Fast MAP-Based Super Resolution Algorithm Using Multilevel Interpolation and Reconstruction

Jiunn-Lin Wu
Institute of Networking and Multimedia
Dept. of Computer Science and Engineering
National Chung Hsing University
Email: jlwu@cs.nchu.edu.tw

Meng-Han Hung
Dept. of Computer Science and Engineering
National Chung Hsing University

***Abstract*—Super-resolution (SR) image reconstruction can produce a higher resolution digital image from multiple low-resolution (LR) photographs. Many applications require high resolution images, including medical imaging, satellite imaging, and video applications. Recovering lost details from down-sampled images is the main challenge of super resolution techniques. However, most researches do not take computational complexity into consideration. Therefore, this paper presents a new fast maximum a posteriori (MAP)-based SR image reconstruction method, it is based on the multilevel algorithm. In particular, this work focuses on the case of the input LR images that are not enough for analysis. A two-step interpolation process is proposed to increase the quality of the constructed SR image. Experimental results show that the proposed algorithm can simultaneously reduce blocking artifacts caused by a lack of LR images and computational complexity.**

***Index Terms*—Super-resolution; Maximum a Posteriori (MAP) method; multilevel algorithm; computational complexity; multiframe image reconstruction.**

I. Introduction

Super-resolution (SR) image reconstruction can reconstruct a high-resolution (HR) image from multiple low-resolution (LR) photographs by exploiting the correlations between those images. When there are sub-pixel shifts between LR images, it is also possible to increase spatial resolution. Typical SR techniques model the process of acquiring LR images from an unknown HR image, and usually consist of two steps: (1) image registration, and (2) combining registered LR images to produce an improved resolution image. The first step is also called motion estimation, and many different registration approaches with fractional pixel accuracy are available. Generally, accurate sub-pixel motion estimation affects the success of SR techniques. The second step can be performed in different algorithms, such as the frequency domain method, the projection onto convex sets (POCS) method, and the maximum a posteriori (MAP) method. Many studies discuss the advantages and disadvantages of each algorithm [1], [5].

Many applications require high resolution images, including medical imaging, satellite imaging, and video applications. How to recover lost details from down-sampled images is the primary challenge in this area, and researchers have proposed many different algorithms to conquer this task. The most well known algorithm is the Bayesian maximum a posteriori (MAP) method. This approach has a variety of outstanding advantages, including robustness and flexibility in modeling noise characteristics in the spatial domain and a priori knowledge about the solution. However, it has a large computational complexity.

To decrease the complexity, conjugate gradient (CG) method [2], [3] and preconditioned conjugate gradient (PCG) method [4], [6] were proposed to fasten the optimization procedure and minimize the cost function used in MAP-based SR image reconstruction. Later, Di Zhang et al. [7] proposed the recursive multilevel reconstruction algorithm, which dramatically reduces computational complexity.

This paper focuses on Zhang's method, showing that when not enough LR images are available (for example, 16 LR images are required to recover a HR image magnified by four), blocking artifacts appear in the reconstructed HR image. This is a very practical problem since capturing adequate and usable LR images can be a difficult task. Some unexpected factors, such as moving objects, illumination changes, or atmosphere perturbation, can influence the available information captured in a single LR image. These destructive LR images must be eliminated from the total number of LR images, leading to an inadequacy of LR images.

This paper proposes a new algorithm that improves the performance of the recursive multilevel reconstruction algorithm [7] as only a few LR images are available. Experimental results show that the proposed method can simultaneously reduce blocking artifacts and computational complexity. Unlike Zhang's algorithm, which assumes LR images are uniformly distributed, the proposed algorithm can also deal with LR images with arbitrary shifts [7].

II. MAP-BASED SUPER-RESOLUTION

A. Formulation

The SR image reconstruction problem can be analyzed by formulating the imaging process in a common imaging system as a linear mapping between a HR input signal x and LR images y_i, $i = 1, \ldots, N$, where N is the total number of LR images. Each LR image can be considered as a noisy and down-sampled version of the HR image that has been shifted and blurred. The imaging process is formulated as

$$y_i = DB_i M_i x + n_i, \quad i = 1, \ldots, N \tag{1}$$

where x is the desired HR image measuring $qL \times qW$ and q represents the down-sampling factor in the horizontal and vertical directions. Thus, each observed LR image y_i measures $L \times W$. M_i is a warp matrix representing the geometric shifts between the i th LR image and the HR image, B_i is a blur matrix simulating the point spread function of the sensor elements, D is a subsampling matrix, and n_i is the zero-mean independent identically distributed (IID) Gaussian noise with a probability density function (PDF) given by

