

## Research Article

# Multiframe Superresolution Reconstruction Based on Self-Learning Method

Shao-Shuo Mu,<sup>1,2</sup> Ye Zhang,<sup>1</sup> Ping Jia,<sup>1</sup> Xun Yang,<sup>1</sup> and Xiao-Feng Qiu<sup>3</sup>

<sup>1</sup>Key Laboratory of Airborne Optical Imaging and Measurement, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

<sup>2</sup>Graduate University of Chinese Academy of Sciences, Beijing 100039, China

<sup>3</sup>Northeast Normal University, Changchun 130033, China

Correspondence should be addressed to Shao-Shuo Mu; 860600894@qq.com

Received 28 August 2014; Revised 11 November 2014; Accepted 11 November 2014

Academic Editor: Daniel Zaldivar

Copyright © 2015 Shao-Shuo Mu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One category of the superresolution algorithms widely used in practical applications is dictionary-based superresolution algorithms, which constructs a single high-resolution (HR) and high-clarity image from multiple low-resolution (LR) images. Despite the fact that general dictionary-based superresolution algorithms obtain redundant dictionaries from numerous HR-LR images, HR image distortion is unavoidable. To solve this problem, this paper proposes a multiframe superresolution reconstruction based on self-learning methods. First, multiple images from the same scene are selected to be both input and training images, and larger-scale images, which are also involved in the training set, are constructed from the learning dictionary. Then, different larger-scale images are constructed via repetition of the first step and the initial HR sets whose scale closely approximates that of the target HR image are finally obtained. Lastly, initial HR images are fused into one target HR image under the NLM idea, while the IBP idea is adopted to meet the global constraint. The simulation results demonstrate that the proposed algorithm produces more accurate reconstructions than those produced by other general superresolution algorithms, while, in real scene experiments, the proposed algorithm can run well and create clearer HR images from input images captured by cameras.

## 1. Introduction

In modern engineering, high-resolution (HR) images are always desirable which provides important information for making important decisions in various practical applications, for example, biometrics, airborne detection system, medical diagnosis, high definition television (HDTV), and remote surveillance. One way to obtain HR images is to increase sensor pixel resolution, which, however, leads to increased cost in sensors. A promising alternative is the superresolution technique, which reconstructs one HR image from one or more low-resolution (LR) images shot from the same scene, providing sufficient relevant information.

In recent years, superresolution techniques have been extensively studied and put into practice. CEVA Company firstly combined superresolution techniques with CEVA-MM3101 low-power imaging and vision platform successfully, which used LR sensors to construct HR image, and have been

applied to camera-enabled mobile devices. French company SPOT applied superresolution techniques to SPOT-5 satellite which can obtain clearer ground scene images and identify targets more accurately, compared with SPOT-4 which did not adopt superresolution techniques. So superresolution technique has been a hot spot to improve images' resolution without changing the optical system of the camera.

Superresolution techniques have three categories in general, namely, interpolation-based, reconstruction-based, and dictionary-based algorithms. Interpolation-based algorithms are simple but ineffective, in which the unknown image pixel values are calculated from the known pixels using interpolation formula [1–4], such as bilinear, bicubic, and Lanczos interpolations. Reconstruction-based algorithms have process inverse to imaging system solving. These algorithms use prior image knowledge and recover high-frequency information via established imaging models [5–10].

Most recently, dictionary-based algorithms have been more widely applied due to the development in compressed sensing [11–15]. This category of algorithms requires a number of external training images, using similarity redundancy structure of natural images to restore missing high-frequency information in input images. Roweis and Saul [16] proposed a nonlinear dimension reduction method, namely, locally linear embedding (LLE), which has been widely applied to image processing. Based on [16], Chang et al. [17] applied LLE technology to superresolution reconstruction. In this method, the single input LR image is divided into blocks with image shift; then, each image block searches for  $k$  nearest blocks in all LR training image blocks whereby the optimal weights are obtained; finally, a HR image is constructed from a linear combination of the  $k$  HR training blocks corresponding to the  $k$  LR training blocks. References [18–20] improved the algorithm in [17] and achieved enhanced reconstruction results. Yang et al. [21] proposed a sparse coding method for dictionary learning. Compared with previous algorithms, their algorithm relies on a larger number of databases, while obtaining missing high-frequency information more effectively. In [22], their sparse coding is improved to strengthen the sparse similarity between LR and HR image blocks. This method is simpler and more effective compared with [21]. He et al. [23] adopted a beta process dictionary learning approach to make the HR and LR dictionaries more accurate and consistent, gaining better results in comparison with [22]. Jeong et al. [24] proposed a superresolution method robust to noise, which involves two phases: the learning phase and the reconstruction phase. In the learning phase, the training images are classified and different dictionaries are constructed according to noise quantity; and, in the reconstruction phase, the input images with various noise quantities adopt appropriate dictionaries to reconstruct HR images.

All the above reconstruction methods require a large number of external training images. When external images cannot be accessed, it is feasible to start from the point of similarity redundancy structure of input images. In [25–28], different self-learning methods are proposed. Glasner et al. [25] realized superresolution reconstruction without external images by adopting the gradual magnification scheme. Chen et al. [26] used high-frequency and low-frequency components of input images to train dictionary image blocks. Then, they used the Approximate Nearest Neighbor Search to reconstruct images and recovered the missing high-frequency information. Zhang et al. [27] proposed a self-similarity redundant algorithm, which selects original LR images and corresponding degraded images as the training images and uses LLE algorithm to obtain one HR image. In [28], Ramakanth and Babu proposed single image learning dictionary and adopted the approximate nearest neighbor fields (ANNF) method to find similarities between input LR blocks and LR dictionary. This method reconstructs HR images more effectively and more efficiently compared with other dictionary-based methods.

Other studies [25–28] have shown that natural images have self-similarity information on different scales and high-frequency information in small-scale images can be obtained

from large-scale images. So this paper proposes a multiple-frame superresolution algorithm based on self-learning, and this algorithm requires no external training image. First, self-similarity information of multiple input images is used to create larger-scale training images through a step by step process. Then, an over-complete dictionary is studied by sparse representations. Finally, the reconstructed images are further improved by exerting the global constraint. Compared with [25–27], the proposed algorithm possesses several unique features as follows.

- (1) Larger-scale images are created from input images. Both input images and larger-scale images are involved in the training set. Sparse representation is adopted to learn the mapping relation of the dictionary.
- (2) One larger-scale image is reconstructed from one input image; apply this reconstruction to all  $n$  input images. Through gradual magnification,  $n$  initial estimates are obtained, each corresponding to one of the input images, with a scale closely approximating that of the target HR image. Then, the NLM method is employed to fuse all estimates into one target HR image.
- (3) In order to restore more high-frequency parts, degraded images of input images are used as LR training set and high-frequency parts of input images are used as HR training set.

Our theoretical analysis demonstrates that the proposed algorithm is feasible, and the simulation results show that the proposed algorithm performs better than general dictionary-based superresolution algorithms. In real scene experiments carried out on a camera system, the proposed algorithm obtains clearer real HR images.

This paper is organized as follows. Section 2 gives a discussion on image sparse models. Section 3 presents the framework of our proposed algorithm as well as its implementation, followed by experimental results in Section 4. Finally, Section 5 gives a conclusion of this paper.

