Advance Access publication on June 24, 2007 doi:10.1093/comjnl/bxm028

Super-Resolution Reconstruction Algorithm To MODIS Remote Sensing Images

HUANFENG SHEN[1], MICHAEL K. NG[2,*], PINGXIANG LI[1] AND LIANGPEI ZHANG[1]

[1]*The State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei, China*
[2]*Centre for Mathematical Imaging and Vision, Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong*
**Corresponding author: mng@math.hkbu.edu.hk*

In this paper, we propose a super-resolution image reconstruction algorithm to moderate-resolution imaging spectroradiometer (MODIS) remote sensing images. This algorithm consists of two parts: registration and reconstruction. In the registration part, a truncated quadratic cost function is used to exclude the outlier pixels, which strongly deviate from the registration model. Accurate photometric and geometric registration parameters can be obtained simultaneously. In the reconstruction part, the L_1 norm data fidelity term is chosen to reduce the effects of inevitable registration error, and a Huber prior is used as regularization to preserve sharp edges in the reconstructed image. In this process, the outliers are excluded again to enhance the robustness of the algorithm. The proposed algorithm has been tested using real MODIS band-4 images, which were captured in different dates. The experimental results and comparative analyses verify the effectiveness of this algorithm.

Keywords: super-resolution; MODIS images; outliers; L_1 norm data fidelity; Huber prior

Received 12 May 2006; revised 23 April 2007

1. INTRODUCTION

High-resolution (HR) images are useful in many applications such as remote sensing, video frame freezing, medical diagnostics and military information gather, etc. However, because of the high cost and physical limitations of the high-precision optics and image sensors, it is not easy to obtain the desired HR images in many cases. Therefore, super-resolution (SR) image reconstruction techniques, which can reconstruct one or a set of HR images from a sequence of low-resolution (LR) images of the same scene, have widely been researched in the last two decades. Multi-frame SR problem was first formulated by Tsai and Huang [1] in the frequency domain. They proposed a formulation for the reconstruction of an HR image from a set of under-sampled, aliased but noise-free LR images. Kim *et al.* [2] extended the formulation to consider observation noise as well as the effects of spatial blurring. They solved the extended formulation by weighted recursive least squares method to improve computational efficiency. Later, Kim and Su [3] extended their work by considering different blurs for each LR image. Rhee and Kang [4] proposed a DCT-based algorithm in which the computational costs were reduced using discrete cosine transform (DCT) instead of discrete fourier transform (DFT). Furthermore, there have been appeared a couple of papers, which concentrate on wavelet SR methods [5–7].

In the spatial domain, Ur and Gross [8] suggested a non-uniform interpolation method based on the generalized multi-channel sampling theorem of Papoulis [9] and Yen [10]. Irani and Peleg [11] proposed an iterative back-projection method adapted from a computer-aided tomography. Stark and Oskoui [12] proposed a noteworthy projection onto convex sets (POCS) based formulation to SR image reconstruction problems. Their approach was extended by Tekalp *et al.* to include the observation noise [13] and motion blur [14]. Patti *et al.* [15] extended the POCS approach accounting for arbitrary sampling lattices and non-zero aperture time. Using statistical methods, Tom and Katsaggelos [16] suggested a maximum-likelihood (ML) formulation for SR reconstruction. Shultz and Stevenson [17] developed a popular maximum *a posteriori* (MAP) formulation to the SR problem. This

MAP estimator uses a motion compensated sub-sampling matrix-based observation model and an edge preserving Huber–Markov random field for the image prior. Hardie *et al.* [18] also presented a MAP SR reconstruction procedure in which the cases of global as well as non-global motion estimation were considered. Ng and Yip [19] used preconditioned conjugate gradient method to speed up the convergence in MAP reconstruction. Elad and Feuer proposed a hybrid ML/MAP/POCS method and an adaptive filtering method in [20] and [21], respectively. Recently, researchers paid much attention to color [22], compressed [23, 24] dynamic [25] and motion-free [26, 27] SR methods.

It is well noted that the first multi-frame SR idea in [1] was motivated by the requirement to improve the resolution of Landsat remote sensing images. In 2002, CNES (National Space Study Center, France) successfully launched the SPOT5 satellite. Using SR technique, SPOT5 can deliver a 2.5 m panchromatic image through the processing of two 5 m images, which are shifted to half a sampling interval by a double CCD linear array [28]. Without the shift-controlled cases, however, the SR techniques are applied to generic camera images, medical images and video sequence much more commonly than remote sensing images. Although some papers in the literature provided the SR results of satellite images, most of them used synthetic images and assumed known motion parameters. The main reason is that it is very difficult to obtain data sources and execute accurate sub-pixel registration. Alternatively, many researchers tackled the image fusion problem of reconstructing an LR image using an HR image. A typical example is the use of panchromatic image for sharpening multi/hyper spectral images [29–34]. However, this type of method often destroys the spectral information, which is very important for the remote sensing application of multi/hyper spectral images.

In this paper, we propose an SR reconstruction algorithm applied to real moderate-resolution imaging spectroradiometer (MODIS) remote sensing images in the same spectral band. MODIS is a key sensor aboard TERRA satellite, which observes every point on our world every 1–2 days, and acquires data in 36 discrete spectral bands at three spatial resolutions (250, 500 and 1000 m) [35]. Our image processing method is concentrated on the fourth spectral band, which has the resolution of 500 m. To get accurate photometric and geometric parameters among the observed images, we used a truncated quadratic cost function to exclude the outliers in the sub-pixel registration part. Then, we use the MAP estimation with robust L_1 norm data fidelity and edge-preserving Huber prior to get the desired HR image in the reconstruction part. The proposed SR algorithm is evaluated visually and quantitatively.

The remainder of this paper is organized as follows. In Section 2, we describe the image observation model. The image registration method and the reconstruction method are presented in Sections 3 and 4, respectively. Experimental results are provided in Section 5. In Section 6, we give some concluding remarks of this paper.

2. IMAGE OBSERVATION MODEL

The Image observation model relates the desired HR image to the observed LR images. Therefore, the first step to comprehensively analyze the SR image reconstruction problem is to understand the image acquisition model. A typical model assumes that the imaging process involves warping followed by blurring and down-sampling to generate LR images from an HR image. Let us denote the underlying HR image in vector form by $\boldsymbol{z} = [z_1, z_2, \ldots, z_{L_1N_1 \times L_2N_2}]^{\mathrm{T}}$, where $L_1N_1 \times L_2N_2$ is the HR image size. Letting L_1 and L_2 denote the down-sampling factors in the horizontal and vertical directions, respectively, each observed LR image is of size $N_1 \times N_2$. Thus, the LR image can be represented as $\boldsymbol{g}_k = [g_{k,1}, g_{k,2}, \ldots, g_{k,\ N_1 \times N_2}]^{\mathrm{T}}$, where $k = 1, 2, \ldots, P$, with P being the number of the LR images. The typical image observation model can be represented as [20, 36]:

$$\boldsymbol{g}_k = \boldsymbol{D}\boldsymbol{B}_k\boldsymbol{M}_k\boldsymbol{z} + \boldsymbol{n}_k, \tag{1}$$

where $\boldsymbol{M}_k$ is the warp matrix with size of $L_1N_1L_2N_2 \times L_1N_1L_2N_2$, $\boldsymbol{B}_k$ represents the blur matrix also of size $L_1N_1L_2N_2 \times L_1N_1L_2N_2$, $\boldsymbol{D}$ is a $N_1N_2 \times L_1N_1L_2N_2$ down-sampling matrix and $\boldsymbol{n}_k$ represents the $N_1N_2 \times 1$ noise vector.

