

OPTIMIZATION OF THE SPARSE REPRESENTATION PARAMETERS FOR THE FUSION OF REMOTELY SENSED SATELLITE IMAGES

A. Asefpour Vakilian *, M. R. Saradjian

School of Surveying and Geospatial Engineering, University College of Engineering, University of Tehran, Tehran, Iran -
(afshin.asefpour, sarajian)@ut.ac.ir

Commission IV, WG IV/3

KEY WORDS: Satellite Image, Fusion, Sparse Representation, Optimization, Regularization, Constraints.

ABSTRACT:

Image fusion methods are widely used in remote sensing applications to obtain more information about the features in the study area. One of the recent satellite image fusion techniques that can deal with noise and reduce computational cost and deal with geometric misregistration is sparse representation model. The important part of creating a generalized sparse representation model for satellite image fusion problems is defining initial constraints and adjusting the corresponding regularization coefficients. Regularization coefficients play an essential role in the performance of the sparse representation model and convergence of the optimization solution. Also, the number and size of sub-images extracted from the dictionary matrix in the sparse representation model, and the number of iterations of the optimization step are important in building a sparse representation model. Therefore, in this research, the four parameters that affect the performance of the sparse representation model were investigated: the number of sub-images, the size of sub-images, regularization coefficients, and the number of iterations. Results obtained from pan-sharpening of OLI-8 images showed that optimal values for the number and size of sub-images, regularization coefficients, and the number of iterations were equal to 150, 9×9 pixels, 10^{-4} , and 4 respectively. Results from this study can be generalized to other satellite image fusion problems using sparse representation models.

1. INTRODUCTION

Image fusion methods were introduced to produce a fine spatial resolution image with the highest possible spectral information preserved from input coarse resolution images (Richards, 1992). In these methods, spatial details were added to the regions with the highest spectral variations to reduce the effect of spectral distortions and loss of spatial details (Ghassemian, 2000). High computational cost, geometrical degradation of the results, and the limited use for specific problems are some of the main drawbacks of well-designed fusion methods. Among these methods, sparse representation (SR) models were introduced to cope with the limitations of fusion methods (Zhang et al., 2017). These models can be combined with other methods (Cheng et al., 2015) or new constraints can be introduced (Liu and Wang, 2013) to create a well-constrained SR framework for the fusion problem. Recent researches on the fusion of bulk satellite images using SR models showed that they could deal with noise and big input data, along with the ability to deal with common geometric mismatch and radiometric incompatibility between input satellite images (Song and Huang, 2013) and (Liu et al., 2014). To cope with the limitations of regular satellite image fusion methods, it is necessary to develop new models for remote sensing problems that can be optimized by defining spatial and spectral constraints (Liu and Wang, 2013). The evolution of SR models for fusion of remote sensing images has been presented in the following.

A variety of SR models were introduced for the fusion of satellite images with specific benefits and drawbacks. Nevertheless, according to the literature, SR models were

widely used for the fusion of satellite images. One of the main reasons is the reduced computational cost due to the ability to deal with noise and reduce the dimensions of input satellite images. SR models can produce a dictionary of endmembers that can reconstruct any input image with the lowest possible elements called atoms. Moreover, SR model can preserve spectral information and enhance spatial details of the input images using a well-constrained model. In an SR model, any input image is decomposed into a base (dictionary) matrix and coefficients (sparse) matrix, so that the input image can be reconstructed only using a few number of atoms from the dictionary matrix. Therefore, the coefficients matrix has to be defined in the sparsest way possible with the lowest number of nonzero elements. In this case, the size of input data for creating the SR model drastically decreases, which is an advantage of using SR models.

