ICIC Express Letters
Volume **16**, Number **10**, October **2022**

ICIC International ©2022 ISSN 1881-803X
pp. **1037–1045**

TOTAL VARIATION AND RANDOM RAW PATCHES FOR RECONSTRUCTING IMAGE FROM INCOMPLETE FREQUENCY

Ratri Dwi Atmaja[1,3,*], Andriyan Bayu Suksmono[1], Donny Danudirdjo[1] and Taufiq Hidayat[2]

[1]School of Electrical Engineering and Informatics
[2]Faculty of Mathematics and Natural Sciences
Institut Teknologi Bandung
Jl. Ganesha, 10, Bandung 40132, Jawa Barat, Indonesia
{ suksmono; donny }@stei.itb.ac.id; taufiq@as.itb.ac.id
*Corresponding author: ratridwiatmaja@students.itb.ac.id

[3]School of Electrical Engineering
Telkom University
Jl. Telekomunikasi, 1, Bandung 40257, Jawa Barat, Indonesia
ratridwiatmaja@telkomuniversity.ac.id

Received October 2021; accepted January 2022

Abstract. *In many measurement applications, the small number of sensors or the limited measurement time causes the measured signal to be incomplete. Therefore, reconstruction is required to obtain the original signal. Total variation (TV) is a reconstruction method that can improve the shape of an object. However, noise still appears around the object. Also, the l_1-norm compressive sensing reconstruction has good estimation accuracy but has a long computation time. This paper proposes a combination of TV and random raw patches for image reconstruction from incomplete frequency. We suspect that using these two methods can improve the object shape and reduce noise even though it requires a longer computation time. Based on the experimental results, the proposed method has a better structural similarity (SSIM) index on ten test images than TV and seven test images than without TV (modified random raw patches alone). Visually, the object image becomes sharper while the noise becomes lower.*
Keywords: Image reconstruction, Compressive sensing, Random raw patches, TV, Incomplete frequency

1. **Introduction.** In digital applications, a digitization step of the input signal is required before entering the signal processing. This digitization step consists of sampling, quantization, and encoding. The sampling theory is first proposed [1], and then expanded by [2]. The minimum sampling frequency is twice the maximum frequency of a signal. This theory still has two weaknesses, i.e., a large number of samples and the high redundancy for certain signals, like on periodic signals. This weakness triggers the proposed compressive sensing (CS). The pioneers of this CS theory are [3] and [4]. CS has the assumption that the signal is sparse, meaning that the signal has only a few significant parts, and the rest is zero, for example, a sinusoidal signal in the frequency region. The sparsity of a signal is denoted by the level of sparsity k. For example, $k = 3$ means that there are three nonzero values in the signal.

Apart from sparsity, CS also contains several other important terms, namely norm, measurement matrix, and restricted isometric property. If $x(n)$ is a signal at the interval

DOI: 10.24507/icicel.16.10.1037

$n = 1, 2, \ldots, N$, then the p-order norm (non-negative p) of $x(n)$ is

$$|x|_p = \sqrt[p]{\sum_{0}^{N-1} |x(n)|^p}. \tag{1}$$

In (1), symbol $|.|$ denotes absolute value. Norms that are often used in CS are the 0-order norm (l_0), the 1st-order norm (l_1), and the 2nd-order norm (l_2). The 0-order norm expresses how many nonzero values in the signal (the level of sparsity k), the 1-order norm expresses the absolute number of nonzero values in the signal, while the 2-order norm expresses the Euclidean distance formed by nonzero values in the signal.

A measurement matrix or sensing matrix is used to decrease the sample number of the original signal. If x consists of n elements, then to decrease the sample of matrix x to a matrix y with the size of m elements ($m < n$), measurement matrix A is needed with size $m \times n$.

$$y = A \times x \tag{2}$$

In (2), x is the original signal with size $n \times 1$ and y is the measurement signal with size $m \times 1$.

Another problem with CS is how to choose the measurement matrix A so the signal x can be retrieved from the signal y. Candes [4] derives the property of the matrix A which is named a restricted isometric property (RIP). The matrix A has RIP if fulfilling the following criteria

$$(1 - \delta_s)|x|_2 \leq |Ax|_2 \leq (1 + \delta_s)|x|_2. \tag{3}$$

In criterion (3), note that we have a small value δ_s. This criterion means that the 2nd-order norm of the original x does not change much after being transformed into the measurement signal y.