$$P(n_i) = \frac{1}{C} \exp\left(-\frac{n_i^2}{2\sigma^2} \right). \tag{2}$$

The MAP estimator of x maximizes the a posteriori PDE $P(x \mid y_i)$ with respect to x and yields:

$$x = \arg\max P(x \mid y_1, y_2, \ldots, y_N). \tag{3}$$

Applying Bayes' theorem to the conditional probability and taking the logarithmic function results in the MAP optimization problem as shown in (4):

$$x = \arg\max \{ \ln P(y_1, y_2, \ldots, y_N \mid x) + \ln P(x) \}. \tag{4}$$

The conditional PDF $P(y_1, y_2, \ldots, y_N \mid x)$ is defined by the statistical distribution of the Gaussian noise n_i as following

$$\begin{aligned} &P(y_1, y_2, \ldots, y_N \mid x) \\ &= \frac{1}{C} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{N} \left\| y_i - DB_i M_i x \right\|^2 \right) \end{aligned} \tag{5}$$

Further, the prior PDF $P(x)$ is defined by a priori knowledge concerning the HR image x. Since images are frequently assumed to be globally smooth, we simply use a Gaussian model

$$P(x) = \frac{1}{Z}\exp\left(-\frac{1}{2\lambda^2}\|x-\mu\|^2\right) \tag{6}$$

where μ represents the mean of image x obtained by interpolating and averaging LR images. By substituting (5) and (6) into (4), the MAP optimization problem can be rewritten as

$$\hat{x} = \arg\min\left\{\frac{1}{2\sigma^2}\sum_{i=1}^{N}\|y_i - DB_iM_ix\|^2 + \frac{1}{2\lambda^2}\|x-\mu\|^2\right\}. \tag{7}$$

From (7), the MAP estimate of x minimizes the cost function:

$$\begin{aligned} E(x) = \frac{1}{2}\sum_{i=1}^{N}(y_i - DB_iM_i)^T W^{-1}(y_i - DB_iM_ix) \\ + \frac{1}{2}(x-\mu)^T C^{-1}(x-\mu) \end{aligned} \tag{8}$$

where W is a diagonal matrix, and its diagonal is equal to σ^2, C serves as a constant diagonal matrix.

Gradient descent techniques can be employed to solve the optimization problem. The image x can be estimated by iteratively updating an initial estimate in the direction of the negative gradient of $E(x)$. At the k th iteration, the estimate is

$$\hat{x}^{(k)} = \hat{x}^{(k-1)} - \alpha\nabla E(\hat{x}^{(k-1)}) \tag{9}$$

where α is the step size and $\nabla E(x)$ can be found as

$$\begin{aligned} \nabla E(x) = -\sum_{i=1}^{N}(DB_iM_i)^T W^{-1}(y_i - DB_iM_ix) \\ + C^{-1}(x-\mu). \end{aligned} \tag{10}$$

In our experiments, α is updated in each iteration using the formula

$$\alpha = \frac{\left(\nabla E\left(x^{(k-1)}\right)\right)^T\left(\nabla E\left(x^{(k-1)}\right)\right)}{\left(\nabla E\left(x^{(k-1)}\right)\right)^T H\left(\nabla E\left(x^{(k-1)}\right)\right)} \tag{11}$$

where H is the Hessian matrix which is defined by

$$H = \sum_{i=1}^{N}(DB_iM_i)^T W^{-1}(DB_iM_i) + C^{-1} \tag{12}$$

B. Recursive multilevel reconstruction algorithm

The MAP-based SR algorithm is a time-consuming task because it exploits every LR images in each iteration. The number of operations is generally based on the LR image size, the magnification factor, and the total number of LR images. Zhang [7] proposed a multilevel algorithm that reduces the number of operations by simply splitting one big task into multiple small tasks. Each small task involves only few LR images and attempts to reconstruct a HR image which is twice the size of LR images. Fig. 1 presents a flowchart of the multilevel SR algorithm.