To date, many SR models have been developed for image fusion problems that differ in the constraints they use and where in the model they use those constraints. Accordingly, various SR models can be classified into different categories as below:

- 1) Traditional SR (Gao et al., 2006), where the sparse matrix only applies to the dictionary matrix,
- 2) Joint SR (Yu et al., 2011), which is a generalized form of the traditional model that applies the sparse matrix to the edge pixels for classification purposes or a combination of input images from different satellite sensors for pan-sharpening or spatiotemporal image fusions,
- 3) Group SR (Li et al., 2012), in which non-zero elements in the dictionary matrix can only occur in blocks and their distribution is not random,

* Corresponding author

- 4) Two-fold SR (Jiang et al., 2014), which uses a tradeoff to determine the ratio between the panchromatic and multispectral images to produce the fused image,
- 5) Coupled SR (Zhu and Bamler, 2013), which employs several explicit non-blind deconvolutions on the dictionary matrix to retrieve the final fused dictionary matrix,
- 6) Nonnegative SR (Wang et al., 2014), in which a positive constraint multiplies by the dictionary matrix as well as the sparse matrix,
- 7) Alternate optimization solution for the SR (Wei et al., 2015), that uses split augmented Lagrangian shrinkage (SALSA), and standard least squares to minimize the objective function in the optimization step,
- 8) Robust SR (Zhang and Levine, 2016), applies a sparse matrix on the sparse reconstruction errors and dictionary matrix,
- 9) Multitask robust SR (Zhang et al., 2018), is an error-based sparse representation model that applies a trained joint SR on the sparse reconstruction errors, considering a connection between all the steps of a sparse representation,
- 10) Local adaptive joint SR (Peng et al., 2019), in which a sparse matrix is reconstructed from a selection of signal sets with the most similarity to the training sample,
- 11) Spatial-spectral SR (Dian et al., 2019), which uses sparsity, nonlocal spectral similarity, and spectral unmixing as three initial constraints to solve the optimization problem and achieve a unique solution.

SR models have been developed to use in different satellite image fusion problems, which is due to the characteristics of SR models such as reducing computational cost, dealing with noise and geometrical misregistration of input images, and radiometric inconsistencies. Compared to other well-known fusion methods, spectral distortion and noise of the fused result from the SR model are lower than others. Moreover, the SR model can be generalized to be used in different image fusion problems by adding different parameters and defining new constraints in the optimization step.

2. MATERIALS AND METHODS

In the following, a description of SR models and implementation and generalization of the results for use in satellite image fusion is presented. SR models employs a space that provides a new look to represent input data with a full-ranked dictionary matrix (Ma et al., 2019). This representation showed a good result in dealing with remote sensing applications such as noise reduction (Dong et al., 2011), and elimination of image distortion (Liu et al., 2014). Despite noisy regions in the input images, noise-free regions can be presented using SR models. Hence, by setting a signal to noise ratio (SNR), noise data can be excluded from the results. Different expressions are presented for continuous signal (image) processing that input continuous signal in R^m space estimated by a linear combination of the columns from dictionary matrix (sub-images) in $R^{m \times n}$ space. For the first time, a study in the field of neuroimaging (Olshausen and Field, 1996), small sub-images with dimensions of 16×16 pixels were used instead of the big input image to design a dictionary matrix for training the model. The size of the sub-images was corresponding to the radiometric resolution of the input images (8 bits). Results from this study showed that their method can automatically detect any interpretable structure in the input data. Formation of a dictionary matrix that represents the input image with a limited number of sub-images becomes a common strategy in many studies in the field of satellite image processing.

2.1 Sparse Representation Model

SR model, as the main concept in this study, introduces a sparse decomposition in the optimization problem to find the sparsest possible representation for any input signal (\mathbf{X}). SR model reduces the number of non-zero elements in a linear decomposition of the input signal to a minimum to reconstruct the input signal with a limited number of elements from the dictionary matrix (\mathbf{D}) and the corresponding elements from the sparse matrix (\mathbf{S}). The most common form of SR models developed for multi-modal problems consisting of images with more than one spectral band is presented in Equation 1 (Dian et al., 2019).

$$\min_{\mathbf{D}, \mathbf{S}} \left(\sum_{m=1}^M \left\| \left(\mathbf{X}^{(m)} \right)^T - \mathbf{D}^{(m)} \mathbf{S} \right\|^2 + \phi(\mathbf{S}) \right), \quad (1)$$

where $\mathbf{X}^{(m)}$ = fused image (N pixels, M spectral bands)
 $\phi(\mathbf{S})$ = regularization function

$\mathbf{X}^{(m)}$ is transposed so that each row is assigned to a pixel, and columns indicate the gray levels for corresponding spectral bands. $\mathbf{D}^{(m)}$ and \mathbf{S} matrices can be retrieved from an iterative optimization procedure (Equation 2). $\phi(\mathbf{S})$ enhances the training performance of the dictionary matrix (Equation 2).

