Signal Processing: Image Communication 50 (2017) 21–33

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Detail enhancement of image super-resolution based on detail synthesis

Jinsheng Xiao[a,b,*], Enyu Liu[a], Ling Zhao[a], Yuan-Fang Wang[b], Wenbin Jiang[c]

[a] School of Electronic Information, Wuhan University, Wuhan, Hubei 430072, China
[b] Department of Computer Science, University of California, Santa Barbara, CA 93106, USA
[c] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430079, China

ARTICLE INFO

Keywords:
Super resolution
Detail enhancement
Singular value decomposition
Local self-similarity

ABSTRACT

In this paper, a detail-enhancement and super-resolution algorithm based on detail synthesis is proposed. The novelty of this algorithm is in combining local self-similarity search and singular value decomposition of patches together to obtain details with more natural high-frequency. The proposed algorithm improves the facet or line phenomenon on edges and areas that have rich texture. The algorithm firstly searches for an image patch and extracts the high-frequency components based on a local self-similarity of the original, low-resolution image. The matrix of the high-frequency block is then decomposed into two sub-spaces by the singular value decomposition and the pseudo high-frequency is removed by a soft threshold. Then, the high-frequency block is reconstructed using effective singular values. The final super-resolution image is restored by the detail synthesis with the initial super-resolution image. The experimental results show that the proposed method can significantly remove the artificial effect of facet or line phenomenon caused by pseudo high-frequency. Moreover, the method is also applicable to other super-resolution algorithm in detail enhancement.

1. Introduction

Loosely speaking, image super resolution (SR) refers to a class of algorithms that synthesize high-resolution (HR) images by estimating the values of unknown pixels from the original low-resolution (LR) images. It is a challenging branch of image science with important theoretical and practical advances made in the past. SR has been widely used in varieties of fields such as HDTV, printing, automatic target tracking, video surveillance, media players, remote sensing and medical image processing [1–3].

Existing SR algorithms can be classified into four categories: interpolation-based, reconstruction-based, edge-directed, and learning-based methods. Interpolation-based approaches [4,5], such as bilinear and bicubic interpolation, are the most commonly used methods in practice to upscale images. However, they tend to produce ringing and jagged artifacts as well as over-smoothing the critical structures such as edges [6]. In order to solve the above drawbacks, Li [7] proposes NEDI model which utilizes the similarity of the covariance matrix between the LR image and corresponding HR one to interpolate the pixels along the edges. But the NEDI algorithm has a high computational cost and not suitable for real-time application. Getreuer [8] introduces contour stencils for estimating the image contours and designing an edge directed color interpolation method. This interpolation-based algorithm preserves the outline and details of image well compared to other interpolation-based algorithms.

By building an imaging model of the LR image, reconstruction-based algorithms [9] consider the image transformation in imaging, and look for the optimal estimation of the true HR image. They assume a simplified continuous imaging process and usually formulate the imaging process as a linear system. Among them, the frequency-domain method is a pioneering approach [10]. The reconstruction-based approaches do not exploit prior knowledge for potential effect improvement. What's more, the high-frequency information often can't be reconstructed well with a large SR scaling factor. Hence, the reconstruction-based algorithms are limited in improving image resolution [11].

Edge-directed methods estimate the target HR image by enforcing some edge knowledge [12–15]. For example, Xie [15] proposes the construction of HR edge map from the edges of the LR depth image through a Markov random field and explores the self-similarity of patches during the edge construction stage. Wang [16] proposes an edge-directed SR approach based on a novel adaptive gradient magnitude self-interpolation. This algorithm estimates a sharp HR gradient field directly from the input LR image, and constructs HR images while preserving the sharp edges. However, the enforced edge knowledge can produce noticeable jaggy or ringing phenomenon, and makes image appearing unnaturally.

Finally, the learning-based approaches [17–24] have become

* Corresponding author at: School of Electronic Information, Wuhan University, Wuhan, Hubei 430072, China.
E-mail address: xiaojs@whu.edu.cn (J. Xiao).

http://dx.doi.org/10.1016/j.image.2016.11.001
Received 14 March 2016; Received in revised form 21 July 2016; Accepted 1 November 2016
Available online 09 November 2016

Fig. 1. Visual comparison with scale factor=4 for 'child'.

popular recently. Such an approach assumes that an inheritance relationship exists between LR and HR image patches. A learning procedure is then used to supply prior knowledge to guide the SR reconstruction [25,26]. This method can solve many defects of the reconstruction-based approaches. But it pays the price of large memory consumption. If an external training database is used, noticeable artifacts often exist in the results when the specific image content does not match well with the training database. An alternative [21,27–32] is to acquire the training set from the image itself based on local self-similarities (LSS), which can reduce memory consumption and achieve good results. LSS is based on research on image statistics that suggests image patches can be well-represented as a sparse linear combination of elements from an appropriately chosen over-complete dictionary [33].

In order to solve many existing problems with SR, a detail–enhancement, super-resolution algorithm based on singular value threshold and LSS is proposed in this paper. The algorithm combines the local, patched-based self-similarity search and the SVD of patch together to obtain details with more natural high-frequency. The proposed algorithm can avoid the pseudo facet or line on edges and areas with rich texture. Furthermore, the detail enhancement step can be applied to many SR algorithms regardless of the original SR process. The remainder of this paper is organized as follows. Section 2 gives a brief introduction about learning-based approaches, LSS algorithm, and theory of image enhancement with SVD. Section 3 describes the proposed algorithm in detail. Section 4 presents the experimental results and analysis. Finally, the conclusion is drawn in Section 5.

2. Related work

Learning-based approaches are becoming popular nowadays. The proposed learning-based algorithm removes the unwanted facet or line phenomenon in the SR algorithm based on LSS. Compared to some state-of-art learning-based approaches, our algorithm can acquire better subjective and objective effects. Some related work is reviewed here.

2.1. The brief summary of learning-based approaches

Yang [34] uses the theory of LSS to seek a sparse representation for each patch of the LR image, and then use the coefficients of this representation to generate the HR output. In particular, [28,29] combine image self-similarities with the theory of sparse representation to build a redundant dictionary. This dictionary is used to guide the HR image reconstruction. For example, Freedman [27] recovers the HR image according to the self-similarities and finds the most similar image patch in the original image using a block-matching search method. It then extracts the precious high-frequency component which is added to the corresponding HR image block. Comparing with other algorithms, the algorithm has a visually better subjective quality. However, the algorithm does have its drawbacks. When the high-frequency component is inappropriately added to the HR image, it might cause visually disturbing facet or line phenomena [35]. What is more, the detail of texture information is lost and the HR image can appear unnaturally. Dong [22] proposes a deep learning method based on convolutional neural network for single image SR. The network learns an end-to-end mapping between low- and high-resolution images with little pre-processing. Compared to other algorithms, the algorithm can achieve better HR results.

2.2. Local self-similarity

The theory of LSS is for texture synthesis based on image feature redundancy. And often times, image patches have similar structures in certain local image regions. The similarity may exist in different image scales [27]. The SR image is thus reconstructed using the high-frequency information extracted from similar image blocks of the LR image. Freedman proposes such an algorithm [27] which reconstructs the SR image based on iterative amplification and the LSS theory. It also shows that local minimum image patches are very similar to themselves upon small scaling factors [27].

The LSS algorithm does not rely on any additional external databases, which is the main difference with the dictionary-based learning algorithms. The missing high-frequency details are estimated from the LR images directly. The HR images are reconstructed efficiently with some simple calculation. The LSS algorithm has better subjective effects. However, the reconstructed images may have significant facet or line phenomenon in edges or regions that have rich details, as shown in the following figures.

Fig. 1 presents different results of the LSS and bi-cubic algorithms. There is a significant facet or line phenomenon on the child's hat in Fig. 1(b) using the LSS algorithm compared with Fig. 1(c) using bi-cubic interpolation. Fig. 2 shows the comparative results of the high-frequency components between the LSS and bi-cubic methods. From the Fig. 2, we can see that the LSS algorithm can produce unnatural line and facet phenomenon because of inaccurate high-frequency information. Our observation is that as the LSS algorithm adds high-frequency details to the initial SR image, it significantly improves the image sharpness, but the facet or line artifacts may appear. This comparison indicates that the facet or line artifacts are introduced as pseudo high-frequency components, which make the SR images appear unnatural. At the same time, the facet or line artifacts are not common high-frequency noise, and cannot be removed by general denoising mechanisms. In our approach, the SVD is used to remove the pseudo high-frequency components. The final SR image is constructed by a detail synthesis.