Although this observation model has been used in many SR papers, this model does not satisfy our application. In remote sensing imaging, the sun zenith angle and atmospheric absorption and scattering affect the amount of radiance received by the sensor. It has been mentioned that the real MODIS images used in this paper were captured on different dates, so the photometric effects of zenith angle and atmosphere should be considered. Fortunately, these effects in many cases can simply be modeled as a linear system [37, 38], with which the following image observation model can be obtained.

$$\boldsymbol{g}_k = h_{1,k}\boldsymbol{D}\boldsymbol{B}_k\boldsymbol{M}_k\boldsymbol{z} + h_{0,k}\boldsymbol{I} + \boldsymbol{n}_k, \tag{2}$$

where $\boldsymbol{I}$ is the $N_1N_2 \times 1$ unit vector, $h_{1,k}$ and $h_{0,k}$ are, respectively, the gain and offset of the photometric parameters, which balance the relative effects of sun zenith angle and atmosphere condition between different observed images. It is worth noting that this model may be unsuitable when the images contain too many clouds or different orientation of very steep terrain slopes. The de-cover of clouds and terrain correction are important and independent research directions in the field of remote sensing image processing, and they are beyond the scope of this paper.

3. GEOMETRIC AND PHOTOMETRIC REGISTRATION

From the observation model in equation (2), we can know that photometric registration and geometric registration play critical roles in the SR reconstruction. In this section, we present a robust registration method with which we can get accurate geometric and photometric parameters simultaneously.

In general, the relationship between the observed kth and lth images can be expressed by:

$$E(\boldsymbol{\theta}) = \|\boldsymbol{O}_k(\boldsymbol{g}_k - \boldsymbol{f}_k^{(l,\boldsymbol{\theta})})\|_2^2 \tag{3}$$

where (x, y) denotes the pixel site, $g_k(x, y)$ is a pixel in frame k, $\boldsymbol{\theta}$ is the vector contains both the photometric parameters and the geometric parameters between the two images, $f_k^{(l,\theta)}(x, y)$ is the predicted pixel of $g_k(x, y)$ from frame l using parameter vector $\boldsymbol{\theta}$, and $\varepsilon_k^{(l,\boldsymbol{\theta})}(x, y)$ denotes the model error. By assuming the linear photometric transform and affine motion model, we obtain

$$f_k^{(l,\boldsymbol{\theta})}(x, y) = h_{1,k} \cdot g_l (a_{0,k} + a_{1,k}x + a_{2,k}y, b_{0,k} + b_{1,k}x + b_{2,k}y) + h_{0,k}. \tag{4}$$

In this model, $\boldsymbol{\theta} = (h_{0,k}, h_{1,k}, a_{0,k}, a_{1,k}, a_{2,k}, b_{0,k}, b_{1,k}, b_{2,k})^{\mathrm{T}}$ contains two photometric and six geometric parameters.

To solve $\boldsymbol{\theta}$, there are two common choices for the estimation criteria. They are based on the minimization of the sum of squared difference and the sum of absolute difference error metric [39], respectively. Obviously, the former is in the L_2 norm and the later is in the L_1 norm. The two minimization functions can be expressed as:

$$E(\boldsymbol{\theta}) = \|\boldsymbol{g}_k - \boldsymbol{f}_k^{(l,\boldsymbol{\theta})}\|_2^2 \text{ and } E(\boldsymbol{\theta}) = \|\boldsymbol{g}_k - \boldsymbol{f}_k^{(l,\boldsymbol{\theta})}\|_1,$$

respectively. Generally speaking, the L_2 norm function can easily be solved by some gradient-based methods, but not robust to outliers. However, the L_1 norm function has some robustness to outliers, the minimization can only be accomplished by some direct search methods, which often gets in local minima. Some sophisticated methods for solving the L_1 norm function are developed, see for instance [40], but the computational cost is more expensive. Here we solve this problem using a simple approach. We use a truncated quadratic cost function described as follows:

$$E(\boldsymbol{\theta}) = \|\boldsymbol{O}_k(\boldsymbol{g}_k - \boldsymbol{f}_k^{(l,\boldsymbol{\theta})})\|_2^2. \tag{5}$$

In the objective function in equation (5), $\boldsymbol{O}_k$ is a diagonal matrix that denotes which pixels are outliers. Each pixel in $\boldsymbol{g}_k$ corresponds to a diagonal element in $\boldsymbol{O}_k$. If a pixel is an outlier, its corresponding diagonal element is one otherwise zero. Using this modified quadratic formulation, the outliers are truncated from the minimization, which can lead to a robust estimation of the parameters in the registration part. Here we remark that the above formulation is used to estimate the photometric parameters and geometric parameters between the observed kth and lth images. The corresponding parameters of other frames can be estimated using the same approach.

Using the Gaussian–Newton method, the objective function (5) can iteratively be solved. We just expand the function by Taylor formula, we obtain

$$E(\boldsymbol{\theta}) \approx E(\boldsymbol{\theta}^n) + \left(\frac{\partial E(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^n}\right)^{\mathrm{T}} (\Delta\boldsymbol{\theta}) + \frac{1}{2}(\Delta\boldsymbol{\theta})^{\mathrm{T}}\boldsymbol{H}^n(\Delta\boldsymbol{\theta}), \tag{6}$$

where n is the iteration number, $\partial E(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^n$ and $\boldsymbol{H}^n$ denote the gradient matrix and Hessian matrix of $E(\boldsymbol{\theta})$ at $\boldsymbol{\theta}^n$, respectively, expressed as:

$$\frac{\partial E(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^n} = 2(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{r}^n, \tag{7}$$

$$\boldsymbol{H}^n = 2(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{J}^n + 2\sum_i r_i^n O_{k,ii}^n \boldsymbol{H}_i^n. \tag{8}$$

In equations (7) and (8), $\boldsymbol{r}^n$ is the residual vector that is equal to $\boldsymbol{g}_k - \boldsymbol{f}_k^{(l,\boldsymbol{\theta}_n)}$, $\boldsymbol{J}^n = \partial\boldsymbol{r}^n/\partial\boldsymbol{\theta}^n$ denotes the gradient matrix of $\boldsymbol{r}^n$, and $\boldsymbol{H}_i^n$ is the Hessian matrix of r_i^n. For small $\boldsymbol{r}^n$, we have the following approximations: $\boldsymbol{H}^n \approx 2(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{J}^n$. Thus, differentiating equation (6) with respect to $\Delta\boldsymbol{\theta}$ and setting the result equal to zero, after some manipulation, we have

$$(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{J}^n(\Delta\boldsymbol{\theta}) = -(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{r}^n \tag{9}$$

and

$$\Delta\boldsymbol{\theta} = [(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{J}^n]^{-1}[-(\boldsymbol{J}^n)^{\mathrm{T}}\boldsymbol{O}_k^n\boldsymbol{r}^n]. \tag{10}$$

Then, the parameter vector can be updated by

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n + \Delta\boldsymbol{\theta}. \tag{11}$$

By considering the above computation, the photometric and geometric model parameters can be updated iteratively, and the residual corrections computed by equation (10) will get small with the increase of iteration numbers. The iteration terminates when the corrections are smaller than a specific stopping criterion.