$$\phi(\mathbf{S}) = \beta \|\mathbf{S}\|_1 + \alpha \sum_{m=1}^M \left\| \mathbf{D}^{(m)} \mathbf{S} - \mathbf{d}^{(m)} \right\|_1^2, \quad (2)$$

where α, β = nonnegative regularization coefficients
 $\|\mathbf{S}\|_1$ = sparsity constraint
 $\mathbf{d}^{(m)}, \mathbf{s}$ = dictionary and sparse matrices

α and β control the tradeoff between the sparsity of the coefficients and reconstruction error (Castro and Sapiro, 2012), and their optimal values are obtained through an iterative process that maximizes the peak signal to noise ratio (PSNR) of the estimated results. $\|\mathbf{S}\|_1$ is equal to the sum of the absolute values of elements of the unknown sparse matrix (\mathbf{S}), and $\mathbf{d}^{(m)}$ and \mathbf{s} are corresponding to the sparse representation produced using proposed constraints. In recent studies, sparsity, local spectral similarity, and spatial distance were used in the optimization step to create the SR model (Dian et al., 2019) and (Asefpour Vakilian and Saradjian, 2022). Therefore, the same constraints are used in this study to represent the sparse problem in remote sensing. Feature sign search (l_1 least squares) and Lagrangian dual method are used for optimization of sparse and dictionary matrix, respectively. Recent studies proposed a simultaneous optimization of the dictionary and sparse matrices (Asefpour Vakilian and Saradjian, 2020).

Three steps are required to establish an SR model for image fusion: extract sub-images from the input image, create a dictionary matrix, and define an optimization problem based on dictionary and sparse matrices to achieve an acceptable estimation of the final fused image. These steps are discussed in the following.

2.2 Extract Sub-images from Input Image

The first step to establish an SR model on the input image is to extract sub-images. The difference in the size of input images does not affect the SR model, as it can be solved by selecting

square sub-images (Liu et al., 2014). Therefore, sub-images with the same length and width were used in this study. To investigate the effect of the size and the number of sub-images extracted from input image on the performance of the SR model in the fusion problem, these two parameters were considered variable and changed during the evaluation process. The size of sub-images was changed so that the size of the largest sub-image does not exceed more than ten percent of the size of input image to eliminate overlap between extracted sub-images. The optimal number and size of sub-images are discussed in the next section.

2.3 Building the Dictionary Matrix

After extraction of sub-images with proper size and number, the dictionary matrix can be built according to the input image. The number of rows and columns of the dictionary matrix was defined based on the product of the number of spectral modalities multiplied by the radiometric resolution and the number of pixels, respectively. After building the dictionary matrix corresponding sparse matrix was built to achieve the fusion result using optimization techniques. The number of rows and columns of the sparse matrix was equal to the number of pixels in the input image and the product of the number of spectral bands multiplied by the radiometric resolution, respectively. An iterative optimization solution was used to achieve optimal values for the sparse matrix.

2.4 Optimization Solution

In the proposed method in this research, a least-squares l_1 method is used to optimize the sparse matrix, and a dual Lagrange method is used to optimize the dictionary matrix. These methods were first introduced in a research done at Stanford University (Lee et al., 2006) and have been used simultaneously in recent researches (Wei et al., 2015). To create the optimization problem, first, an arbitrary number was considered as the number of elements in each sub-image in the dictionary matrix. This number was changed during the evaluation of the sparse matrix to obtain the optimal size for the sub-images. A random matrix was formed with columns equal to the number of sub-images and was used as the initial dictionary matrix. An eight-step algorithm was then used to obtain the sparse and dictionary matrix using the input image.

- 1) First, all the elements in the sparse matrix were set to zero.
- 2) Then, substituting element number i from the sparse matrix (s_i) in Equation 2 when the condition in Equation 3 is met.

$$\arg \max_i \left| \frac{\partial \|x - Ds_i\|^2}{\partial s_i} \right|, \quad (3)$$

Partial derivative obtains the difference between the s_i with neighbouring elements in vertical and horizontal directions. According to the conditions mentioned in Equation 4, θ_i is either +1 or -1.

$$\begin{aligned} & \text{if } \frac{\partial \|x - Ds_i\|^2}{\partial s_i} > \beta \Rightarrow \theta_i = -1 \\ & \text{if } \frac{\partial \|x - Ds_i\|^2}{\partial s_i} < -\beta \Rightarrow \theta_i = 1 \end{aligned}, \quad (4)$$

where β = regulation coefficient

β converges the solution of the optimization problem, and the optimal value for it is calculated based on an iterative process. The effect of this coefficient on the performance of the estimation of the final fused result in the SR model is discussed in the next section.