(a) Original image (b) LSS algorithm (c) bicubic

Fig. 2. Comparison of the high-frequency.

2.3. The theory of image enhancement with SVD

SVD is a kind of nonlinear filter, and has a clear physical meaning. A matrix containing image information is decomposed into a series of sub-spaces represented by singular values and singular vectors. Generally speaking, large singular values of the corresponding singular vectors indicate useful signals of the image, while the small singular values of the corresponding singular vectors represent noises. Image denoising is achieved by removing components along singular vectors of small singular values [36].

For image$A \in R^{m\times n}$, with the SVD of the matrix we can obtain:

$$A = U\begin{bmatrix}\Sigma & 0_1\\ 0_2 & 0_3\end{bmatrix}V^T \tag{1}$$

Where $U \in R^{m\times m}$, $V \in R^{n\times n}$ are orthogonal matrices and $0_1 \in R^{r\times(n-r)}$, $0_2 \in R^{(m-r)\times r}$, $0_3 \in R^{(m-r)\times(n-r)}$ are zero matrices. $\Sigma \in R^{r\times r}$ is a diagonal matrix, where r is the rank of matrix A. The diagonal elements of $\sum$ are the nonzero singular values of A which are sorted in descending order. Generally speaking, A is made up of the signals contaminated by noise. It can be expressed as the sum of the signal subspace $\bar{A} \in R^{m\times n}$ and the noise subspace $N \in R^{m\times n}$.

$$A = \bar{A} + N \tag{2}$$

The recovery and enhancement of images can be formulated as a denoising problem, namely, how to find the best approximation of $\bar{A}$ given A [37]. With an appropriate threshold τ, we can obtain $\sum_\tau \in R^{r\times r}$ by thresholding the singular value matrix. $\sum_\tau$ is the approximation singular value matrix of the denoised matrix of A. The reconstructed denoising image can be expressed as the following formula

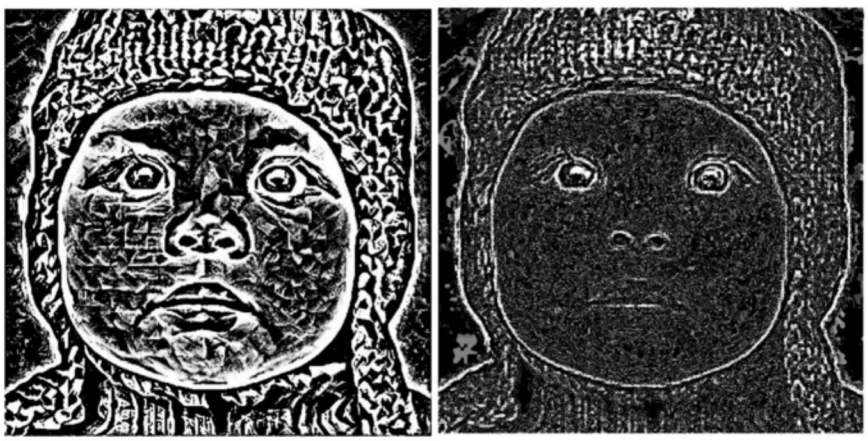
$$\bar{A} = U\begin{bmatrix}\Sigma_\tau & 0_1\\ 0_2 & 0_3\end{bmatrix}V^T \tag{3}$$

In SR algorithms based on LSS, the high-frequency information superimposed on the HR image can be inaccurate. The pseudo high-frequency noise can produce significant facet or line phenomenon on edges and areas that have rich texture. Using the algorithm of image enhancement with SVD thresholding, firstly high-frequency image block matrix is decomposed into a series of singular values and corresponding singular vectors. Then the pseudo high-frequency components can be removed or reduced by SVD thresholding. Finally, we can acquire more accurate high-frequency information by reconstructing the image matrix with the effective singular values, as shown in the figure below.

Fig. 3 presents the enhancement of the LSS using the SVD. The LSS result on the left appears to contain inaccurate high-frequency information with unwanted and unnatural lines and facets. As shown in Fig. 3(b), most of the high-frequency artifacts are removed after post-processed by SVD. Furthermore, SVD does not blindly remove high-frequency components, as can be seen in Fig. 3(b) that major contours (outlines of faces, eyes, nose, etc.) are preserved. It can be observed that SVD processing reduces the blocking phenomenon, and obtains more accurate high-frequency information.

3. Detail enhancement for image super-resolution

In order to remove the unwanted facet or line phenomenon in the SR algorithm based on LSS, a detail-enhancement SR algorithm based on singular-value threshold and LSS is proposed here. The unwanted



(a) Obtained by LSS (b) Processed by SVD

Fig. 3. Comparison of the high-frequency.

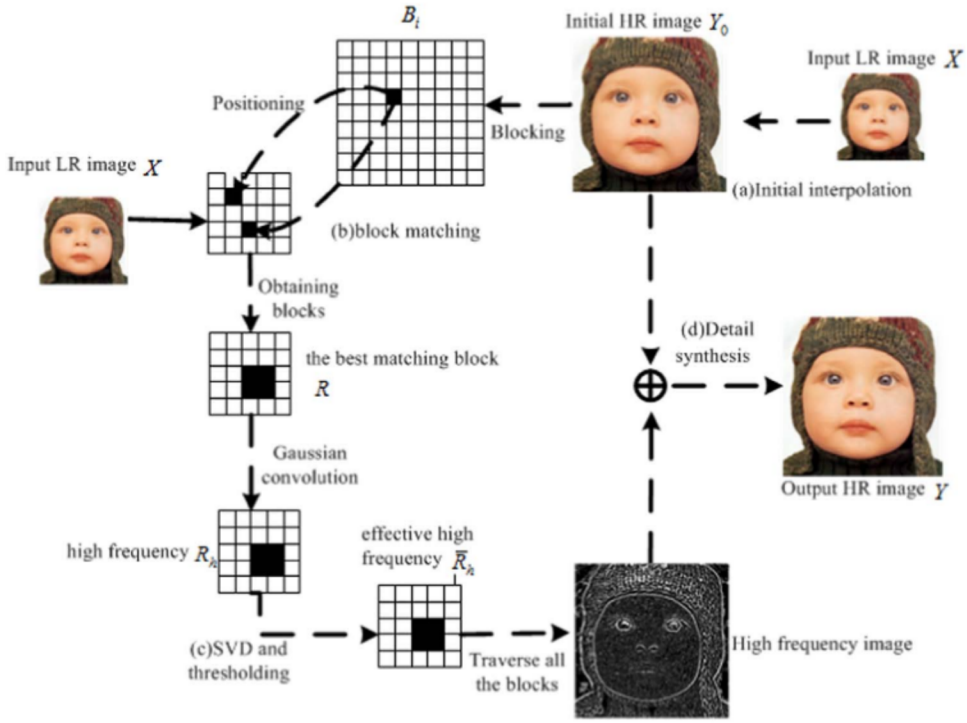


Fig. 4. Flowchart of proposed algorithm.

facet or line phenomenon is caused by inaccurate high-frequency information, and it is the pseudo high-frequency components. The pseudo high-frequency can be considered as noise, but not as conventional Gaussian noise or salt-and-pepper noise [13]. The SR algorithm based on singular-value threshold does not require the prior knowledge of noise, and can successfully remove all kinds of noise. What is more, it has low computational complexity.

In more detail, first, the initial SR image is obtained by the contour stencils interpolation [8]. The prior knowledge is captured based on the LSS of the original LR image. Then for each block in the SR image, the most similar image patch is found in the original image with a block-matching method. The high-frequency detail information of the best-matched block is extracted and the information matrix of the high-frequency block is decomposed into two sub-spaces using the SVD. The pseudo high-frequency is removed by a soft threshold. Finally, the high-frequency block matrix is reconstructed by using the high-frequency components corresponding to the effective singular values only. The finally SR image is restored by a detail synthesis with the initial SR image. The details of the proposed algorithm are shown in Fig. 4.

To expand even further, the proposed detail-enhancement image SR algorithm comprises four major steps: initial interpolation, block matching, SVD and thresholding, detail synthesis (Fig. 4).

(a) Initial interpolation: The selection of the interpolation algorithm will have a major impact on the final result. A good interpolation algorithm is bound to get better results. We have selected the contour stencils interpolation algorithm [8] proposed by Getreuer to obtain the initial SR image. Because of the similarity between HR and LR image on the contour, the contour stencils interpolation algorithm estimates the shape of natural images contour in advance using a set of 57 contour stencils. It calculates neighborhood contour stencil values using each pixel of the original low resolution image as the center of the stencil. Then it chooses the appropriate stencils for interpolation, to obtain the final HR image. It is a new method for estimating the image contours based on total variation along curves. And it is able to distinguish lines of different orientations, curves, corners, and other geometric features with a computationally efficient formula. What's more, this method has linear complexity in the number of pixels. It has a strong advantage in maintaining the whole contour and details of the HR image.

