

RESEARCH

Open Access



Online multi-frame super-resolution of image sequences

Jieping Xu^{1*} , Yonghui Liang², Jin Liu², Zongfu Huang² and Xuewen Liu²

Abstract

Multi-frame super-resolution recovers a high-resolution (HR) image from a sequence of low-resolution (LR) images. In this paper, we propose an algorithm that performs multi-frame super-resolution in an online fashion. This algorithm processes only one low-resolution image at a time instead of co-processing all LR images which is adopted by state-of-the-art super-resolution techniques. Our algorithm is very fast and memory efficient, and simple to implement. In addition, we employ a noise-adaptive parameter in the classical steepest gradient optimization method to avoid noise amplification and overfitting LR images. Experiments with simulated and real-image sequences yield promising results.

Keywords: Super-resolution, Image processing, Image sequences, Online image processing

1 Introduction

Images with higher resolution are required in most electronic imaging applications such as remote sensing, medical diagnostics, and video surveillance. For the past decades, considerable advancement has been realized in imaging system. However, the quality of images is still limited by the cost and manufacturing technology [1]. Super-resolution (SR) is a promising digital image processing technique to obtain a single high-resolution image (or sequence) from multiple blurred low-resolution images.

The basic idea of SR is that the low-resolution (LR) images of the same scene contain different information because of relative subpixel shifts; thus, a high-resolution (HR) image with higher spatial information can be reconstructed by image fusion. Subpixel motion can occur due to movement of local objects or vibrating of imaging system, or even controlled micro-scanning [2, 3]. Numerous SR algorithms have been proposed since the concept was introduced by Tsai and Huang [4] in the year of 1984. Most of them operate in batch mode, i.e., a sequence of images are co-processed at the same time. Thus, these algorithms require a high memory resource to store the LR images and temporary data, and need a high computing resource as well. These disadvantages limit their practical application.

There are a variety of SR techniques, including multi-frame SR and single-frame SR. Readers can refer to Refs. [1, 5, 6] for an overview of this issue. Our discussion below is limited to work related to fast multi-frame SR method, as it is the focus of our paper.

One SR category close to our algorithm is dynamic SR which means estimating a sequence of HR images from a sequence of LR images. A natural approach is to make an extension of the static SR methods, as is adopted by some video SR algorithms [7, 8]. Although these batch-based dynamic SR algorithms have the ability to enhance the resolution of images/videos, the high memory and computational requirements restrict their practical use. Farsiu et al. [9] proposed a fast and memory-efficient algorithm for dynamic demosaicing and color SR of video sequences. They first got a blurry HR image by shifting and adding each LR frame recursively based on the Kalman filter, then the blurry HR image was deconvolved based on maximum a posteriori (MAP) technique. To make Farsiu's algorithm more robust, Kim et al. [10] proposed a measurement validation method based on Mahalanobis distance to eliminate the LR images whose motion estimation errors were larger than a threshold. Recently, researchers pay more attention to SR methods based on neural network (NC)[11, 12], which can be trained from data to approximate complex nonlinear functions, while these algorithms require significant computational resources to achieve real-time throughput. Graphic processing unit (GPU) is used to accelerate, as done by Hu et al. [13–16].

*Correspondence: civftor@163.com

¹Space Engineering University, No. 1, Bayi Road, Beijing, 101416 China
Full list of author information is available at the end of the article

Stefan Harmeling et al. [17] are the first to propose a blind deconvolution algorithm for astronomical imaging working in online mode. Instead of minimizing the total overall cost function for all blurry images, they just minimized the cost function for one image at each time and the current deblurred image was used as the initial estimate for the next time recursively. Thus, the deblurred image was refined gradually along with the image capturing. Hirsch et al. [18] extended the online blind deconvolution algorithm by incorporating super-resolution. The different information contained in different LR images results from blurring with different blur kernels. Both online blind deconvolution (OBD) and its SR extension (OBDSR) get the HR image estimate by minimizing an auxiliary function rather than minimizing the cost function directly to avoid image overfitting, while sometimes the strategy failed, especially when the noise of LR images is large.

Stimulated by the online blind deconvolution algorithm proposed by Stefan Harmeling et al. [17], we propose a multi-frame super-resolution algorithm operating in online mode which we name as online SR (OSR). Different from the OBDSR, the different information provided by different LR images in our algorithm results from relative motion between each other instead of different blur kernels. Another difference is that we employ a modified steepest gradient optimization method to estimate the HR image instead of an auxiliary function, which is proved to be more robust to noise. In this paper, the LR images are assumed to be obtained in a streaming fashion with a common space-invariant blur, and the relative motion between each other can be modeled as translation. At each time, we get an LR image, and we immediately use it to improve the current HR image estimate.

The rest of this paper is organized as follows. Section 2 addresses the formulation of the SR problem. In Section 3, we lay out the details of our algorithm and make it compatible with color images. In Section 4, we present experimental results with simulated and real-image sequences. The paper will conclude in Section 5 with a brief summary and future work plan.

2 Problem statement

To study the super-resolution problem, we first formulate the observation model. For notational simplicity, the image in our method is represented as a lexicographically ordered vector. Let us consider the desired HR image of size $LM \times LN$, where L denotes the down-sampling factor in the observation model and M and N are the number of rows and columns of the LR images. The most commonly used observation model is given by

$$\mathbf{y}_t = \mathbf{D}_t \mathbf{B}_t \mathbf{M}_t \mathbf{x} + \mathbf{n}_t, \quad \text{for } t = 1, 2, \dots, T \quad (1)$$

where \mathbf{x} denotes the lexicographically ordered HR image of size $L^2MN \times 1$, \mathbf{M}_t is the subpixel motion matrix of

size $L^2MN \times L^2MN$, \mathbf{B}_t represents blur matrix of size $L^2MN \times L^2MN$, \mathbf{D}_t denotes the down-sampling matrix of size $MN \times L^2MN$, \mathbf{n}_t is the lexicographically ordered noise vector, and \mathbf{y}_t represents the lexicographically ordered LR image vector of size $MN \times 1$.