- 3) D' is built as a part of D , in which columns correspond to the non-zero elements of θ .
- 4) The new value for s' (s'_{new}) and θ' matrices were also created from corresponding non-zero elements of θ .
- 5) s'_{new} is obtained using Equation 5.

$$s'_{\text{new}} = (D')^T D')^{-1} \left((D')^T x - \frac{\beta \theta'}{2} \right), \quad (5)$$

- 6) Then, Equation 6 was used on s'_{new} and every element in the s' .

$$\|x - D's'\|^2 + \beta(\theta')^T s', \quad (6)$$

where

$$\beta(\theta')^T s' = \text{the regularization term}$$

The regularization term improves the learning rate of the dictionary matrix to converge faster.

- 7) The new value for s'_{new} leads to the lowest value for Equation 6.

- 8) The values from s' corresponding to non-zero elements in θ were selected to build S and other elements were set to zero.

After obtaining the sparse matrix using the above eight-step algorithm, the dictionary matrix was then optimized using the following algorithm by considering the input image and the sparse matrix as initials (Lee et al., 2006).

Thus, after solving the Lagrangian minimization problem, Equation 7 was obtained.

$$\text{trace}(X^T X - XS^T(SS^T + \Lambda)^{-1}(XS^T)^T - c\Lambda), \quad (7)$$

where

$$c = \text{constant}$$

c is greater than the sum of squared elements in the dictionary matrix. In the proposed method, the elements in the dictionary matrix were normalized between 0 and 1 and c was set to 1. After estimation of Λ using Equation 7, dictionary matrix can be approximated using Equation 8.

$$D^T = (SS^T + \Lambda)^{-1}(XS^T)^T, \quad (8)$$

After the first approximation of dictionary and sparse matrix, the optimization process was repeated until optimal values for both matrices were achieved. The optimal number of iterations was determined based on the PSNR index calculated from fused result and is discussed in the next section. The optimal sparse matrix can express the data structure in the final fused image.

3. RESULTS AND DISCUSSION

In this section, optimal parameters of the SR model for the satellite image fusion problem were investigated and reported. Multispectral and panchromatic OLI-8 images from Alberta province (55° 0' N, 115° 0' W) in Canada during the summer of 2020 and Tehran province (35° 43' N, 51° 24' E) in Iran during the summer of 2019 taken by Landsat-8 satellite were used to

implement an SR model for pan-sharpening. To implement the SR model on the input images, it is necessary to create the dictionary and sparse matrices using appropriate constraints and estimate their values using an optimization process. Different parameters can affect the optimization process in an SR model: the number of random sub-images extracted from input images, the size of extracted sub-images, α and β regularization coefficients, and the number of iterations in optimization step. The effect of each parameter was investigated on the performance of satellite image fusion to provide an efficient and robust SR model. To obtain optimal values for each parameter, one parameter was changed while others were considered constants. Also, the computational complexity of the fusion method was calculated for different values of each parameter based on the required processing time. All codes were programmed using MATLAB R2018b on a desktop PC equipped with an Intel® Core™ i7-6700K processor (8 MB Cache, up to 4.20 GHz), and 12 GB of RAM.

3.1 Number of Sub-images

The effect of the number of extracted random sub-images on the fused result is shown in Figure 1. Sub-images were extracted randomly from the rows of the dictionary matrix. The number of random sub-images changed from 50 to 250 to investigate the PSNR index on the fused result in a satellite image fusion problem. On one hand, by extracting less than 50 sub-images, a significant part of the input image may not be covered and useful information in the study area may be lost. On the other hand, if the number of extracted sub-images exceeds 250, the sub-images will overlap. In this case, redundant features enter the approximation procedure that is not desirable and increase the computational cost. The possibility of overlap cannot be reduced to zero but can be minimized by selecting the appropriate number of sub-images.