## 2. Image Sparse Model

Superresolution reconstruction is a reverse evaluation problem. Dictionary-based superresolution reconstruction trains the corresponding LR dictionary and HR dictionary by mapping the relationship between LR images and HR images. Then, the relationship between LR dictionary and HR dictionary is solved, and a HR image is reconstructed with recovery of the missing high-frequency information. Imaging model between LR image and HR image is as shown in (1): when the real scene is captured through the optical system, we obtain a group of LR image scenes due to the impact of optical blur and downsampling. Consider

$$Y = SHX. \quad (1)$$

$Y$  is a LR image captured by the camera.  $X$  is the real scene image.  $H$  is the optical blur matrix.  $S$  is the downsampling matrix.

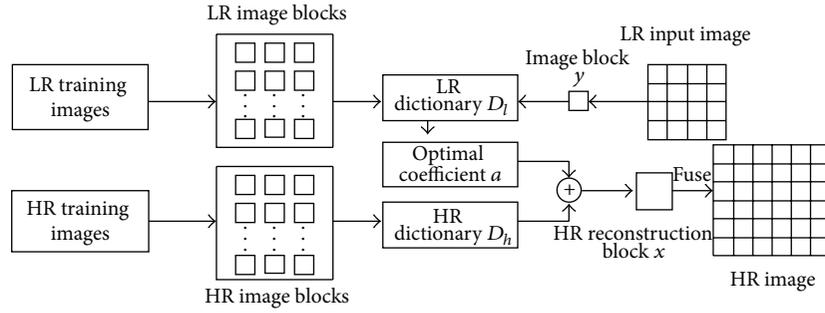


FIGURE 1: Simple diagram of general dictionary-based reconstruction algorithm.

Superresolution reconstruction is a highly ill-condition problem. Given one LR input image  $Y$ , many HR images  $X$  are suitable for formula (1). With the development of compressed sensing, dictionary-based superresolution methods have been applied more widely in image restoration areas. We can divide image into blocks  $y$  of the same dimension, every image block is represented by both sparse representation “ $a$ ” and sparse matrix  $D$ , such as  $y = D \times a$ , where  $D$  is a dictionary matrix and “ $a$ ” is sparse coefficients. The basic process of dictionary-based reconstruction is shown in Figure 1.

First, we require lots of training images and train the corresponding LR dictionary and HR dictionary by learning LR training images and HR training images. Second, the optimal sparse coefficients “ $a$ ” of LR input image block are solved according to the equation  $\min(y - D_l \times a)$ . Then, we get the corresponding HR image block  $x$  via solving the formula  $D_h \times a$  and fuse all HR image blocks into one HR image  $X$ .

### 3. The Proposed Method

When one LR image is given to reconstruct one HR image, the main problem of dictionary-based methods is how to find lots of HR-LR training images and utilize self-similarity information of external images. But it will cause the reconstructed HR image distortion, given the absence of the input images’ frequency information in the selected HR-LR training set. In this paper, we propose a multiple superresolution reconstruction method based on self-learning dictionary. We collect multiple images of the same scene as both input images and training image set, and larger-scale images, which are constructed from last training set, are also involved in the training set constantly. After that, we get the initial reconstruction images whose scale closely approximates target HR image. Then, we use NLM method to fuse the initial reconstruction images into one HR image. Lastly, IBP idea is adopted to satisfy the global reconstruction constraint. Here, four major parts are detailed in the following. In Section 3.1, we mainly discuss how to use input images to gradually magnify image and create a large number of different larger-scale images as training set. In Section 3.2, sparse representation is adopted to jointly learn redundant dictionary. In Section 3.3, we employ NLM method to fuse initial reconstruction images

into one HR image and introduce the importance of global reconstruction constraint. In Section 3.4, the working process of the proposed algorithm is introduced. In Section 3.5, we discuss how to simplify the operation.

**3.1. Self-Learning Method.** As we have said above, dictionary-based reconstruction algorithms need lots of training images to learn the overcomplete dictionary for constructing HR image. Literatures [25–28] point out that input image has enough self-similarity information and the high-frequency information of small-scale image can be obtained from large-scale image. They can use multiframe input images to reconstruct images accurately. Therefore, when external images cannot be accessed, this paper proposes a self-learning pyramid method which constructs larger-scale images gradually to join in training set. And it is feasible. We can fully learn all the redundant information from different scale images. First, assuming that the target HR image resolution is  $P$  times that of the input LR image resolution, we will enlarge input images resolution  $p_i$  times step by step, as the following formula:

$$\begin{aligned} p_1 \times p_2 \times \cdots \times p_n \times s &= P, \\ s < p_i, \quad i &= 1, \dots, n, \\ p_1 \times p_2 \times \cdots \times p_n &\leq P. \end{aligned} \quad (2)$$

Assuming that the input images are  $N \times N$  dimension, there are two cases for the principle of  $p$  as follows:

- $p_1 = p_2 = \cdots = p_n = f/g$ ;  $g^n \times w = N$  ( $w < g$  and  $f, g$  all are integers.);
- $p_1 = f_1/g_1, \dots, p_n = f_n/g_n$ ;  $f_i/g_j > 1$  ( $f_i, g_j$  all are integers,  $s > p_i$ , and  $i = 1, 2, \dots, n$ ).

We discuss this paper based on condition (a).

Then, the degraded images  $Y_1^\wedge$  are solved via the formula  $Y_1 \downarrow_p \uparrow_p$  ( $\downarrow_p$  is  $p$  times downsampling;  $\uparrow_p$  is  $p$  times upsampling), which together constitute the HR-LR training set  $Q_1$  with high-frequency part of input images  $Y_1$ , as the formula (3). (The degraded images  $Y_1^\wedge$  are LR training set. High-frequency part ( $Y_1 - Y_1^\wedge$ ) of input images  $Y_1$  is HR training set.) This is the difference between this paper and [25–27], and our method can recover the edge high-frequency better. We train  $Q_1$  to get dictionary  $D_1$  and construct the corresponding

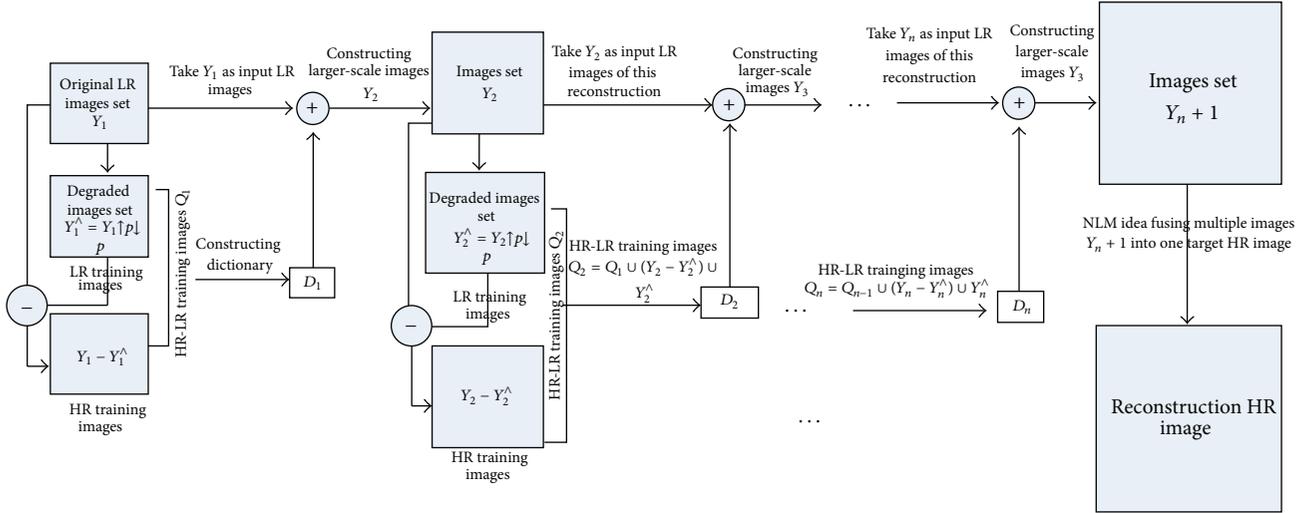


FIGURE 2: Process of training images.