Mathematically, the SR problem is an inverse problem whose objective is to estimate an HR image from multiple observed blurry LR images. MAP estimator is a powerful tool to solve this type of problem. The MAP optimization problem can be represented by

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max \{ \ln P(\{\mathbf{y}_t\} | \mathbf{x}) + \ln P(\mathbf{x}) \} \quad (2)$$

where $\{\mathbf{y}_t\}$ means all LR images, and $P(\mathbf{x})$ represents a priori knowledge of the HR image. However, some complex priori models used by the MAP estimator may increase the computational burden remarkably. If we omit the priori term, the MAP estimator degenerates to a maximum likelihood (ML) estimator. Under the assumption of independent and identical Gaussian distributed image noise with zero mean, ML optimization problem can be simplified as a least square problem, i.e.,

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \min \sum_{t=1}^T \|\mathbf{y}_t - \mathbf{D}_t \mathbf{B}_t \mathbf{M}_t \mathbf{x}\|^2 \quad (3)$$

Generally, due to the ill-posed nature of SR inverse problem, the MAP estimator is superior to the ML estimator. For both MAP estimator and ML estimator, all LR images should be stored in memory; thus, a high memory resource is required.

3 The proposed method

3.1 The loss function

In our OSR algorithm, when we get a new LR image in the data stream, we retrieve the new information immediately and add it to the HR image estimate. The process is realized by minimizing the loss function incurred at time t instead of overall loss. The problem can be expressed as a non-negatively constrained problem:

$$\hat{\mathbf{x}}_t = \min_{\mathbf{x} \geq 0} J_t(\mathbf{x}; \mathbf{y}_t) = \min_{\mathbf{x} \geq 0} \|\mathbf{y}_t - \mathbf{D}_t \mathbf{B}_t \mathbf{M}_t \mathbf{x}\|^2 \quad (4)$$

The motion matrix \mathbf{M}_t could vary in time, while the down-sampling matrix \mathbf{D}_t and blur matrix \mathbf{B}_t remain constant over time for most situations (i.e., $\mathbf{B} = \mathbf{B}_t$, $\mathbf{D} = \mathbf{D}_t$). We further assume the relative motion between LR images can be modeled as pure translation. Above assumptions are valid in staring imaging and some videos during a certain period of time. The matrices \mathbf{D} , \mathbf{B} , and \mathbf{M} are quite huge with few non-zero elements, thus can be constructed as sparse matrices. A faster and more memory-efficient way is to interpret them as image operators without explicitly constructing the sparse matrices [19, 20]. Setting $\mathbf{W}_t = \mathbf{D} \mathbf{B} \mathbf{M}_t$, Eq. (4) can be rewritten as

$$\hat{\mathbf{x}}_t = \min_{\mathbf{x} \geq 0} J_t(\mathbf{x}; \mathbf{y}_t) = \min_{\mathbf{x} \geq 0} \|\mathbf{y}_t - \mathbf{W}_t \mathbf{x}\|^2 \quad (5)$$

Blurring may be caused by optical aberrations, out of focus, diffraction, finite detector size, etc. The blur kernel in this paper is assumed to be of Gaussian shape, and its standard deviation is estimated by a blind deconvolution algorithm based on regularized ML [21] in advance.

Essentially, the loss function is in the form of maximum likelihood (ML). We just decompose the overall ML lost function to many parts to save memory and computation resources. The drawback is the reconstructed image may be not the optimal because the optimization of the loss function for one LR image should be terminated earlier to avoid overfitting.

3.2 Update strategy

Equation (5) can be easily solved by some gradient-based optimization methods. But if it is not handled properly, it may result in overfitting \mathbf{y}_t . To solve this problem, we can introduce an auxiliary function similar to OBDSR [18]:

$$L_t(\mathbf{x}, \tilde{\mathbf{x}}) = \mathbf{y}_t^T \mathbf{y}_t - 2\mathbf{y}_t^T \mathbf{W}_t \mathbf{x} + \tilde{\mathbf{x}}^T \mathbf{W}_t^T \mathbf{W}_t \left(\frac{\mathbf{x} \odot \mathbf{x}}{\tilde{\mathbf{x}}} \right) \quad (6)$$

where \odot denotes elementwise product and division is elementwise as well.

The HR image estimate \mathbf{x}_t at time t can be obtained by solving $\nabla_{\mathbf{x}} L_t(\mathbf{x}, \mathbf{x}_{t-1})|_{\mathbf{x}=\mathbf{x}_t} = 0$, which yields a simple multiplicative update,

$$\mathbf{x}_t = \mathbf{x}_{t-1} \odot \frac{\mathbf{W}_t^T \mathbf{y}_t}{\mathbf{W}_t^T \mathbf{W}_t \mathbf{x}_{t-1}} \quad (7)$$

The update equation naturally fulfills non-negative constraint on \mathbf{x} if a non-negative initial image estimate is used. The usage of auxiliary function decreases the convergence rate of \mathbf{x} , thus reducing the probability of overfitting one specific LR image.