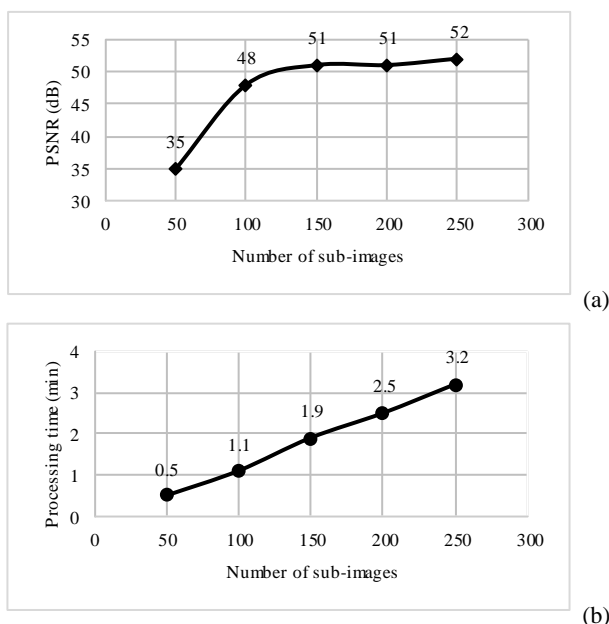


Figure 1. The effect of the number of extracted random sub-images on (a) PSNR, and (b) processing time.

Figure 1 (a) shows improvement in the quality of the fused result with increasing the number of sub-images to 150, however, PSNR remains almost the same with the number of

sub-images more than 150. Therefore, a large number of sub-images will not necessarily improve the performance of the fusion method but increases the computational complexity and the required processing time as shown in Figure 1 (b). Thus, extracting 150 sub-images from the dictionary matrix ensures the best performance for the fusion method. Sub-images are mainly selected from the rows of the dictionary matrix to contain endmembers from all object classes in the input image equally. In other words, 150 sub-images corresponding to pure pixels were added to the SR model to approximate the input image with five different object classes. The number of sub-images increases, if the number of object classes increases.

3.2 Size of Sub-images

After determining the appropriate number of sub-images, they were used to create the dictionary and sparse matrices. Another parameter that affects the performance of the SR model is the size of sub-images. In this research, the number of elements in a sub-image that is corresponding to the number of columns in the dictionary matrix was considered as a square number so that it can be represented as a square matrix. The appropriate size of sub-images is investigated in the following.

Sub-images smaller than a specific size may cover a small or insignificant part of features in the input image. However, larger sub-images may overlap and contain background information along with a specific feature. Therefore, a tradeoff between the size of sub-images and the average size of the object classes is necessary. In this study, the size of sub-images was changed from 4×4 pixels to 32×32 pixels to obtain the performance of the SR model. Sub-images smaller than 4×4 pixels or bigger than 32×32 pixels were not considered in the computation of the size of sub-images as they lack important features or inject background information to the fused result, respectively. Figure 2 (a) shows the impact of the size of sub-images on the performance of the SR model.

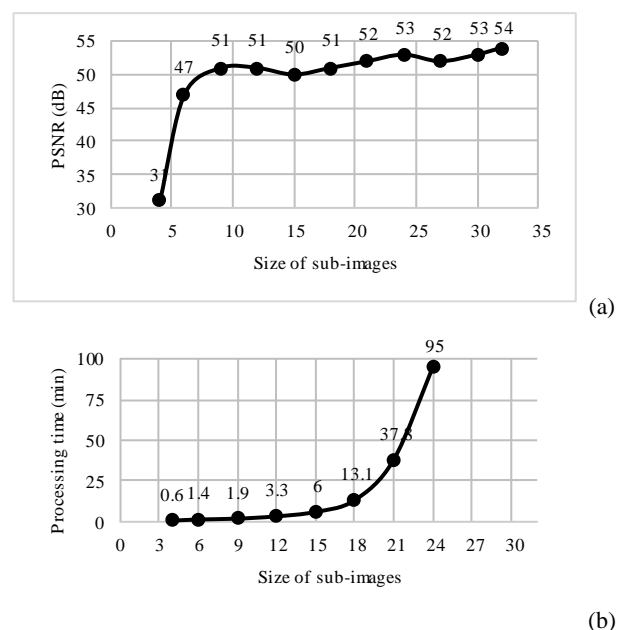


Figure 2. The effect of the size of extracted random sub-images on (a) PSNR, and (b) processing time.