It must be noted the matrix $\boldsymbol{O}_k$, which denotes which pixels are outliers should be updated at each step of the iteration. To detect the outliers, we use two criteria. The first is a *geometrical* criterion that requires the predicted location of a pixel using the current motion parameters is still in the image field. The horizontal and vertical predicted locations are

judged by equations (12) and (13), respectively.

$$0 \le a_{0,k} + a_{1,k}x + a_{2,k}y \le N_1, \tag{12}$$

$$0 \le b_{0,k} + b_{1,k}x + b_{2,k}y \le N_2. \tag{13}$$

Here, N_1 and N_2 are, respectively, the horizontal and vertical size of the observed image. The second criterion is a *photometrical* one. An absolute displaced frame difference criterion $|g_k(x, y) - f_k^{(l,\boldsymbol{\theta}_n)}(x, y)| < d$ is often employed to determine whether a pixel is outlier [17, 41]. In this criterion, the left-hand side $|g_k(x,y) - f_k^{(l,\boldsymbol{\theta}_n)}(x, y)|$ represents the mapping error using the current registration parameters. It is clear that if the mapping error is too large the corresponding pixel should be regarded as an outlier. So d in the right-hand side acts as a scale threshold. However, we note that the mapping error in sharp regions is more possible to be larger than that in smooth regions with the same registration error. To balance this difference, we use the following relative criterion:

$$|g_k(x, y) - f_k^{(l,\boldsymbol{\theta}_n)}(x, y)| < d * \sigma_k(x, y), \tag{14}$$

where $\sigma_k(x, y)$ is the standard deviation at site (x, y) of the kth image. Using equation (14), we can obtain more desired outlier distribution in the experiments.

Because registration parameters are significantly over-determined by the data and outliers are excluded in the minimization, we believe that the use of modified L_2 norm in equation (5) is robust enough. To prove our approach, we have tested the L_1 norm using the Matlab 'fminsearch' function with the L_2 norm solution as an initial guess, and got very similar results with the the L_2 norm. However, the L_1 norm function takes more considerable computational time than our proposed method.

4. IMAGE RECONSTRUCTION ALGORITHM

In many situations, the problem of SR is an ill-posed inverse problem because the information contained in the observed LR images is not sufficient to solve the HR image. Thus, the solution for the HR image can be constructed by applying the MAP technique, which has the following minimization cost function:

$$\hat{\boldsymbol{z}}_{\text{MAP}} = \arg\ \min\left[\sum_k L(\boldsymbol{g}_k, \boldsymbol{z}) + \lambda R(\boldsymbol{z})\right], \tag{15}$$

where the first term $\sum_k L(\boldsymbol{g}_k, \boldsymbol{z})$ is the data fidelity term that provides a force of the conformance of the present HR image to the observed LR images according to the image observation model; the second term $R(\boldsymbol{z})$ is the regularization term that penalizes the estimation by some *a priori* constraint; and λ is the regularization parameter that balances the two competing terms.

For the data fidelity, the linear least square term which is in the L_2 norm is widely used. However, there has been a growing interest in using L_1 norm for image restoration and image reconstruction in the literature [40, 42–44]. It has been proved that the L_1 norm method is more effective than the L_2 norm when the images have random noises [40], inevitable motion estimation error [42] and data outliers [42, 44]. Although we use the modified L_2 norm method in the registration, we have ample reasons to use the L_1 norm method here such as the inevitable motion error and under-determined condition. According to the observation model (2), we can get the straightforward L_1 norm data fidelity term,

$$\sum_k L(\boldsymbol{g}_k, \boldsymbol{z}) = \sum_k \|\boldsymbol{g}_k - h_{1,k}\boldsymbol{D}\boldsymbol{B}_k\boldsymbol{M}_k\boldsymbol{z} - h_{0,k}\boldsymbol{I}\|_1. \tag{16}$$

In order to increase the effectiveness of the SR algorithm, the outliers should also be excluded in the image reconstruction part as in the image registration part. Although the outliers in the registration part and in the reconstruction part maybe have some differences, we ignore such differences as they are usually small. Thus, we can directly obtain the information about the outliers from the registration results. In this manner, we rewrite the data fidelity term as follows:

$$\sum_k L(\boldsymbol{g}_k, \boldsymbol{z}) = \sum_k \|\boldsymbol{O}_k(\boldsymbol{g}_k - h_{1,k}\boldsymbol{D}\boldsymbol{B}_k\boldsymbol{M}_k\boldsymbol{z} - h_{0,k}\boldsymbol{I})\|_1, \tag{17}$$

where $\boldsymbol{O}_k$ has been defined in Section 2.

For the regularization term, Tikhonov regularization and Gauss–Markov prior are commonly used. Tikhonov regularization has the following form

$$R(\boldsymbol{z}) = \|\boldsymbol{Q}\boldsymbol{z}\|_2^2, \tag{18}$$

where $\boldsymbol{Q}$ is often chosen as a linear high-pass operation that penalizes the estimation which is not smooth. Differently, the regularization term using Gauss–Markov prior can be expressed as:

$$R(\boldsymbol{z}) = \sum_{x,y}\sum_{c\in C}\rho(d_c(z_{x,y})). \tag{19}$$

In this expression, c is a clique within the set of all image cliques C, the quantity $d_c(z_{x,y})$ is a spatial activity measure to pixel $z_{x,y}$, which is often formed by first-order or second-order differences, and $\rho(\cdot)$ is a quadratic potential function