larger-scale HR images  $Y_2$  whose resolution is  $p$  times that of the input images. Then, degraded images  $Y_2^\wedge$  are solved, which together constitute new HR-LR training set  $Q_2$  with high-frequency part of  $Y_2$  and  $Q_1$  again. It is as shown in (3) that we reconstruct final HR images whose resolution is closest to target HR images, followed by  $n$  repeated learning. Process of training images is as shown in Figure 2. We enlarge every input image step by step just like pyramid. Consider

$$\begin{aligned}
 Q_1 &= Y_1^\wedge \cup (Y_1 - Y_1^\wedge), \\
 Q_2 &= Y_2^\wedge \cup (Y_2 - Y_2^\wedge) \cup Q_1, \\
 &\vdots \\
 Q_m &= Y_m^\wedge \cup (Y_m - Y_m^\wedge) \cup Q_{m-1}.
 \end{aligned} \tag{3}$$

**3.2. Learning Dictionary.** In the previous section, we have discussed the self-learning pyramid method. Chen et al. [26] and Zhang et al. [27] use image blocks manifold to learn dictionary. They solve the optimal weight between input image block and several nearest local training blocks. But it cannot effectively utilize global image information. Yang et al. [22] point out that sparse representation can utilize global image information better and construct universal dictionary with a better reconstructed image. So we propose jointly learning dictionary method based on sparse representation in this section. As shown in Section 3.1 that  $Q_m$  is the training set, we use images  $Y_m$  of  $Q_m$  as input images of next reconstruction.  $x$  and  $y$  are HR training blocks and LR training blocks respectively, which are represented by atomic linear combination of matrix  $D$ :  $x = D_h \times \alpha$ ;  $y = D_l \times \beta$  ( $\alpha$  and  $\beta$  are sparse coefficients.  $D_h$  and  $D_l$  are HR dictionary and LR dictionary, resp.). In this section, the premise of learning dictionary is that sparse coefficients  $\alpha$  and  $\beta$  are the same. It is as shown in (4). Consider

$$\begin{aligned}
 y &= D_l \times \alpha, \\
 x &= D_h \times \alpha.
 \end{aligned} \tag{4}$$

By solving the sparse coefficients  $\alpha$  of LR dictionary  $D_l$ , we can get HR image blocks according to  $x = D_h \times \alpha$ . This is the basic idea of training dictionary. Before reconstructing HR image, we must learn dictionaries  $D_l$  and  $D_h$  which must satisfy constraints of (4) firstly. So we can convert (4) into the following optimization problem, as shown in (5). Consider

$$\begin{aligned}
 \min \|\alpha\|_0, \quad \text{s.t. } \|y - D_l \times \alpha\|_2^2 &\leq \varepsilon, \\
 \min \|\alpha\|_0, \quad \text{s.t. } \|y - D_h \times \alpha\|_2^2 &\leq \varepsilon.
 \end{aligned} \tag{5}$$

As long as the sparse coefficients  $\alpha$  are sparse enough, (5) is solved by solving  $L^1$  norm problem, as shown in (6) and (7). Consider

$$\operatorname{argmin} \frac{1}{2} \|x - D_h \times \alpha\|_2^2 + \mu \|\alpha\|_1, \tag{6}$$

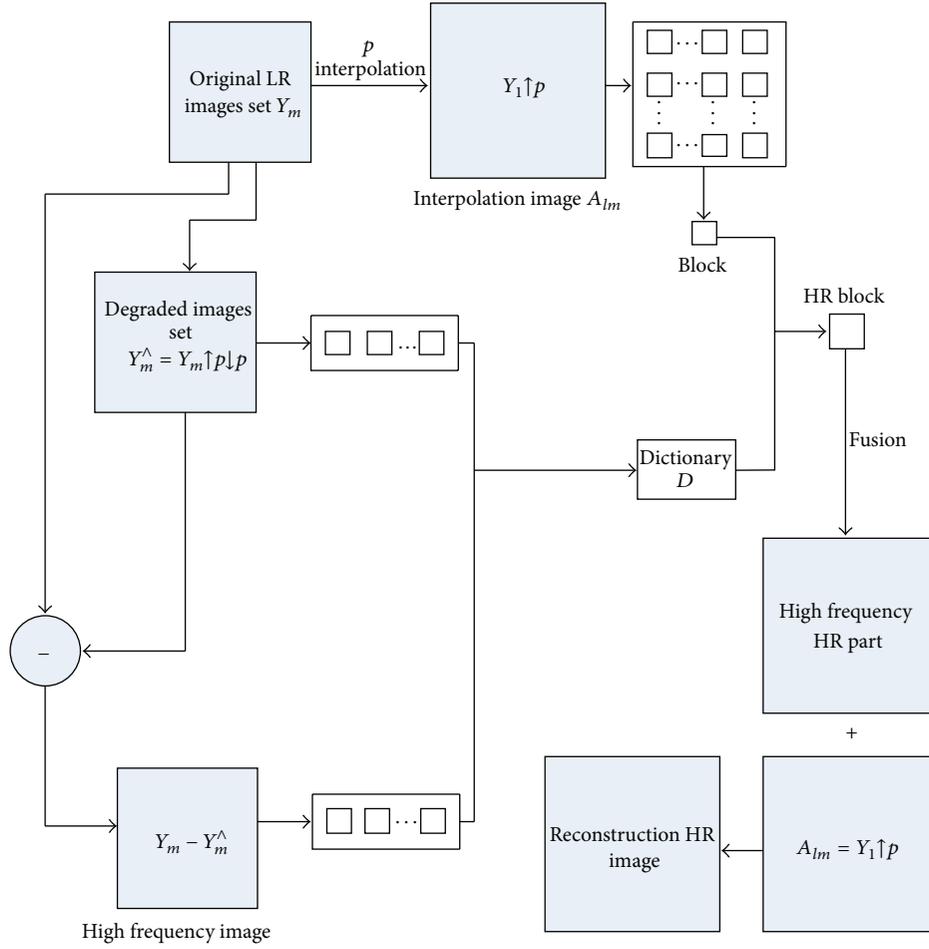
$$\operatorname{argmin} \frac{1}{2} \|y - D_l \times \alpha\|_2^2 + \mu \|\alpha\|_1. \tag{7}$$