However, experiments show that the strategy of ‘‘auxiliary function’’ may fail, especially when the LR images have a low signal to noise ratio (SNR). An optional solution to Eq. (5) is using the steepest descent method. We impose the non-negative constraint by reparameterizing \mathbf{x}_t as

$$\mathbf{x}_t = \boldsymbol{\phi}_t^2 \quad (8)$$

The gradient of the loss function with respect to $\boldsymbol{\phi}_t$ is given by

$$\nabla_{\boldsymbol{\phi}} J_t(\mathbf{x}; \mathbf{y}_t)|_{\boldsymbol{\phi}=\boldsymbol{\phi}_t} = -4\boldsymbol{\phi}_t \odot \left(\mathbf{W}_t^T \mathbf{y}_t - \mathbf{W}_t^T \mathbf{W}_t \boldsymbol{\phi}_t^2 \right) \quad (9)$$

The gradient descent update for $\boldsymbol{\phi}$ can be expressed as

$$\boldsymbol{\phi}_{t,k} = \boldsymbol{\phi}_{t,k-1} - \alpha \nabla_{\boldsymbol{\phi}} J_t(\mathbf{x}; \mathbf{y}_t)|_{\boldsymbol{\phi}=\boldsymbol{\phi}_{t,k-1}} \quad (10)$$

for $k = 1, 2, \dots, K$, and $\boldsymbol{\phi}_{t,0} = \boldsymbol{\phi}_{t-1,K}$. The parameter α represents the step size, which should be small enough to

prevent divergence and large enough to provide reasonable convergence rate. The maximum iteration index K should have an appropriate value for the same reason. In our paper, K is set to 3 empirically.

The update strategy given by Eq. (10) can yield an approving SR result if the step size is selected carefully. However, two main drawbacks are obvious as well. One drawback is that the step size is constant for each iteration; thus, it may be too large or too small for some iterations. The other drawback is the step size is selected manually, which is not friendly to non-professional users. To avoid these drawbacks, we propose an adaptive step size strategy. According to the steepest gradient optimization algorithm, the optimal step size can be calculated by minimizing

$$J_t(\mathbf{x}_t; \mathbf{y}_t) = J_t \left(\mathbf{x}_{t-1} - \alpha \nabla_{\mathbf{x}} J_t(\mathbf{x}; \mathbf{y}_t)|_{\mathbf{x}=\mathbf{x}_{t-1}}; \mathbf{y}_t \right) \quad (11)$$

with respect to α . The gradient of the loss function with respect to \mathbf{x} is given by

$$\nabla_{\mathbf{x}} J_t(\mathbf{x}; \mathbf{y}_t) = -2\mathbf{W}_t^T \mathbf{y}_t + 2\mathbf{W}_t^T \mathbf{W}_t \mathbf{x} \quad (12)$$

By solving $\nabla_{\alpha} J_t(\mathbf{x}_t; \mathbf{y}_t) = 0$, we get

$$\alpha = - \frac{(\mathbf{y}_t - \mathbf{W}_t \mathbf{x}_{t-1})^T \mathbf{W}_t \mathbf{g}_t}{(\mathbf{W}_t \mathbf{g}_t)^T \mathbf{W}_t \mathbf{g}_t} \quad (13)$$

$$\mathbf{g}_t = \nabla_{\mathbf{x}} J_t(\mathbf{x}; \mathbf{y}_t)|_{\mathbf{x}=\mathbf{x}_{t-1}} \quad (14)$$

We find that if α is used as the update step size directly, the solution may overfit LR images. Thus, we employ a parameter to control the convergence rate. When the SNR of the LR images is low, the parameter should decrease the convergence rate to avoid noise amplification. On the other hand, when the SNR of the LR images is high, the parameter should be selected to increase the convergence rate. The parameter is empirically given by

$$\beta = -0.006\sigma_n^2 + 0.5 \quad (15)$$

where σ_n^2 is the noise variance, which can be estimated by a noise level estimation algorithm proposed by Liu et al. [22] in advance. The modified update expression of the HR image converts to

$$\mathbf{x}_t = \mathbf{x}_{t-1} - \beta \alpha \mathbf{g}_t \quad (16)$$

It can be seen that the noise-adaptive step size update strategy is fully automatic. However, the update expression cannot guarantee the non-negativity of the HR image; thus, the pixels whose intensity is less than 0 are set to zero after each time step.

3.3 Color SR

Color SR can be achieved by simply incorporating each channel into the loss function. The modified loss function is given by



Fig. 1 Reconstructed HR images for different noise levels. Top row: one bicubic-interpolated LR image with up-sampling factor equal to 3 for each noise level, the noise variances are 0, 10, 20, 30, 40, and 50 respectively from left to right. Bottom row: corresponding final HR images reconstructed by OSR with up-sampling factor equal to 3 for each noise level

$$J_t(\mathbf{x}; \mathbf{y}_t) = \sum_{l=R,G,B} \|\mathbf{y}_{t,l} - \mathbf{DB}_{t,l}\mathbf{M}\mathbf{x}_t\|^2 \quad (17)$$

The blur kernel may be different for each color channel, so it may need to be estimated separately.

3.4 Algorithm

A summary of the OSR algorithm using adaptive step size is provided in Algorithm 1. Other update strategies can also be used within a similar framework. The motion vector is estimated by a very fast subpixel image registration technique based on cross-correlation which is proposed by Feng et al. [23]. The registration result is verified by a measurement validation method based on Mahalanobis distance [10]; thus, the LR images with large registration error are eliminated. At each time step t , we align the current LR image with the initial LR image rather than the current HR image estimate because the first several HR image estimates may contain large amount of artifacts. Experiments show that multiple iterations on \mathbf{x}_t at each time step do not yield much improvement, so \mathbf{x}_t is updated only once when adaptive step size update is adopted.