(b) Block matching: This paper adopts the minimum SAD (Sum of absolute difference) between the corresponding pixel values for two image blocks as the matching criterion.

The initial SR image is divided into blocks point by point. Namely, denote Y_0 the initial HR image. For each pixel $b \in Y_0$ with coordinates (b_x, b_y), we construct a reference block $B_i(b_x, b_y)$ with size of $i \times i$ by making b the top left corner "anchor" vertex. The schematic of a local block for $i = 4$ is shown as Fig. 5.

Here each small square represents a pixel. The bold box represents a 4×4 block with the small black square as the anchor. The corresponding position pixel p of the anchor point b is found in the original LR image. The initial search location of pixel p is (p_x, p_y) which is defined as follows:

$$p_x = b_x/SF, \quad p_y = b_y/SF \tag{4}$$

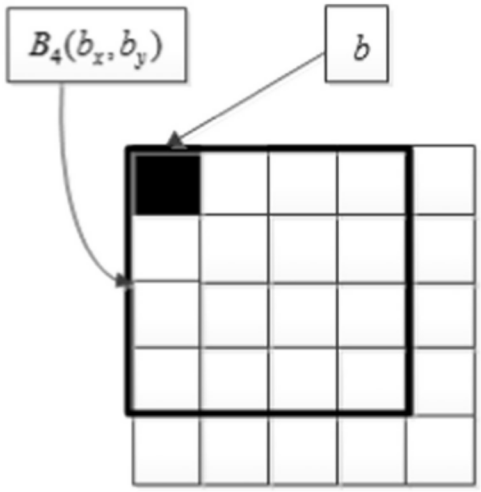


Fig. 5. Schematic of local block.

Table 1
Comparison of objective indicators.

images	Objective indexes	[8] Getreuer	[16] Wang	[34] Yang	[27] Freedman	[22] Dong	proposed
Lena	PSNR	28.7313	28.8379	28.9647	25.1916	**29.3817**	29.0970
	SSIM	0.8968	0.8932	0.9013	0.8486	**0.9309**	0.9191
Parrot	PSNR	28.8170	28.3181	28.4976	26.6355	28.7333	**29.0625**
	SSIM	0.9306	0.9156	0.9196	0.8801	**0.9472**	0.9331
Koala	PSNR	39.0635	38.6047	**41.6668**	26.0218	41.0425	41.3921
	SSIM	0.9850	0.9841	0.9927	0.7841	0.9910	**0.9929**
Child	PSNR	38.1641	38.6223	**41.0802**	25.9091	39.9646	40.4418
	SSIM	0.9863	0.9867	0.9938	0.8673	0.9921	**0.9943**
Girl	PSNR	40.2762	39.9788	41.5014	25.5399	**41.7116**	41.5465
	SSIM	0.9879	0.9873	0.9935	0.7589	0.9928	**0.9936**
Chip	PSNR	41.0968	39.8155	41.9456	34.0812	41.3595	**43.3545**
	SSIM	0.9964	0.9937	0.9945	0.9839	0.9977	**0.9983**
Peppers	PSNR	27.1736	27.2387	27.2423	24.8153	**27.5922**	27.4748
	SSIM	0.9142	0.9121	0.9129	0.8742	**0.9530**	0.9381
Painted face	PSNR	27.1080	27.5142	27.3912	23.7416	**27.7756**	27.7595
	SSIM	0.8691	0.8683	0.8720	0.8396	**0.9153**	0.8747

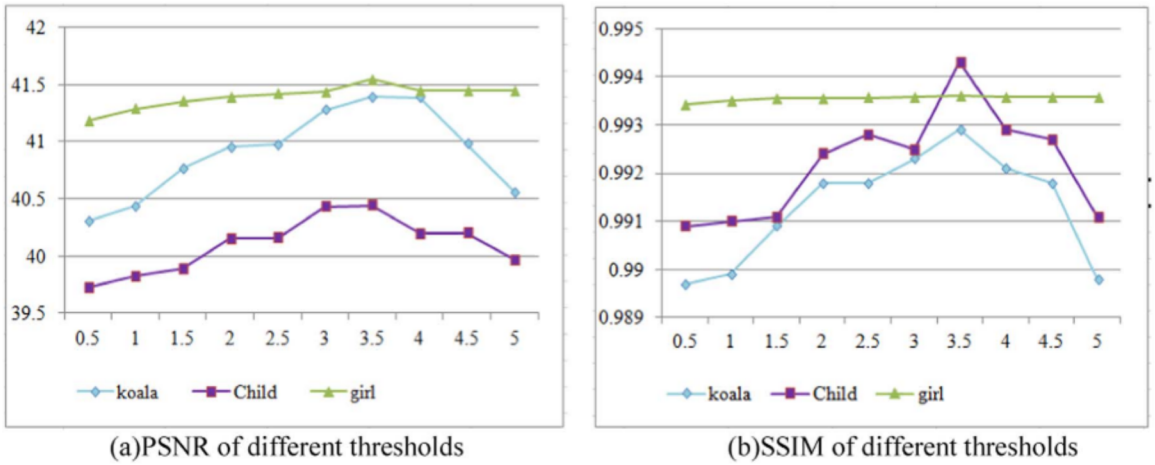


Fig. 6. PSNR and SSIM of different thresholds.

Where SF denotes scale factor, with b being the anchor point. The program uses a 5×5 search window around p to perform the full search for each pixel b. For each potential matching piece, the SAD value is calculated between the matching block and the reference block. The block with the minimum SAD value is selected as the best matching block F(similar block).

Note that the matching is done on a single scale in the proposed algorithm. The initial SR image and the original LR image respectively provide the reference and learning blocks. Freedman [27] and Daniel [30] have used a series of LR images at different scales by down-sampling the original LR image. The initial SR image and the series of scaled LR images provide the learning blocks. While the algorithm in this paper directly uses the information of original LR image and initial SR image which has lower computational complexity.

(c) SVD and thresholding: after the step (b), we find the best matching block F to the reference block $B_1(h_x, h_y)$ in the original LR image. The high-frequency information R_h of F is computed using the following formula:

$$R_h = R - R * G_\sigma \tag{5}$$

Where "*" is the convolution symbol, $G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-b)^2}{2\sigma^2}}$ is the Gaussian kernel function [38]. The high-frequency image block R_h is processed with the singular value threshold as follows:

1) Perform a SVD of $R_h = U\Sigma V^T$, and the singular value matrix can be written as :

$$\Sigma = \mathrm{diag}(\lambda_1, \lambda_2 \cdots \lambda_r),\ \lambda_i \geq \lambda_2 \geq \cdots \lambda_r \geq 0 \tag{6}$$

2) With a selected appropriate threshold value, the pseudo high-frequency components are removed from the singular value matrix. The effective singular value matrix of the high-frequency image block R_h is $\Sigma_\tau = \mathrm{diag}(\lambda'_1, \cdots \lambda'_r)$, where

$$\lambda'_i = \begin{cases} \lambda_i - \tau, \lambda_i > \tau \\ 0, \lambda_i \leq \tau \end{cases} \tag{7}$$