$$\rho(i) = i^2. \tag{20}$$

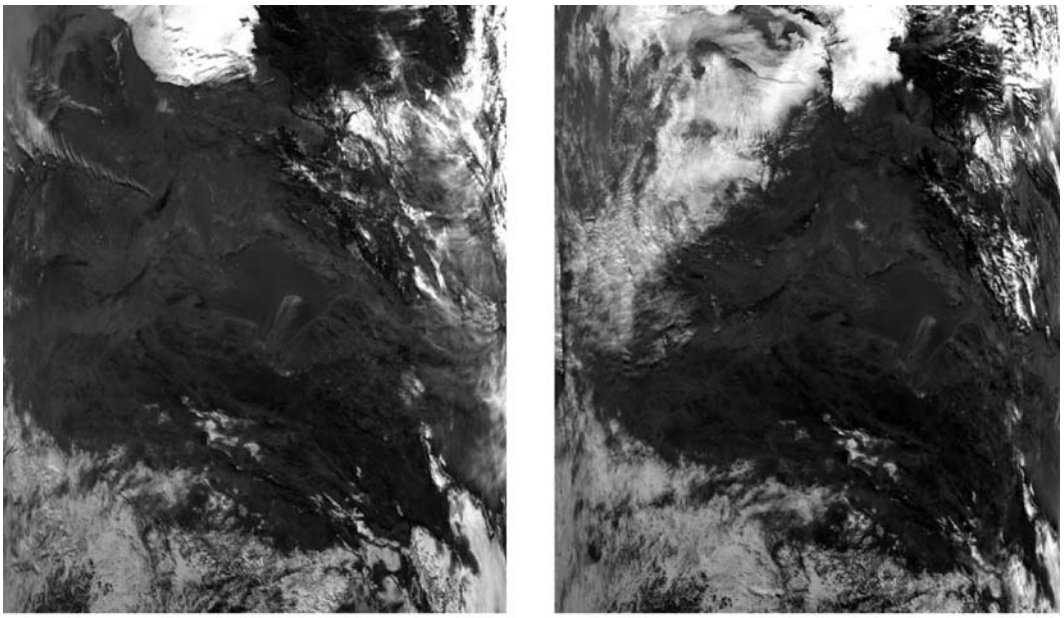

FIGURE 1: Two of the MODIS band-4 full-view images. (a) Captured on 6 January 2004. (b) Captured on 8 January 2004.

A common criticism to these two regularizations is that the sharp edges in the SR estimates tend to be overly smoothed. To solve this problem, some edge-preserving regularization has been used in the literature. For instance, the Huber–Markov prior was used in [17, 45], and a robust regularization function called bilateral-TV was used in [22, 43]. Both the two regularization functions perform better than the non-edge-preserving ones. In this paper, we select the Huber–Markov prior model as regularization. The difference between Huber–Markov prior and Gauss–Markov prior is only on the potential function $\rho(\cdot)$. The Huber function is defined as:

$$\rho(i) = \begin{cases} i^2, & |i| \le \mu \\ 2\mu|i| - \mu^2, & |i| > \mu, \end{cases} \tag{21}$$

where μ is a threshold parameter separating the quadratic and linear regions [17]. It is easy to see that the Gauss–Markov prior can be regarded as a special case of the Huber–Markov prior with $\mu \to \infty$. As for the $d_c(z_{x,y})$, we compute the following finite second-order differences in four adjacent cliques for every location (x, y) in the SR image

$$d_c^1(z_{x,y}) = z_{x-1,y} - 2z_{x,y} + z_{x+1,y}, \tag{22}$$

$$d_c^2(z_{x,y}) = z_{x,y-1} - 2z_{x,y} + z_{x,y+1}, \tag{23}$$

$$d_c^3(z_{x,y}) = \frac{1}{\sqrt{2}}\left[z_{x-1,y-1} - 2z_{x,y} + z_{x+1,y+1}\right], \tag{24}$$

$$d_c^4(z_{x,y}) = \frac{1}{\sqrt{2}}\left[z_{x-1,y+1} - 2z_{x,y} + z_{x+1,y-1}\right]. \tag{25}$$

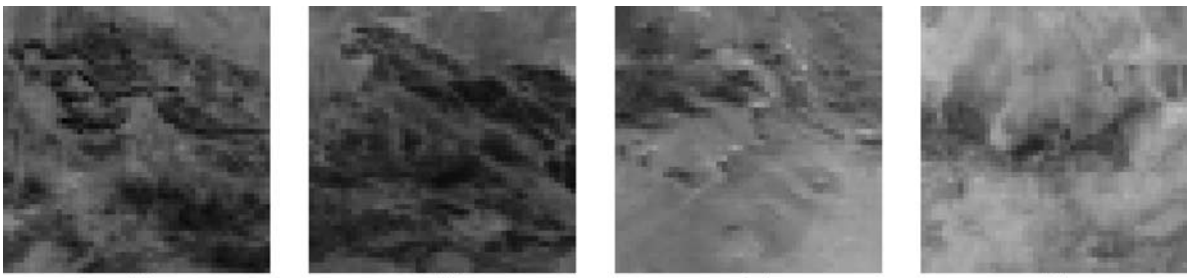

FIGURE 2: The four 50×50 interest regions cropped from the full-view band-4 MODIS image captured on 6 January 2004. (a) Region 1. (b) Region 2. (c) Region 3. (d) Region 4.

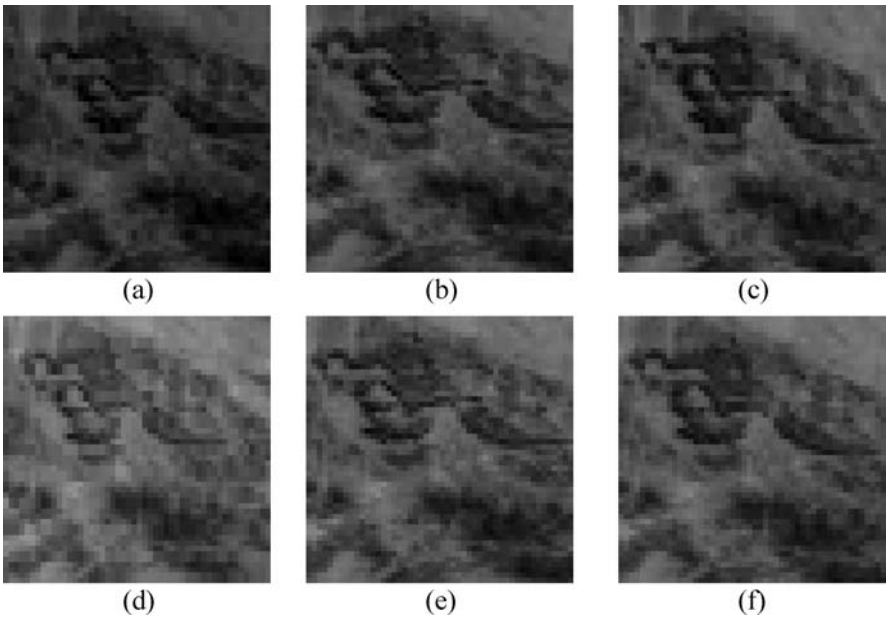

FIGURE 3: Block images of Region 1 captured on: (a) 28 December 2003. (b) 30 December 2003. (c) 1 January 2004. (d) 4 January 2004. (e) 6 January 2004. (f) 8 January 2004.

TABLE 1: The estimated photometric and geometric registration parameters in the experiment of region 1.

