

INTEGRATED FUSION METHOD FOR MULTIPLE TEMPORAL-SPATIAL-SPECTRAL IMAGES

Huanfeng Shen

School of Resource and Environmental Science, Wuhan University, P.R. China
shenhf@whu.edu.cn

KEY WORDS: Data fusion, remote sensing, multiple temporal-spatial-spectral images

ABSTRACT:

Data fusion techniques have been widely researched and applied in remote sensing field. In this paper, an integrated fusion method for remotely sensed images is presented. Differently from the existed methods, the proposed method has the performance to integrate the complementary information in multiple temporal-spatial-spectral images. In order to represent and process the images in one unified framework, two general image observation models are firstly presented, and then the maximum a posteriori (MAP) framework is used to set up the fusion model. The gradient descent method is employed to solve the fused image. The efficacy of the proposed method is validated using simulated images.

1. INTRODUCTION

In order to get more information, image fusion techniques are often used to integrate the complementary information among different remote sensing images. By far, a great number of fusion methods for remote sensing images have been developed (Luo et al., 2002; Pohl and van Genderen, 1998). Classical remote sensing image fusion techniques include panchromatic(PAN) / multi-spectral(MS) fusion (Joshi and Jalobeanu, 2010; Li and Leung, 2009), MS / hyper-spectral(HS) fusion (Eismann and Hardie, 2005) and multi-temporal (MT) fusion (Shen et al., 2009) etc. However, most fusion methods were developed to fuse images from two sensors, and little work attempted to solve the fuse problem of more sensors. In this paper, we propose an integrated fusion method for multiple temporal-spatial-spectral scales of remote sensing images. This method is based on the maximum a posteriori (MAP) framework, which has the performance to fuse images from arbitrary number of optical sensors.

2. IMAGE OBSERVATION MODELS

The image observation models relate the desired image to the observed images. Let $\mathbf{x} = [x_1, x_2, \dots, x_{B_x}]^T$ denote the desired image with B_x being the total band number. Generally, the band numbers of the observed images are less than or equal to B_x . Here we use \mathbf{y} to denote the images whose band number is equal to B_x and use \mathbf{z} to denote the images whose band number is less than B_x . Thus, the b th band of the k th image of \mathbf{y} can be denoted as $y_{k,b}$, and the b th band of the k th image of \mathbf{z} can be denoted as $z_{k,b}$.

The observation model in terms of $y_{k,b}$ is represented as

$$y_{k,b} = \mathbf{D}_{y,k} \mathbf{M}_{y,k} \mathbf{S}_{y,k,b} \mathbf{x}_b + \mathbf{n}_{y,k,b} \quad (1)$$

where $\mathbf{S}_{y,k,p}$ represents the blur matrix, $\mathbf{M}_{y,k}$ is the motion matrix, $\mathbf{D}_{y,k}$ is down-sampling matrix, and $\mathbf{n}_{y,k,b}$ represents the noise vector. For convenience, equation (1) can be rewritten

as (2) by substituting the product of matrices $\mathbf{S}_{y,k,b}$, $\mathbf{M}_{y,k}$ and $\mathbf{D}_{y,k}$ with $\mathbf{A}_{y,k,b}$

$$y_{k,b} = \mathbf{A}_{y,k,b} \mathbf{x}_b + \mathbf{n}_{y,k,b} \quad (2)$$

The second image observation model relates the desired image \mathbf{x} to the observed image \mathbf{z} . Generally the band of \mathbf{z} is wider than that of \mathbf{x} . It has been proved that a wide-band image is almost a linear combination of several narrow-band images when the wide band approximately covers the narrow bands (Boggione et al., 2003; Li and Leung, 2009; Vega et al., 2009). Thus, if the spatial resolutions of \mathbf{x} and \mathbf{z} are same, the spectral combination model can be denoted as

$$z_{k,b}(i, j) = \sum_{p=1}^{B_x} c_{k,b,p} x_p(i, j) + \tau_{k,b} + n_{k,b}(i, j) \quad (3)$$

where $c_{k,b,p}$ is the corresponding weight of the p th band value $x_p(i, j)$, $\tau_{k,b}$ is an offset, and $n_{k,b}(i, j)$ is the noise. It can be expressed in matrix vector form as

$$\mathbf{z}_{k,b} = \mathbf{C}_{z,k,b} \mathbf{x} + \tau_{k,b} \mathbf{I} + \mathbf{n}_{z,k,b} \quad (4)$$

