Article

Multi-Frame Super-Resolution of Gaofen-4 Remote Sensing Images

Jieping Xu * [ID]**, Yonghui Liang, Jin Liu and Zongfu Huang**

National University of Defense Technology, College of Opto-Electronic Science and Engineering, Deya Road 109, Changsha 410073, China; yonghuiliang@sina.com (Y.L.); liujin1344@sina.com (J.L.); hzf_huang@126.com (Z.H.)
*   Correspondence: xiaotianxia@nudt.edu.cn; Tel.: +86-158-7317-1748

Abstract: Gaofen-4 is China's first geosynchronous orbit high-definition optical imaging satellite with extremely high temporal resolution. The features of staring imaging and high temporal resolution enable the super-resolution of multiple images of the same scene. In this paper, we propose a super-resolution (SR) technique to reconstruct a higher-resolution image from multiple low-resolution (LR) satellite images. The method first performs image registration in both the spatial and range domains. Then the point spread function (PSF) of LR images is parameterized by a Gaussian function and estimated by a blind deconvolution algorithm based on the maximum a posteriori (MAP). Finally, the high-resolution (HR) image is reconstructed by a MAP-based SR algorithm. The MAP cost function includes a data fidelity term and a regularized term. The data fidelity term is in the L_2 norm, and the regularized term employs the Huber-Markov prior which can reduce the noise and artifacts while preserving the image edges. Experiments with real Gaofen-4 images show that the reconstructed images are sharper and contain more details than Google Earth ones.

Keywords: super-resolution; blind deconvolution; remote sensing; staring imaging; MAP

1. Introduction

Images with higher resolution are required in most electronic imaging applications, such as remote sensing, medical diagnostics, and video surveillance. During the past decades, considerable advances have been realized in imaging systems. However, the quality of images is still limited by the cost and the manufacturing technology [1]. Super-resolution (SR) is a promising digital image processing technique to construct high-resolution (HR) images from several blurred low-resolution (LR) images.

The basic idea of SR is that the LR images of the same scene contain different information because of relative sub-pixel shifts, thus, an HR image with higher spatial information can be reconstructed by image fusion. Sub-pixel motion can occur due to movement of local objects or vibrating of the imaging system, or even controlled micro-scanning [2,3]. Numerous SR algorithms have been proposed since the concept of SR was introduced by Tsai and Huang [4] in the frequency domain. However, most of researchers nowadays address the problem mainly in the spatial domain, because it is more flexible to model all kinds of image degradations [5]. Ur and Gross [6] proposed a non-uniform interpolation of multiple spatially-shifted LR images based on the generalized multichannel sampling theorem, followed by a deblurring process. Stark et al. [7] were the first to suggest an SR algorithm based on the projection onto convex sets approach (POCS). Patti et al. [8] extended the method to handle arbitrary sampling lattices and non-zero aperture time. The advantage of the POCS method is that it allows a convenient inclusion of prior knowledge into the reconstruction process. Irani and Peleg [9] estimated the HR image by back projecting the error between re-projected LR images and the observed LR images. For the ill-posed nature of the SR problem, regularization methods are widely used in SR

reconstruction. Schultz and Stevenson [10] reconstructed a single HR video frame from a short image sequence using the maximum a posteriori (MAP) approach with a discontinuity-preserving image prior. They used a block-matching method to estimate the sub-pixel displacement vectors between LR images. A MAP framework for jointly estimating image registration parameters and the HR image was presented by Hardie et al. [11]. The sub-pixel displacement vectors were iteratively updated along with the HR image in a cyclic coordinate-descent optimization procedure. Zomet et al. [12] proposed a robust SR algorithm in which a median estimator was used as the gradient of the L_2 norm cost function. Farsiu et al. [13] used a cost function in the L_1 norm rather than modified the gradient directly to further increase the robustness.

For remote sensing applications, researchers mainly focused on single-frame SR based on learning or sparse representation for lack of multiple images of the same scene [14–17]. Nevertheless, some researchers have made their contributions. Merino and Núñez [18] proposed a method named Super-Resolution Variable-Pixel Linear Reconstruction (SRVPLR) to reconstruct an HR image from multiple LR images acquired over a long period. This algorithm was derived from drizzle [19] which was designed to work with astronomical dithered under-sampled images. A prior histogram matching on LR images and high geometrical corrections were introduced to make it compatible with remotely-sensed images. Shen et al. [20] proposed an SR method to MODIS remote sensing images, where image registration in the range and spatial domains were performed before image reconstruction. Li et al. [21] proposed a MAP-based SR algorithm with a universal hidden Markov tree model, and tested it with Landsat7 panchromatic images captured on different days. Fan et al. [22] proposed a POCS-based SR algorithm with a point spread function (PSF) estimated by a slant knife-edge method. The method was tested on Airborne Digital Sensor 40 (ADS40) three-line array images and the overlapped areas of two adjacent Gaofen-2 images.

This paper presents a technique that performs multi-frame super-resolution of Gaofen-4 satellite images. Gaofen-4 is a staring imaging satellite with high temporal resolution, and its high orbit limits spatial resolution, therefore, multi-frame SR is possible and essential. First, we build the observation model for the Gaofen-4 satellite, then we propose the pipeline of our MAP-based SR algorithm. Our algorithm consists of three steps: image registration, PSF estimation, and image fusion. Image registration is performed both in the range and spatial domains, and the Gaussian PSFs of LR images are estimated via a modified MAP-based blind deconvolution algorithm. Finally, The HR image is estimated by an image fusion method based on the MAP method with a Huber image prior, while the registration parameters are further refined along with the HR image.

The rest of this paper is organized as follows: Section 2 addresses the SR problem and details our algorithm; Section 3 presents the experimental results with synthetic and real data; Section 4 concludes the paper.

2. Materials and Methods

2.1. Observation Model

To analyze the SR problem comprehensively, an image observation model must be formulated first. For notational simplicity, the image in our method is represented as a lexicographically-ordered vector. Let us consider the desired HR image **o** of size $LM \times LN$, where L denotes the downsampling factor in the observation model, and M and N are the number of rows and columns of the LR images, respectively. Generally, the observation model-related HR image and the LR images is given by [1]:

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

where $\mathbf{M}_t$ is the subpixel motion matrix of size $L^2 MN \times L^2 MN$; $\mathbf{B}_t$ represents the blur matrix of size $L^2 MN \times L^2 MN$; $\mathbf{D}_t$ denotes the downsampling matrix of size $MN \times L^2 MN$; $\mathbf{n}_t$ is the lexicographically-ordered additive Gaussian noise vector with zero mean; and $\mathbf{y}_t$ represents the lexicographically-ordered LR image vector of size $MN \times 1$.