	$a_{0,k}$	$a_{1,k}$	$a_{2,k}$	$b_{0,k}$	$b_{1,k}$	$b_{2,k}$	$h_{0,k}$	$h_{1,k}$
28 December 2003	−0.710	1.002	0.023	−5.739	−0.020	1.104	−26.078	0.943
30 December 2003	1.594	1.003	−0.019	−0.153	0.001	0.971	−13.011	0.981
1 January 2004	1.730	0.984	−0.069	4.437	0.028	1.126	−14.824	1.022
4 January 2004	−1.141	1.013	0.037	−2.950	−0.027	1.269	18.142	0.991
8 January 2004	0.201	0.998	−0.038	−1.167	0.031	1.011	−6.293	1.034

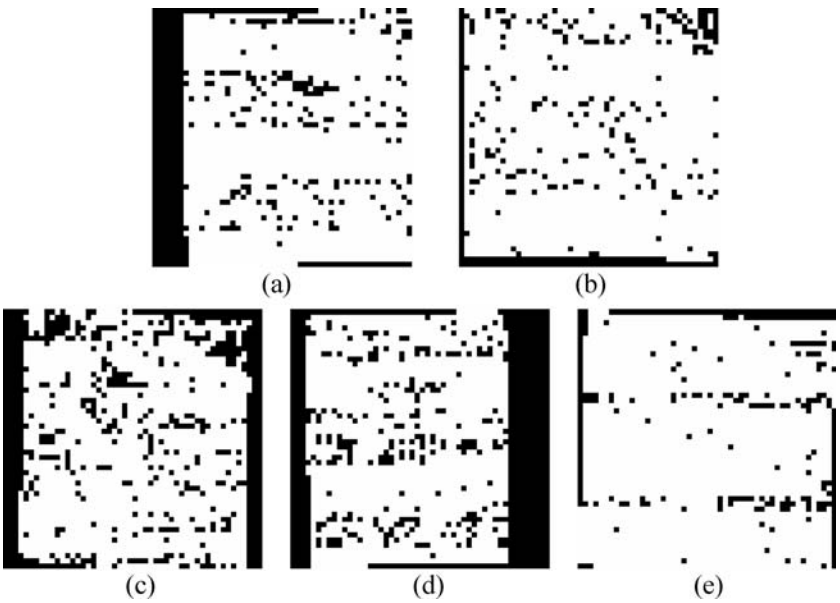

FIGURE 4: The outlier maps of the unreferenced images in the Region 1 experiment. (a) 28 December 2003. (b) 30 December 2003. (c) 1 January 2004. (d) 04 January 2004. (e) 8 January 2004.

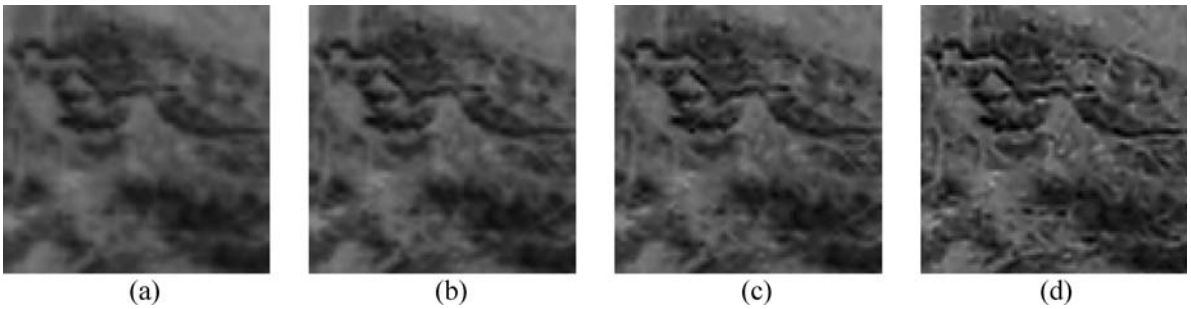

FIGURE 5: Experimental results of Region 1. (a) Bilinear interpolation. (b) Cubic interpolation. (c) MAP interpolation. (d) The proposed SR algorithm.

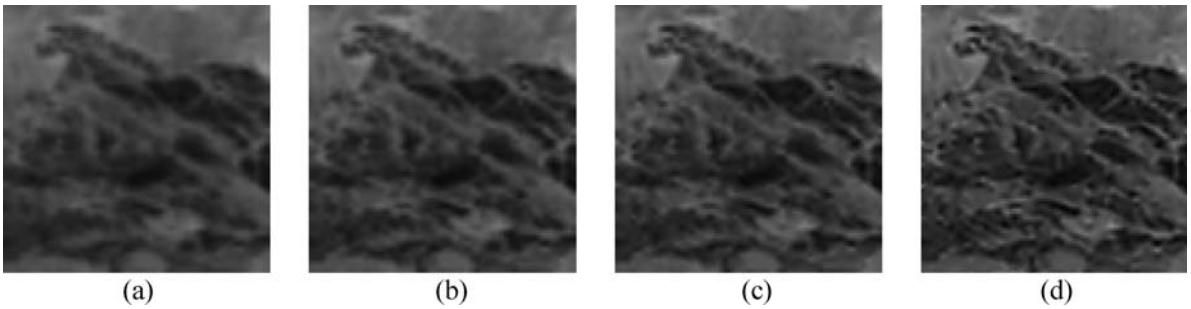

FIGURE 6: Experimental results of Region 2. (a) Bilinear interpolation. (b) Cubic interpolation. (c) MAP interpolation. (d) The proposed SR algorithm.

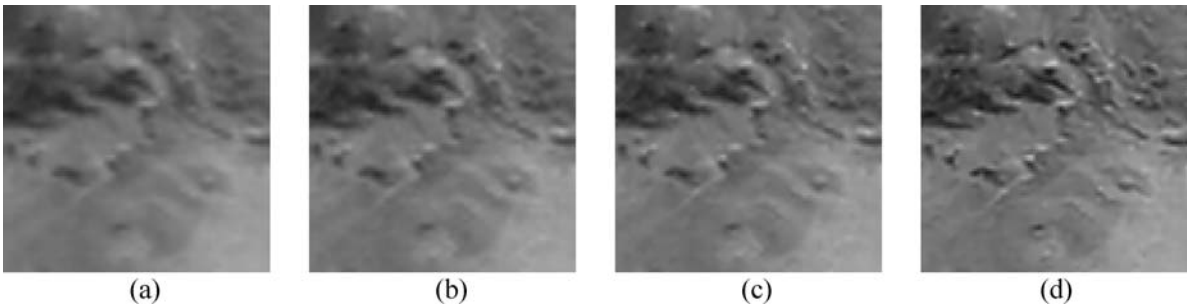

FIGURE 7: Experimental results of Region 3. (a) Bilinear interpolation. (b) Cubic interpolation. (c) MAP interpolation. (d) The proposed SR algorithm.