It is obvious that the performance of the model improves significantly by increasing the size of sub-images up to 9×9

pixels; However, by increasing the size of sub-images, PSNR slightly decreased and then slightly increased. This explains that 4×4 pixel sub-images may not include important features, but as the size of sub-images increases, the possibility of containing important features increases too. In other words, most of the important features in an OLI image were placed within 9×9 pixel sub-images. PSNR increases with the size of sub-images, although according to Figure 2 (b) the processing time increases exponentially. This is because the number of columns in the dictionary matrix increases with the increase in the size of sub-images. Due to the large size of satellite images, any increase in computational complexity is highly time-consuming. Thus, 9×9 pixel sub-images were selected as the most appropriate sub-images to use in the SR model to fuse images.

3.3 α and β Regularization Coefficients

A common SR model optimization problem presented in Equation 1, seeks to minimize the difference between the product of dictionary and sparse matrices and the unknown fused image. $\phi(S)$ is the regularization function that improves the learning performance of the dictionary matrix (Equation 2). α and β are regularization coefficients and were changed from 10^{-6} to 1 to obtain their optimal values for creating the SR model using PSNR index; These coefficients adjust the difference between the product of dictionary and sparse matrices and the fused result.

According to Figure 3, the value of the PSNR index for the fused result corresponding to α and β coefficients between 10^{-6} and 10^{-4} is approximately constant and then significantly decreases for coefficients higher than 10^{-3} . β is the regularization coefficient that reduces the spectral difference between the fused result and input images, and α is the regularization coefficient that adjusts the sparsity constraint in the SR model.

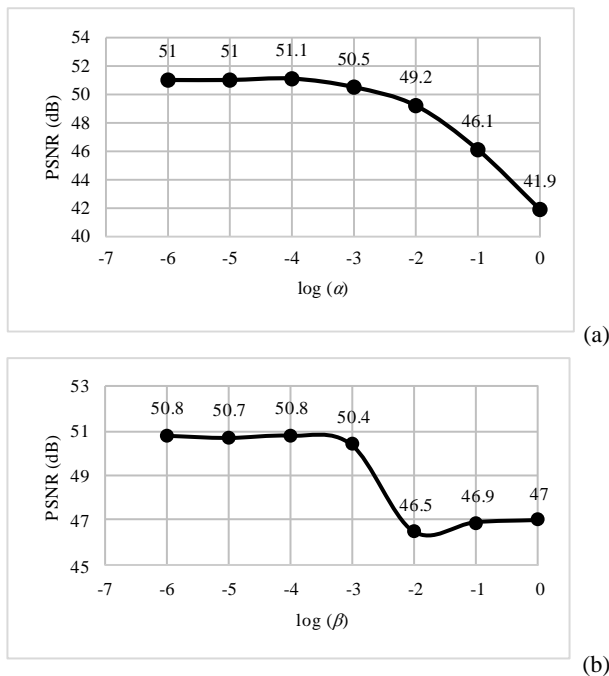


Figure 3. The effect of the (a) α and (b) β regularization coefficients on PSNR. Due to the small values of α and β , they were represented in the logarithm of 10.

Figure 3 (a) shows a direct relationship between the PSNR index and sparsity constraint. Figure 3 (b) indicates that a slight change in the β coefficient may cause drastic changes in the fused result from the SR model. Thus, spectral differences can be minimized to a certain extent, otherwise, spectral distortions in the fused result would increase drastically. Optimal values for α and β were then obtained comparing Figures 3 (a) and 3 (b) and were set to 10^{-4} .

3.4 Number of Iterations

Another parameter that directly affects the performance of the SR model is the number of iterations in the optimization process. Dictionary and sparse matrices were optimized simultaneously in the optimization process. As the number of iterations in the optimization solution increases, the sparsity of the sparse matrix increases, and the dictionary matrix more appropriately models the input data. In this study, the number of iterations was changed from 1 to 9 to obtain the optimal value. Figure 4 (a) shows the PSNR index has reached an acceptable value after four iterations and then remained the same with more iterations. Figure 4 (a) also indicates that after four iterations of the optimization process, the sparsity of the sparse matrix was sufficiently improved, and then according to Figure 4 (b), increasing the number of iterations only increased the processing time with no effect on the PSNR. Finally, the number of iterations of the optimization process was set to four in the proposed SR model to obtain the optimal results for the final dictionary and sparse matrices.