3) The effective high-frequency image block is obtained by a reconstruction using only the effective singular values:

$$\overline{R}_h = U\Sigma_\tau V^T \tag{8}$$

(d) Detail synthesis: the effective high-frequency image block $\overline{R}_h$

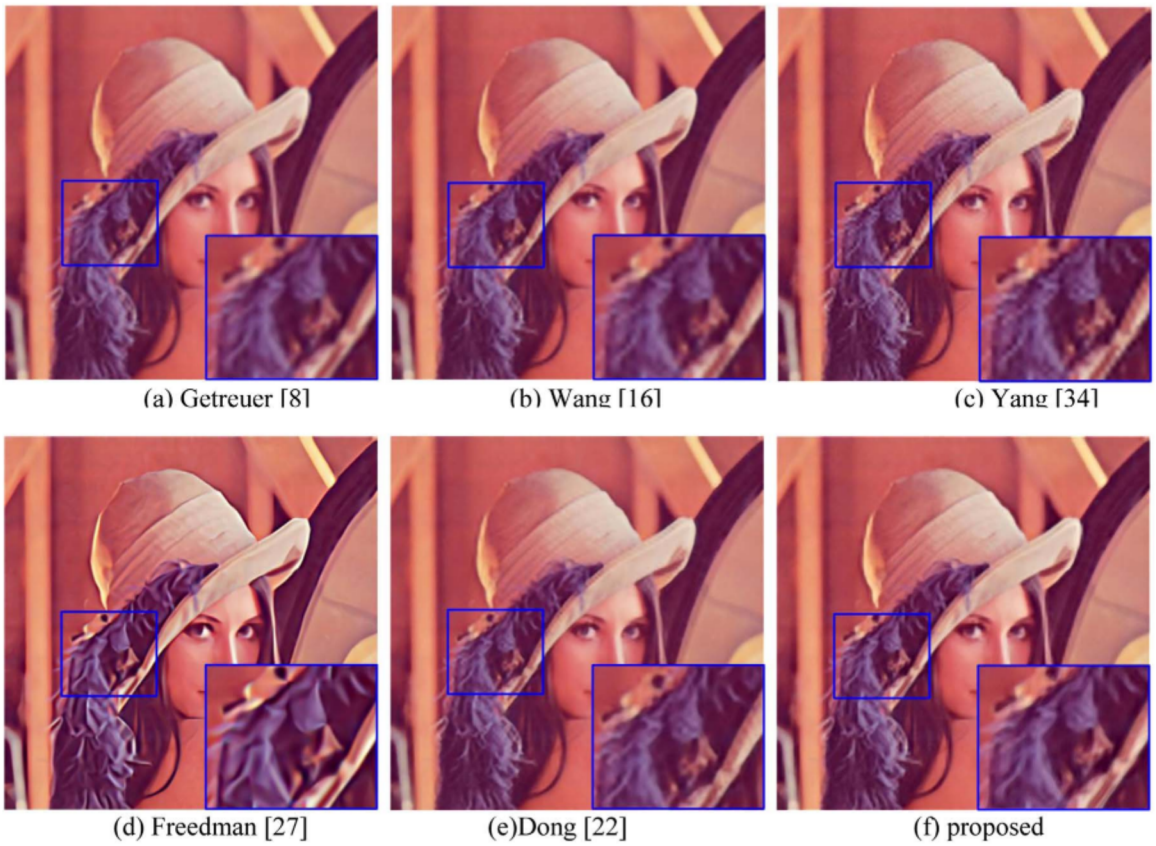


Fig. 7. Visual effect comparison of the image Lena. (a) Getreuer [8], (b) Wang [16], (c) Yang [34], (d) Freedman [27], (e) Dong [22], (f) proposed.

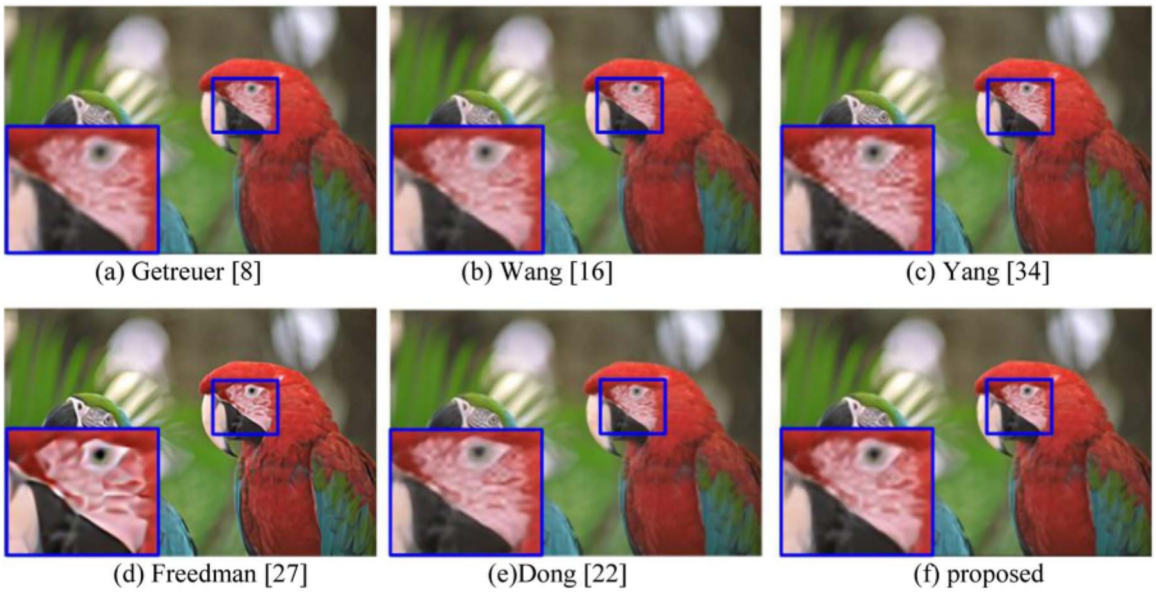


Fig. 8. Visual effect comparison of the image parrot. (a) Getreuer [8], (b) Wang [16], (c) Yang [34], (d) Freedman [27], (e)Dong [22], (f) proposed.

without pseudo high-frequency noise should not be simply added to the initial SR image reference block B_i. Instead, a window is added on the image block by the center symmetric Gaussian function in order to suppress the effects of block overlap. We got $\bar{R}_{rs}$ after adding window.

$$R_{rs} = R_n G_\sigma(\|x - x_c\|) \tag{9}$$

Where $G_\sigma(\|x - x_c\|) = \frac{1}{2\pi\sigma^2} e^{-\frac{\|x-x_c\|^2}{2\sigma^2}}$ is the function of Gaussian window. The SR image blocks with texture features and details of the LR image are synthesized as:

$$B_n^i = B_i + \bar{R}_{rs} \tag{10}$$

In summary, the proposed algorithm proceeds as follows:
Input: The LR image X
Output: The SR image Y

1. The initial SR image Y_0 is generated from X by the contour stencils interpolation algorithm.
2. Y_0 is divided into blocks E;
3. X is decomposed into X_h and X_l by a high-pass filter;

Fig. 9. Visual effect comparison of the image koala. (a) Getreuer [8] (b) Wang [16], (c) Yang [34], (d) Freedman [27] (e)Dong [22] (f) proposed.

4. The anchor point of $B_i(B_i \in B)$ is found in X. The most similar matching patch $R(R \in X)$ of B_i is found by the local searching and matching.
5. The high-frequency image block R_h of R is found in the high-frequency component X_h. Then R_h is processed with singular value soft threshold. R_h is computed after the pseudo high-frequency components being removed.
6. R_h is windowed to obtain R_{hs} according to the formula (9). Then R_{hs} is added to the original high resolution image block B_i according to the formula (10) to obtain the SR image blocks B_h^i;
7. Steps 4–6 are repeated until all image blocks are processed and Y is generated.

4. Experimental results and analysis

To demonstrate the performance of the proposed algorithm, it was compared with many popular methods published recently [16,22,27,34]. Furthermore, we applied the proposed algorithm to other SR algorithms. That is, we firstly obtained the initial SR image using other SR algorithms. Next, we performed steps of detail enhancement, namely block matching, SVD, thresholding, and detail synthesis. The final SR images were compared with the SR images before enhancement.

The algorithm in the experiment was implemented in MATLAB 2012a and run on a PC with a 3.3 GHz dual-core CPU and 4GB RAM.

The SR images were evaluated with the objective evaluation index and subjective visual examination. Here, we tested a large number of images which were selected from the Berkeley Segmentation Database [39]. The standard test images Lena (128×128), Child (128×128), Parrot (192×128), and koala (161×241) were also used here. In order to make the algorithm contrast effect more apparently, we enlarged the original image by 16 (4×4). And they are compared with the HR version of the original image. So the SR images have the same size with the original HR image. We chose some fragments of the high resolution image (in the blue box) for detail comparison. The objective evaluation index is shown in Table 1 and the subjective visual comparison is shown in Figs. 7–11.