In more general case, the model can be rewritten as

$$\mathbf{z}_{k,b} = \mathbf{D}_{z,k} \mathbf{M}_{z,k} \mathbf{S}_{z,k,b} (\mathbf{C}_{z,k,b} \mathbf{x} + \tau_{z,k,b} \mathbf{I}) + \mathbf{n}_{z,k,b} \quad (5)$$

Simplifying this equation by multiplying corresponding matrices and vectors

$$\mathbf{z}_{k,b} = \mathbf{A}_{z,k,b} \mathbf{x} + \tau_{z,k,b} \mathbf{B}_{z,k,b} + \mathbf{n}_{z,k,b} \quad (6)$$

3. THE FUSION METHOD

The proposed method is based on the maximum a posteriori (MAP) framework. For the MAP model, given the images \mathbf{y} and \mathbf{z} , the desired image can be estimated as:

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{y}, \mathbf{z}) \quad (7)$$

Applying Bayes' rule, equation (7) becomes:

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} \frac{p(\mathbf{x}) p(\mathbf{y}, \mathbf{z} | \mathbf{x})}{p(\mathbf{y}, \mathbf{z})} \quad (8)$$

Since $p(\mathbf{y}, \mathbf{z})$ is independent of \mathbf{x} , it can be considered a constant and removed from the maximum function:

$$\begin{aligned} \hat{\mathbf{x}} &= \arg \max_{\mathbf{x}} p(\mathbf{x}) p(\mathbf{y}, \mathbf{z} | \mathbf{x}) \\ &= \arg \max_{\mathbf{x}} p(\mathbf{x}) p(\mathbf{y} | \mathbf{x}) p(\mathbf{z} | \mathbf{x}, \mathbf{y}) \\ &= \arg \max_{\mathbf{x}} p(\mathbf{x}) p(\mathbf{y} | \mathbf{x}) p(\mathbf{z} | \mathbf{x}) \end{aligned} \quad (9)$$

Since \mathbf{z} and \mathbf{y} are both known quantities, so it is tenable for $p(\mathbf{z} | \mathbf{x}, \mathbf{y}) = p(\mathbf{z} | \mathbf{x})$ in (9).

The function $p(\mathbf{y} | \mathbf{x})$ provides a measure of the conformance of the estimated image \mathbf{x} to the observed image \mathbf{y} according to the observation model (2). Assuming that the noise is zero-mean Gaussian noise, and each image is independent

$$p(\mathbf{y} | \mathbf{x}) = \prod_{k=1}^{K_y} \prod_{b=1}^{B_{y,k}} p(\mathbf{y}_{k,b} | \mathbf{x}_b) \quad (10)$$

where K_y is the image number of \mathbf{y} , $B_{y,k}$ is the band number of $\mathbf{y}_{k,b}$, and

$$p(\mathbf{y}_{k,b} | \mathbf{x}_b) = \frac{1}{(2\pi\alpha_{y,k,b})^{N_{y,k,h}N_{y,k,v}/2}} \exp\left\{-\frac{\|\mathbf{y}_{k,b} - \mathbf{A}_{y,k,b}\mathbf{x}_b\|^2}{2\alpha_{y,k,b}}\right\} \quad (11)$$

where $\alpha_{y,k,b}$ is the variance of the noise $\mathbf{n}_{y,k,b}$, $N_{y,k,h}$ and $N_{y,k,v}$ are the image dimensions of $\mathbf{y}_{k,b}$.

