S. Kindermann, S. Osher, P.W. Jones, Deblurring and denoising of images by nonlocal functionals, Multiscale Modeling & Simulation 4(4)(2005) 1091-1115.

[23] A. Langer, S. Osher, C.B. Schönlieb, Bregmanized domain decomposition for image restoration, Journal of Scientific Computing 54(2)(2013) 549-576.

[24] M. Protter, M. Elad, H. Takeda, P. Milanfar, Generalizing the nonlocal-means to super-resolution reconstruction, IEEE

Transactions on Image Processing 18(1)(2009) 36-51.

[25] W. Dong, G. Shi, L Zhang, X Wu, Super-resolution with nonlocal regularized sparse representation, in: Proc. of Visual Communications and Image Processing, 2010.

[26] R. Fransens, C. Strecha, L.V. Gool, Optical flow based super-resolution: a probabilistic approach, Computer Vision and Image Understanding 106(1)(2007) 106-115.

[27] G. Peyré, Manifold models for signals and images, Computer Vision and Image Understanding 113(2)(2009) 249-260.

[28] M. Hein, M. Maier, Manifold denoising as preprocessing for finding natural representations of data, in: Proc. of the National Conference on Artificial Intelligence, 2007.

[29] R. Sparks, A Madabhushi, Statistical shape model for manifold regularization: gleason grading of prostate histology, Computer Vision and Image Understanding 117(9)(2013) 1138-1146.

[30] Z. Wang, A.C. Bovik, H. R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transaction on Image Processing 13(4)(2004) 600-612.