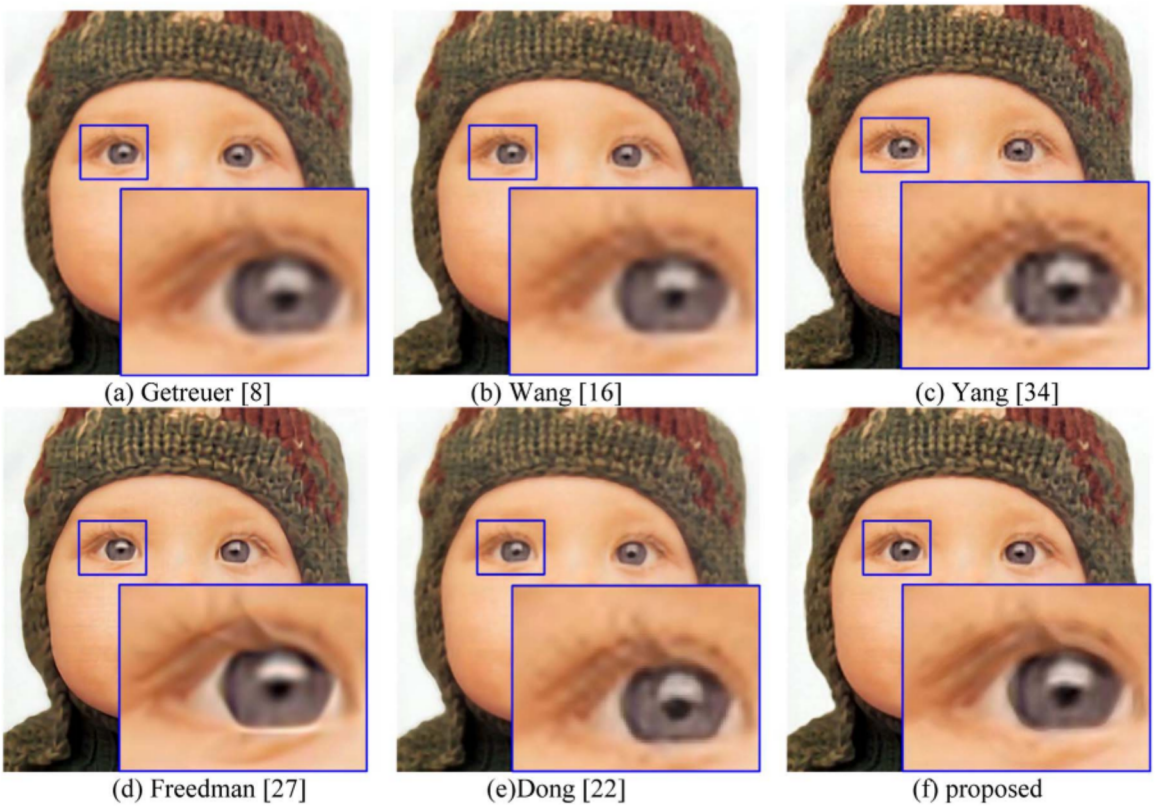


Fig. 10. Visual effect comparison of the image Child. (a) Getreuer [8] (b) Wang [16] (c) Yang [34], (d) Freedman [27] (e)Dong [22] (f) proposed.

4.1. Objective comparison

Peak signal-to-noise (PSNR) and structural similarity (SSIM) were employed to measure the SR performance of these methods. The PSNR quantitatively calculated the discrepancy between the reconstructed image and the original image. The SSIM was used to measure the similarity of two images. Specifically, the PSNR is defined as follows:

$$E_{PSNR} = -10\log_{10}\left[\sum_{i=1}^{N}\left(\frac{(X(i)-Y(i))}{255}\right)^2/N\right] \tag{19}$$

where X and Y denote two images of the same size respectively. In this paper, the images are down-sampled to 1/16 of the original size to serve as experimental inputs. The down-sampled images are then reconstructed to the same size of the original images using these different SR methods.

The SSIM is used to measure the similarity of two images, which is defined in the following formula:

$$E_{SSIM} = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \tag{20}$$

where μ_X,μ_Y are separately mean values of two images respectively, σ_X ,σ_Y are standard deviation of two images, respectively, σ_{XY} is the covariance of the two images, and C_1, C_2 are constants, x and y are of the same sizes [40].

Different parameter settings can lead to different results. It is difficult to set values patch-by-patch and set proper values in an automatic way at present. So by learning the parameter weights based on the tradeoff between the speed and effect of algorithm on sample training images [41], we set the parameters in the proposed algorithm as follows: The radius and variance of Gaussian function were selected as 3 and 0.5, respectively. The size of the reference block in block matching was 4. The threshold value in singular value thresholding was 3.5. These parameters were used consistently in all experiments. Some justification of the parameter settings is given below.

It is critical to set appropriate threshold in the part of singular values. To choose the optimal threshold, we varied the threshold in the experiment from 0.5 to 5 to compute the objective Indicators of super-resolution in case of different thresholds. We tested three sample images and the results were summarized in Fig. 6.

From Fig. 6, when threshold equaled to 3.5, the PSNR and SSIM achieved the maximum at the same time, so we set soft threshold as 3.5 in the paper.

We used these two indicators to evaluate the results of these algorithms objectively. Part of the test results are shown in Table 1.

From the results listed in Table 1, the proposed method surpassed all the compared methods in terms of SSIM. The PSNR of the proposed method was also higher than other comparison algorithms except for two images of the Yang's algorithm [34]. Even so, the result of Yang's algorithm had obvious staircase effect, and the visual effect was worse than that of the proposed method. At the same time, PSNR and SSIM of Dong's algorithm were higher than proposed algorithm in three test images. As Dong's algorithm is based on a novel method named Deep Convolutional Networks which is time-consuming. The proposed algorithm based on singular-value threshold does not require the prior knowledge of noise. Thus it can successfully remove all kinds and has low computational complexity. Above all, our algorithm is competitive.

From what has been discussed above, we can conclude that the detail-enhancement super-resolution algorithm based on singular value threshold and local self-similarity in this paper is superior to other algorithms, for both the objective indexes used.

4.2. Comparison of subjective visual results

We chose some fragments of the HR image (in the blue box) for detail comparison. The objective evaluation index is shown in Table 1 and the subjective visual comparisons are shown in Figs. 7–11.

As what can be seen from Figs. 7–11, the detail of Getreuer method [8] could still be poor, as shown in the unclear texture in the fur of

Fig. 11. Visual effect comparison of the image Girl. (a) Getreuer [8] (b) Wang [16] (c) Yang [34], (d) Freedman [27] (e)Dong [22] (f) proposed.

koala. Wang's algorithm [16] also had the problem of low clarity. What's more, noticeable jaggy phenomenon existed in the results of Wang's algorithm, such as the fur of the koala, the hat of the Child and the hat edge in the lower right corner of Lena's local contrast part. The reconstruction images by Yang's algorithm [34] had rich texture and details. However it still had obvious noise and staircase effect, such as the staircase effect in the hat edge of Lena, and the noise in the hair of the Girl. Freedman's algorithm [27] had high clarity and sharpness. But it appeared somewhat faceted in the edge and regions with fine detail. The detail information loses can be serious, such as the facet in the hair of Lena, the fur of the koala and the feather in the face of Parrot. There are obvious lines on the eyeball of Child, which make the eyeball polygonal and unnatural. Because of the facet or line phenomenon, the detail and texture information of image lost seriously. Dong algorithm (SRCNN) could maintain the detail and edge information of image compared to above algorithms, but it still lose some high-frequency component. The proposed algorithm improved the facet or line phenomenon which caused by the pseudo frequency component to make the HR image more natural and closer to the true image. What's more, the result of our method had fewer jaggy effects. And the proposed algorithm had good clarity, clear texture structure and edge character, and particularly outstanding texture details. Compared to the truth image, proposed algorithm has higher PSNR and SSIM. The truth image looks a bit blurry, zoomed region from the proposed methods looks a bit blurry too. Overall, the proposed algorithm has better visual effect than other competing methods from the visual contrast.