Our online SR algorithm has two main advantages: (1) The algorithm is very fast and memory efficient. During the process of HR image reconstruction, only the initial LR image, the current LR image, and the current estimated HR image need to be stored, which need much less memory and computing resource compared to batch-mode algorithms. (2) The algorithm is very simple and friendly to users. In our algorithm, we only use one parameter to control the fitting degree of each LR image, and the parameter is noise-adaptive; thus, our algorithm is easy to handle.

4 Experimental results and discussion

We first test our OSR method with simulated data to study the performance under different conditions. Then,

Algorithm 1: Proposed OSR algorithm.

Input: LR images $\{\mathbf{y}_t\}$, blur kernel h for each color channel, SR scale L , noise variance σ_n^2 .

Output: HR image.

$t \leftarrow 1$;

$\mathbf{y}_0 \leftarrow \mathbf{y}_1$;

$\mathbf{x}_0 \leftarrow$ up-sampled \mathbf{y}_1 with scale L ;

$\beta \leftarrow -0.006\sigma_n^2 + 0.5$;

while the LR image \mathbf{y}_{t+1} available **do**

$t \leftarrow t + 1$;

$\mathbf{x}_t \leftarrow \mathbf{x}_{t-1}$;

estimate motion vector $[d_x, d_y]$ between \mathbf{y}_0 and \mathbf{y}_t ;

if registration invalid **then**

continue;

end

$\mathbf{g}_t \leftarrow \nabla_{\mathbf{x}} J_t(\mathbf{x}; \mathbf{y}_t)|_{\mathbf{x}=\mathbf{x}_t}$;

update α using Eq. (13);

$\mathbf{x}_t \leftarrow \mathbf{x}_t - \beta\alpha\mathbf{g}_t$;

$\mathbf{x}_t \leftarrow \max(\mathbf{x}_t, 0)$

end

we apply our method to real-image sequences generated by different cameras. Our algorithm is implemented using MATLAB R2014b, and all the experiments are carried out using an Intel Core i7-4790 CPU PC with 16GB RAM.

4.1 Simulated data tests

The first experiment is to test the noise robustness of our algorithm. The simulated LR images are generated following the observation model given by Eq. (1). The subpixel translation along horizontal and vertical direction is uniformly distributed within $(-2.5, 2.5)$ in high-resolution grid, and the blur kernel is of Gaussian shape with the standard deviation equal to 1. To test

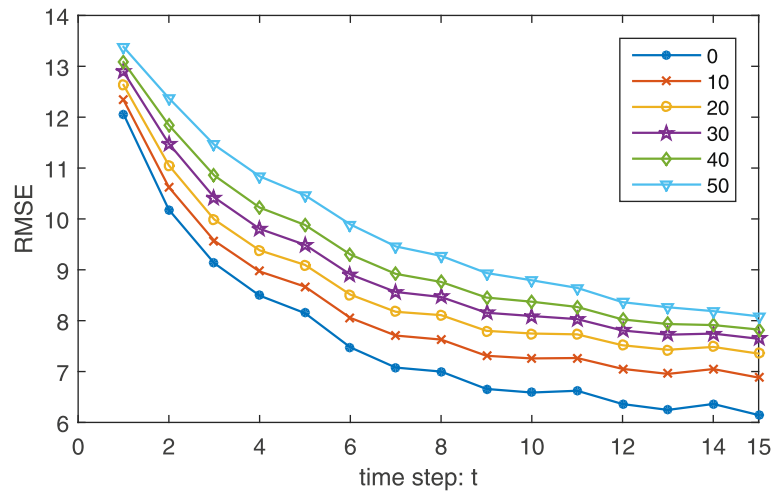


Fig. 2 RMSE of HR images reconstructed by OSR for different noise levels. The data along with the legend are noise variances

the robustness of our method under different noise levels, we generate 15 LR images with varying noise variances $\sigma_n^2 \in \{0, 10, 20, 30, 40, 50\}$ respectively. The HR image is a standard test image “goldhill” [24] of size 576×720 in gray mode, and the size of LR images is 153×192 after clipping the boundary with a down-sampling scale equal to 3. The estimated standard deviation of the Gaussian blur kernel is 1.07, and the estimated noise variances are 0.8, 10.2, 20.3, 32.4, 42.4, and 51.7 for each noise level respectively. The top row of Fig. 1 shows one of the bicubic-interpolated LR images with increasing σ_n^2 from left to right, and the bottom row

shows the final reconstructed HR image for each noise level. Note that we just display a common part of each image to see the detail more clearly. For each noise level, the reconstructed HR image is more clear than the bicubic-interpolated image, and the SNR is higher as well. Figure 2 shows how root mean square error (RMSE) evolves as the time step t for each noise level. From a general view, the RMSE reduces as t increases, which means the HR images are improved as the algorithm progresses. Experimental results also show that the algorithm’s performance is better with lower noise level.