The process of retrieving the original signal x is called reconstruction. In general, there are three methods we can use for reconstruction, namely the Greedy algorithm [5-8], convex optimization [9,10], and weight point algorithm [11]. The Greedy algorithm has a fast computation time but results in less accuracy. On the other hand, convex optimization can provide good accuracy but requires a long computation time. As a middle ground, the weight point algorithm has better accuracy than the Greedy algorithm and the computation time is faster than convex optimization.

With a focus on the quality of the reconstructed results and ignoring computation time, we combined TV and random raw patches. The reconstruction using the random raw patches method [12] is a convex optimization group using the l_1-norm minimization technique. Appearing before CS, TV [13] is also a minimization technique that is still popularly used in the field of image reconstruction [14,15]. In this paper, we apply the TV before acting the training process to enhancing the quality of the low-resolution image. We suspect that if the quality of the low-resolution image is close to the high-resolution image, it can give better reconstruction results. As the confirmation of the proposed results, we compare its performance with TV and without TV (modified random raw patches alone).

The organization of this paper is as follows. In Section 2, we explain the proposed combination of TV and random raw patches. In Section 3, we give the results of the proposed method compared to TV and without TV. We then conclude and discuss the future research in Section 4.

2. **Proposed Method.** Figure 1 represents the proposed method that is modified from the random raw patches [12] by removing bicubic interpolation and iterative back projection and adding TV and highest magnitude attenuation (HMA). The TV and HMA are in white in Figure 1. The proposed method is composed of a training and reconstruction process. A high-resolution and low-resolution dictionary is the output of the training. Then, the dictionary is used in the reconstruction to get an estimated image.

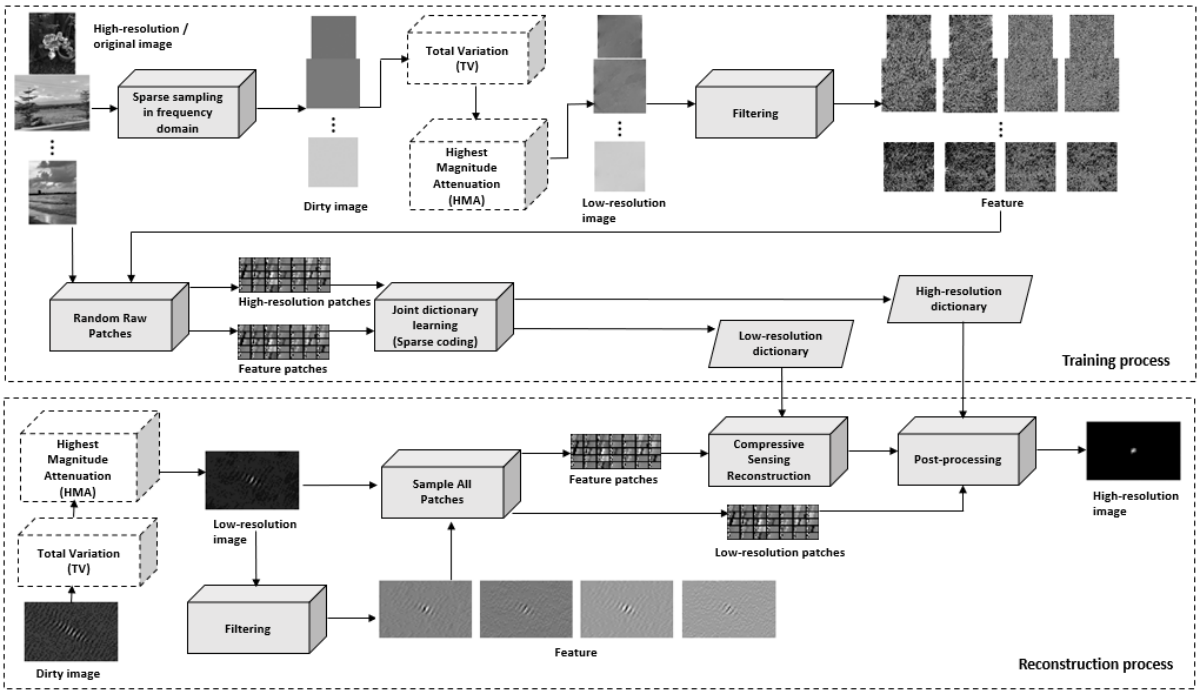


FIGURE 1. The proposed method, modified from [12]