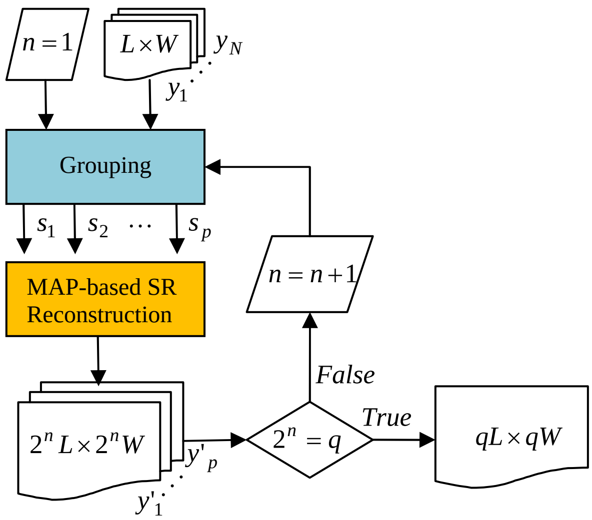

Fig. 1. The flowchart of the multilevel super-resolution image reconstruction algorithm.

Given N LR images y_i of the same $L \times W$ size, the proposed SR algorithm attempts to reconstruct a HR image of size $qL \times qW$. Ideally, N equals q^2 and q is a power of two ($q = 2^n$, and n is a positive integer which represents the total number of levels).

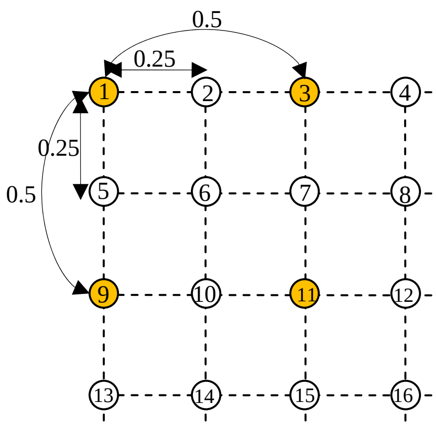

Fig. 2. An example of the grouping process in the proposed method.

The recursive multilevel algorithm can be stated as following:

1. Estimate spatially transformations between LR images and the reference LR image by image registration.
2. On the first level ($n = 1$), N LR images y_i are set as input images.
3. Classify input images into multiple subsets $\{s_1, s_2, \ldots s_p\}$ based on a grouping principle. p is the total number of subsets. In each subset, if the translation of a LR image is (k, l), the translations of other LR images should be $(k + 0.5, l)$, $(k, l + 0.5)$ or $(k + 0.5, l + 0.5)$. Fig. 2 shows that each circle represents an input image, and 16 input images are transferred onto the reference grid. Based on the grouping principle, input images $\{1, 3, 9, 11\}$ belong to the same subset.
4. Apply the MAP-based SR reconstruction to p subsets and obtain HR images $\{y'_1, y'_2, \ldots y'_p\}$ of size $2^n L \times 2^n W$.
5. If $2^n = q$, then stop; otherwise set $n = n + 1$, and set HR images $\{y'_1, y'_2, \ldots y'_p\}$ as input images. Return to Step 3.

In the ideal case, if the translation of a LR image is (k, l), the principle of grouping searches for the other three LR images with the translation being $(k + 0.5, l)$, $(k, l + 0.5)$, or $(k + 0.5, l + 0.5)$ and classifies all of them into the same subset. However, this principle is difficult to implement in real world situation because LR images usually have arbitrary translations. These arbitrary translations result in more computational complexity in the grouping process.

If $N < q^2$, some subsets classified in the ideal case may be null or may have less than four images. In the absence of sufficient LR images, MAP-based super-resolution cannot recover the high-frequency components; furthermore, image interpolation methods used in SR reconstruction can create blocking artifacts in diagonal edges or lines.

The proposed fast MAP-based SR algorithm is presented in Section 3, which implements the grouping method for real world cases and reduce the blocking artifacts when there are not enough LR images.

III. Proposed Fast MAP-Based SR Algorithm

Suppose that N LR images of the same $L \times W$ size are obtained to reconstruct a HR image measuring $qL \times qW$. The number of new set of LR images is equal to $q \times q$.

Fig. 3 shows that two steps should be performed successively to obtain a new set of LR images before applying recursive multilevel reconstruction algorithm.