The function $p(\mathbf{z} | \mathbf{x})$ is determined by the probability density of the noise vector $\mathbf{n}_{z,k,b}$ in (6), and is expressed as:

$$p(\mathbf{z} | \mathbf{x}) = \prod_{k=1}^{K_z} \prod_{b=1}^{B_{z,k}} p(\mathbf{z}_{k,b} | \mathbf{x}) \quad (12)$$

$$p(\mathbf{z}_{k,b} | \mathbf{x}) = \frac{1}{(2\pi\alpha_{z,k,b})^{N_{z,k,h}N_{z,k,v}/2}} \exp\left\{-\frac{\|\mathbf{z}_{k,b} - \mathbf{A}_{z,k,b}\mathbf{x} - \tau_{z,k,b}\mathbf{B}_{z,k,b}\|^2}{2\alpha_{z,k,b}}\right\} \quad (13)$$

where $\alpha_{z,k,b}$ is the variance of the noise $\mathbf{n}_{z,k,b}$, $B_{y,k}$ is the band number of $\mathbf{z}_{k,b}$, and $N_{z,k,h}$ and $N_{z,k,v}$ are the image dimensions of $\mathbf{z}_{k,b}$.

An edge-preserving Huber-Markov image model (Schultz and Stevenson, 1996; Shen and Zhang, 2009) is employed for density function $p(\mathbf{x})$

$$p(\mathbf{x}) = \prod_{b=1}^{B_x} \frac{1}{(2\pi\alpha_{z,b})^{N_{x,h}N_{x,v}/2}} \exp\left\{-\sum_{i,j} \sum_{\xi \in \psi} \rho(d_{\xi}(x_b(i,j))) / 2\alpha_{z,b}\right\} \quad (14)$$

where $\alpha_{z,b}$ is the model parameter of the b th band, ξ is a local group of pixels called a clique, and ψ is the set of all the cliques, $N_{x,h}$ and $N_{x,v}$ are the image dimensions of \mathbf{x} . The quantity $d_{\xi}(x_b(i,j))$ is a spatial activity measure to pixel $x_b(i,j)$, and the following finite second-order differences are computed in two adjacent cliques for every location (i,j) in the image.

$$d_{\xi}^1(x_b(i,j)) = x_b(i-1,j) - 2x_b(i,j) + x_b(i+1,j) \quad (15)$$

$$d_{\xi}^2(x_b(i,j)) = x_b(i,j-1) - 2x_b(i,j) + x_b(i,j+1) \quad (16)$$

In (13), $\rho(\cdot)$ is the Huber function defined as:

$$\rho(h) = \begin{cases} h^2 & |h| \leq \mu \\ 2\mu|h| - \mu^2 & |h| > \mu \end{cases} \quad (17)$$

where μ is a threshold parameter separating the quadratic and linear regions. When μ approaches $+\infty$, the prior becomes the Gauss-Markov, which has similar spatial constraints to the Laplacian prior.

Substituting (10)–(14) in (9) and implementing the monotonic logarithm function, after some manipulation, $N_{x,h}$, $N_{x,v}$, $N_{y,k,h}$, $N_{y,k,v}$, $N_{z,k,h}$ and $N_{z,k,v}$ can be safely dropped, and the maximization of this posterior probability distribution is equivalent to the following regularized minimum problem:

$$\hat{\mathbf{x}} = \arg \min [E(\mathbf{x})] \quad (18)$$

where

$$E(\mathbf{x}) = \sum_{k=1}^{K_y} \sum_{b=1}^{B_{y,k}} \frac{1}{2\alpha_{y,k,b}} \|\mathbf{y}_{k,b} - \mathbf{A}_{y,k,b}\mathbf{x}_b\|^2 + \sum_{k=1}^{K_z} \sum_{b=1}^{B_{z,k}} \frac{1}{2\alpha_{z,k,b}} \|\mathbf{z}_{k,b} - \mathbf{A}_{z,k,b}\mathbf{x} - \tau_{z,k,b}\mathbf{B}_{z,k,b}\|^2 + \sum_{b=1}^{B_x} \frac{1}{2\alpha_{z,b}} \sum_{i,j} \sum_{\xi \in \psi} \rho(d_{\xi}(x_b(i,j))) \quad (19)$$