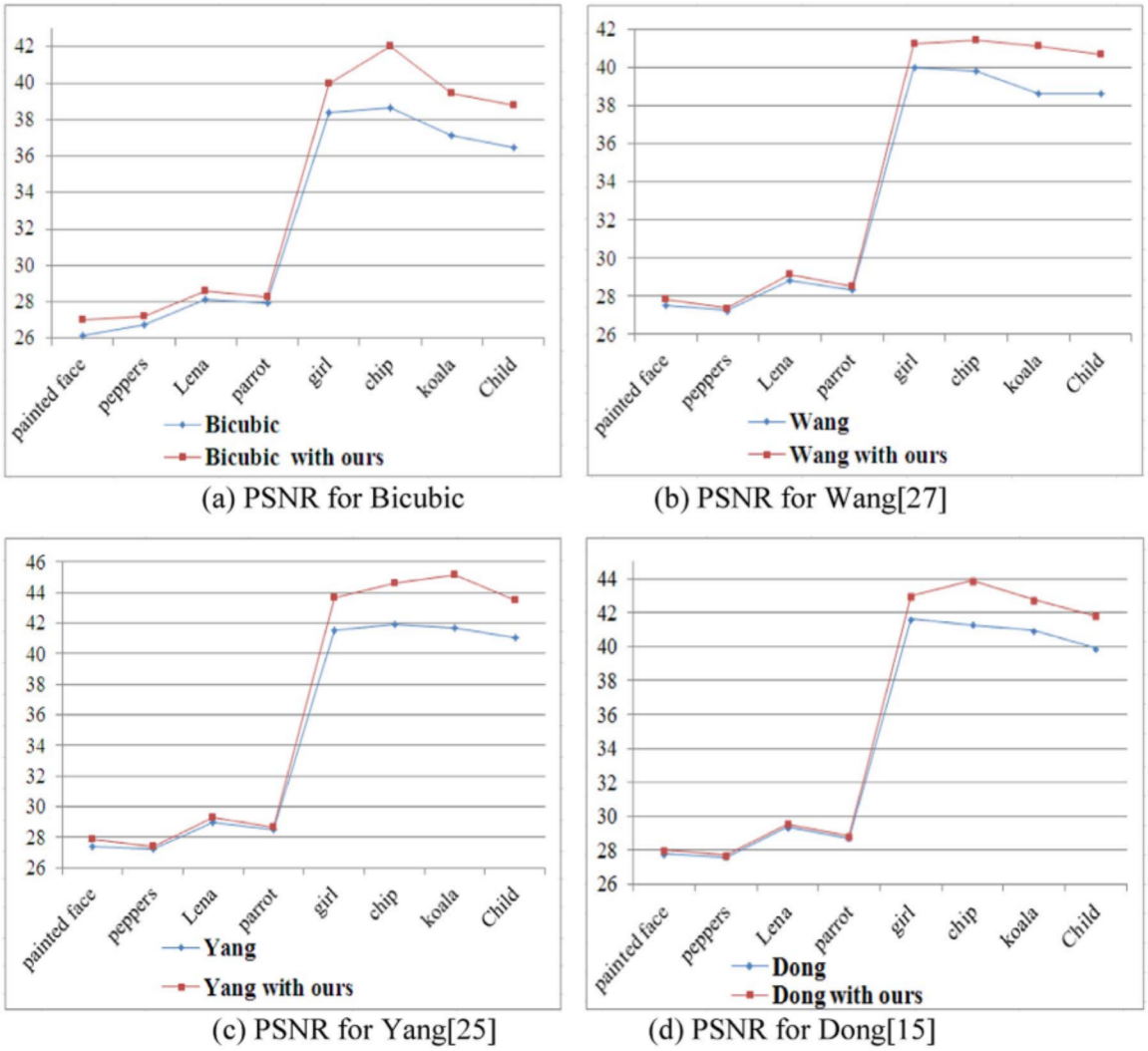
4.3. Detail enhancement analysis

In the previous part of this paper, the contour stencils interpolation algorithm was used to obtain the initial SR image. The prior knowledge can be gained based on the local self-similarity of the original image. The finally SR images were generated with the subsequent steps of detail enhancement, namely, block matching, SVD thresholding, and detail synthesis. From the comparison of subjective and objective results, the proposed detail-enhancement algorithm can achieve good results.

To demonstrate the effectiveness and applicability of our method, other SR algorithms were used to obtain the initial SR images. The subsequent steps of detail enhancement with our algorithm were used to give a final 'polish' of the SR results. Due to the space limitations,

Table 2
The comparison of objective indicators.

Images	Objective indicators	Bicubic	Bicubic With ours	Wang [16]	Wang with ours	Yang [34]	Yang with ours	Dong [22]	Dong with ours
Lena	PSNR	28.1443	**28.5921**	28.8379	**29.1267**	28.9647	**29.2707**	29.3817	**29.5359**
	SSIM	0.8816	**0.9042**	0.8932	**0.9170**	0.9013	**0.9249**	0.9309	**0.9338**
Parrot	PSNR	27.9253	**28.2714**	28.3181	**28.5146**	28.4976	**28.6838**	28.7333	**28.8529**
	SSIM	0.9076	**0.9127**	0.9156	**0.9296**	0.9196	**0.9336**	0.9472	**0.9504**
Koala	PSNR	37.1294	**39.4604**	38.6047	**41.1176**	41.6668	**45.1562**	40.9614	**42.7392**
	SSIM	0.9850	**0.9872**	0.9841	**0.9922**	0.9927	**0.9975**	0.9908	**0.9941**
Child	PSNR	36.4381	**38.7846**	38.6223	**40.6617**	41.0802	**43.4779**	39.8986	**41.7968**
	SSIM	0.9801	**0.9888**	0.9867	**0.9931**	0.9938	**0.9975**	0.9920	**0.9953**
Girl	PSNR	38.3842	**39.9716**	39.9788	**41.2256**	41.5014	**43.6527**	41.6193	**42.9873**
	SSIM	0.9826	**0.9886**	0.9873	**0.9924**	0.9935	**0.9965**	0.9926	**0.9945**
Chip	PSNR	38.6593	**42.0313**	39.8155	**41.4509**	41.9456	**44.6442**	41.2746	**43.8973**
	SSIM	0.9943	**0.9978**	0.9937	**0.9977**	0.9945	**0.9985**	0.9975	**0.9985**
Peppers	PSNR	26.7824	**27.1892**	27.2387	**27.3628**	27.2423	**27.3996**	27.5922	**27.6859**
	SSIM	0.9129	**0.9428**	0.9121	**0.9518**	0.9129	**0.9532**	0.9530	**0.9562**
Painted face	PSNR	26.1839	**27.0291**	27.5142	**27.8318**	27.3912	**27.8837**	27.7756	**28.0095**
	SSIM	0.8400	**0.8959**	0.8683	**0.9142**	0.8720	**0.9193**	0.9153	**0.919**



(a) PSNR for Bicubic (b) PSNR for Wang[27]

(c) PSNR for Yang[25] (d) PSNR for Dong[15]

Fig. 12. Comparison of PSNR between enhanced and unenhanced images. (a) PSNR for Bicubic (b) PSNR for Wang [27], (c) PSNR for Yang [25] (d) PSNR for Dong [15].