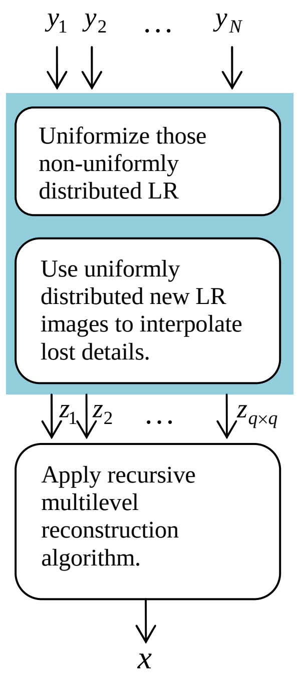

Fig. 3. The block diagram of the proposed fast SR image reconstruction method.

The first step is to convert original LR images $\{y_1, y_2, \ldots, y_N\}$ with arbitrary shifts to a new set of LR images $\{z_1, z_2, \ldots, z_{q\times q}\}$ with uniform shifts. Motion information obtained by registration is used to estimate the relative shifts between each LR image and the reference LR image. According to the relative shifts, each LR image is determined to participate in making up four nearest new LR images. In Fig. 4, the green circles are pixels in the reference LR image. The red circle represents one LR image, which plays a part in the construction of images $\{z_2, z_3, z_6, z_7\}$. The blue circle is another LR image, which influences the images $\{z_{11}, z_{12}, z_{15}, z_{16}\}$. The construction of a new image z_i can be expressed by

$$z_i(n_1, n_2) = \frac{\sum_{j\in\Omega} \omega_j y_j(m_1, m_2)}{\sum_{j\in\Omega} \omega_j}, \qquad 1 \le i \le q \times q, \quad (13)$$

where Ω denotes the set $\{y_j \mid j \in \Omega\}$ and z_i is the weighted average of LR images in the set.

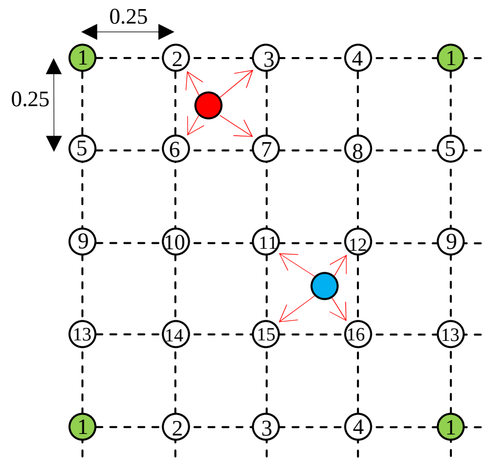

Fig. 4. The first step in the interpolation process of the proposed method: distribute non-uniformly distributed LR images to obtain new set of LR images.

Since the transformations between image z_i and y_j are already known, the nearest neighbor search can be employed to determine the pixel value $y_j(m_1, m_2)$.

Fig. 5 shows that the pixel $y_j(n'_1, n'_2)$ is transformed from pixel $z_i(n_1, n_2)$; thus, pixel $y_j(m_1, m_2)$ is its nearest neighbor. By using (13), a new LR image can be obtained by weighted combination of its nearest neighbors, where ω_j is the weight estimated by a Gaussian function based on the distance from $y_j(n'_1, n'_2)$ to $y_j(m_1, m_2)$.

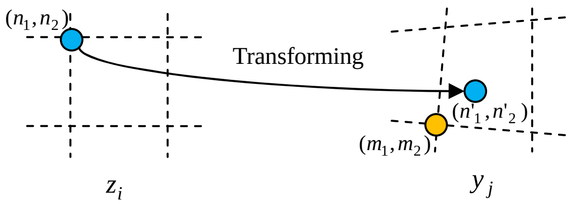

Fig. 5. Nearest neighbor search.

In the second step, if any LR images in the set $\{z_1, z_2, \ldots, z_{q\times q}\}$ are still null, these null images can be constructed by other non-null LR images. In Fig. 6, the white circle labeled as 8 is a null image, and it can be interpolated by non-null images $\{z_1, z_3, z_7, z_{11}, z_{12}\}$. The interpolation can be formulated as

$$z_i(n_1,n_2) = \frac{\sum_{j\in\Theta}\omega_j z_j(n_1+l, n_2+k)}{\sum_{j\in\Theta}\omega_j}, \quad i\in\Phi, \quad l,k\in[-1,0,1] \tag{14}$$

where Φ denotes the set in which images are null, and Θ represents those non-null images in a 3×3 neighborhood of image z_i. The pixel $z_i(n_1,n_2)$ can be obtained by merging the closest pixels in $\{z_j \mid j\in W\}$. For example, as shown in Fig. 6, the pixel $z_8(n_1,n_2)$ is a weighted average of pixels $\{$ $z_1(n_1+1,n_2)$, $z_3(n_1,n_2)$, $z_7(n_1,n_2)$, $z_{11}(n_1,n_2)$, $z_{12}(n_1,n_2)$ $\}$.