The observation model represented by Equation (1) assumes an illumination invariant environment. However, this assumption may not be valid due to the variation of light conditions over time. The modified observation model is given by:

$$\mathbf{y}_t = \lambda_{1,t}\mathbf{D}_t\mathbf{B}_t\mathbf{M}_t\mathbf{o} + \lambda_{0,t}\mathbf{I} + \mathbf{n}_t, \quad \text{for } t = 1, 2, \ldots, T, \qquad (2)$$

where $\lambda1, t$ and $\lambda0, t$ denote the contrast gain and illumination offset, respectively, and $\mathbf{I}$ is a unit vector of size $MN \times 1$. The motion matrix $\mathbf{M}t$ could vary in time, while the downsampling matrix $\mathbf{D}t$ and blur matrix $\mathbf{B}t$ remain constant over time for most situations. Therefore, $\mathbf{D}$ and $\mathbf{B}$ are used instead of $\mathbf{D}t$ and $\mathbf{B}t$, respectively. In the following of this section, we build the specific model for the Gaofen-4 satellite.

Remote sensing sensors can be modeled by rational polynomial coefficients (RPC) geometrically. However, for registering images taken by the Gaofen-4 staring imaging satellite, the scene can be effectively modeled as a planar region for its high orbit when the region of interest (ROI) is not too large and, hence, a simple affine camera model is sufficient to describe the image geometry. The affine transformation contains six parameters and can be written as:

$$T = \begin{bmatrix} a_1 & a_2 & a_0 \\ b_1 & b_2 & b_0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad (3)$$

where a_0 and b_0 denote translation along x and y directions respectively, and a_1, a_2, b_1 and b_2 are the parameters representing rotation, scale and deformation. Using matrix notation, the entire affine motion parameters are represented by $\theta = [a_0, a_1, a_2, b_0, b_1, b_2]^T$.

We assume a shift-invariant, convolutional blur, which can be described by the PSF in the spatial domain or the modulation transfer function (MTF) in the frequency domain. The remote sensing images can be blurred by the atmosphere, the optical system, relative motion, and pixel integration. The overall PSF is the successive convolution of all of the individual terms:

$$\text{PSF} = \text{PSF}_{\text{atm}} \otimes \text{PSF}_{\text{opt}} \otimes \text{PSF}_{\text{mot}} \otimes \text{PSF}_{\text{sen}}, \qquad (4)$$

which respectively summarize the atmospheric, optical system, motion and sensor contributions.

The atmosphere contributes through turbulence, light scattering, and attenuation. The spatial resolution of Gaofen-4 in the visible and near-infrared band is 50 m, and the contribution of turbulence at this spatial resolution is insignificant. The scattering and absorption of energy by the aerosols affects all spatial frequencies, therefore, causing edges in the image to be blurred. The aerosol MTF can be approximated by a Gaussian form $\exp\left[-k(r/r_c)^2\right]$ for r below the cut-off frequency r_c [23].

A limited aperture size of the optical system causes a diffraction blur, and the corresponding PSF is given by the Airy disk. Other optical aberrations, such as defocus and spherical aberrations, also contribute to the optical system, which can be seen as perturbations to the Airy disk. A general optical system can be expressed as a Gaussian function:

$$\text{PSF}_{\text{opt}}(x, y) = \frac{1}{2\pi ab} \exp\left(-\frac{x^2}{2a^2}\right) exp\left(-\frac{y^2}{2b^2}\right), \qquad (5)$$

where a and b denote the deviations of the two-dimensional Gaussian function along the x and y direction, respectively.

The motion-related blur results from shift and vibrations. For pushbroom and whiskbroom sensors, the PSF can be characterized by a rectangular function along the scanning direction. We assume the motion blur of the Gaofen-4 satellite can be ignored for it is a staring imager equipped with focal-plane arrays.

The finite size of the sensor pixels results in sensor blur which presents the integration of irradiance over pixels. The PSF is given by:

$$\text{PSF}_{\text{det}}(x, y) = \text{rect}(x/w)\text{rect}(y/w), \tag{6}$$

where w is the size of a pixel.

All PSF terms can are gathered into a single Gaussian term, which is given by:

$$\text{PSF}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right], \tag{7}$$

where we further assume the deviations along the x and y direction are equal and denoted by σ.

2.2. Image Registration

The success of the SR technique relies on the subpixel motion between images, and the image registration accuracy affects the quality of the reconstructed HR image significantly. Geometric registration and photometric registration are interactive. Most photometric registration approaches require accurate geometric registration, and geometric registration may fail when images are not aligned photometrically. We first perform geometric registration coarsely using speeded up robust features (SURF) [24], which is not invariant to image contrast gain and illumination bias. Then photometric and geometric registration are jointly performed using alternative minimization.

The relationship between the target image $f_T(x, y)$ and the reference image $f_R(x, y)$ is given by the following model:

$$f_R(\mathbf{x}) = P(G(f_T(\mathbf{x}; \boldsymbol{\theta})); \boldsymbol{\lambda}) + n(\mathbf{x}), \tag{8}$$

where $\mathbf{x}$ denotes the spatial coordinates; $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ are geometric and photometric parameters in matrix form, respectively; G and P are geometric and photometric models, respectively; and $n(\mathbf{x})$ is the additive Gaussian noise. In this paper, we focus on the linear photometric model and affine motion. Thus, Equation (8) becomes:

$$f_R(x, y) = \lambda_1 f_T(a_1 x + a_2 y + a_0, b_1 x + b_2 y + b_0) + \lambda_0 + n(x, y), \tag{9}$$

where λ_1 and λ_0 denote the contrast gain and illumination bias, respectively. The goal of the image registration is to estimate these parameters via minimization of an objective function in the least square sense:

$$\{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\lambda}}\} = \text{argmin}F(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \text{argmin}\|f_R(x, y) - g(x, y)\|^2_{(x,y)\in\Omega}, \tag{10}$$

where $g(x, y)$ is the transformation of the target image which is expressed as:

$$g(x, y) = \lambda_1 f_T(a_1 x + a_2 y + a_0, b_1 x + b_2 y + b_0) + \lambda_0, \tag{11}$$

and Ω represents the region of inliers. To detect inliers, we use a geometric criterion and a photometric criterion. Geometric criterion requires that the pixels of the warped image should have the same geometric field with the reference image, i.e.,

$$-N/2 + 1 \leq a_1 x + a_2 y + a_0 \leq N/2, \tag{12}$$

$$-M/2 + 1 \leq b_1 x + b_2 y + b_0 \leq M/2, \tag{13}$$

where M and N are the number of rows and columns of the reference image, respectively. The photometric criterion requires that the intensity difference between the reference image and the registered target image should be small enough, that is to say:

$$|f_R(x, y) - g(x, y)| \leq d, \tag{14}$$

where d is the intensity threshold and is set to 8.