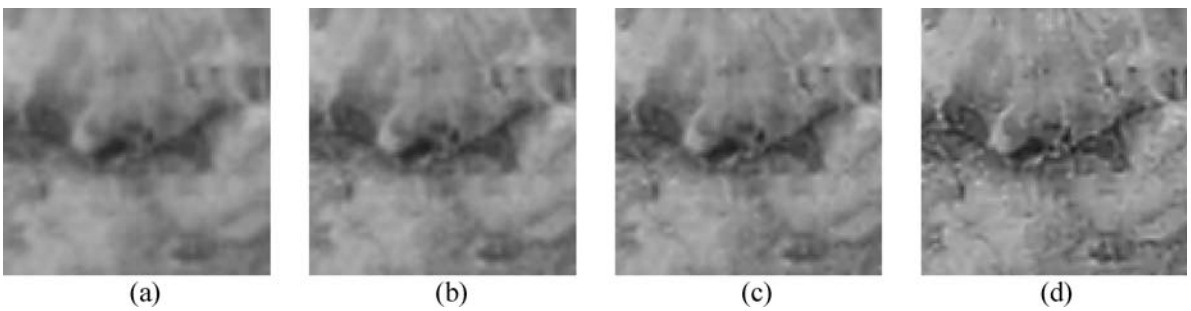

FIGURE 8: Experimental results of Region 4. (a) Bilinear interpolation. (b) Cubic interpolation. (c) MAP interpolation. (d) The proposed SR algorithm.

Using the L_1 data fidelity term and the Huber–Markov prior model, we can rewrite the minimization cost function (15) as:

$$\hat{\boldsymbol{z}}_{\mathrm{MAP}} = \arg\min\left[\sum_k \|\boldsymbol{O}_k(\boldsymbol{g}_k - h_{1,k}\boldsymbol{D}\boldsymbol{B}_k\boldsymbol{M}_k\boldsymbol{z} - h_{0,k}\boldsymbol{I})\|_1 + \lambda\sum_{x,y}\sum_{c\in C}\rho(d_c(z_{x,y}))\right]. \quad (26)$$

The steepest descent optimization is used to minimize this cost function. Differentiating equation (26) with respect to $\boldsymbol{z}$, we have:

$$\boldsymbol{r} = -h_{1,k}\boldsymbol{M}_k^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{O}_k^{\mathrm{T}} \operatorname{sign}(\boldsymbol{g}_k - h_{1,k}\boldsymbol{D}\boldsymbol{B}_k\boldsymbol{M}_k\boldsymbol{z} - h_{0,k}\boldsymbol{I}) + \lambda\boldsymbol{r}', \quad (27)$$

where $\boldsymbol{r}''$ is the derivative of the regularization term that can be solved on a pixel-by-pixel basis. Thus, the HR image is solved using the successive approximations iteration

$$\hat{\boldsymbol{z}}^{n+1} = \hat{\boldsymbol{z}}^n - \beta\boldsymbol{r}^n, \quad (28)$$

where β is the step size and n is the iteration number.

It is worth noting that the matrices $\boldsymbol{D}$, $\boldsymbol{B}_k$, $\boldsymbol{M}_k$ and $\boldsymbol{O}_k$, and their corresponding transposes can be interpreted as direct image operators in the computation process [25, 46]. By this means, it is not necessary to generate big matrices. The proposed algorithm can be implemented very efficiently.

5. EXPERIMENTAL RESULTS

In the experiments, we used six MODIS band-4 images with ground resolution of 500 m. These images were captured on 28, 30 December 2003 and 1, 4, 6, 8 January 2004, respectively. We obtained these images from the Satellite Remote Sensing Receiving Station of Wuhan University. As the original data type is SHORT INTEGER (2 byte), we firstly converted them into the BYTE type (1 byte) for convenience. To retain the original relations between the observed images, we used the same linear transformation to all images in the conversion. Two of the transformed full-view images which were captured on 6 and 8 January 2004 are, respectively, shown in Fig. 1a and b as illustration examples. The original size of the full-view images is 2078×3260, but our data processing was restricted to typical block images to reduce computational cost and avoid the effect of the existing clouds in the images. Furthermore, we found the global motion model may be unsuitable if the block size was chosen too big. One main reason is that there are considerable scanning and mosaicing errors in the MODIS images. With this in mind, we chose four 50×50 interest regions to test the proposed algorithm independently. The four corresponding regions cropped from the full-view image captured on 6 January 2004 are shown in Fig. 2a–d, respectively. Besides, we also demonstrate the six multi-temporal block images of region 1 in Fig. 3.

It is clear from Fig. 3a–f that these block images have different luminance and geometric sampling because they are captured in different dates and with different angles of view. This is also true for the other three sequences of block images. Therefore, the photometric registration and the geometric registration are both necessary. In all the experiments of the four selected regions, the registration method presented in Section 3 was implemented regarding the 6 January 2004 as the referenced image, and the parameter d in equation (14) was set to 1.2. Here, we demonstrate the estimated registration parameters and the detected outliers in the experiment of region 1 in Table 1 and Fig. 4, respectively. In the outlier maps, the uniform outliers around the image verge are detected by conditions (12) and (13), and the other discrete ones are detected by condition (14). Besides the verge ones, the outliers distribute somewhat along the horizontal direction. The special example is Fig. 4e, in which some equal-interval outlier lines exist. Actually, this outlier distribution is just consistent with the remote sensor imaging because MODIS scans the surface across the flight track (vertical direction in the image). Moreover, the scan interval is 20 pixels (in 500 m resolution bands), which is equal to the outlier line interval in Fig. 4e. Therefore, these horizontally distributed outliers can be ascribed to the abnormal sensor scanning and/or image mosaicing. This also proves that the detected outliers are necessary and useful.

The next issue is the evaluation and analysis of the reconstructed HR images. In the reconstruction part, the corresponding parameters were set as: $\mu = 1$, $\lambda = 0.01$ and $\beta = 4$. A 3×3 Gaussian blur kernel with unit variance was assumed and was commonly used. We assume the downsampling factors in both the horizontal and vertical directions have a value of two. The initial HR image is obtained using the bi-cubic interpolation, and the iteration was terminated when the iteration numbers get to 60. The proposed SR algorithm was compared with the traditional bilinear interpolation, cubic interpolation and a Huber MAP interpolation [47], which can be regarded as the special single-frame case of the proposed algorithm. The experiment results of the four selected regions are, respectively, shown in Figs. 5–8. By visual comparison, it is seen that the results of the proposed SR algorithm are much clearer than those of the single interpolation algorithms, the reason for which is that these results fused the complementary information in different observed images.

The quantitative evaluation is not easy because the real HR band-4 MODIS images are not known. A commonly used strategy to solve the problem that no real HR images exist is to degrade the original images to an inferior resolution level and reconstruct the degraded images, and then treat the

TABLE 2: Quantitative evaluation results of Region 1.