In the training, high-resolution images are transformed into the frequency domain by utilizing the Fourier transform. Then, the results are multiplied with a sparse sampling operator to get its dirty beams. The sparse sampling indicates incomplete frequency measurement. Then, the inverse Fourier transform is performed on the dirty beams to get its dirty images. Subsequently, the dirty images are processed with TV and HMA first before using them as low-resolution images.

Assuming a_{ij} are pixels in coordinate (i, j) of a 2-D image a with size $m \times m$, then determine the operators

$$G_{h;ij}a = \begin{cases} a_{i+1,j} - a_{ij}, & i < m \\ 0, & i = m \end{cases}, \tag{4}$$

$$G_{v;ij}a = \begin{cases} a_{i,j+1} - a_{ij}, & j < m \\ 0, & j = m \end{cases}, \tag{5}$$

and

$$G_{ij}a = \begin{pmatrix} G_{h;ij}a \\ G_{v;ij}a \end{pmatrix}. \tag{6}$$

The TV of a is defined by [13]

$$\|a\|_{TV} = \sum_{ij} \sqrt{(G_{h;ij}a)^2 + (G_{v;ij}a)^2} = \sum_{ij} \|G_{ij}a\|_2, \tag{7}$$

or

$$\|a\|_{TV} = \sum_{i} \sum_{j} \sqrt{|a_{i+1,j} - a_{ij}|^2 + |a_{i,j+1} - a_{ij}|^2}. \tag{8}$$

So we can recover the image a using TV minimization with equality constraints [16]:

$$\min \mathrm{TV}(a) \ \text{ subject to } Aa = b, \tag{9}$$

with A as the measurement matrix, and b as the observation result. Meanwhile, HMA is performed with the following steps: do Fourier transform to the output of TV, separate the magnitude and phase, attenuate the highest magnitude, unite back the magnitude and phase, and do the inverse Fourier transform.

The next step is filtering. We can conduct the filtering using one or more filters. The filtering is a 2-D convolution between the low-resolution image and the filters to obtain some features. Then, we take the random raw patches from the low-high resolution images. After that, the patches are converted to 1-D. A joint learning process is used to connect the low-high resolution patches. We apply a sparse coding method in [17] to the joint learning. Finally, two dictionaries (low-resolution dictionary and high-resolution dictionary) are the training output and will be used in the reconstruction process.

In reconstruction, all the processes are almost the same as in the training, except in the sample all patches, the compressive sensing reconstruction, and the post-processing. The sample of all patches is performed by taking all the patches from left to right and top to bottom. Then, the feature patches are converted into 1-D and normalized, divided by the patch norm. In the compressive sensing reconstruction, the normalized feature patches and the low-resolution dictionary are used to solve a convex problem using Lasso method [18]. Lasso is a linear regression that works based on l_1-norm and used to find the sparse coefficients based on the following equation

$$m = K_{lr}\beta. \tag{10}$$

In Equation (10), m is the normalized feature patch, β is the sparse coefficient, and K_{lr} is the low-resolution dictionary. To find the sparse coefficient β, we can use the following form

$$\min_{\beta} \left(\|m - K_{lr}\beta\|_2^2 + \lambda \|\beta\|_1 \right). \tag{11}$$

In form (11), the sparsity of β can be tuned with the parameter λ. The user can set the value of λ to balance the solution sparsity and accuracy of the estimation to m. In the post-processing, we can get an estimated 1-D high-resolution patch $\hat{m}$ using the following equation

$$\hat{m} = K_{hr}\beta n. \tag{12}$$

In Equation (12), K_{hr} is the high-resolution dictionary and n is the norm of the feature patch. Finally, we need to add the mean of the 1-D low-resolution patch. Then, rearrange all estimated 1-D high-resolution patches to 2-D before writing to the estimated high-resolution image.