In this paper, we assume $\alpha_{y,k,b}$, $\alpha_{z,k,b}$ and $\alpha_{x,b}$ are invariable. Thus, minimum function can be simplified as

$$E(\mathbf{x}) = \lambda_1 \sum_{k=1}^{K_y} \sum_{b=1}^{B_{y,k}} \|\mathbf{y}_{k,b} - \mathbf{A}_{y,k,b}\mathbf{x}_b\|^2 + \sum_{k=1}^{K_z} \sum_{b=1}^{B_{z,k}} \|\mathbf{z}_{k,b} - \mathbf{A}_{z,k,b}\mathbf{x} - \tau_{z,k,b}\mathbf{B}_{z,k,b}\|^2 + \lambda_2 \sum_{b=1}^{B_x} \sum_{i,j} \sum_{\xi \in \psi} \rho(d_{\xi}(x_b(i,j))) \quad (20)$$

where λ_1 and λ_2 are two regularization parameters. At last, the steepest gradient descent method (Shen and Zhang, 2009; Shen et al., 2010) is employed to solve the fusion images.

4. EXPERIMENTAL RESULTS

The proposed method was tested using simulated images, and the experimental images and results are illustrated in Fig.1. We used one HS image to simulate one PAN image, one MS image (four bands) and four degraded HS images. Fig.1(a)-(c) show the PAN image, the cubic interpolated versions of the MS image and HS image respectively. The fusion method were implemented in four cases that the input images are respectively four HS images, HS image + MS image, one HS image + PAN image, and all the simulated images. The fused results are shown in Fig.1(d)-(g). By visual inspection, each of the fusion results enhances the spatial resolution. Moreover, the image of the integrated fusion method has the best visual quality.

The fused images are evaluated using five quality indices. These are the root mean square error (RMSE), correlation coefficient (CC), universal image quality index (UIQI), means relative dimensionless global error (ERGAS) and spectral angle (SA). They are defined by equations (21)–(25), respectively.