	Bilinear interpolation	Cubic interpolation	MAP interpolation	Proposed SR algorithm
MSE	101.999	78.930	77.242	59.372
CC	0.935	0.948	0.949	0.961
UIQI	0.924	0.944	0.948	0.960

TABLE 3: Quantitative evaluation results of Region 2.

	Bilinear interpolation	Cubic interpolation	MAP interpolation	Proposed SR algorithm
MSE	64.124	52.650	52.005	35.990
CC	0.950	0.958	0.958	0.971
UIQI	0.945	0.956	0.958	0.971

TABLE 4: Quantitative evaluation results of Region 3.

	Bilinear interpolation	Cubic interpolation	MAP interpolation	Proposed SR algorithm
MSE	129.800	103.806	101.067	77.890
CC	0.918	0.932	0.933	0.949
UIQI	0.901	0.926	0.931	0.947

TABLE 5: Quantitative evaluation results of Region 4.

	Bilinear interpolation	Cubic interpolation	MAP interpolation	Proposed SR algorithm
MSE	49.198	37.782	37.735	31.776
CC	0.963	0.971	0.971	0.975
UIQI	0.958	0.970	0.971	0.975

original images as the real HR images to compare with the reconstructed images [29–33]. We also adopt this strategy in this paper. As for the evaluation criteria, the mean square error (MSE) and correlation coefficient (CC) are two popular ones. Besides, Zhou and Bovik [48] proposed a universal image quality index (UIQI) for the evaluation of single band image data. This index has been used in [29, 49]. Here, we use these three criteria to evaluate the proposed SR algorithm. They are, respectively, defined by equations (29)–(31).

$$\mathrm{MSE} = \frac{\|\hat{z} - z\|^2}{N}, \tag{29}$$

$$\mathrm{CC} = \frac{\sigma_{\hat{z}z}}{\sigma_{\hat{z}}\sigma_z}, \tag{30}$$

$$\mathrm{UIQI} = \frac{4\sigma_{\hat{z}z} m_{\hat{z}} m_z}{(\sigma_{\hat{z}}^2 + \sigma_z^2)(m_{\hat{z}}^2 + m_z^2)}. \tag{31}$$

Here, $\hat{z}$ and z represent the reconstructed HR image and the original image, respectively, and N is the total number of pixels in the image. $\sigma_{\hat{z}z}$ is the covariance between and $\hat{z}$ and z, $m_{\hat{z}}$ and m_z their means, and $\sigma_{\hat{z}}$ and σ_z their standard deviations, respectively. The ideal values of the MSE, CC and UIQI are, respectively, 0, 1 and 1. The evaluation results of the four selected regions are, respectively, shown in Tables 2–5. It is seen that the proposed algorithm produced much better quantitative evaluation results than the single frame interpolation algorithms. This agrees with the visual judgment.

6. CONCLUSIONS

In this paper, we have proposed an SR image reconstruction algorithm to real MODIS remote sensing images. To increase the robustness and the effectiveness, the outliers are truncated in both the registration and reconstruction processes. The proposed reconstruction algorithm simultaneously uses the robust L_1 norm data fidelity and edge-preserving Huber prior. Experiment results validated the proposed algorithm performs better than single frame interpolation methods in terms of both the quantitative measurements and visual evaluation.

ACKNOWLEDGEMENTS

H.S. would like to thank M.K.N for his hospitality during his visit to Center for Mathematical Imaging and Vision, Hong Kong Baptist University from March 2006–March 2007. This work was done during his visit to Hong Kong Baptist University. Research was supported in part by RGC 7035/04P and 7035/05P, HKBU FRGs.

REFERENCES

[1] Tsai, R. Y. and Huang, T. S. (1984) Multi-frame image restoration and registration. *Adv. Comput. Vis. Image Process.*, **1**, 317–339.

[2] Kim, S. P., Bose, N. K. and Valenzuela, H. M. (1990) Recursive reconstruction of high-resolution image from noisy undersampled multiframes. *IEEE Trans. Acoust. Speech Signal Process.*, **38**, 1013–1027.

[3] Kim, S. P. and Su, W. Y. (1993) Recursive high-resolution reconstruction of blurred multiframe images. *IEEE Trans. Image Process.*, **2**, 534–539.

[4] Rhee, S. and Kang, M. G. (1999) Discrete cosine transform based regularized high-resolution image reconstruction algorithm. *Opt. Eng.*, **38**, 1348–1356.

[5] Chan, R., Chan, T., Shen, L. and Shen, Z. (2003) Wavelet algorithms for high-resolution image reconstruction. *SIAM J. Sci. Comput.*, **24**, 1408–1432.

[6] Lertrattanapanich, S. and Bose, N. K. (2002) High-resolution image formation from low-resolution frames using Delaunay triangulation. *IEEE Trans. Image Process.* **11**, 1427–1441.

[7] Ng, M. K., Sze, C. K. and Yung, S. P. (2004) Wavelet algorithms for deblurring models. *Int. J. Imag. Syst. Technol.*, **14**, 113–121.

[8] Ur, H. and Gross, D. (1992) Improved resolution from sub-pixel shifted pictures. *CVGIP: Graph. Models Image Process*, **54**, 181–186.

[9] Papoulis, A. (1977) Generalized sampling expansion. *IEEE Trans. Circuits Syst.*, **24**, 652–654.

[10] Yen, L. J. (1956) On non-uniform sampling of bandwidth limited signals. *IRE Trans. Circuits Theory*, **3**, 251–257.

[11] Irani, M. and Peleg, S. (1991) Improving resolution by image registration. *CVGIP: Graph. Models Image Process.*, **53**, 231–239.

[12] Stark, H. and Oskoui, P. (1989) High-resolution image recovery from image plane arrays, using convex projections. *J. Opt. Soc. Am. A: Opt. Image Sci. Vis.*, **6**, 1715–1726.

[13] Tekalp, A. M., Ozkan, M. K. and Sezan, M. I. (1992) High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. *Proc. IEEE Int. Conf. of Acoustics, Speech and Signal Processing*, San Francisco, CA, USA, March 23–26, 169–172, IEEE.

[14] Patti, A. J., Sezan, M. I. and Tekalp, A. M. (1994) High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur. *Proc. IEEE Int. Conf. on Image Processing*, Austin, TX, USA, pp. 343–347, IEEE, Los Alamitos, CA, USA.

[15] Patti, A. J., Sezan, M. I. and Tekalp, A. M. (1997) Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Trans. Image Process.*, **6**, 1064–1076.

[16] Tom, B. C. and Katsaggelos, A. K. (1994) Reconstruction of a high-resolution image from multiple-degraded misregistered low-resolution images. *Proc. SPIE—The International Society for Optical Engineering*, Chicago, IL, USA, pp. 971–981, Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, USA.

[17] Schultz, R. R. and Stevenson, R. L. (1996) Extraction of high-resolution frames from video sequences. *IEEE Trans. Image Process.*, **5**, 996–1011.