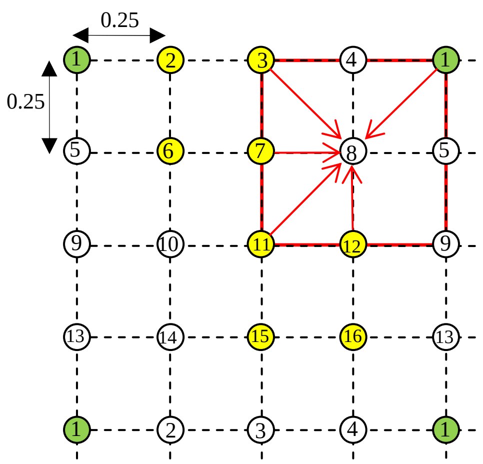

Fig. 6. The second step in the interpolation process of the proposed method: interpolate null images by its neighbors.

In summary, the proposed algorithm turns arbitrarily distributed LR images into a new set of LR images with a uniform distribution, it makes the grouping process simple and easy to implement. Next, when input images in a subset is not enough to recover high frequency components in a HR image, a simple interpolation in (14) is proposed to avoid blocking artifacts. For example, as shown in Fig. 7, only 9 LR images (represented by orange circles) are available for magnifying a HR image by a factor of four. If Zhang's method [7] is applied, subsets $\{s_2, s_3, s_4\}$ will fall short of the necessary number of LR images, as Fig. 7(a) indicates. Unlike Zhang's work, the proposed algorithm estimates the lost LR images (represented by blue circles) in $\{s_2, s_3, s_4\}$ before grouping, as Fig. 7(b) shows.

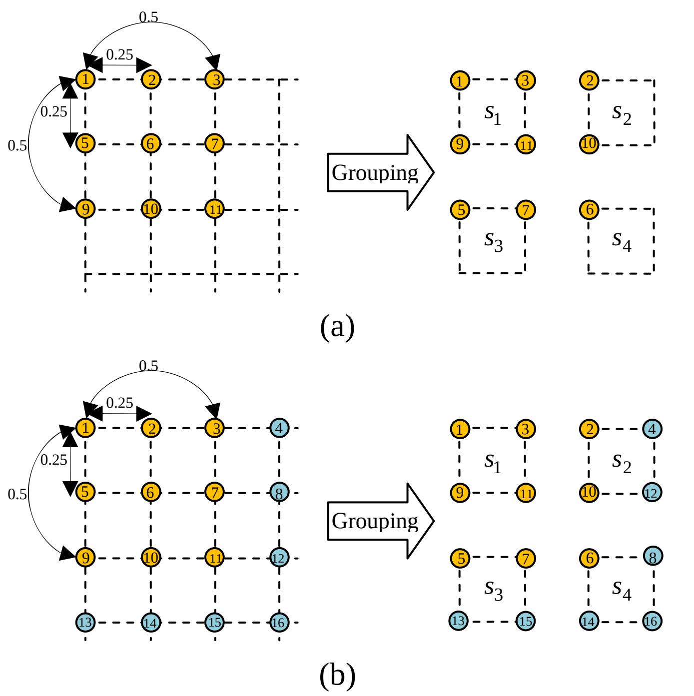

Fig. 7. The improvement in the grouping of the proposed method. (a) The pictorial grouping example for Zhang's work [7]; (b) The grouping result of the proposed method.

IV. Experimental Results

A. Simulations

This study includes a controlled experiment to test the performance of the proposed algorithm. The *Lena*, *Baboon*, and *Airplane* images are horizontally and vertically down-sampled by four to create 16 LR images measuring 128×128. Nine of these 16 LR images were used as input images to model the inadequacy of LR images. Fig. 7(a) shows how these 9 LR images were uniformly distributed. The following four different algorithms for reconstructing a HR image are compared:

1. Bilinear interpolation.
2. MAP-based reconstruction algorithm [8].
3. Recursive multilevel reconstruction approach [7].
4. The proposed fast SR reconstruction algorithm.

This study uses the PSNR between the original image and the reconstructed image as a measure to provide a quantitative comparison of reconstruction quality. To compare computational complexity, we recorded the CPU runtime of each algorithm except the bilinear interpolation. All the tests were run on a personal computer with a Pentium 4 3-GHz CPU and 1G RAM.