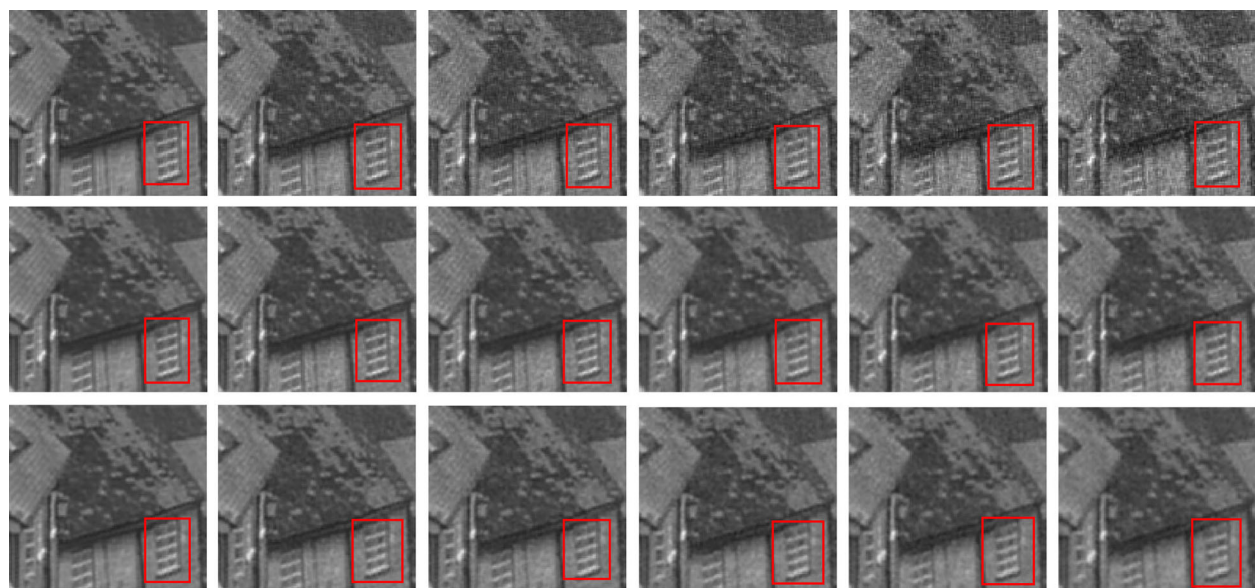


Fig. 3 The final reconstructed HR images using different update strategies. Top row: using multiplicative update strategy. Middle row: using manual step size update strategy. Bottom row: adaptive step size update strategy. The noise variances of corresponding LR images are 0, 10, 20, 30, 40, and 50 respectively from left to right

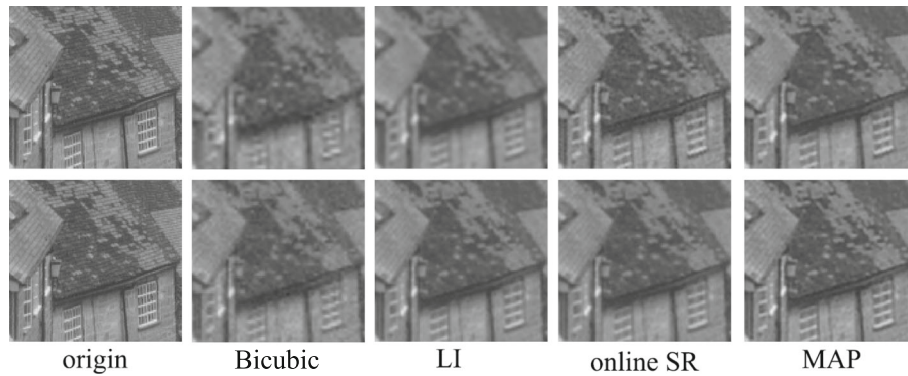


Fig. 4 Comparison of results of different reconstruction methods. From left to right: origin original HR image, reconstructed HR image using bicubic interpolation, LI, OSR, and MAP. Top row: images reconstructed with a up-sampling factor of 3 (except for the origin original image). Bottom row: images reconstructed with a up-sampling factor of 2 (except for the origin original image)

The second experiment conducts a comparison of the three update strategies mentioned in Section 3. For convenience, the update strategies using Eqs. (7), (10), and (16) are named as multiplicative update, manual step size update, and adaptive step size update, respectively. Figure 3 shows the final HR images reconstructed by different update strategies. The step size for manual step size strategy is set to {0.004, 0.004, 0.003, 0.002, 0.002, 0.002} as the noise variance increases from 0 to 50. The step size is selected via multiple trials, thus is nearly optimal. Compared to adaptive step size update, the HR images reconstructed by manual step size update have similar visual effects, while the HR images reconstructed by multiplicative update have larger noise which we believe is caused by overfitting LR images.

The third experiment studies the performance of our algorithm compared with lucky imaging (LI) and MAP-based SR algorithm [25, 26]. The lucky imaging algorithm first aligns and adds all up-sampled LR images with registration validation, then deconvolves the resulting image using Wiener filter. We generate 15 LR images with the same noise variance equal to 20, and other experimental settings are the same with the first experiment. Experimental results are shown in the top row of Fig. 4. We generate another 15 LR images with a down-sampling factor of 2 and use them to reconstruct an HR image with an up-sampling factor of 2. The results are shown in the bottom row of Fig. 4. Note that Fig. 4 just shows

a common selected region of all images in order to be seen more clearly. For these data sets, we have visual quality $MAP > OSR > LI > Bicubic$. MAP-based method gets better HR image than OSR method because it tries to find a global optimal solution with all LR images processed at one time. Table 1 shows the performance of different methods evaluated in RMSE and elapsed time. It takes our method much less time to reconstruct an HR image compared with MAP-based SR. Additionally, the elapsed time of our method is related to the size of the HR image, rather than the LR image, for most of image operators act on the HR image, or up-sampled LR images. Our method can process at a speed over 11 fps to reconstruct a HR image of size 459×576 in gray mode. Note that our MATLAB-implemented algorithm has not been optimized for speed. The computation speed can be further improved by using other more efficient programming languages or GPU-based computing technique, while it is beyond the scope of this article.