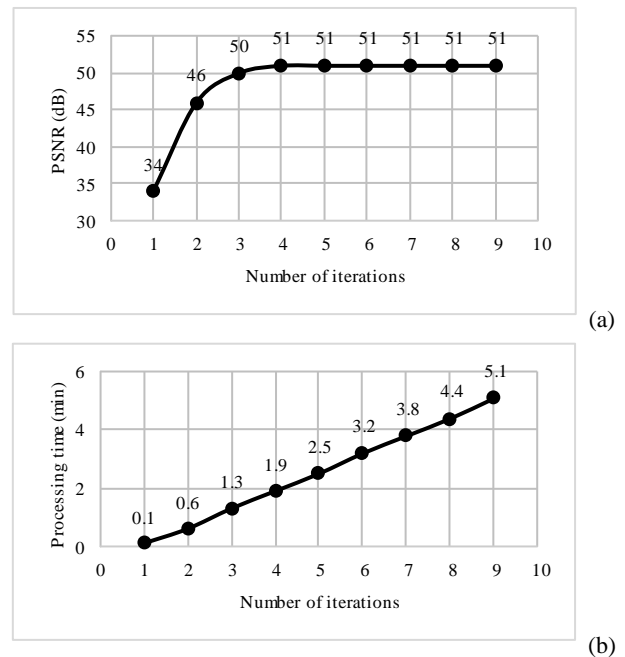


Figure 4. The effect of the number of iterations on (a) PSNR and (b) processing time.

Figure (5) shows a set of sub-images selected for building the dictionary matrix after zero, one, and four iterations. Initial values of the dictionary matrix (in the zero iteration) were selected randomly, and the dictionary matrix for the zero iteration is only a random set of numbers with no optimizations. The sub-images are more smoothed with fewer variations after four iterations, as shown in Figure (5). The sparsity of the

sparse matrix would be optimal by using smoothed sub-images to create the sparse matrix.

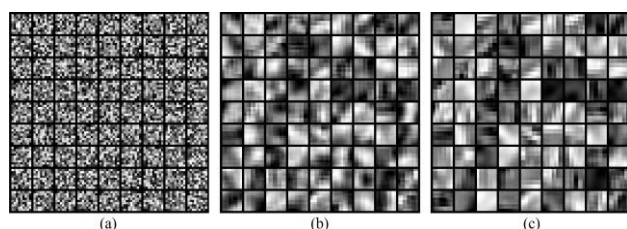


Figure 5. The effect of the number of iterations on sub-images with a size of 9×9 pixels; (a) zero iteration, (b) one iteration, and (c) four iterations.

Finally, the best approximation of the fused result was obtained using optimal values for the parameters of the SR model, including the number of random sub-images (150), the size of random sub-images (9×9 pixels), the α and β regularization coefficients (10^{-4} for both), and the number of iterations (4). In this case, PSNR was equal to 50.84, and the processing time was 1.9 minutes.

4. CONCLUSION

In this research, optimal values of the parameters required for optimization of the general SR model for pan-sharpening of the OLI-8 images were approximated so that the obtained results can be generalized to use in other image fusion problems. To approximate the optimal value of each parameter using an optimization method, three other parameters were considered constant. The optimal value for each parameter was obtained using a tradeoff between the PSNR index and the processing time. Therefore, the improvement of PSNR for each parameter was compared to the processing time to obtain the best value for each parameter. Usually, 9×9 pixel sub-images could capture almost all the objects in input images. But in smooth areas with a lower number of features that are bigger, and the possibility of capturing important features in 9×9 pixel sub-images on the OLI-8 images with a spatial resolution equal to 30 m is low, the size of sub-images has to increase.