Fig. 8 shows the reconstructed results of the four algorithms on the "Lena" image. The MAP-based algorithm reconstructs the HR image with sharper edges and requires more operation time. The recursive multilevel SR algorithm loses some edge information and suffers from blocking artifacts; however, it is the fastest method. On the other hand, the proposed algorithm simultaneously recovers high-frequency components and reduces artifacts. Its complexity lies between the MAP-based and recursive multilevel algorithm. Table 1 and 2 give the comparisons of the PSNR performance and processing time respectively between the proposed method and other approaches in the literature. Fig. 9 shows the reconstructed results of the four algorithms on the image "Airplane".

Table 1 PSNR(dB) performance comparison of the proposed method with other approaches in the literature on test images *Lena*, *Baboon* and *Airplane*.

Test Image	Bilinear Interpolation	MAP	Recursive multilevel	Proposed
Lena	26.67	28.21	27.28	27.89
Baboon	19.21	20.69	20.04	20.48
Airplane	25.48	27.11	26.54	26.87

Table 2 Comparison of processing time (seconds) of the proposed method with other approaches in the literature on test images *Lena*, *Baboon* and *Airplane*.

Test Image	Bilinear Interpolation	MAP	Recursive multilevel	Proposed
Lena	0	29	12	17
Baboon	0	40	16	22
Airplane	0	28	12	18

B. Real data set

In addition to the simulated data, we also tested the algorithm with a real image sequence *books*, which captured by a consumer digital camera. There are 60 LR images measuring 160×120 with arbitrary shifts which are used to magnify a HR image by a factor of 8. The reconstructed results by using the simplest bilinear interpolation algorithm, MAP-based approach and the proposed algorithm are shown in Fig. 10. It is obvious that the MAP-based algorithm has the best visual performance, however it requires 2325 seconds for reconstruction, while the proposed algorithm only needs 137 seconds. This dramatic reduction of 94% in CPU runtime is achieved but sacrificing little quality of the HR image.

V. CONCLUSION

This paper has presented a new fast MAP-based SR reconstruction algorithm, it has lower computational burden compared with the conventional Bayesian super-resolution approach. By using a simple two-step interpolation method, the proposed method reduces obvious blocking artifacts in the reconstructed SR image, it enhances the performance of the recursive multilevel algorithm effectively, especially in the case when there are not many LR images are available. And the proposed method is also able to deal with the problem of LR images with arbitrary transformations. As a future work, we will investigate the higher-level interpolation method to improve the quality of the reconstructed SR images.

(a)

(b)

(c)

(d)

Fig. 8. Comparison of reconstructed *Lena* with bilinear interpolation, MAP-based reconstruction algorithm and the recursive multilevel SR algorithm [7]. (a) Result using bilinear interpolation; (b) Result recovered by the MAP-based reconstruction algorithm. (c) Result recovered by the recursive multilevel SR algorithm [7]. (d) Result recovered by our algorithm.