Experiments with another 3 standard test images (Lenna, Monarch, and Boats) are also performed. Fifteen LR images are used to reconstruct an HR image respectively, and the SR scales are set to 3. The reconstructed results are shown in Fig. 5, and the RMSE and elapsed time are shown in Table 2. Also, the MAP-based SR yields better results in image quality than OSR, while much less computation time is needed by OSR.

Table 1 Comparison of performance of different SR methods in RMSE and elapsed time

| <i>L</i> | LR image size | RMSE | | | | Average elapsed time (ms/frame) | | | |
|----------|---------------|---------|-------|------|------|---------------------------------|----|-----|-----|
| | | Bicubic | LI | OSR | MAP | Bicubic | LI | OSR | MAP |
| 3 | 153 × 192 | 12.60 | 12.37 | 7.27 | 6.58 | – | 58 | 88 | 688 |
| 2 | 230 × 288 | 9.49 | 8.96 | 6.30 | 5.88 | – | 57 | 88 | 979 |

The RMSE of bicubic-interpolated images is the average of 15 bicubic-interpolated images. *L* is the up-sampling factor



Fig. 5 Comparison of results of different reconstruction methods with different test images. From left to right: origin original HR image, reconstructed HR image using bicubic interpolation, MAP, and OSR; From top to bottom: Lenna, Monarch, and Boats

To study the influence of the initial values and the order of the LR frames, we randomly permute the order of obtaining the input frames. The LR images and other experiment parameters are the same with the first experiment. Figure 6 shows the RMSE curves evolving with the time step for a fixed noise variance of 20. As can be seen, all RMSE curves have a same evolving trend and converge to a similar value finally.

4.2 Real data tests

The first real data set is a color video with a size of 91×121 [27]. These frames approximately follow the global

translation motion model. The estimated standard deviation of Gaussian blur kernel is 1.6 for each color channel, and the estimated noise variance is 2.3 which is the average of three color channels. It takes 3.0 s to reconstruct the HR image from 20 LR frames with an up-sampling factor equal to 4, and image registration occupies 1.4 s. Figure 7 shows one bicubic-interpolated LR image and the final reconstructed HR image. The improved definition and sharpness demonstrate the good performance of our method.

The second real data set is a color video captured by a digital single lens reflex (DSLR). We extract a region of interest (ROI) of size 134×182 . These LR frames approximately follow the global translation motion model. The estimated standard deviation of Gaussian blur kernel is 0.8 for each color channel, and the estimated noise variance is 1.3 which is the average of three color channels. It takes 0.65 s to reconstruct the HR image from 10 LR frames with an up-sampling factor equal to 2, and image registration occupies 0.3 s. Figure 8 shows one bicubic-interpolated LR image, the final reconstructed HR image and the ground-truth image. It can be seen that the aliasing effect is decreased by our OSR method.

The third real data set is ten 5-band Gaofen-4 remote sensing images without ortho-rectification [28]. The images were captured on March 3, 2017, from 11:10:20 to 11:20:21. We chip crop the ROI with a size of 44×78 . The estimated standard deviation of Gaussian blur kernel is 1.2, and the estimated noise variance is 1.1. It takes 59.7 ms to reconstruct the HR image, and image registration occupies 26.9 ms. To validate our reconstructed

Table 2 Comparison of performance of different SR methods in RMSE and elapsed time with different test images

| Image name | Resolution of LR images | Methods | RMSE | Elapsed time (ms/frame) |
|------------|-------------------------|---------|------|-------------------------|
| Lenna | 136×136 | Bicubic | 9.2 | – |
| | | MAP | 4.7 | 454 |
| | | OSR | 6.1 | 58.9 |
| Monarch | 204×136 | Bicubic | 12.2 | – |
| | | MAP | 5.3 | 750 |
| | | OSR | 7.6 | 88.9 |
| Boats | 192×153 | Bicubic | 9.3 | – |
| | | MAP | 4.4 | 739 |
| | | OSR | 6.2 | 98.9 |

The RMSE of bicubic-interpolated images is the average of 15 bicubic-interpolated images. The SR scales are all set to 3

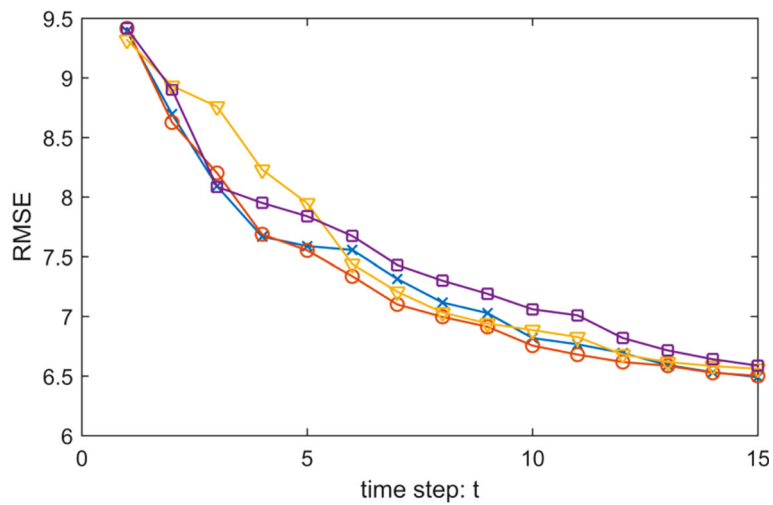


Fig. 6 The RMSE evolving with the time steps in random orders. The noise variances of all LR images are 20. Each line corresponds to a different order of the LR images