3. **Results.** We train 60 high-resolution images in grayscale and perform sparse sampling operation using the sampling pattern in Figure 2. The high-resolution images are natural images and in different sizes to enrich the dictionaries. Then, high-pass filters ($[-1 \quad 0 \quad 1]$, $[-1 \quad 0 \quad 1]^T$, $[1 \quad 0 \quad -2 \quad 0 \quad 1]$, and $[1 \quad 0 \quad -2 \quad 0 \quad 1]^T$) are used to obtain the corresponding features [12]. High-pass filter is chosen because it is more sensitive to the human eye. One pair of feature patches and high-resolution patches are sampled at the same position. Then, the other pairs are sampled at different positions randomly.

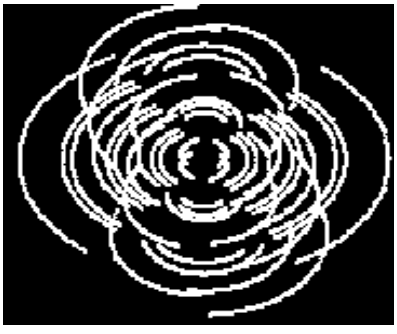


FIGURE 2. The binary image in pixels (sampling pattern to obtain the dirty image), modified from [19]

The 3×3 patch is chosen because it gives the best results. A high-resolution dictionary K_{hr} of size 9×300 is generated from sampling 300 pairs of patches.

In the reconstruction, we use astronomical images from the VLBA-BU Blazar Monitoring Program [20]. These images are used as the original image. Meanwhile, its dirty images are obtained by applying the same sparse sampling pattern in Figure 2. After reconstruction, the estimated high-resolution images are compared with the original image using SSIM index [21].

To determine the HMA performance, we applied HMA to 33 TV output images. All images experienced SSIM improvement except for the 7th image. The best improvement was in the 24th image (see Figure 3). Visually, HMA can reduce noise around objects. Figure 4 shows the image before and after the HMA, and the plot of magnitude before doing HMA (it appears that there is the highest magnitude before being damped).