In order to get HR dictionary and LR dictionary of the same coefficient  $\alpha$ , (6) and (7) can be converted into (8). Consider

$$\operatorname{argmin} \frac{1}{2} \|P - D \times \alpha\|_2^2 + \mu \|\alpha\|_1. \tag{8}$$

In (8),

$$P = \begin{pmatrix} \frac{x}{\sqrt{N}} \\ \frac{y}{\sqrt{M}} \end{pmatrix}, \quad D = \begin{pmatrix} \frac{D_h}{\sqrt{N}} \\ \frac{D_l}{\sqrt{M}} \end{pmatrix}, \tag{9}$$


 FIGURE 3: Process of constructing dictionary and reconstructing image (the upscale factor is  $p$ ).

where  $N$  is dimension of HR image blocks and  $M$  is dimension of LR image blocks. In order to solve (8), we set the initial value of matrix  $D$  and (8) can be converted into (10). Consider

$$\min \frac{1}{2} (P - D \times \alpha)^T (P - D \times \alpha) + \mu |\alpha|. \quad (10)$$

Then, the transformation of (10) is as shown in (11). Consider

$$\min \frac{1}{2} Q' \alpha Q + p' \alpha + \mu |\alpha|, \quad (11)$$

where  $Q = D^T D$  and  $p = D^T y$ . After obtaining  $\alpha$ , we can solve (12) to get the value of matrix  $D$ . Consider

$$\operatorname{argmin} \|P - D \times \alpha\|_2^2. \quad (12)$$

Therefore, dictionaries  $D_l$  and  $D_h$  are obtained. Then, we enlarge every input image of  $Y_m$  one by one, as shown in Figure 3. First, every input image is enlarged to  $A_{lm}$  by interpolation and the upscale factor is  $p$ . Second,  $A_{lm}$  is divided into blocks  $y$  of the same size. According to dictionaries  $D_l$  and  $D_h$ , we solve (7), (8), and (10) to get  $\alpha$ . Then, HR image

blocks are obtained according to the formula  $x = D_h \times \alpha$ . When all of the HR image blocks are gained, they are merged into one HR image  $A_{hlm}$  according to the fixed order.  $A_{hlm} + A_{lm}$  is the larger-scale reconstructed image and also is one of reconstruction images  $Y_{m+1}$ , as shown in Figure 3. Lastly, we obtain the reconstruction images set  $Y_{m+1}$  by using dictionary to construct every image in  $Y_m$ . Then, image sets  $Y_{m+1}$  and  $Q_{m+1}$  will become the input images and training images set of next reconstruction enlargement, respectively.  $Q_{m+1}$  is as shown in (13). Consider

$$Q_{m+1} = Y_{m+1}^{\wedge} \cup (Y_{m+1} - Y_{m+1}^{\wedge}) \cup Q_m. \quad (13)$$

**3.3. Reconstruction Fusion.** In Sections 3.1 and 3.2, we use self-learning pyramid method to construct training images and get reconstruction set  $Y_{n+1}$ . But the scale of reconstruction images is  $p^n$  times that of the original input images' scale which may not be the same scale with target HR image. So we use NLM idea to fuse all images of  $Y_{n+1}$  into one HR image which is the same scale with target HR image. Meanwhile, it can effectively eliminate fusion noise.

Nonlocal mean (NLM) filter is a very effective denoising filter, which is based on an assumption that each pixel is a repetition of surrounding pixels. It can be used to obtain the

target pixel value with a weighted sum of the surrounding pixels. Now, NLM is widely used in the field of super-resolution reconstruction. Just as [9], when we use NLM idea to reconstruct multiframe images, the smaller the magnification is, the smaller the reconstruction error is. So when image dimension of  $Y_{n+1}$  is closest to target HR image, using NLM idea to process  $Y_{n+1}$  is feasible. In this section, we adopt (14) to fuse images of  $Y_{n+1}$  into target HR image  $X_0$ . Consider

$$X = \operatorname{argmin} \sum_{(k,l) \in \omega} \sum_{t=1}^{\rho} \sum_{(i,j) \in N_{k,l}} \omega(k, l, i, j, t) \times \|D_{k,l} E_{k,l}^H X - E_{k,l}^L y_t\|, \quad (14)$$

where  $\omega(k, l, i, j, t)$  is the weight coefficient, usually calculated by the following formula. Consider

$$\omega(k, l, i, j, t) = \exp\left(-\frac{\|P_{k,l} y_t - P_{i,j} y_t\|^2}{2\sigma^2}\right). \quad (15)$$

The  $(k, l)$  is a pixel of the target HR image.  $y_t$  represents one of the image sets,  $Y_{n+1}$ .  $(i, j)$  is a pixel of image  $y_t$ .  $P_{k,l} y_t$  represents extracting image block from image  $y_t$  whose center is  $(k, l)$ . After amplification and denoising of reconstruction image, we get the prime estimation image  $X_0$ . Finally, we must add global reconstruction constraint to meet  $Y = SHX$ , as shown in (16). Consider

$$X = \operatorname{argmin} \|y - SHX\|_2^2 + \mu \|X - X_0\|_2^2. \quad (16)$$

Turn into the following formula (17). Consider

$$X = X_0 + SH^T (y - SHX) + \mu (X - X_0). \quad (17)$$

We can solve (17) to obtain the optimal estimated HR image  $X$  via IBP idea.

**3.4. The Proposed Algorithm.** The process of the proposed algorithm is as shown in the following writing based on the above sections. Suppose that we have four LR images of the same scene.

*Objective: Reconstruct HR Image from Multiple LR Images*

- (1) Initialization: input four LR images  $Y_1$  and solve the corresponding degrade images by formula  $Y_1 \downarrow_p \uparrow_p$ ; construct the first training image set  $Q_1$ ; set magnification  $p$  and iterations  $n$ .
- (2) For  $m = 1 : n$ , consider the following.
  - (a) HR image and LR image of the HR-LR training set  $Q_m$  are divided into image blocks  $x$  and  $y$ , respectively. Training dictionaries  $D_h$  and  $D_l$  are solved according to section (b) and the following formula. Consider

$$\begin{aligned} \operatorname{argmin} \frac{1}{2} \|x - D_h \times \alpha\|_2^2 + \mu \|\alpha\|_1, \\ \operatorname{argmin} \frac{1}{2} \|y - D_l \times \alpha\|_2^2 + \mu \|\alpha\|_1. \end{aligned} \quad (18)$$

- (b) Take  $Y_m$  as new input LR images to be reconstructed. And solve new HR images  $Y_{m+1}$  from dictionaries  $D_h$  and  $D_l$ . Then, we get four reconstruction images  $Y_{m+1}$ , in which resolution is  $p^m$  times of the original input images.
- (c) Put high-frequency of  $Y_{m+1}$  and corresponding degrade images into the HR training set and LR training set, respectively. So we get new training image set according to section (a) and (3). Take  $Y_{m+1}$  as new LR images to be reconstructed. Consider

$$\begin{aligned} Y_m^\wedge &= ((Y_m) \downarrow_p) \uparrow_p, \\ Q_{m+1} &= Y_{m+1}^\wedge \cup (Y_{m+1} - Y_{m+1}^\wedge) \cup Q_m. \end{aligned} \quad (19)$$

- (3) End
- (4) Adopt NLM idea to fuse images of  $Y_{n+1}$  into one the same scale HR image  $X_0$  with target image.
- (5) Add global reconstruction constraint to meet  $Y = SHX$ . Use IBP idea and gradient descent algorithm to solve (16) and obtain the optimal reconstruction HR image  $X$ . Consider

$$X = X_0 + SH^T (y - SHX) + \mu (X - X_0). \quad (20)$$

- (6) Output the final HR image  $X$ .