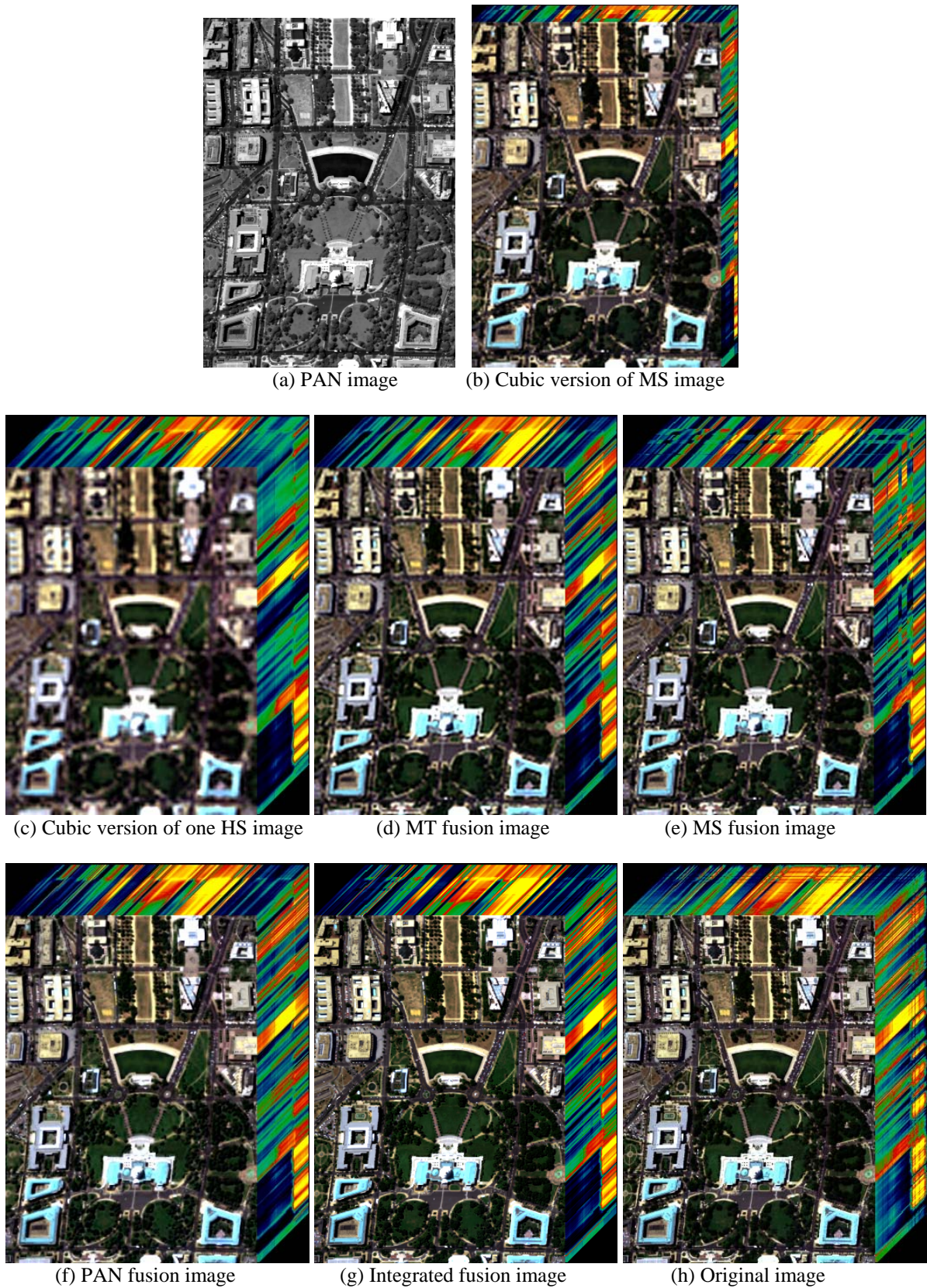


Fig.1 Simulated experiment of different fusion methods.

$$RMSE_b = \sqrt{\frac{\|\hat{x}_b - x_b\|^2}{N_{x,h}N_{x,v}}} \quad (21)$$

$$CC_b = \frac{\sigma_{\hat{x}_b x_b}}{\sigma_{\hat{x}_b} \sigma_{x_b}} \quad (22)$$

$$UIQI_b = \frac{4\sigma_{\hat{x}_b x_b} m_{\hat{x}_b} m_{x_b}}{(\sigma_{\hat{x}_b}^2 + \sigma_{x_b}^2)(m_{\hat{x}_b}^2 + m_{x_b}^2)} \quad (23)$$

$$ERGAS = 100 \cdot L \cdot \sqrt{\frac{1}{B} \sum_b \frac{MSE_b}{m_{x_b}^2}} \quad (24)$$

$$SA = \frac{1}{N_{x,h}N_{x,v}} \sum_{i=1}^{N_{x,h}N_{x,v}} \cos^{-1} \frac{\sum_{b=1}^B (x_{i,b} \cdot \hat{x}_{i,b})}{\sqrt{\sum_{b=1}^B x_{i,b}^2} \cdot \sqrt{\sum_{b=1}^B \hat{x}_{i,b}^2}} \quad (25)$$