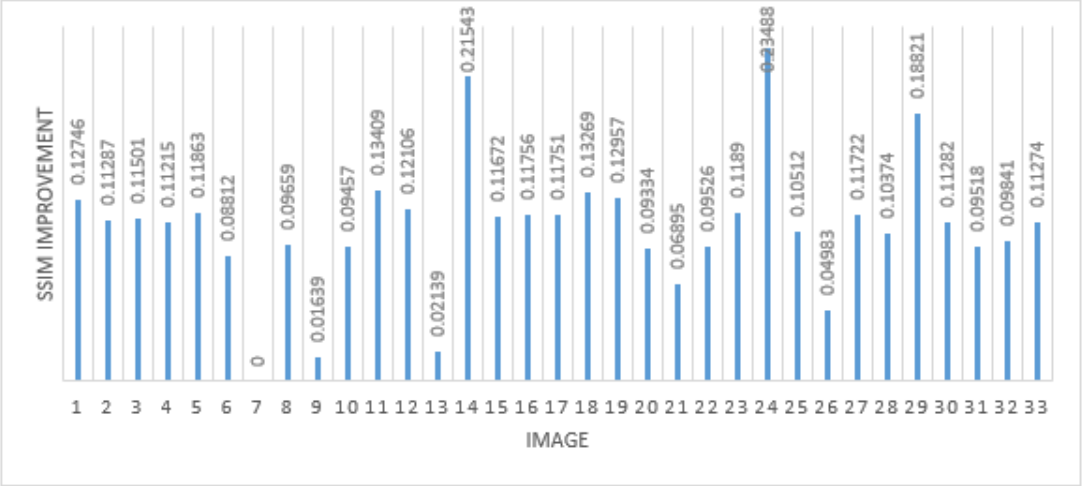


FIGURE 3. The performance of HMA: SSIM improvements on 33 images

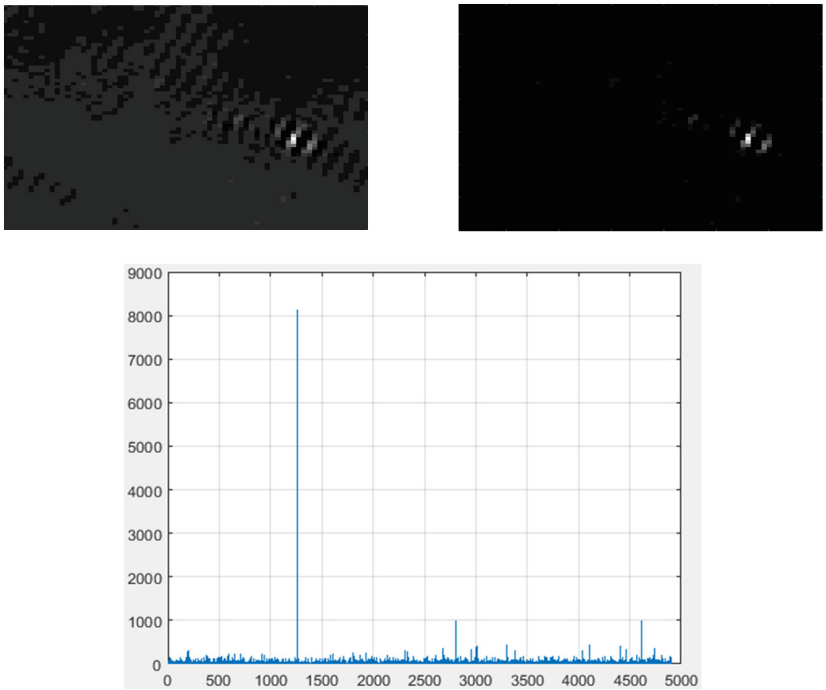
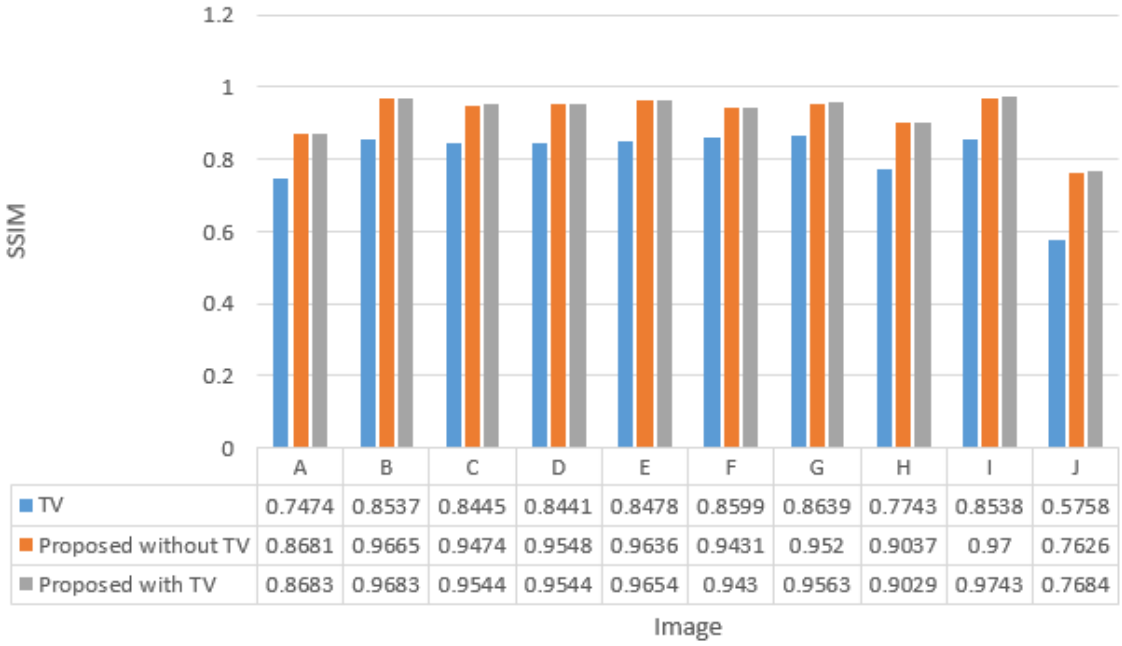