The solution of Equation (10) can be obtained by an optimization algorithm, such as the conjugate gradient descent algorithm or the Gaussian-Newton method. However, it is easy to fall into a local optimal solution when the misalignment between two images is large. Thus, SURF is employed to estimate the geometric parameters coarsely, then an iterative updating strategy is used to estimate the photometric and geometric parameters more precisely. Specifically, at each iteration, the affine motion parameters are firstly updated by a new variation of the Marquardt-Levenberg algorithm, which was proposed by Thévenaz et al. [25], given the photometric parameters are fixed, then the affine motion parameters are kept fixed, and the photometric parameters are updated by the conjugate gradient method. The gradients of $F(\boldsymbol{\theta}, \boldsymbol{\lambda})$ with respect to λ_1 and λ_0 are given by:

$$\frac{\partial F(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \lambda_1} = \sum_{(x,y) \in \Omega} -2[f_R(x,y) - \lambda_1 f_W(x,y) - \lambda_0] f_W(x,y), \tag{15}$$

$$\frac{\partial F(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \lambda_0} = \sum_{(x,y) \in \Omega} -2[f_R(x,y) - \lambda_1 f_W(x,y) - \lambda_0], \tag{16}$$

where $f_W(x,y)$ is the target image after geometric registration, but without photometric registration.

The image registration algorithm flow is shown in Algorithm 1. In the feature-matching step, the random sample consensus (RANSAC) method [26] is used to eliminate matching outliers. The stop criterion of this image registration algorithm is that values of the estimated parameters do not change anymore, which can be expressed by:

$$\sum \left| \frac{\hat{\boldsymbol{\theta}}_{k+1} - \hat{\boldsymbol{\theta}}_k}{\hat{\boldsymbol{\theta}}_k} \right|^2 + \sum \left| \frac{\hat{\boldsymbol{\lambda}}_{k+1} - \hat{\boldsymbol{\lambda}}_k}{\hat{\boldsymbol{\lambda}}_k} \right|^2 < \varepsilon, \tag{17}$$

where $\sum$ means the sum of all elements; $\hat{\boldsymbol{\theta}}_k$ represents the estimated geometric parameters after k times cycles, and so on. ε is the precision threshold; here, 10^{-4} is used.

Algorithm 1. Image registration algorithm.

1　　***Input:*** LR images $\{\mathbf{y}_t\}$.
2　　***Output:*** Geometric and photometric parameters.
3　　Select one image $\mathbf{y}_k$ as the reference image, and detect its SURF features.
4　　***for*** $l = 1, 2, \dots, T (l \neq k)$, ***do***
5　　　　Detect SURF features of the LR image $\mathbf{y}_l$.
6　　　　Match features of $\mathbf{y}_l$ and $\mathbf{y}_k$, then estimate geometric parameters $\boldsymbol{\theta}$ based on RANSAC.
7　　　　***do***
8　　　　　　Estimate photometric parameters $\boldsymbol{\lambda}$ using conjugate gradient method for fixed geometric
　　　　　　　parameters $\boldsymbol{\theta}$.
9　　　　　　Estimate geometric parameters $\boldsymbol{\theta}$ using the Marquardt-Levenberg method for fixed photometric
　　　　　　　parameters $\boldsymbol{\lambda}$.
10　　　　***until*** $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ do not change anymore.
11　　***endfor***

2.3. PSF Estimation

The PSF of remote sensing images can be measured after being launched, or estimated via sharp edge prediction. Sometimes, appropriate sharp edges do not exist, so the PSF estimation method should be employed. We propose a modified MAP-based multi-frame blind deconvolution algorithm to estimate the PSF, which was originally proposed by Matson et al. [27]. The forward model used in this algorithm is:

$$g_t(\mathbf{x}) = o(\mathbf{x}) * h_t(\mathbf{x}) + n_t(\mathbf{x}), \quad \text{for } t = 1, 2, \dots, T, \tag{18}$$

where $g_t(\mathbf{x})$ is the t-th measurement image; $o(\mathbf{x})$ is the true object; $h_t(\mathbf{x})$ is the t-th PSF; $n_t(\mathbf{x})$ is the t-th additive Gaussian noise; and $*$ denotes convolution. The cost function to generate estimates of the PSFs is given by:

$$U\left[\hat{\mathbf{o}}(\mathbf{x}), \{\hat{h}_t(\mathbf{x})\}\right] = \sum_t \frac{\|g_t(\mathbf{x}) - \hat{o}(\mathbf{x}) * \hat{h}_t(\mathbf{x})\|^2}{\sigma_{n,t}^2} + \gamma \hat{o}(\mathbf{x})^2, \tag{19}$$

where γ is the regularized parameter and $\sigma_{n,t}^2$ is the image noise variance of the t-th frame. The minimization of the cost function is performed via the conjugate gradient method in a cyclic coordinate-descent optimization procedure. Readers can refer to [27] for more details of this blind deconvolution algorithm.

To make the above algorithm compatible with our application, we use the bicubic upsampled registered LR images as the measurement images. In order to save computational resources, no more than four measurement images are used. The estimated PSFs are reparameterized by Gaussian PSFs, which are given by Equation (7). The mean deviation of these PSFs is used in image fusion. We use the mean deviation because we assume there are no significant differences between the PSFs.

2.4. Image Fusion

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

$$\hat{\mathbf{x}}_{\mathrm{MAP}} = \mathrm{argmax}\, p(\{\mathbf{o}|\{\mathbf{y}_t\}) = \mathrm{argmax} \sum_t \frac{p(\mathbf{y}_t|\mathbf{o})p(\mathbf{o})}{p(\mathbf{y}_t)}. \tag{20}$$

Under the assumption of additive Gaussian noise, the likelihood term $p(\mathbf{y}_t|\mathbf{o})$ can be expressed as:

$$p(\mathbf{y}_t|\mathbf{o}) = \frac{1}{\sqrt{2\pi}\sigma_{n,t}} \exp\left[-\frac{\|\mathbf{y}_t - \lambda_{1,t}\mathbf{DBM}_t\mathbf{o} - \lambda_{0,t}\|^2}{2\sigma_{n,t}^2}\right]. \tag{21}$$

The term $p(\mathbf{o})$ represents the probability density of the HR image, which restricts the solution space and introduces prior knowledge to the reconstruction. In particular, we employ a Huber Markov random field (HMRF) as the image prior. HMRF encourages piecewise smoothness, and can preserve image edges as well. The expression of $p(\mathbf{o})$ is given by:

$$p(\mathbf{o}) = \frac{1}{Z} \exp\left[-\frac{v}{2}\rho(\mathbf{So}, \alpha)\right], \tag{22}$$

where $\mathbf{S}$ denotes a gradient operator; v denotes the prior strength; α is a parameter of the Huber function; and Z is a normalization factor. The Huber function ρ is defined as:

$$\rho(\mathbf{z}, \alpha) = \sum_i \rho(z_i, \alpha), \tag{23}$$

$$\rho(z_i, \alpha) = \begin{cases} z_i^2, & |z_i| < \alpha \\ 2\alpha|z_i|, & \text{otherwise} \end{cases}. \tag{24}$$

where z_i is the i-th element of the vector $\mathbf{z}$.