[18] Hardie, R. C., Tuinstra, T. R., Bognar, J., Barnard, K. J. and Armstrong, E. (1997) High-resolution image reconstruction from digital video with global and non-global scene motion. *Proc. IEEE Int. Conf. on Image Processing*, Santa Barbara, CA, USA, pp. 153–156, IEEE Comp Society, Los Alamitos, CA, USA.

[19] Ng, M. K. and Yip, A. M. (2001) A fast MAP algorithm for high-resolution image reconstruction with multisensors. *Multidimens. Syst. Signal Process.*, **12**, 143–164.

[20] Elad, M. and Feuer, A. (1997) Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Trans. Image Process.*, **6**, 1646–1658.

[21] Elad, M. and Feuer, A. (1999) Superresolution restoration of an image sequence: adaptive filtering approach. *IEEE Trans. Image Process.*, **8**, 387–395.

[22] Farsiu, S., Elad, M. and Milanfar, P. (2006) Multiframe demosaicing and super-resolution of color images. *IEEE Trans. Image Process.*, **15**, 141–159.

[23] Segall, C. A., Katsaggelos, A. K., Molina, R. and Mateos, J. (2004) Bayesian resolution enhancement of compressed video. *IEEE Trans. Image Process.*, **13**, 898–910.

[24] Segall, C. A., Molina, R. and Katsaggelos, A. K. (2003) High-resolution images from low-resolution compressed video. *IEEE Signal Process. Mag.*, **20**, 37–48.

[25] Farsiu, S., Robinson, D., Elad, M. and Milanfar, P. (2004) Dynamic demosaicing and color super-resolution of video sequences. *Proceeding of SPIE—The International Society for Optical Engineering*, Denver, CO, USA, pp. 169–178, International Society for Optical Engineering, Bellingham, WA, USA.

[26] Rajagopalan, A. N. and Phani Kiran, V. (2003) Motion-free superresolution and the role of relative blur. *J. Opt. Soc. Am. A: Opt. Image Sci. Vis.*, **20**, 2022–2032.

[27] Chaudhuri, S. and Manjunath, J. (2005) *Motion-Free Super-Resolution.* Springer-Verlag New York, USA.

[28] Latry, C. and Rouge, B. (2003) Super resolution: quincunx sampling and fusion processing. *Proc. Int. Geoscience and Remote Sensing Symposium (IGARSS)*, Toulouse, France, pp. 315–317, Institute of Electrical and Electronics Engineers Inc.

[29] Wang, Z. J., Ziou, D., Armenakis, C., Li, D. and Li, Q. Q. (2005) A comparative analysis of image fusion methods. *IEEE Trans. Geosci. Remote Sens.*, **43**, 1391–1402.

[30] Ranchin, T., Aiazzi, B., Alparone, L., Baronti, S. and Wald, L. (2003) Image fusion—the ARSIS concept and some successful implementation schemes. *ISPRS J. Photogramm. Remote Sens.*, **58**, 4–18.

[31] Gonzalez-Audicana, M., Otazu, X., Fors, O. and Alvarez-Mozos, J. (2006) A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1683–1691.

[32] Joshi, M. V., Bruzzone, L. and Chaudhuri, S. (2006) A model-based approach to multiresolution fusion in remotely sensed images. *IEEE Trans. Geosci. Remote Sens.*, **44**, 2549–2562.

[33] Park, J. H. and Kang, M. G. (2004) Spatially adaptive multi-resolution multispectral image fusion. *Int. J. Remote Sens.*, **25**, 5491–5508.

[34] Vega, M., Mateos, J., Molina, R. and Katsaggelos, A. K. (2005) Super resolution and pansharpening of multispectral images. *Proc. First Int. Conf. on Super Resoltuion Imaging*, Hong Kong. Multimensional Systems and Signal Processing, **18**, 2–3 September, 2007.

[35] http://terra.nasa.gov/About/MODIS/index.php.

[36] Park, S. C., Park, M. K. and Kang, M. G. (2003) Super-resolution image reconstruction: a technical overview. *IEEE Signal Process. Mag.*, **20**, 21–36.

[37] Conel, J. E. (1985) Calibration of data using ground-based spectral reflectance measurements. *Proc. Airborne Imaging Spectrometer Data Analysis Workshop*, CA, USA, April, pp. 184–185, JPL Publication 85-41, Jet Propulsion Laboratory, Pasadena.

[38] Roberts, D., Yamaguchi, Y. and Lyon, R. (1985) Calibration of airborne imaging spectrometer data to percent reflectance using field spectral measurements. *Proc. Nineteenth Int. Symp. on Remote Sensing of Environment*, MI, USA, October 21–25, pp. 679–688, International Committee on Remote Sensing of Environment.

[39] Kumar, S., Biswas, M. and Nguyen, T. (2004) Global motion estimation in frequency and spatial domain. *Proc. IEEE Int. Conf. On Acoustics Speech And Signal Processing*, pp. 333–336, IEEE.

[40] Fu, H., Ng, M. K., Nikolova, M. and Barlow, J. L. (2006) Efficient minimization methods of mixed L1–L1 and L2–L1 norms for image restoration. *SIAM J. Sci. Comput.*, **27**, 1881–1902.

[41] Alvarez, L. D., Mateos, J., Molina, R. and Katsaggelos, A. K. (2004) High-resolution images from compressed low-resolution video: motion estimation and observable pixels. *Int. J. Imag. Syst. Technol.*, **14**, 58–66.

[42] Farsiu, S., Robinson, D., Elad, M. and Milanfar, P. (2003) Robust shift and add approach to super-resolution. *Proc. SPIE—The International Society for Optical Engineering*, San Diego, CA, USA, pp. 121–130, The International Society for Optical Engineering, SPIE.

[43] Farsiu, S., Robinson, M. D., Elad, M. and Milanfar, P. (2004) Fast and robust multiframe super resolution. *IEEE Trans. Image Process.*, **13**, 1327–1344.

[44] Nikolova, M. (2004) A variational approach to remove outliers and impulse noise. *J. Math. Imag. Vis.*, **20**, 99–120.

[45] Capel, D. and Zisserman, A. (2003) Computer vision applied to super resolution. *IEEE Signal Process. Mag.*, **20**, 75–86.

[46] Zomet, A. and Peleg, S. (2000) Efficient super-resolution and applications to mosaics. *Proc. Int. Conf. on Pattern Recognition*, September, pp. 579–583, The International Association for Pattern Recognition.

[47] Schultz, R. R. and Stevenson, R. L. (1994) A Bayesian approach to image expansion for improved definition. *IEEE Trans. Image Process.*, **3**, 233–242.

[48] Zhou, W. and Bovik, A. C. (2002) A universal image quality index. *IEEE Signal Process. Lett.*, **9**, 81–84.

[49] Kumar, A. S., Kumar, A. S. K. and Navalgund, R. R. (2006) Selection of IRS-P6 LISS-4 MO mode band for producing band-sharpened multispectral imagery. *IEEE Geosci. Remote Sens. Lett.*, **3**, 32–35.