REFERENCES

- Asefpour Vakilian, A., Saradjian, M.R., 2022. An object-based sparse representation model for spatiotemporal image fusion. *Scientific Reports*, 12, 5021.
- Asefpour Vakilian, A., Saradjian Maralan, M.R., 2020. Optimizing the regularization parameters of prior information in sparse coding-based multispectral image fusion. *EOGE Earth Observation and Geomatics Engineering*, 4(2), 169-179.
- Castrodad, A., Sapiro, G., 2012. Sparse modeling of human actions from motion imagery. *Int. J. Comput. Vision*, 100, 1-15.
- Cheng, J., Liu, H., Liu, T., Wang, F., Li, H., 2015. Remote sensing image fusion via wavelet transform and sparse representation. *ISPRS J. Photogramm.*, 104, 158-173.
- Dian, R., Li, S., Fang, L., Wei, Q., 2019. Multispectral and hyperspectral image fusion with spatial-spectral SR. *Inform. Fusion*, 49, 262-270.
- Dong, W., Li, X., Zhang, L., Shi, G., 2011. Sparsity-based image denoising via dictionary learning and structural clustering. *CVPR 2011*, Colorado Springs, CO, USA, 457-464.
- Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE T. Geosci. Remote*, 44(8), 2207-2218.
- Ghassemian, H., 2000. Multi-sensor image fusion by inverse subband coding. *Int. Arch. Photogramm. Remote Sens.*, XXXIII, Supplement B2. Amsterdam, 20-27.
- Jiang, C., Zhang, H., Shen, H., Zhang, L., 2014. Two-step sparse coding for the pan-sharpening of remote sensing images. *IEEE J. Sel. Top. Appl.*, 7(5), 1792-1805.
- Lee, H., Battle, A., Raina, R., Ng, A.Y., 2006. Efficient sparse coding algorithms. in *Advances in neural information processing systems (NIPS '06)*, Computer Science Department, Stanford University, CA 94305, 19, 801-808.
- Li, S. Yin, H., Fang, L., 2012. Group-sparse representation with dictionary learning for medical image denoising and fusion. *IEEE T. Bio-Med. Eng.*, 59(12), 3450-3459.
- Liu, W., Tao, D., Cheng, J., Tang, Y., 2014. Multiview hessian discriminative sparse coding for image annotation. *Comput. Vis. Image Und.*, 118, 50-60.
- Liu, Y.Y., Wang, Z., 2013. A practical pan-sharpening method with wavelet transform and sparse representation. *2013 IEEE International Conference on Imaging Systems and Techniques (IST)*, Beijing, China, 288-293.
- Ma, J., Ma, Y., Li, C., 2019. Infrared and visible image fusion methods and applications: A survey. *Inform. Fusion*, 45, 153-178.
- Olshausen, B.A., Field, D.J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 607- 609.
- Peng, J., Jiang, X., Chen, N., Fu, H., 2019. Local adaptive joint sparse representation for hyperspectral image classification. *Neurocomputing*, 334, 239-248.
- Richards, J.A., 1992. Remote sensing digital image analysis. Springer-Verlag, Cambridge, UK.
- Song, H., Huang, B., 2013. Spatiotemporal satellite image fusion through one-pair image learning. *IEEE T. Geosci. Remote*, 51(4), 1883-1896.
- Wang, J., Peng, J., Feng, X., He, G., Fan, J., 2014. Fusion method for infrared and visible images by using non-negative sparse representation. *Infrared Phys. Techn.*, 67, 477-489.
- Wei, Q., Bioucas-Diaz, J., Dobigeon, N., Tourneret, J., 2015. Hyperspectral and multispectral image fusion based on a sparse representation. *IEEE T. Geosci. Remote*, 53(7), 3658-3668.
- Yu, N., Qiu, T., Bi, F., Wang, A., 2011. Image features extraction and fusion based on joint sparse representation. *IEEE J-STSP*, 5(5), 1074-1082.

Zhang, Q., Levine, M.D., 2016. Robust multi-focus image fusion using multi-task sparse representation and spatial context. *IEEE T. Image Process.*, 25(5), 2045-2058.

Zhang, Q., Liu, Y., Blum, R.S., Han, J., Tao, D., 2017. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review. *Inform. Fusion*, 40, 57-75.

Zhu, X.X., Bamler, R., 2013. A sparse image fusion algorithm with application to pan-sharpening. *IEEE T. Geosci. Remote*, 51(5), 2827-2836.