By substituting Equations (21)–(24) to Equation (20) and neglecting the constant term, the MAP estimator can be converted to a minimizing cost function, which is given by:

$$J(\mathbf{o}, \{\boldsymbol{\theta}_t\}, \{\boldsymbol{\lambda}_t\}) = \sum_{t=1}^{T} \frac{\|\mathbf{y}_t - \lambda_{1,t}\mathbf{DBM}_t\mathbf{o} - \lambda_{0,t}\|^2}{\sigma_{n,t}^2} + v \sum_{z \in \mathbf{So}} \rho(z, \alpha), \tag{25}$$

The accuracy of the geometric and photometric registration is limited by the low quality of the LR images. Thus, an iterative update of the registration parameters, along with the HR image, is required. For the discrepancy of convergence rate of **x**, $\{\theta_t\}$, and $\{\lambda_t\}$, a cyclic coordinate-descent optimization procedure is adopted. At each iteration, the HR image is updated by minimizing Equation (25) with respect to **o** using a scaled conjugate gradient optimization method, given the registration parameters are fixed; then the registration parameters are updated by minimizing the cost function with respect to the registration parameters, given the HR image estimate is fixed. This minimization is completed by an image registration process, which is the same with the registration method proposed in Section 2.2. However, feature-based registration is not needed anymore, and the current HR image estimate is used as the reference image instead. The algorithm stops when the value of the function does not change. The proposed algorithm as a whole system is illustrated in Figure 1. The algorithm consists of three stages: image registration, PSF estimation and image fusion.
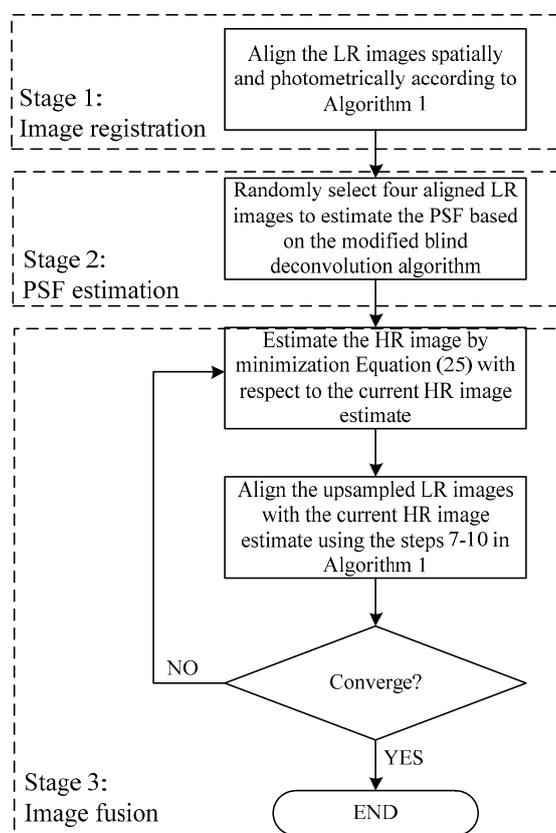

Figure 1. Diagram of the proposed algorithm.

3. Results and Discussion

3.1. Simulation Images

We first test our algorithm using synthetic images. The quantitative evaluation of the algorithm is achieved by computing the root mean square error (RMSE) and Structural Similarity (SSIM). RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{m}\|\hat{\mathbf{o}} - \mathbf{o}\|^2},\tag{26}$$

where **o** and **ô** are the original image and the reconstructed HR image respectively, and m is the total number of pixels of the original image. The SSIM index is defined as: [28]

$$SSIM(\mathbf{o}, \hat{\mathbf{o}}) = \frac{(2\mu_{\mathbf{o}}\mu_{\hat{\mathbf{o}}} + c_1)(2\sigma_{\mathbf{o}\hat{\mathbf{o}}} + c_2)}{(\mu_{\mathbf{o}}^2 + \mu_{\hat{\mathbf{o}}}^2 + c_1)(\sigma_{\mathbf{o}}^2 + \sigma_{\hat{\mathbf{o}}}^2 + c_2)} \tag{27}$$

where $\mu_{\mathbf{o}}$ and $\mu_{\hat{\mathbf{o}}}$ are the means and $\sigma_{\mathbf{o}}$ and $\sigma_{\hat{\mathbf{o}}}$ are the standard deviations of the original and the reconstructed HR image, $\sigma_{\mathbf{o}\hat{\mathbf{o}}}$ is the covariance of **o** and **ô**, $c_1 = 60$ and $c_2 = 180$.

A sequence of LR images are generated by the following procedure:

(i)   Apply random affine transformations generated by their SVD decomposition $R(\beta)R(-\phi)\Lambda R(\phi)$ with translations added, where $R(\beta)$ and $R(\phi)$ are rotation matrices with β and ϕ uniformly distributed in the range of $(-5°, 5°)$; Λ is a diagonal matrix with diagonal elements uniformly distributed in the range of (0.95, 1.05); and the additive horizontal and vertical displacements are uniformly distributed in the range of (–5, 5) pixels.

(ii) Blur the images with a Gaussian point spread function (PSF) with the standard deviation equal to 1.

(iii) Clip images to 0.8 times of the size both horizontally and vertically, so that only overlapped regions are left.

(iv) Downsample the images by a factor of L.

(v)   Apply photometric parameters with uniformly-distributing contrast gain in the range of (0.9, 1.1) and bias in the range of $(-0.1, 0.1) \times \max(\mathbf{o})$.

(vi) Make the SNR of the images become 30 dB by adding white Gaussian noise with zero mean.

A QuickBird 2 remote sensing image with a size of 256 × 256 is used as the origin HR image, and the downsampling factor is equal to 2. We use six LR images to reconstruct an HR image. The Huber function parameter α is set to 0.04, and the prior strength parameter v is set to 10. The estimated deviation of the total PSF is 1.12. Our method is compared with the bicubic interpolation, the iterative back projection (IBP) approach proposed by Irani and Peleg [9] and Farsiu's robust SR algorithm [13]. Note that the original IBP and Farsiu's algorithm do not consider the variations of the illumination in the image registration stage, thus we use our image registration algorithm instead. Figure 2 shows the original image and the reconstructed images by different methods. Compared with the bicubic upsampled image, the HR images reconstructed by all SR methods are sharper and contain more details. The reconstructed images are further compared quantitatively in RMSE and SSIM, as is listed in Table 1. The performance of our method and IBP is close, and better than the bicubic interpolation and Farsiu's algorithm. Additionally, without photometric registration, many artifacts appear in the reconstructed HR image, as is shown in Figure 2d.

Table 1. The RMSE and SSIM between the original and reconstructed images obtained from bicubic interpolation, IBP, Farsiu's method and the proposed method.