**3.5. The Discussion of Reconstruction Speed.** When we apply the NLM method to adjust reconstruction image in Section 3.3, it will spend much operation time by solving (14) directly. So, in this section, we will accelerate reconstruction speed by simplifying section (c) in this section. Set derivation of the (14) to be equal to zero. We can get (21). Consider

$$\begin{aligned} X &= \left( \sum_{(k,l) \in \omega} \sum_{t=1}^{\rho} \sum_{(i,j) \in N_{k,l}} \omega(k, l, i, j, t) \right. \\ &\quad \left. \times (E_{k,l}^H)^T D_{k,l}^T D_{k,l} E_{k,l}^H \right)^{-1} \\ &\quad \times \left[ \sum_{(k,l) \in \omega} \sum_{t=1}^{\rho} \sum_{(i,j) \in N_{k,l}} \omega(k, l, i, j, t) \times (E_{k,l}^H)^T D_{k,l}^T E_{k,l}^H y_t \right]. \end{aligned} \quad (21)$$

$E_{k,l}^L$  represents extracting one pixel from the HR image. Equation (21) becomes (22). Consider

$$X = \frac{\sum_{t=1}^{\rho} \sum_{(i,j) \in N_{k,l}} \omega(k, l, i, j, t) y_t(i, j)}{\sum_{t=1}^{\rho} \sum_{(i,j) \in N_{k,l}} \omega(k, l, i, j, t)}. \quad (22)$$



FIGURE 4: The input LR images. (a) is one of several Lena images. (b) is one of several building images. (c) is one of several flower images.

As described above, it will reduce operation time and avoid complicated matrix operations. Meanwhile, this algorithm can be parallelized to deal with such steps as 2(a), 2(b), and (4) of Section 3.4. Therefore, we can take advantage of GPU to further improve the reconstruction speed. As shown in Section 4.2, we use Gforce 780 M GPU to process the captured images in camera system experimentation.

Some blocks of the input LR image are smooth with few details, as shown in Figure 4 (the black box). These blocks contain little high-frequency information and have no effect for reconstruction HR image. So this section proposes interpolation algorithm to enlarge these blocks, and the algorithm can run faster. Smooth image block is defined by the following formula:

$$A = \sqrt{\frac{1}{p^2} (\bar{z} - z_i^l)^2}, \quad (23)$$

where  $z_i^l$  is the pixel value of image block and  $\bar{z}$  is the average value of block pixel. We set a fixed value  $A^0$ . If  $A < A^0$ , we use interpolation method to reconstruct the block.

## 4. Experimental Results and Analysis

**4.1. Simulation Experimentation.** In this section, we perform 2x magnification experiments on three test images from Figure 4 to validate the effectiveness of the proposed method. These images cover a lot of contents including Lena, flower, and building, as shown in Figure 4. The compared group includes five other methods: the interpolation-based method [4], the NLM method based on reconstruction [8], Glassner's method [25], Yang's method [23], and Zhang's method [27]. For example, the input LR Lena images are simulated from a  $256 \times 256$  normative HR image revolved, translated randomly, and downsampled by a factor of two. The image block size is set to  $5 \times 5$  with overlap of 4 pixels between adjacent blocks. And upscale factor  $P$  is 2. The gradual enlargement  $p$  is set to 1.25,  $n$  is 3, and  $s$  is 128/125. Although the proposed method has a certain similarity in the idea of dictionary-based reconstruction [23], it takes the following unique features.

- (1) This paper proposes self-learning pyramid method to construct larger-scale image step by step. And it does

not need to search lots of external HR-LR images as a training set.

- (2) The reconstructed result depends on the selected training images. It will cause the reconstructed HR image distortion, given the absence of the input images' frequency information in the selected HR-LR training set. This method puts different scale reconstructed image into the training set, and it can provide much real similarity information to ensure the authenticity of the reconstruction HR image.
- (3) Because the reconstruction results have fusion noise and may not be the same scale with target HR image in Section 3.3, we use NLM idea to meet both cases and merge them into one HR image, which contains more target image information. Lastly, we use IBP idea to meet global reconstruction constraint.

Figures 5, 6, and 7 are the reconstruction results of the testing images Lena, flower, and building.

We prove the superiority of the proposed algorithm from the visual quality and objective quality of the reconstructed HR image. And peak signal to noise ratio (PSNR), structural similarity (SSIM), mean squared error (MSE), and mean absolute error (MAE) are employed to evaluate the objective quality of the reconstructed HR image. Definitions are as follows.

Mean squared error is as follows:

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (g(i, j) - f(i, j))^2. \quad (24)$$

Peak signal to noise ratio is as follows:

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}}. \quad (25)$$

Mean absolute error is as follows:

$$\text{MAE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |g(i, j) - \bar{g}|. \quad (26)$$

The  $g(i, j)$  and  $f(i, j)$  are pixels of the reconstructed HR image and the original HR image, respectively. The  $M$  and  $N$