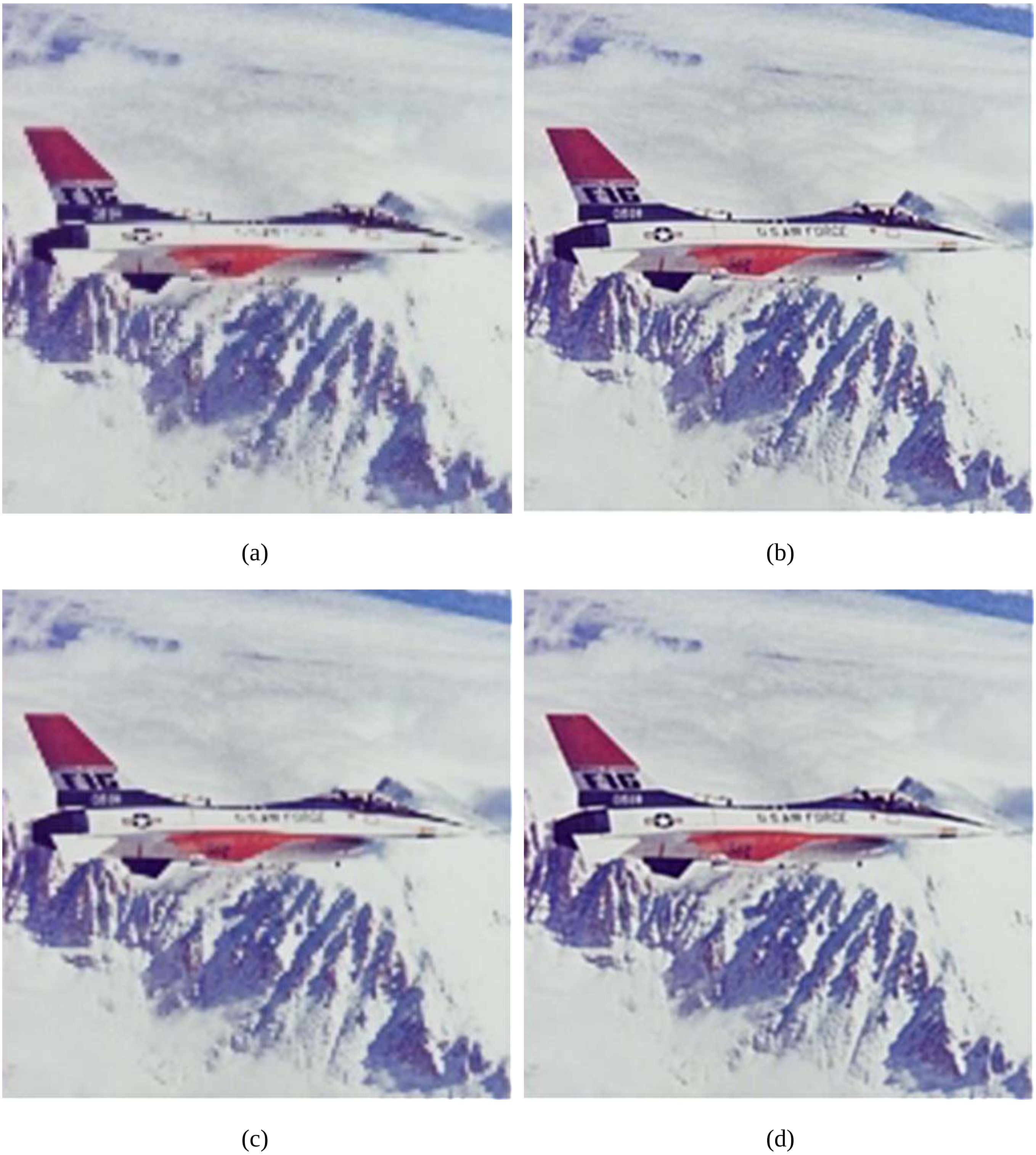

Fig. 9. Comparison reconstructed *Airplane* with bilinear interpolation, MAP-based reconstruction algorithm and the recursive multilevel SR algorithm [7]. (a) Result using bilinear interpolation; (b) Result recovered by the MAP-based reconstruction algorithm. (c) Result recovered by the recursive multilevel SR algorithm [7]. (d) Result recovered by our algorithm.