Fig. 7 Super-resolution of bookcase images. **a** One bicubic-interpolated LR image with an up-sampling scale equal to 4. **b** Reconstructed HR image by OSR with an up-sampling scale equal to 4

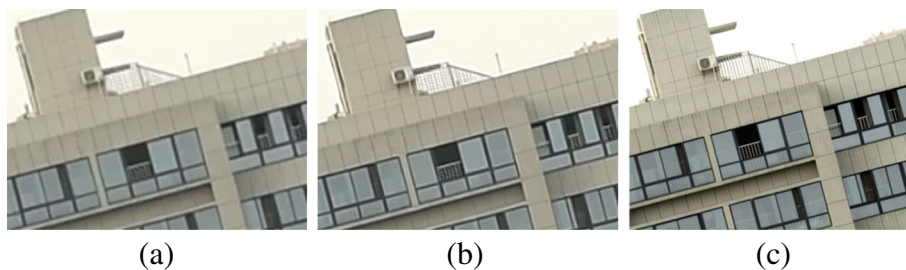


Fig. 8 Super-resolution of building images. **a** One bicubic-interpolated LR image with an up-sampling factor equal to 2. **b** Reconstructed HR image by OSR with an up-sampling factor equal to 2. **c** Ground-truth image

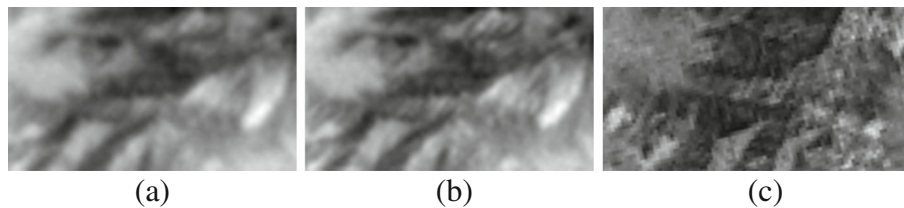


Fig. 9 Super-resolution of Gaofen-4 images. **a** One bicubic-interpolated LR image with an up-sampling factor equal to 2. **b** Reconstructed HR image by OSR with an up-sampling factor equal to 2. **c** Image chipped cropped from Google Earth

HR image, we chip crop an image of the same zone from Google Earth, which was captured on September 16, 2015. Figure 9 shows one bicubic-interpolated LR image, the final reconstructed HR image by OSR and MAP, and the image chipped cropped from Google Earth. It can be seen that the reconstructed HR image is clearer and contains more details which corresponds to Google Earth.

5 Conclusions and future work

In this paper, we propose a multi-frame super-resolution algorithm which operates in online mode. The algorithm is simple and memory efficient and needs much less computing resource compared to batch-mode methods. Additionally, we employ a noise-adaptive parameter in classical steepest gradient optimization algorithm to avoid noise amplification and the overfitting of LR images. Our method is also compatible with color images. Experimental results on simulated and real-image sequences show that our online SR method has a good performance in restoring the details and missing information in LR images and has a real-time application prospect.

Image super-resolution naturally requires large computing resources. A good choice is to just process the region of interest which can also simplify the motion model. The work to incorporating a tracking system and more complex motion model into the online SR framework is ongoing.

Abbreviations

DSLR: Digital single lens reflex; GPU: Graphic processing unit; HR: High-resolution; LI: Lucky imaging; LR: Low resolution; MAP: Maximum a posteriori; ML: Maximum likelihood; OBD: Online blind deconvolution; OBDSR: Online blind deconvolution with super-resolution; OSR: Online super-resolution; RMSE: Root mean square error; ROI: Region of interest; SNR: Signal to noise ratio; SR: Super-resolution

Acknowledgements

The authors would like to thank the anonymous reviewers for their helpful comments.

Funding

This research is funded by the National Science Fund for Outstanding Young Scholars (no. 2017-JCJQ-ZQ-005).

Availability of data and materials

The data used and/or analyzed during the current study are available from the reference websites.

Authors' contributions

JX is the first author and corresponding author of this paper. His main contributions include the basic idea, computer simulations, and writing of this paper. The main contributions of YL and JL include analyzing the basic idea and checking simulations. The main contribution of ZH is experimental design. XL captured the real data and refined the paper. All authors have read and approved the final manuscript.

Authors' information

Jieping Xu is a lecturer at Space Engineering University. He received his bachelor's degree, master's degree, and doctor's degree in Optical Engineering from National University of Defense Technology in 2010, 2012, and 2017, respectively. His current research interests include image reconstruction and object recognition.

Yonghui Liang received the Ph.D. in Optical Engineering from National University of Defense Technology in 2000. She is currently a professor of College of Opto-electronic Science and Engineering, National University of Defense Technology. She worked as a visiting academic at Durham University in 2009. Her current research interests include adaptive optics and digital image processing.

Jin Liu received the Ph.D. in Statistics from Tsinghua University in 2013. He is currently a lecturer of College of Opto-electronic Science and Engineering, National University of Defense Technology. His current research interests include image reconstruction and image enhancement.

Zongfu Huang received the Ph.D. in Electronic Science and Technology from National University of Defense Technology in 2012. He is currently a lecturer of College of Opto-electronic Science and Engineering, National University of Defense Technology. His current research interests include image reconstruction and opto-electronic target detection.

Xuewen Liu is a PhD student at National University of Defense Technology. She received his bachelor's degree in Optical Engineering from National University of Defense Technology in 2014. Her current research interests include adaptive optics and image reconstruction.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Space Engineering University, No. 1, Bayi Road, Beijing, 101416 China.