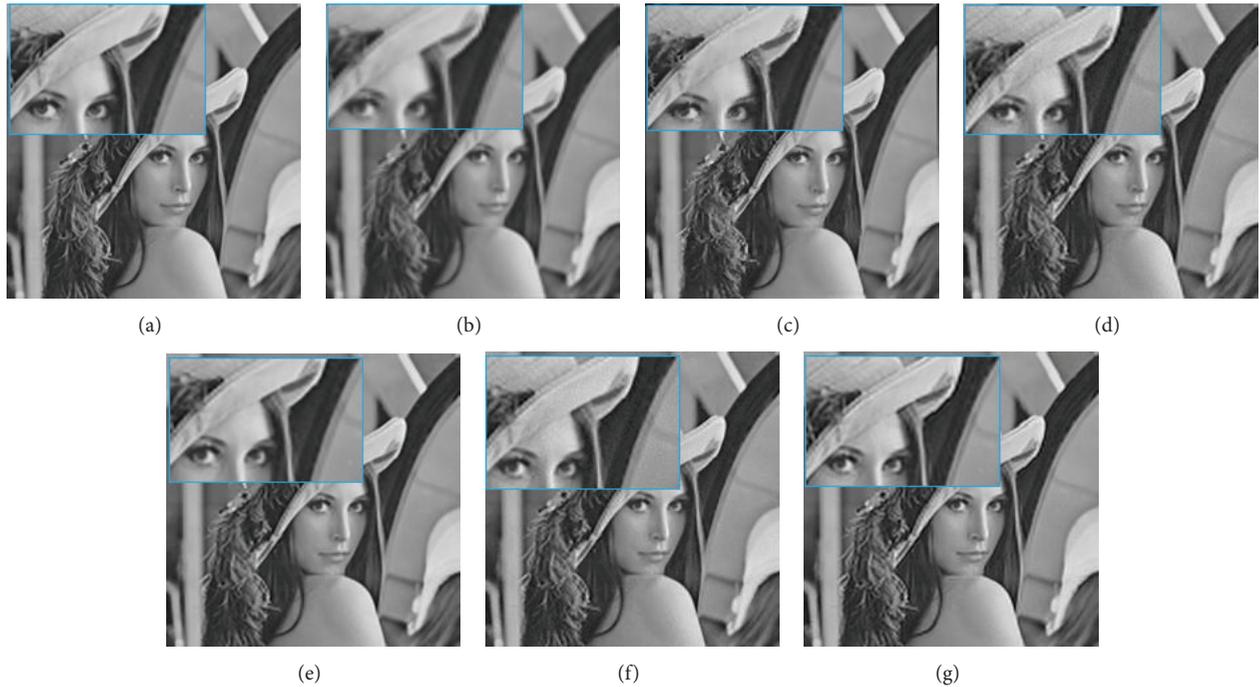


FIGURE 5: Comparison of reconstructed HR images of Lena. (a) Original HR image. (b) Interpolation-based method results [4]. (c) NLM results [8]. (d) Glassner's method [25]. (e) Yang's method [23]. (f) Zhang's method [27]. (g) The proposed method results.

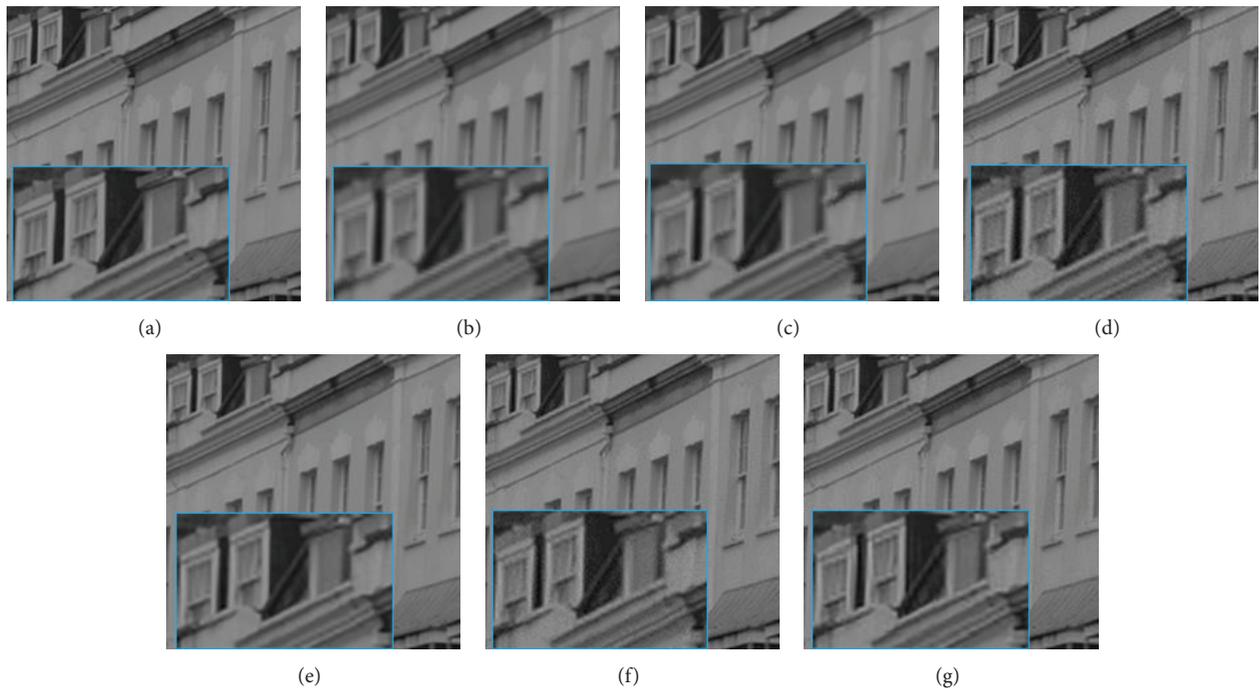


FIGURE 6: Comparison of reconstructed HR images of building. (a) Original HR image. (b) Interpolation-based method results [4]. (c) NLM results [8]. (d) Glassner's method [25]. (e) Yang's method [23]. (f) Zhang's method [27]. (g) The proposed method results.

are dimensions of reconstructed HR image. The  $\bar{g}$  is the mean value of all pixels of reconstructed HR image. When PSNR is larger and MAE is smaller, the quality of the reconstructed image is better. SSIM is to show the reconstructed quality by comparing image intensity, image contrast, and structural

similarity based on the visual system. The larger the SSIM is, the clearer the image is. Values of PSNR, SSIM, and MAE are shown in Table 1. The first, second, and third rows for each test image indicate PSNR, MAE, and SSIM values, respectively. Table 1 suggests that the proposed algorithm

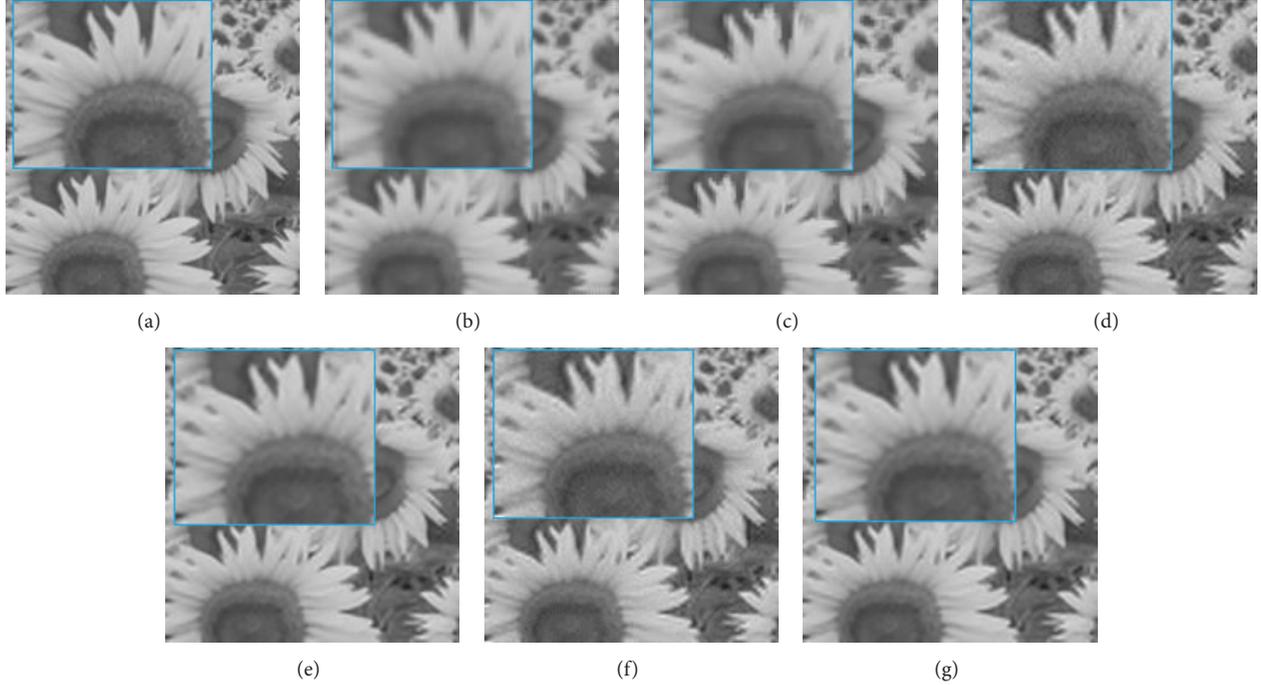


FIGURE 7: Comparison of reconstructed HR images of flower. (a) Original HR image. (b) Interpolation-based method results [4]. (c) NLM results [8]. (d) Glassner's method [25]. (e) Yang's method [23]. (f) Zhang's method [27]. (g) The proposed method results.

can achieve better reconstructed result and recover more missing high-frequency information and clearer edge details in comparison with other five methods.