Methods	Bicubic	Our Method	IBP	Farsiu
RMSE	16.44	13.36	13.94	15.90
SSIM	0.917	0.941	0.940	0.928

3.2. Real Gaofen-4 Images

The second experiment is carried out with ten 5-band Gaofen-4 images without ortho-rectification. The images were all captured on 3 March 2017, from 11:10:20 to 11:20:21. The original size of full-view images is 10,240 × 10,240, and we chip the region of interest (ROI) with a size of 256 × 256. This region is an offshore aquaculture area, and the images captured at 11:10:20 and 11:20:21 are shown in Figure 3.
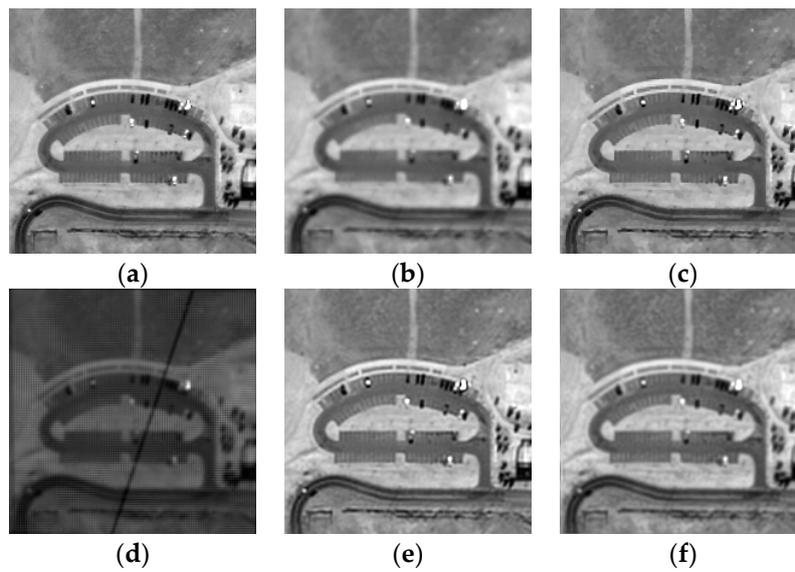
Figure 2. Simulation results of different methods. (**a**) Origin HR image; (**b**) the bicubic upsampled reference image; (**c**) the proposed method; (**d**) the proposed method without photometric registration; (**e**) IBP; (**f**) Farsiu's algorithm.
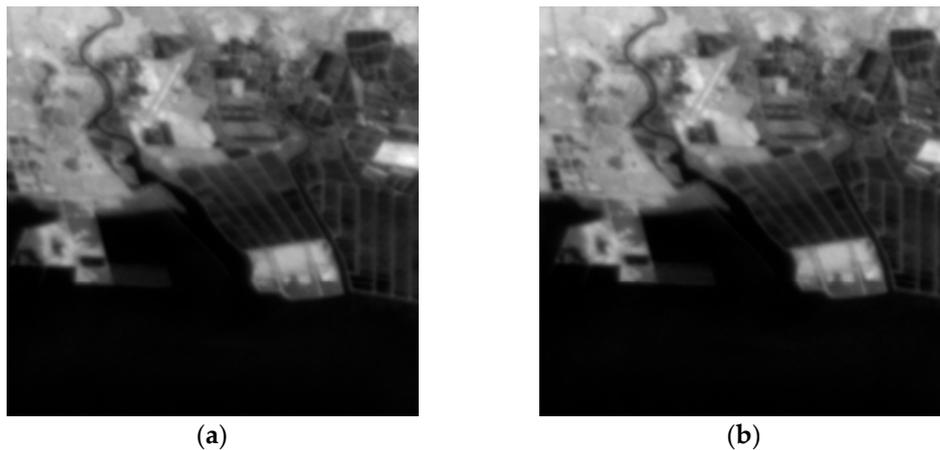

Figure 3. Two LR images: (**a**) captured at 11:10:20; and (**b**) captured at 11:20:21.

The image captured at 11:10:20 is used as the reference image, and the other images are aligned to it. Although the images were captured over a short period, photometric registration is needed. The criterion used to evaluate the registration accuracy is signal to noise ratio (SNR), which is defined as:

$$\text{SNR} = 10\lg\frac{\|f_R(\mathbf{x})\|^2}{\|f_R(\mathbf{x}) - g(\mathbf{x})\|^2}, \tag{28}$$

where $g(\mathbf{x})$ is the target image after registration, and $f_R(\mathbf{x})$ is the reference image. The registration results are listed in Table 2. The images are numbered according to their capture time and the first LR image is used as the reference image. All images have similar registration accuracy with both geometric and photometric registration while, without photometric registration, the SNRs are lower, and become lower and lower as time progresses.

Table 2. Registration results. The row of SNR 1 lists the registration results with both geometric and photometric registration, and the row of SNR 2 lists the registration results with only geometric registration.

Image Number	2	3	4	5	6	7	8	9	10
SNR 1	40.5	39.7	41.0	40.3	40.6	40.1	39.7	38.5	39.1
SNR 2	39.9	38.1	38.0	36.4	35.5	34.4	33.4	32.2	31.7

The Huber function parameter α is set to 0.04, and the prior strength parameter v is set to 10. The estimated deviation of the total PSF is 1.21. The experimental results are shown in Figure 4.
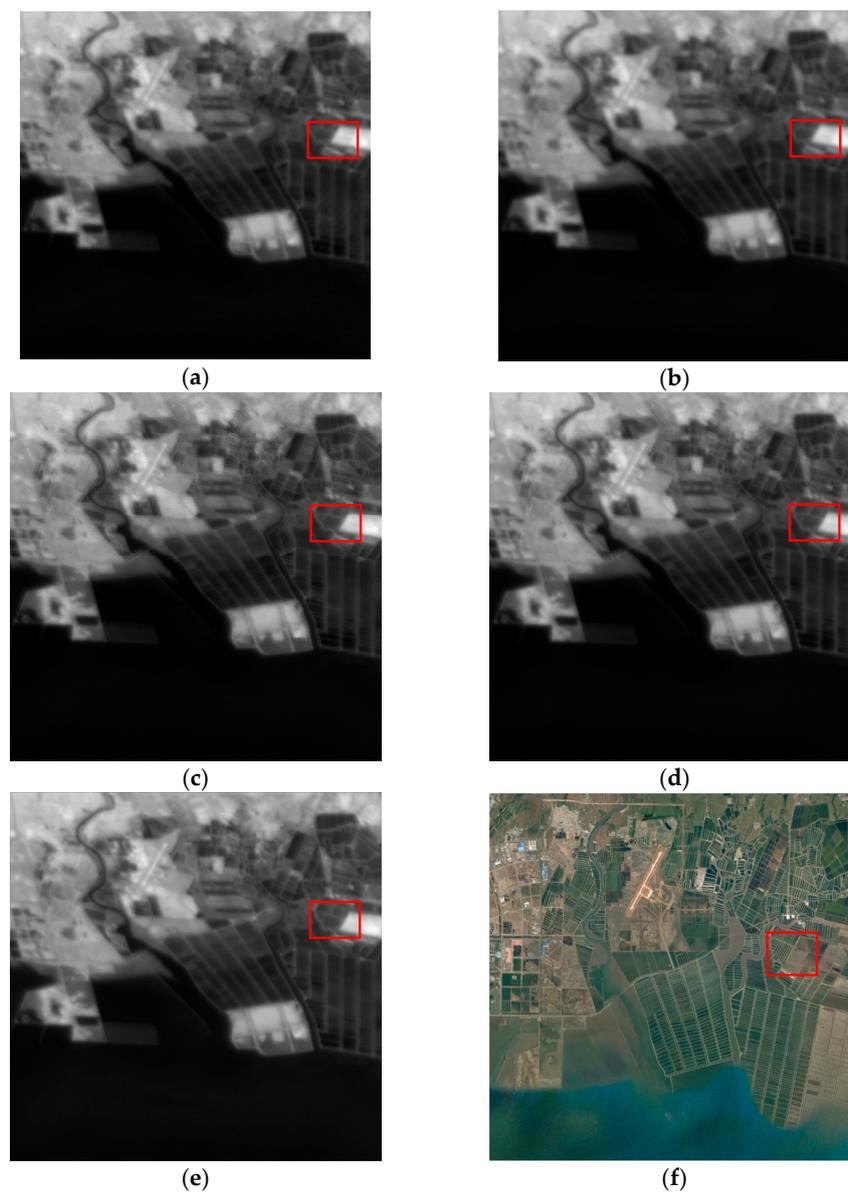
Figure 4. Experimental results of an offshore aquaculture area. (**a**) The bicubic upsampled reference image; (**b**) the added image after registration; (**c**) the HR image reconstructed by the proposed method; (**d**) the HR image reconstructed by IBP; (**e**) the HR image reconstructed by Farsiu's algorithm; and (**f**) the image chipped from Google Earth.