²College of Opto-electronic Science and Engineering, National University of Defense Technology, No. 109, Deya Road, Changsha, 410073 China.

Received: 24 May 2017 Accepted: 8 November 2018

Published online: 04 December 2018

References

1. S. C. Park, M. K. Park, M. G. Kang, Super-resolution image reconstruction: a technical overview. *IEEE Signal Process. Mag.* **20**(3), 21–36 (2003)
2. M. S. Alam, J. G. Bogner, R. C. Hardie, B. J. Yasuda, in *Proc. SPIE*, vol. 3063. High-resolution infrared image reconstruction using multiple randomly shifted low-resolution aliased frames, (1997), pp. 102–122
3. M. S. Alam, J. G. Bogner, R. C. Hardie, B. J. Yasuda, Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames. *IEEE Trans. Instrum. Meas.* **49**(5), 915–923 (2000)
4. R. Tsai, T. S. Huang, Multiframe image restoration and registration. *Adv. Comput. Vis. Image Proc.* **1**(2), 317–339 (1984)
5. J. Ouwerkerk, Image super-resolution survey. *Image Vis. Comput.* **24**(10), 1039–1052 (2006)
6. L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, L. Zhang, Image super-resolution: the techniques, applications, and future. *Signal Process.* **128**, 389–408 (2016)
7. Z. Jiang, T. Wong, H. Bao, in *Computer Vision and Pattern Recognition*, vol. 2. Practical super-resolution from dynamic video sequences, (2003)
8. R. Prendergast, T. Nguyen, in *15th IEEE International Conference on Image Processing, 2008*. A block-based super-resolution for video sequences, (2008), pp. 1240–1243
9. S. Farsiu, D. Robinson, M. Elad, P. Milanfar, Dynamic demosaicing and color super-resolution of video sequences. *Proc. SPIE.* **5562**(1), 169–178 (2004)
10. M. Kim, B. Ku, D. Chung, H. Shin, B. Kang, Han, D.K., H. Ko, in *Advanced Video and Signal Based Surveillance*. Robust dynamic super resolution under inaccurate motion estimation, (2010), pp. 323–328
11. A. Kappeler, S. Yoo, Q. Dai, A. K. Katsaggelos, Video super-resolution with convolutional neural networks. *IEEE Trans. Comput. Imaging.* **2**(2), 109–122 (2016)
12. Y. Huang, W. Wang, L. Wang, Video super-resolution via bidirectional recurrent convolutional networks. *IEEE Trans. Pattern. Anal. Mach. Intell.* **40**(4), 1015–1028 (2018)
13. J. Hu, H. Li, Y. Li, in *IEEE International Conference on Orange Technologies*. Real time super resolution reconstruction for video stream based on GPU, (2014), pp. 9–12
14. W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, Z. Wang, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, (2016), pp. 1874–1883
15. D. Liu, Z. Wang, Y. Fan, X. Liu, Z. Wang, S. Chang, T. Huang, in *2017 IEEE International Conference on Computer Vision (ICCV)*. Robust video super-resolution with learned temporal dynamics, (2017), pp. 2526–2534
16. D. Y. Lee, J. Lee, J.-H. Choi, J.-O. Kim, H. Y. Kim, J. S. Choi, GPU-based real-time super-resolution system for high-quality UHD video up-conversion. *J. Supercomput.* **74**(1), 456–484 (2018)
17. S. Harmeling, M. Hirsch, S. Sra, B. Schölkopf, in *Proceedings of the IEEE Conference on Computational Photography*. Online blind deconvolution for astronomical imaging, (2009)
18. M. Hirsch, S. Harmeling, S. Sra, B. Schölkopf, Online multi-frame blind deconvolution with super-resolution and saturation correction. *Astron. Astrophys.* **531**, A9 (2011)
19. A. Zomet, S. Peleg, in *15th International Conference on Pattern Recognition*, vol. 1. Efficient super-resolution and applications to mosaics, (2000), pp. 579–583
20. S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar, Fast and robust multiframe super resolution. *IEEE Trans. Image Process.* **13**(10), 1327–1344 (2004)
21. C. L. Matson, K. Borelli, S. Jefferies, C. C. Beckner, Fast and optimal multiframe blind deconvolution algorithm for high-resolution ground-based imaging of space objects. *Appl. Opt.* **48**(1), A75–A92 (2009)
22. C. Liu, W. T. Freeman, R. Szeliski, S. B. Kang, in *Computer Vision and Pattern Recognition*, vol. 1. Noise estimation from a single image, (2006), pp. 901–908
23. S. Feng, L. Deng, G. Shu, F. Wang, H. Deng, K. Ji, in *2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI)*. A subpixel registration algorithm for low PSNR images, (2012), pp. 626–630
24. Set of Classic Test Still Images. <http://www.hlevkin.com/TestImages/classic.htm>. Accessed 5 May 2017
25. M. Petrou, M. H. Jaward, S. Chen, M. Briers, Super-resolution in practice: the complete pipeline from image capture to super-resolved subimage creation using a novel frame selection method. *Mach. Vis. Appl.* **23**, 441–459 (2012)
26. J. Xu, Y. Liang, J. Liu, Z. Huang, Multi-frame super-resolution of Gaofen-4 remote sensing images. *Sensors.* **17**(9), 2142 (2017)
27. MDSP super-resolution and demosaicing datasets :: Peyman Milanfar. <https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html>. Accessed 5 May 2017
28. Rscloudmart. <http://www.rscloudmart.com/dataProduct/datacenterStandardData>. Accessed 1 Feb 2018

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