To further assess the visual quality obtained by different methods, we compare reconstructed image details of different methods with original HR image, as shown in Figures 5, 6, and 7. When we contrast eye and the brim of a hat in Figure 5, the result of interpolation-based method is vaguer and NLM method's reconstructed result appears distortion and mosaic. On the contrary, the proposed method does not cause distortion and vague compared with interpolation-based and NLM cases. Moreover, reconstructed details of our method are more delicate than other algorithms. In Figures 6 and 7, the proposed method can also produce better reconstructed details, eaves and flower border, both accurately and visually in comparison with other superresolution algorithms.

**4.2. Camera System Experimentation.** In this section, we conduct research on real scene reconstruction to prove the superiority of the proposed algorithm. It is the camera system assembled by ourselves, which consists of a microdisplacement lens, PI turntable, and Gforce 780 M GPU. By controlling the PI turntable, we can continuously receive multiple real images, as shown in Figure 8(a) (one of multiple real images). The images are taken to be reconstructed by our method, and result is as shown in Figure 8(f). Figures 8(b)–8(e) are the reconstruction results of the NLM method based on reconstruction [9], Glassner's method [25], Yang's method, and Zhang's method [27]. Signal noise ratio (SNR), average gradient (AG), information entropy (IE), and standard deviation (SD), which do not require reference HR

image, are employed to evaluate the objective quality of the reconstructed real HR image in real scene reconstruction of camera system. Definitions are as follows. Consider

$$\text{SNR} = \frac{M}{\text{STD}_{\max}}. \quad (27)$$

$M$  represents image mean value. The  $\text{STD}_{\max}$  represents the max value of local standard deviation. Consider

$$\text{IE} = \sum_{i=0}^{L-1} P_i \log_2 P_i. \quad (28)$$

$L$  is the total image gray level, and  $P_i$  represents the ratio of pixels in gray value  $i$  to pixels in all image gray values. Consider

$$\begin{aligned} \text{AG} &= \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \sqrt{\left( \left( \frac{\partial g(x_i, y_j)}{\partial x_i} \right)^2 + \left( \frac{\partial g(x_i, y_j)}{\partial y_j} \right)^2 \right)} \times \frac{1}{2} \\ &\quad \times ((M-1)(N-1))^{-1}, \\ \text{SD} &= \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (g(x_i, y_j) - \bar{g})^2}{M \times N}}. \end{aligned} \quad (29)$$

The average gradient represents clarity. Information entropy can indicate image information abundance. Standard deviation represents discrete situation of image gray level. When SNR, SD, IE, and AG are larger, the reconstructed

TABLE 1: The results of PSNR, MAE, and SSIM. For every image, there are three rows. The first, second, and third rows for each test image indicate PSNR, MAE, and SSIM values.

|          | Interpolation [4] | NLM method [8] | Glassner's method [25] | Yang's method [23] | Zhang's method [27] | The proposed method |
|----------|-------------------|----------------|------------------------|--------------------|---------------------|---------------------|
| Lena     | 27.3993           | 27.4644        | 30.6150                | 31.0123            | 31.5623             | <b>32.6016</b>      |
|          | 6.0320            | 5.8437         | 4.2831                 | 3.9561             | 3.8601              | <b>3.7735</b>       |
|          | 0.8631            | 0.8643         | 0.9051                 | 0.9168             | 0.9250              | <b>0.9538</b>       |
| Building | 31.2571           | 31.2962        | 37.1000                | 37.0549            | 36.9531             | <b>37.6990</b>      |
|          | 4.1998            | 4.0993         | 2.3300                 | 2.3254             | 2.5365              | <b>2.0510</b>       |
|          | 0.8746            | 0.8753         | 0.9537                 | 0.9524             | 0.9124              | <b>0.9653</b>       |
| Flower   | 26.3318           | 25.7424        | 30.5731                | 31.2439            | 31.3826             | <b>32.1054</b>      |
|          | 8.0847            | 8.3127         | 5.240                  | 4.6710             | 4.6425              | <b>4.4371</b>       |
|          | 0.8176            | 0.7931         | 0.9031                 | 0.9280             | 0.9315              | <b>0.9501</b>       |
| Average  | 28.3294           | 28.1676        | 32.76                  | 33.1037            | 33.2993             | <b>34.1354</b>      |
|          | 6.1055            | 6.0852         | 3.9510                 | 3.6508             | 3.6797              | <b>3.4205</b>       |
|          | 0.8517            | 0.8442         | 0.9206                 | 0.9324             | 0.9229              | <b>0.9564</b>       |