To validate our reconstructed HR image, we chip an image of the same zone from Google Earth, which was captured on 16 September 2015. The reconstructed HR image is clearer and contains more details which corresponds to Google Earth. To see more clearly, the regions denoted by the rectangles in Figure 4 are shown in Figure 5 without zooming. It is clear that the HR image reconstructed by our method is the sharpest.
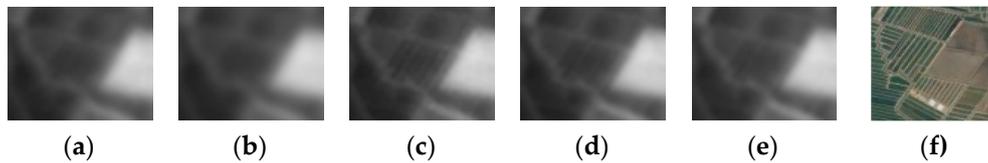

| (a) | (b) | (c) | (d) | (e) | (f) |

Figure 5. The regions denoted by the rectangles in Figure 5. (**a**) The bicubic upsampled reference image; (**b**) the added image after registration; the HR images reconstructed by (**c**) the proposed method, (**d**) IBP and (**e**) Farsiu's algorithm; (**f**) the image chipped from Google Earth.

Quantitatively evaluating the performance of the proposed algorithm is not easy in real image applications, for there are no true original images. A commonly used strategy to solve the problem is to treat the image of an inferior resolution level such as the panchromatic image as the true original image. But the strategy is not applicable for Gaofen-4 images, because the panchromatic and multi-spectral images of Gaofen-4 have the same spatial resolution. Theoretically, the reconstructed HR image should be sharper than the bicubic upsampled image. Thus we assume the upsampled image is the blurred version of the reconstructed HR image and estimate the PSF between them. The PSF estimation is completed by the PSF estimation algorithm proposed by Matson et al. which has been described in Section 2.3. Then the PSF is fitted to the Gaussian function. A larger deviation of the Gaussian function means a sharper reconstructed HR image. The estimated PSFs for IBP, Farsiu's algorithm and our MAP-based SR algorithm are shown in Figure 6. The deviations of the fitted PSFs are 1.0, 0.9 and 1.0 respectively, which means our method has a close performance with IBP, and is better than Farsiu's algorithm.
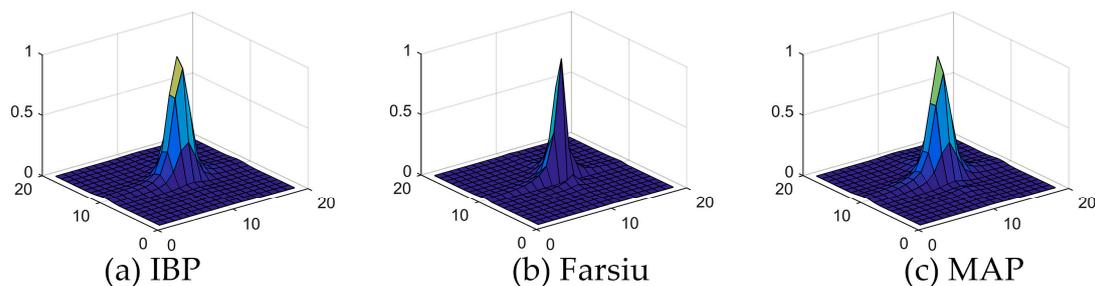

Figure 6. The PSFs between the upsampled reference image and the reconstructed HR images of the offshore aquaculture area.