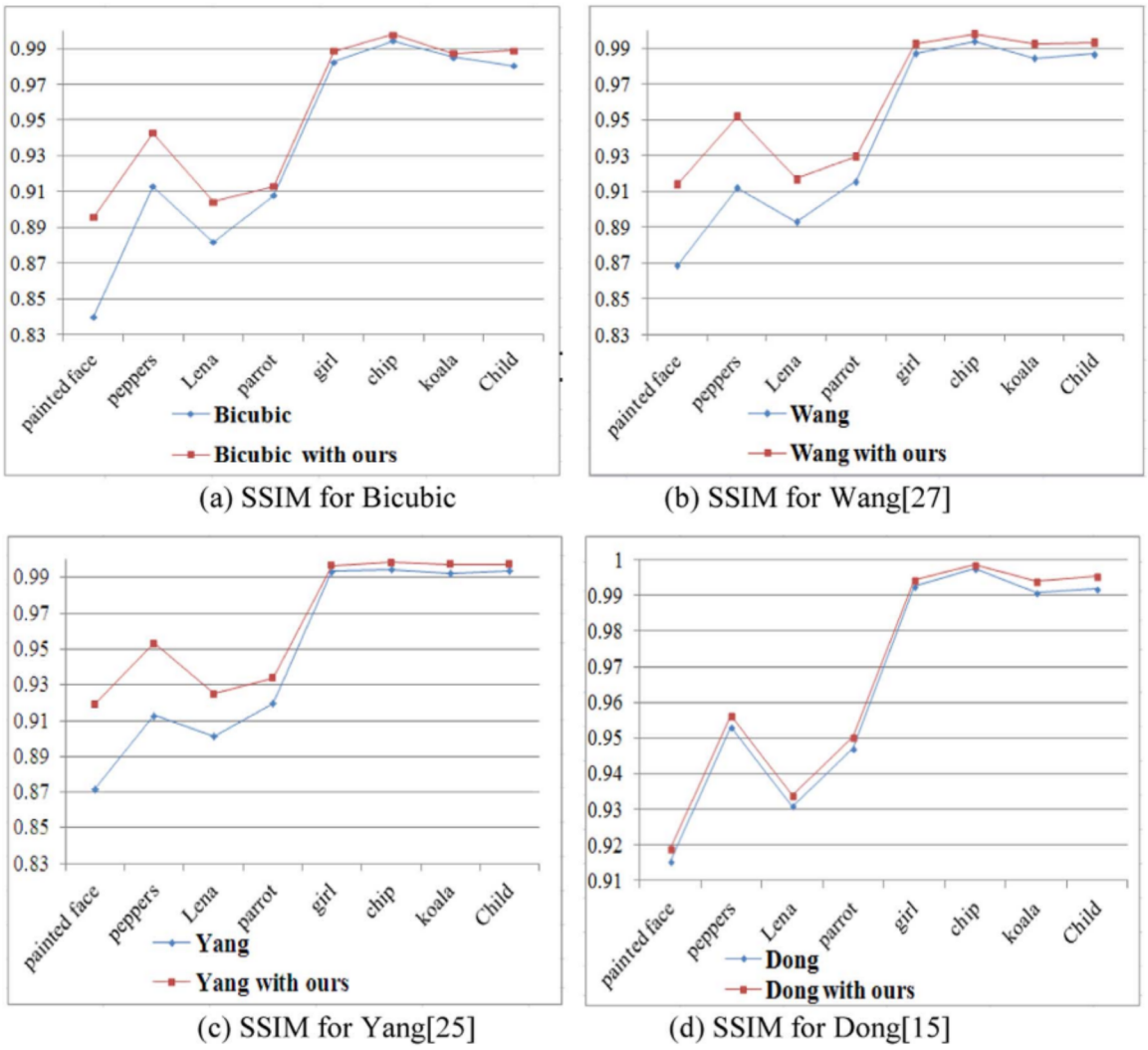


Fig. 13. Comparison of PSNR and SSIM between enhanced and unenhanced images. (a) SSIM for Bicubic (b) SSIM for Wang [27], (c) SSIM for Yang [25] (d) SSIM for Dong [15].

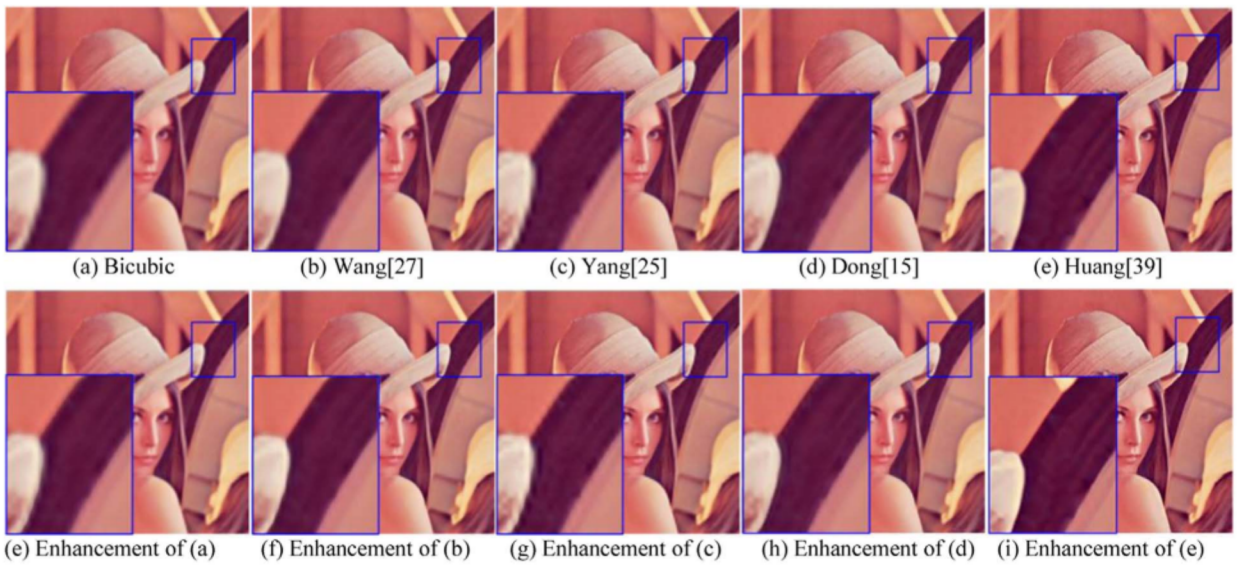


Fig. 14. Visual effect comparison of the image Lena before and after enhancement with ours. (a) Bicubic (b) Wang [27] (c) Yang [25] (d) Dong [15] (e) Huang [39], (e) Enhancement of (a) (f) Enhancement of (b) (g) Enhancement of (c) (h) Enhancement of (d) (i) Enhancement of (e).

only parts of the visual contrast figures are shown here.

To show the change of the objective indexes before and after enhancing more intuitively, Table 2 shows the objective test results of other existing algorithms before and after detail enhancement.

From the Table 2, the objective indicators of other SR algorithms had increased to some extent after detail enhancement. In order to make a clear contrast, Figs. 12 and 13 show the change contrast curve of PSNR and SSIM for each test image.

From the Fig. 12 and Fig. 13, it is found that the PSNR and SSIM index of Bicubic, Wang's algorithm [16], Yang's algorithm [34] and SRCNN algorithm of Dong [22] were all improved after using the detail-enhancement algorithm in this paper. Some increase of PSNR is close to 4 dB and SSIM increase to 0.05.

In Fig. 14, the top row from left to right were the results of Bi-cubic, Wang [16], Yang [34], Dong [22] and Huang [32]. The bottom row was the final SR images where the images in top row were used as the initial SR images with subsequent steps of detail enhancement of the proposed algorithm. Comparing figures in a column, it can be seen that image clarity was improved and the detail information became more abundant. For example, the sharpness of the enhanced image is improved obviously in the edge of hat and hair at the lower right corner of Lena. Because of abnormal operation of the code of Huang [32], so we do not test all images. We take the result of 'lena' processed by Huang [32] as the initial super-resolution images, and the final HR result is indeed improved.

From the subjective visual inspection and objective index comparison above, we can see that the detail-enhancement algorithm in this paper can achieve good results when it is combined as a post-processing step to "polish" the results from the other SR algorithms. Moreover, the clarity and edge sharpness of the enhanced images are improved visually. The details and texture are richer, and the objective indexes are improved significantly. All the above prove that the proposed method has strong applicability as either a stand-alone algorithm or a post-processing step to many other SR algorithms.

5. Conclusion

In this paper, we propose a detail-enhancement SR algorithm based on singular value threshold and LSS. The key point is to combine the local self-similarity search in patch level and the SVD of patches together to obtain details and improve the facet or line phenomenon on edges and areas that have rich texture. First, the LSS of the original low resolution image is used as a priori knowledge source. The best matching block is used to extract useful high-frequency components for enhancements. Then the information matrix of the best matching block is decomposed with SVD, and pseudo high-frequencies that can cause undesirable facet or line phenomenon are removed by a the soft threshold. Finally, the high-frequency detail information without pseudo high-frequency noise is added to the initial SR image. The experimental results show that the proposed method significantly reduces the facet or line phenomenon caused by pseudo high-frequencies, and can well restore the edges and texture structures. At the same time, the algorithm proposed in this paper has good detail enhancement effect and strong applicability to enhance many existing SR algorithms.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant nos. 61471272 and 61672250), and the State Scholarship Fund of China (201306275044). The Titan X used for this research was donated by the NVIDIA Corporation.