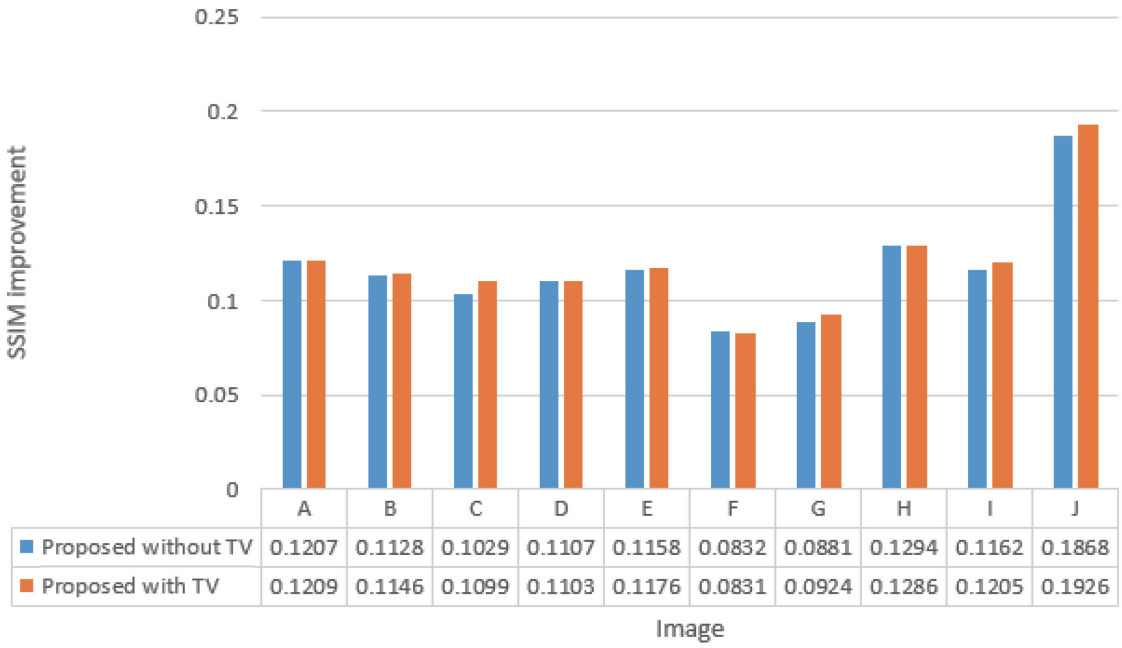
FIGURE 4. Top left: TV output image (before HMA). Top right: the image after HMA. Bottom: plot of magnitude before HMA (the vertical axis is the magnitude and the horizontal axis is the i-th pixel).

The next experiment is to find out the performance of the proposed method when compared to TV. We reconstruct ten images measuring 70×70. Figure 5 shows the SSIM comparison between the proposed method with TV, the proposed without TV, and TV. On the proposed without TV (modified random raw patches alone), the dirty image will directly be input to the HMA without going through the TV. In these 10 test images, the proposed method with TV and that without TV have a higher SSIM than the TV. However, if compared to without TV, the proposed method with TV has a higher SSIM except for the D, F, and H images (see Figure 6). Visually, the proposed method with TV is of better quality than without TV. The proposed method with TV can sharpen the object on all test images (see Figures 7 and 8). If we examine these results, in the



	A	B	C	D	E	F	G	H	I	J
TV	0.7474	0.8537	0.8445	0.8441	0.8478	0.8599	0.8639	0.7743	0.8538	0.5758
Proposed without TV	0.8681	0.9665	0.9474	0.9548	0.9636	0.9431	0.952	0.9037	0.97	0.7626
Proposed with TV	0.8683	0.9683	0.9544	0.9544	0.9654	0.943	0.9563	0.9029	0.9743	0.7684

FIGURE 5. SSIM of reconstruction results: TV, proposed without TV, and with TV



	A	B	C	D	E	F	G	H	I	J
Proposed without TV	0.1207	0.1128	0.1029	0.1107	0.1158	0.0832	0.0881	0.1294	0.1162	0.1868
Proposed with TV	0.1209	0.1146	0.1099	0.1103	0.1176	0.0831	0.0924	0.1286	0.1205	0.1926

FIGURE 6. SSIM improvement for TV: proposed method with TV and without TV

proposed method with TV, TV output that is better than the dirty image affects the quality of the low-resolution image. Meanwhile, in the proposed without TV, the dirty image did not have improvement. These results strengthen the hypothesis. It says that the closer the quality of the low-resolution and the high-resolution images in the training process, the better the reconstruction results.