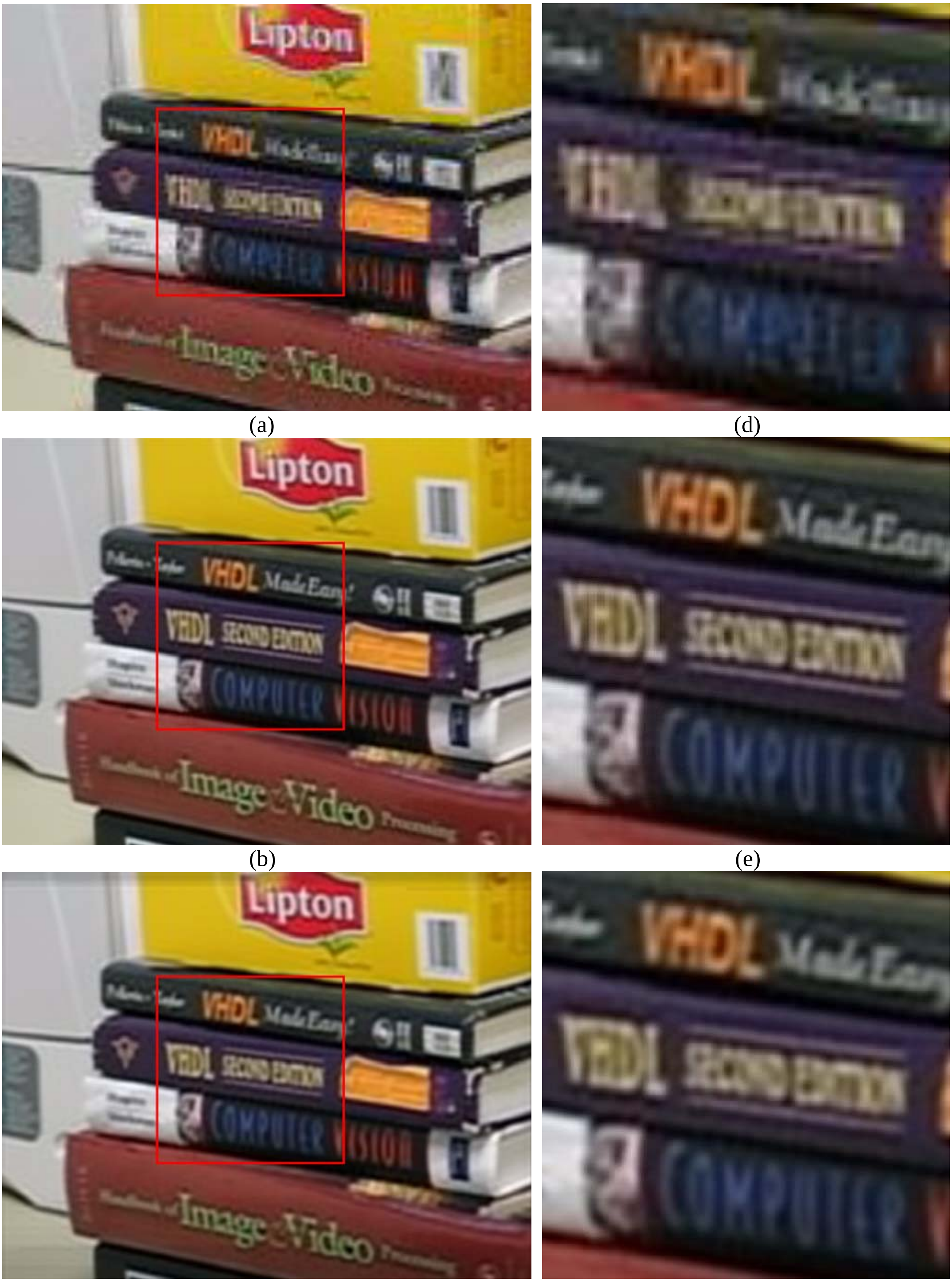

Fig. 10. Comparison of a real case example *Books*. (a) Result using bilinear interpolation. (b) Result reconstructed by the MAP-based reconstruction algorithm. (c) Result reconstructed by our algorithm. (d) , (e) and (f) Zoom-in on the areas of interest in (a), (b) and (c) respectively.

VI. Acknowledgment

This research is partially supported by National Science Council Grant NSC98-2221-E-005-082

References

[1] S. Baker, T. Kanade, “Limits on super-resolution and how to break them,” IEEE Trans. Pattern Analysis and Machine. Intelligence.vol. 24, Issue 9, pp. 1167–1183, September 2002.

[2] T.J. Connolly, R.G. Lane, “Gradient methods for superresolution,” in Proceedings of the International Conference on Image Processing. vol. 1, pp. 917–920, 1997.

[3] R.C. Hardie, K.J. Barnard, J.G. Bognar, E.E. Armstrong, E.A. Watson, “High resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system,” Optical Engineering. vol. 37, no. 1, pp. 247–260, 1998.

[4] N. Nguyen, P. Milanfar, G. Golub, “A computationally efficient superresolution image reconstruction algorithm,” IEEE Trans. Image Processing. vol. 10, Issue 4, pp. 573–583, April 2001.

[5] S.C. Park, M.K. Park, M.G. Kang, “Super-resolution image reconstruction: A technical overview,” IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21–36, May 2003.

[6] J.H. Shin, J.S. Yoon, J.K. Paik, M. Abidi, “Fast superresolution for image sequences using motion adaptive relaxation parameters,” in Proceedings of the 1999 International Conference on Image Processing. vol. 3, pp. 676–680, 1999.

[7] D. Zhang, H. Li and M. Du, “Fast MAP-based multiframe super-resolution image reconstruction,” Image and Vision Computing, vol. 23, no. 7, pp. 671-679, July 2005.

[8] A. Zomet, A. Rav-Acha, S. Peleg, “Robust super-resolution,” Computer Vision and Pattern Recognition, vol. 1, pp. 645-650, December 2001.