References

[1] A. Lapini, F. Argenti, A. Piva, L. Bencini, Comparison of super-resolution methods for quality enhancement of digital biomedical images, in: Proceedings of 8th International Symposium on Medical Information and Communication Technology (ISMICT), 2014, pp.1–5.
[2] C.-T. Shen, H.-H. Liu, M.-H. Yang, Viewing -distance aware super-resolution for high-definition display,, IEEE Trans. Image Process. 24 (1) (2015) 403–418.
[3] C.S. Balure, M.R. Kini, A survey – super resolution techniques for multiple, single, and stereo images, IEEE Electron. Syst. Des. (ISED) (2014) 215–216.
[4] H. Hou, H. Andrews, Cubic splines for image interpolation and digital filtering, IEEE Trans. Speech Signal Process. 26 (6) (1978) 508–517.
[5] K. Guo, X. Yang, H. Zha, W. Lin, S. Yu, Multiscale semi-local interpolation with anti-aliasing, IEEE Trans. Image Process. 21 (2) (2012) 615–625.
[6] P. Thevenaz, T. Blu, M. Unser, Image interpolation and resampling, in: Handbook of Medical Imaging, Processing and Analysis, Academic Press Inc., Orlando, FL, USA, 2000, pp. 393–420.
[7] X. Li, M.T. Orchard, New edge-directed interpolation, IEEE Trans. Image Process. 10 (10) (2001) 1521–1527.
[8] P. Getreuer, Contour stencils: total variation along curves for adaptive image interpolation, SIAM J. Imaging Sci. 4 (3) (2011) 954–979.
[9] Z. Lu, C. Wu, D. Chen, et al., Overview on image super resolution reconstruction, in: Control and Decision Conference (CCDC), IEEE, 2014, pp. 2009–2014.
[10] R. Khan, R. Sablatnig, A. Bais, et al., Comparison of reconstruction and example-based super-resolution,, IEEE Emerg. Technol. (ICET) (2011) 1–6.
[11] S. Baker, T. Kanade, Limits on super-resolution and how to break them, IEEE Trans. Pattern Anal. Mach. Intell. 24 (9) (2002) 1167–1183.
[12] Y. Tai, S. Liu, M.S. Brown, et al., Super resolution using edge prior and single image detail synthesis,, IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) (2010) 2400–2407.
[13] J. Xiao, G. Pang, Y. Zhang, et al., Adaptive shock filter for image super-resolution and enhancement, J. Vis. Commun. Image Represent. 40 (2016) 168–177.
[14] J.Sun, J.Zhu, M.F.Tappen, Context-constrained hallucination for image super-resolution, in: Proceedings IEEE Conference Comput. Vis. Pattern Recognit. (CVPR), 2010, pp. 231–238.
[15] J. Xie, R. Feris, M.T. Sun, Edge-guided single depth image super resolution, IEEE Trans. Image Process. (TIP) 25 (1) (2016) 428–438.
[16] L. Wang, S. Xiang, G. Meng, et al., Edge-directed single-image super-resolution via adaptive gradient magnitude self -interpolation, IEEE Trans. Circuits Syst. Video Technol. 23 (8) (2013) 1289–1299.
[17] C. Kim, K. Choi, J.B. Ra, Example-based super-resolution via structure analysis of patches, IEEE Signal Process. Lett. 20 (4) (2013) 407–410.
[18] C. Dong, C. Loy, K. He, X. Tang, Learning a deep convolutional network for image super-resolution, Eur. Conf. Comput. Vis. (ECCV) (2014) 184–199.
[19] Q. Ning, K. Chen, L. Yi, et al., Image super-resolution via analysis sparse prior, IEEE Signal Process. Lett. Vol. 20 (4) (2013) 399–402.
[20] Y. Zhang, J. Liu, W. Yang, Z. Guo, Image super-resolution based on structure-modulated sparse representation, IEEE Trans. Image Process. 24 (9) (2015) 2797–2810.
[21] Y.-N. Liu, Y.-C. Lin, Y.-L. Huang, et al., Algorithm and architecture design of high-quality video upscaling using database-free texture synthesis, IEEE Trans. Circuits Syst. Video Technol. 24 (7) (2014) 1221–1234.
[22] C. Dong, C. Loy, K. He, et al., Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2) (2016) 295–307.
[23] S.Wang, D.Zhang, Y.Liang, Q.Pan, Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis, in: Proceedings IEEE Conference Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 2216–2223.
[24] K. Guo, X. Yang, Weiyao Lin, R. Zhang, S. Yu, Learning based super-resolution method with a combining of both global and local constraints, IET Image Process. 6 (4) (2012) 337–344.
[25] W.T. Freeman, T.R. Jones, E.C. Pasztor, Example-based super-resolution, IEEE Comput. Graph. Appl. 22 (2) (2002) 56–65.
[26] Y. HaCohen, R. Fattal, D. Lischinski, Image upsampling via texture hallucination, IEEE Int. Conf. Comput. Photogr. (ICCP) (2010) 1–8.
[27] G. Freedman, R. Fattal, Image and video upscaling from local self-examples, ACM Trans. Graph. 30 (2) (2011) 12.
[28] R.Zeyde, M.Elad, M.Protter, On single image scale-up using sparse-representations, in: Proceedings of the 7th International Conference on Curves and Surfaces, Jun. 2010, vol. 6920, pp. 711–730.
[29] C.Yang, J.Huang, M.Yang, Exploiting self-similarities for single frame super-resolution, in: Proceedings of the 10th Asian Conference on Computer Vision (ACCV), Nov. 2010, vol. 6494, pp. 497–510.
[30] D. Glasner, S. Bagon, M. Irani, Super-resolution from a single image, in: IEEE 12th International Conference on Computer Vision. Kyoto, Japan: IEEE Computer Society, 2009, pp. 349–356.
[31] W. Dong, L. Zhang, G. Shi, et al, Nonlocal back-projection for adaptive image enlargement, in: 16th IEEE International Conference on Image Processing (ICIP), Nov. 2009, pp.349–352.
[32] J.-B.Huang, A.Singh, N.Ahuja, Single image super-resolution from transformed self-exemplars, in: Proceedings IEEE Conference Comput. Vis. Pattern Recognit. (CVPR), Jul. 2015, pp. 5197–5206.
[33] W. Lin, S. Lai, Single image super-resolution based on local self-similarity, in: Proceedings of the 2nd IAPR Asian Conference on Pattern Recognition (ACPR), 2013, IEEE, 2013, pp.191–195.
[34] J. Yang, J. Wright, T. Huang, et al., Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2861–2873.
[35] H. Jiang, Z. Cong, Z. Gao, X. Zhang, Image super-resolution with facet improvement and detail enhancement based on local self example, IEEE Wirel. Commun. Signal Process. (WCSP) (2013) 1–6.
[36] Y. Zhang, Y. Ding, J. Xiao, et al., Visibility enhancement using an image filtering

approach, EURASIP J. Adv. Signal Process. 2012 (220) (2012) 1–6.
[37] G.H. Golub, C.F. Van, V.V. Loan, Matrix Computations, 3rd edition, John Hopkins Studues in Mathematical Sciences, Maryland, 1996.
[38] J. Xiao, T. Liu, Y. Zhang, et al., Multi-focus image fusion based on depth extraction with inhomogeneous diffusion equation, Signal Process. 125 (2016) 171–186 (August).
[39] Berkeley Segmentation Database (BSD) [EB/OL], ⟨http://www.eecs.berkeley.edu/Research/ Projects/CS/vision/grouping/ segbench/⟩, 2015, pp. 10–14.
[40] Z. Wang, A.C. Bovik, H.R. Sheikh, et al., Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[41] W. Yang, X. Yin, G. Xia, Learning high-level features for satellite image classification with limited labeled samples, IEEE Trans. Geosci. Remote Sens. 53 (8) (2015) 4472–4482.

Jinsheng Xiao received the B.S. degree in computational mathematics at Wuhan University, China in 1996 and the Ph.D. degree in mathematics from Wuhan University, China in 2001. Between Aug. 2014 and Aug. 2015, he was a Visiting Scholar with University of California, Santa Barbara, USA. He has authored or co-authored more than 50 scientific articles in journals, books, and conference proceedings. He is currently an associate professor in Information and Communication Engineering at School of Electronic Information, Wuhan University, China. His work focuses on Image and Video Processing, Computer Vision, Image and video Enhancement.

Enyu Liu got her bachelor's degree in electronic information from Wuhan University, China in 2016. She is currently learning as a graduate student in Information and Communication Engineering at School of Electronic Information, Wuhan University. She is engaged in Image and video Enhancement, Computer Vision.

Ling Zhao got her bachelor's degree in information technology from Tiebt University, China in 2016. She is currently learning as a graduate student in Electronic and Communication Engineering at School of Electronic Information, Wuhan University. She is engaged in Image and video Enhancement, Computer Vision.

Yuan-Fang Wang received the B.S. degree in electrical engineering from National Taiwan University, Taiwan, in 1981 and the M.S. and Ph.D. Degrees in computer engineering at the University of Texas at Austin in 1983 and 1987. He is a Professor at UCSB and has also played many consulting roles to private industries over the years. Dr. Wang is an academic researcher with a passion for entrepreneurship and has been involved in a number of start-up activities. Dr. Wang has published over 120 peer-reviewed papers and hold three patents. Dr. Wang has received funding support from many federal agencies including NSF, NASA, DARPA, US Army, and US Navy, and also from many private companies. Currently, Dr. Wang's research focuses on 3D computer vision, especially in developing robust, efficient and easy-to-use CV algorithms and systems, with a research objective to benefit the society at large instead of just for the CV community.

Wenbin Jiang is an Associate Professor at School of Computer Science and Technology of Huazhong University of Science and Technology. He received his Ph.D. degree from Huazhong University of Science and Technology (HUST) in 2004. He was a visiting scholar at UCLA (University of California, Los Angeles (UCLA)) in 2014 for one year. His research spans distributed computing, machine learning, multimedia computing, and computer vision.