Here, \hat{x}_b and x_b represent the b th bands of the fused image and original image, $\sigma_{\hat{x}_b x_b}$ is the covariance between \hat{x}_b and x_b , $m_{\hat{x}_b}$ and m_{x_b} their means, and $\sigma_{\hat{x}_b}$ and σ_{x_b} their standard deviations. The ideal values of the RMSE, CC, UIQI, ERGAS and SA are, respectively, 0, 1, 1, 0 and 0. The evaluation results are shown in Table 1. It is seen that the integrated fusion method obtains the best evaluation values in terms of all the indices. This verifies the proposed method has the performance to integrate the complementary information in multiple temporal-spatial-spectral images.

Table 1. Evaluation of the fusion results

| | MT fusion | MS/HS fusion | PAN/HS fusion | Integrated fusion |
|-------|-----------|--------------|---------------|-------------------|
| RMSE | 20.040 | 20.460 | 16.012 | 8.818 |
| CC | 0.952 | 0.945 | 0.969 | 0.990 |
| UIQI | 0.950 | 0.943 | 0.967 | 0.990 |
| ERGAS | 6.365 | 6.570 | 4.916 | 2.662 |
| SA | 5.659 | 8.749 | 8.230 | 5.509 |

5. CONCLUSIONS

This paper presents a fusion method for multiple temporal-spatial-spectral images based on the maximum a posteriori framework. Simulated experiments validated that the proposed method has good performance in terms of both visual inspection and quantitative evaluation. Future works would be carried out to test the proposed method using real remote sensing images.

ACKNOWLEDGEMENTS

This work was supported by the Major State Basic Research Development Program (973 Program) under Grant 2011CB707103, the National Natural Science Foundation under Grant 40971220, 41071269, the Hubei Natural Science Foundation under Grant 2011CDA096, and the Foundation of State Key Laboratory of Remote Sensing Science under Grant OFSLRSS201114

REFERENCES

- Boggione, G.A., Pires, E.G., Santos, P.A. and Fonseca, L.M.G., 2003. Simulation of a panchromatic band by spectral combination of multispectral ETM+ bands. In: *International Symposium on Remote Sensing of Environmental (ISRSE), Hawaii*.
- Eismann, M. and Hardie, R., 2005. Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3): 455-465.
- Joshi, I. and Jalobeanu, A., 2010. MAP Estimation for Multiresolution Fusion in Remotely Sensed Images Using an IGMRF Prior Mode. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3): 1245 -1255.
- Li, Z. and Leung, H., 2009. Fusion of Multispectral and Panchromatic Images Using a Restoration-Based Method. *IEEE Transactions on Geoscience and Remote Sensing*, 47: 1482-1491.
- Luo, R.C., Chih-Chen, Y. and Kuo Lan, S., 2002. Multisensor fusion and integration: approaches, applications, and future research directions. *IEEE Sensors Journal*, 2(2): 107-119.
- Pohl, C. and van Genderen, J.L., 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19(5): 823-854.
- Schultz, R.R. and Stevenson, R.L., 1996. Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6): 996-1011.
- Shen, H., Ng, M.K., Li, P. and Zhang, L., 2009. Super Resolution Reconstruction Algorithm to MODIS Remote Sensing Images. *The Computer Journal*, 52(1): 90-100.
- Shen, H. and Zhang, L., 2009. A MAP-Based Algorithm for Destriping and Inpainting of Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing*, 47(5): 1492-1502.
- Shen, H.F., Liu, Y.L., Ai, T.H., Wang, Y. and Wu, B., 2010. Universal reconstruction method for radiometric quality improvement of remote sensing images. *International Journal of Applied Earth Observation and Geoinformation*, 12(4): 278-286.
- Vega, M., Mateos, J., Molina, R. and Katsaggelos, A.K., 2009. Super-Resolution of Multispectral Images. *The Computer Journal*, 52(1): 153.