We extract another hilly region of interest with a size of 256×256 to further test our algorithm. The parameters are set the same as the previous experiment. The estimated deviation of the total PSF is 1.19, which is almost same with the previous experiment. The experimental results are shown in Figure 7. The Google Earth image was captured on 20 April 2017. To see more clearly, the regions denoted by the rectangles in Figure 7 are shown in Figure 8. Compared with other methods, our method results in the sharpest edges.
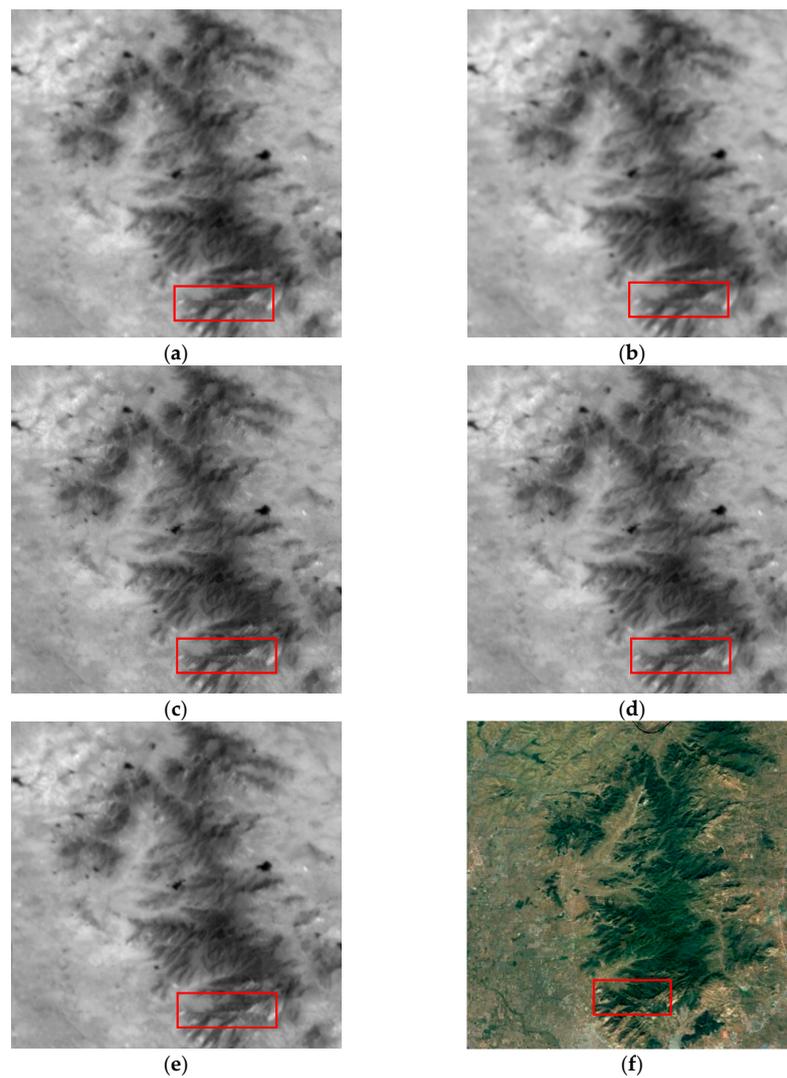
Figure 7. Experimental results of the hilly area. (**a**) The bicubic upsampled reference image; (**b**) the added image after registration; the HR images reconstructed by (**c**) the proposed method, (**d**) IBP and (**e**) Farsiu's algorithm; (**f**) the image chipped from Google Earth.
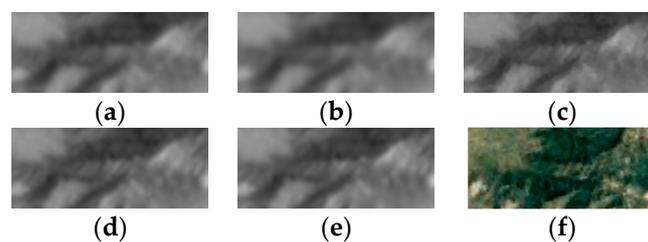

Figure 8. The regions denoted by the rectangles in Figure 7. (**a**) The bicubic upsampled reference image; (**b**) the added image after registration; the HR images reconstructed by (**c**) the proposed method, (**d**) IBP and (**e**) Farsiu's algorithm; (**f**) the image chipped from Google Earth. The Google Earth image is rotated to be compared with the Gaofen-4 images without ortho-rectification.

The PSFs between the HR images and the upsampled reference image of the hilly area are shown in Figure 9. The deviations of the fitted PSFs for IBP, Farsiu's algorithm and our MAP-based SR algorithm are 0.98, 0.82 and 1.0 respectively, which means our method yields the best SR result.
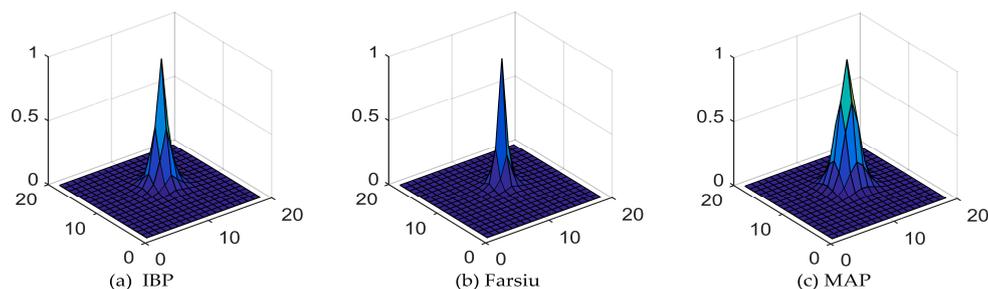
Figure 9. The PSFs between the upsampled reference image and the reconstructed HR images of the hilly area.

3.3. Further Discussion

The deviation of the PSF between the HR image and the upsampled image is less than the deviation estimated in the PSF estimation stage, i.e., the blur is not removed completely. A visually better image can be generated by contrast manipulation plus sharpening after the SR reconstruction. We take the real Gaofen-4 images of the offshore aquaculture area as an example. The result of enhancing the HR image reconstructed by our SR method (Figure 4c) is shown in Figure 10a. For comparison, the reference LR image is enhanced directly without SR reconstruction, and the result is shown in Figure 10b. The regions denoted by the rectangles in Figure 10a,b are shown in Figure 10c,d. Comparing Figures 4, 5 and 10, we note that the enhanced LR image is visually better than the LR image and the reconstructed HR image, but contains less details. That is to say, the SR algorithm reveals more information contained in the LR images.
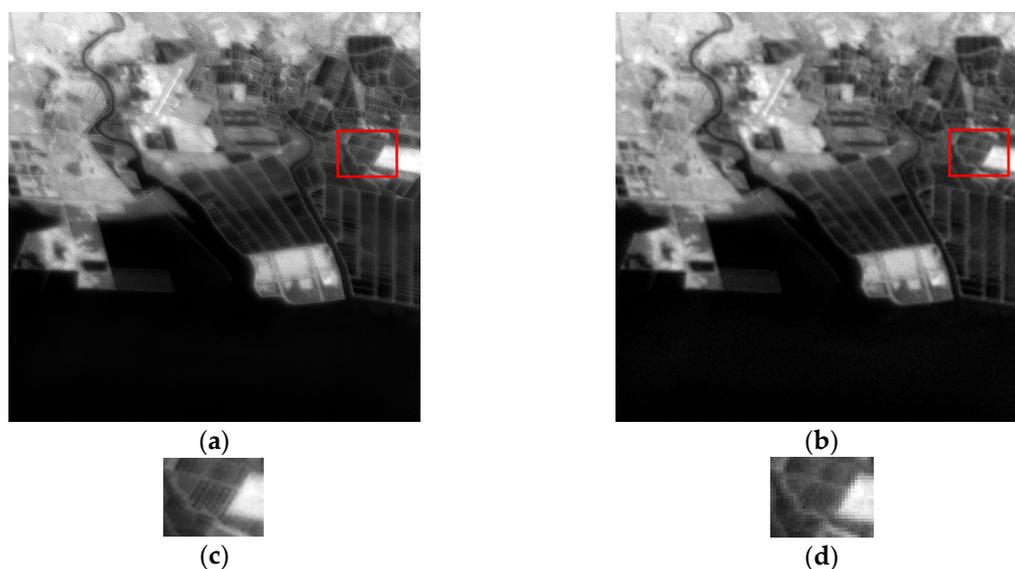

Figure 10. Results of image enhancement. (**a**) SR + image enhancement; (**b**) image enhancement of the reference image; (**c**) and (**d**) are the regions denoted by the rectangles in (**a**) and (**b**) respectively.