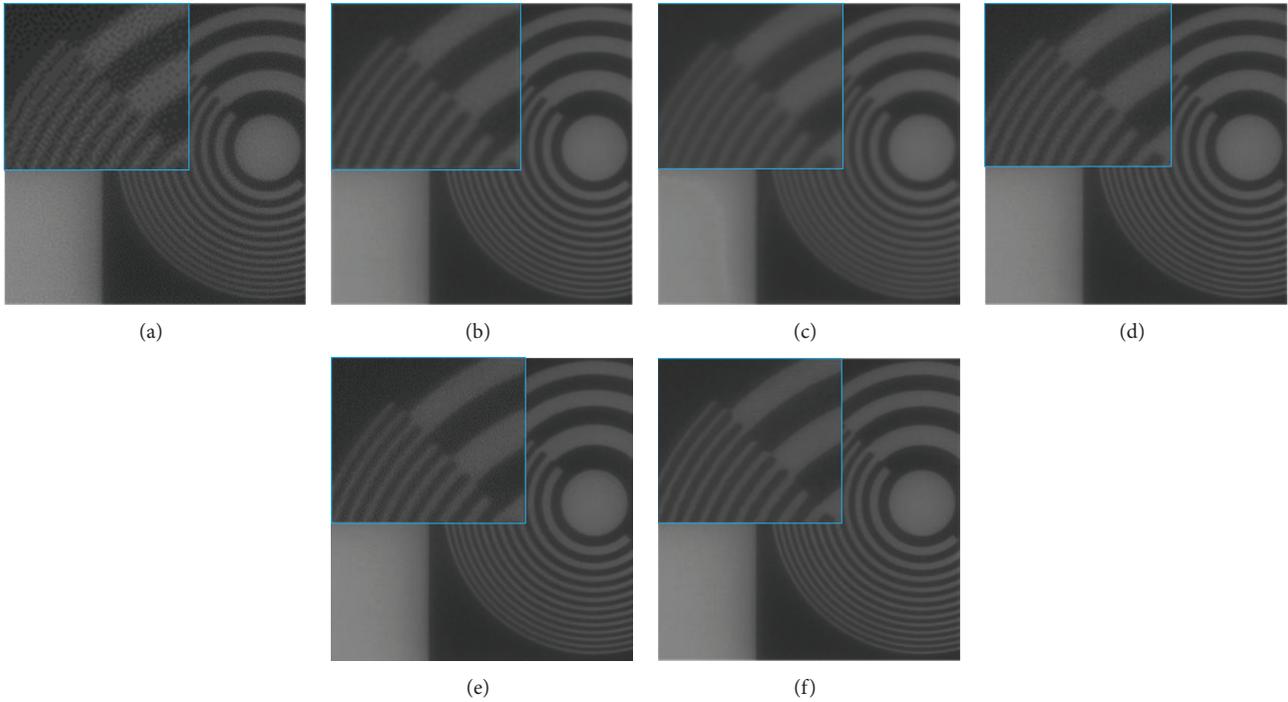


FIGURE 8: Real scene reconstruction. (a) One of the captured images. (b) NLM results [8]. (c) Glassner's method [25]. (d) Yang's method [23]. (e) Zhang's method [27]. (f) The proposed method results.

image is clearer and contains richer information. Therefore, the numeric results of proposed method are better than those of other algorithms as shown in Table 2. Visually, in comparison of contour outline of the reconstructed image detail, the reconstructed image in Figure 8(f) is significantly clearer than the other reconstructed image. Consequently, the proposed method can work out well in the camera system experimentation and be capable of reproducing plausible details and sharp edges with minimal artifacts.

## 5. Conclusion

It is known that, in superresolution reconstruction, distortion of reconstructed images occurs when frequency information

of target images is not contained in training images. To solve this problem, this paper proposes a multiple super-resolution reconstruction algorithm based on self-learning dictionary. First, images on a series of scales created from multiple input LR images are used as training images. Then, a learned overcomplete dictionary provides sufficient real image information that ensures the authenticity of the reconstructed HR image. In both simulation and real experiments, the proposed algorithm achieves better results than that achieved by other algorithms, recovering missing high-frequency information and edge details more effectively. Finally, several ideas towards higher computation speed are suggested.

TABLE 2: The results of SNR, AG (average gradient), IE (information entropy), and SD (standard deviation).

|     | Real scene | NLM method [8] | Glassner's method [25] | Yang's method [23] | Zhang's method [27] | The proposed method |
|-----|------------|----------------|------------------------|--------------------|---------------------|---------------------|
| SNR | 5.3115     | 6.7257         | 6.6978                 | 6.7899             | 6.8049              | <b>7.1254</b>       |
| AG  | 1.1059     | 1.1895         | 1.7399                 | 1.7296             | 1.7327              | <b>1.8138</b>       |
| IE  | 6.3191     | 6.3151         | 6.3536                 | 6.3411             | 6.3609              | <b>6.4141</b>       |
| SD  | 26.2563    | 26.2807        | 26.2999                | 26.3164            | 26.3043             | <b>26.3733</b>      |

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the project of National Science Fund for Distinguished Young Scholars of China (Grant no. 60902067) and the key Science-Technology Project of Jilin province (Grant no. 11ZDGG001).

## References

- [1] P. Thevenaz, T. Blu, and M. Unser, *Image Interpolation and Resampling*, Academic Press, New York, NY, USA, 2000.
- [2] H. S. Hou and H. C. Andrews, "Cubic spline for image interpolation and digital filtering," *IEEE Transactions on Signal Process*, vol. 26, no. 6, pp. 508–517, 1978.
- [3] N. A. Dodgson, "Quadratic interpolation for image resampling," *IEEE Transactions on Image Processing*, vol. 6, no. 9, pp. 1322–1326, 1997.
- [4] V. Patrick, S. Sabine, and V. Martin, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 71459, 14 pages, 2006.
- [5] M. Irani and S. Peleg, "Improving resolution by image registration," *Graphical Models and Image Processing*, vol. 53, no. 3, pp. 231–239, 1991.
- [6] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," in *Maximum Entropy and Bayesian Methods: Santa Barbara, California, U.S.A., 1993*, vol. 62 of *Fundamental Theories of Physics*, pp. 293–308, Springer, 1996.
- [7] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of under-sampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, 1997.
- [8] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.
- [9] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 36–51, 2009.
- [10] J.-D. Li, Y. Zhang, and P. Jia, "Super-resolution reconstruction using nonlocal means," *Optics and Precision Engineering*, vol. 21, no. 6, pp. 1576–1585, 2013.
- [11] C. Yang, J. Huang, and M. Yang, "Exploiting self-similarities for single frame super-resolution," in *Proceedings of the 10th Asian Conference on Computer Vision (ACCV '10)*, vol. 3, pp. 497–510, 2010.
- [12] K. Zhang, X. Gao, X. Li, and D. Tao, "Partially supervised neighbor embedding for example-based image super-resolution," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 2, pp. 230–239, 2011.
- [13] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*, pp. 711–730, Springer, 2010.
- [14] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2400–2407, San Francisco, Calif, USA, June 2010.
- [15] J. Sun, J. Zhu, and M. F. Tappen, "Context-constrained hallucination for image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 231–238, IEEE, San Francisco, Calif, USA, June 2010.
- [16] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [17] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 1, pp. 275–282, 2004.
- [18] Q. Liu, L. Liu, Y. Wang, and Z. Zhang, "Locally linear embedding based example learning for pan-sharpening," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR '12)*, pp. 1928–1931, Tsukuba, Japan, 2012.
- [19] Y. Tao, Z. L. Gan, and X. C. Zhang, "Novel neighbor embedding super resolution method for compressed images," in *Proceedings of the 4th International Conference on Image Analysis and Signal Processing (IASP '12)*, pp. 1–4, November 2012.
- [20] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi Morel, "Neighbor embedding based single-image super-resolution using semi-nonnegative matrix factorization," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '12)*, vol. 1, pp. 1289–1292, Kyoto, Japan, March 2012.
- [21] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.
- [22] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [23] L. He, H. Qi, and R. Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 345–352, Portland, Ore, USA, June 2013.
- [24] S.-C. Jeong, Y.-W. Kang, and B. C. Song, "Super-resolution algorithm using noise level adaptive dictionary," in *Proceedings of the 14th IEEE International Symposium on Consumer Electronics (ISCE '10)*, June 2010.

- [25] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proceedings of the 12th International Conference on Computer Vision (ICCV '09)*, pp. 349–356, Kyoto, Japan, October 2009.
- [26] S. Chen, H. Gong, and C. Li, "Super resolution from a single image based on self similarity," in *Proceedings of the International Conference on Computational and Information Sciences (ICCIS '11)*, pp. 91–94, Chengdu, China, October 2011.
- [27] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with multiscale similarity learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 10, pp. 1648–1659, 2013.
- [28] S. A. Ramakanth and R. V. Babu, "Super resolution using a single image dictionary," in *Proceedings of the IEEE International Conference on Electronics, Computing and Communication Technologies (IEEE CONECCT '14)*, pp. 1–6, Bangalore, India, January 2014.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