4. Conclusions

In this paper, we propose an SR algorithm to reconstruct an HR image from multiple LR images captured by the Gaofen-4 staring imaging satellite. An appropriate observation model with photometric parameters is built. The registration with photometric adjustment improves the registration accuracy and the quality of the final reconstructed HR image. A MAP-based blind deconvolution algorithm is modified to estimate the imaging PSF which is reparameterized by

a Gaussian function. Finally, the PSF is used to estimate the HR image. Experimental results with synthetic and real images show that the reconstructed HR image reveals more details which cannot be seen in LR images.

Author Contributions: Jieping Xu is the first author and the corresponding author of this paper. His main contributions include the basic idea and writing of this paper. The main contributions of Yonghui Liang and Jin Liu include analyzing the basic idea and checking the experimental results. The main contribution of Zongfu Huang was in providing the real Gaofen-4 images and refining the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Park, S.C.; Park, M.K.; Kang, M.G. Super-resolution image reconstruction: a technical overview. *IEEE Signal Process. Mag.* **2003**, *20*, 21–36. [CrossRef]
2. Alam, M.S.; Bognar, J.G.; Hardie, R.C.; Yasuda, B.J. High resolution infrared image reconstruction using multiple, randomly shifted, low resolution, aliased frames. *Proc. SPIE* **1997**, *3063*, 102–112.
3. Alam, M.S.; Bognar, J.G.; Hardie, R.C.; Yasuda, B.J. Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames. *IEEE Trans. Instrum. Meas.* **2000**, *49*, 915–923. [CrossRef]
4. Tsai, R.Y. Multiframe image restoration and registration. *Adv. Comput. Vis. Image Process.* **1984**, *1*, 317–339.
5. Yang, J.; Huang, T. Image super-resolution: Historical overview and future challenges. In *Super-Resolution Imaging*; Milanfar, P., Ed.; CRC Press: Boca Raton, FL, USA, 2010; pp. 1–33, ISBN 978-1-4398-1931-9.
6. Ur, H.; Gross, D. Improved resolution from subpixel shifted pictures. *CVGIP Graph. Models Image Process.* **1992**, *54*, 181–186. [CrossRef]
7. Stark, H.; Oskoui, P. High-resolution image recovery from image-plane arrays, using convex projections. *JOSA A* **1989**, *6*, 1715–1726. [CrossRef]
8. Patti, A.J.; Sezan, M.I.; Tekalp, A.M. Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Trans. Image Process.* **1997**, *6*, 1064–1076. [CrossRef] [PubMed]
9. Irani, M.; Peleg, S. Improving resolution by image registration. *CVGIP Graph. Models Image Process.* **1991**, *53*, 231–239. [CrossRef]
10. Schultz, R.R.; Stevenson, R.L. Extraction of high-resolution frames from video sequences. *IEEE Trans. Image Process.* **1996**, *5*, 996–1011. [CrossRef] [PubMed]
11. Hardie, R.C.; Barnard, K.J.; Armstrong, E.E. Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. Image Process.* **1997**, *6*, 1621–1633. [CrossRef] [PubMed]
12. Zomet, A.; Rav-Acha, A.; Peleg, S. Robust Super-Resolution. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 8–14 December 2001.
13. Farsiu, S.; Robinson, M.D.; Elad, M.; Milanfar, P. Fast and robust multiframe super resolution. *IEEE Trans. Image Process.* **2004**, *13*, 1327–1344. [CrossRef] [PubMed]
14. Yang, S.; Sun, F.; Wang, M.; Liu, Z.; Jiao, L. Novel Super Resolution Restoration of Remote Sensing Images Based on Compressive Sensing and Example Patches-aided Dictionary Learning. In Proceedings of the International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM), Xiamen, China, 10–12 January 2011.
15. Sreeja, S.J.; Wilscy, M. Single Image Super-Resolution Based on Compressive Sensing and TV Minimization Sparse Recovery for Remote Sensing Images. In Proceedings of the IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, India, 19–21 December 2013.
16. Yang, Q.; Wang, H.; Luo, X. Study on the super-resolution reconstruction algorithm for remote sensing image based on compressed sensing. *Int. J. Signal Process. Image Process. Pattern Recognit.* **2015**, *8*, 1–8. [CrossRef]
17. Pan, Z.; Yu, J.; Huang, H.; Hu, S.; Zhang, A.; Ma, H.; Sun, W. Super-resolution based on compressive sensing and structural self-similarity for remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4864–4876. [CrossRef]
18. Merino, M.T.; Nunez, J. Super-resolution of remotely sensed images with variable-pixel linear reconstruction. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1446–1457. [CrossRef]

19. Fruchter, A.S.; Hook, R.N. Drizzle: A method for the linear reconstruction of undersampled images. *Publ. Astron. Soc. Pac.* **2002**, *114*, 144. [CrossRef]
20. Shen, H.; Ng, M.K.; Li, P.; Zhang, L. Super-resolution reconstruction algorithm to MODIS remote sensing images. *Comput. J.* **2007**, *52*, 90–100. [CrossRef]
21. Li, F.; Jia, X.; Fraser, D.; Lambert, A. Super resolution for remote sensing images based on a universal hidden Markov tree model. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1270–1278.
22. Fan, C.; Wu, C.; Li, G.; Ma, J. Projections onto convex sets super-resolution reconstruction based on point spread function estimation of low-resolution remote sensing images. *Sensors* **2017**, *17*, 362. [CrossRef] [PubMed]
23. Jalobeanu, A.; Zerubia, J.; Blanc-Féraud, L. Bayesian estimation of blur and noise in remote sensing imaging. In *Blind Image Deconvolution: Theory and Applications*; Campisi, P., Egiazarian, K., Eds.; CRC Press: Boca Raton, FL, USA, 2007; pp. 239–275.
24. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded Up Robust Features. In Proceedings of the European Conference on Computer Vision 2006 (ECCV 2006), Graz, Austria, 12–18 October 2006; pp. 404–417.
25. Thévenaz, P.; Ruttimann, U.E.; Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **1998**, *7*, 27–41. [CrossRef] [PubMed]
26. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM* **1981**, *24*, 381–395. [CrossRef]
27. Matson, C.L.; Borelli, K.; Jefferies, S.; Beckner, C.C., Jr.; Hege, E.K.; Lloyd-Hart, M. Fast and optimal multiframe blind deconvolution algorithm for high-resolution ground-based imaging of space objects. *Appl. Opt.* **2009**, *48*, A75–A92. [CrossRef] [PubMed]
28. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2009**, *13*, 600–612. [CrossRef]