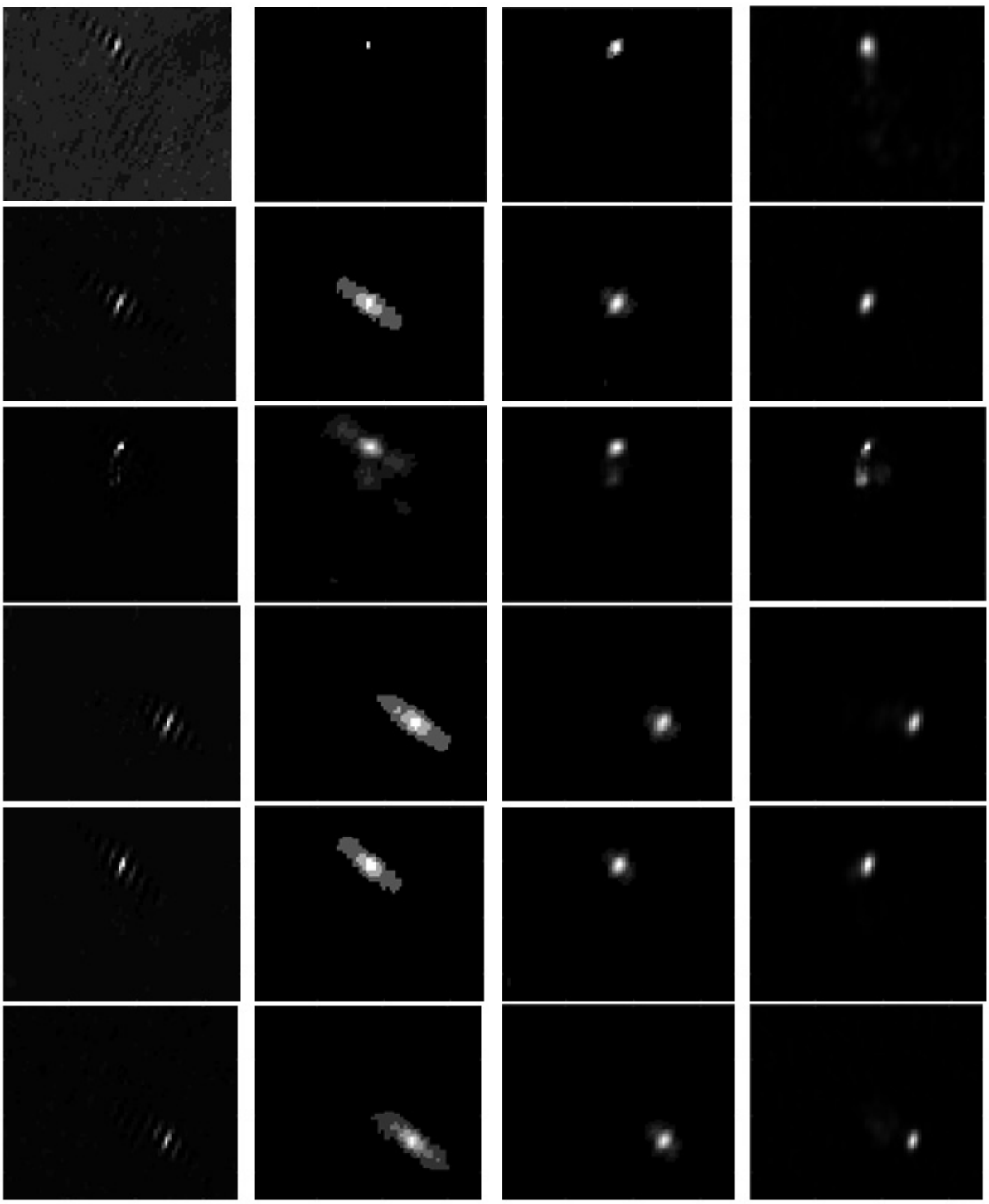


FIGURE 7. Reconstruction results. Left to right: TV, proposed without TV, proposed method with TV, and the original image. Top to bottom: image A, B, C, D, E, and F.

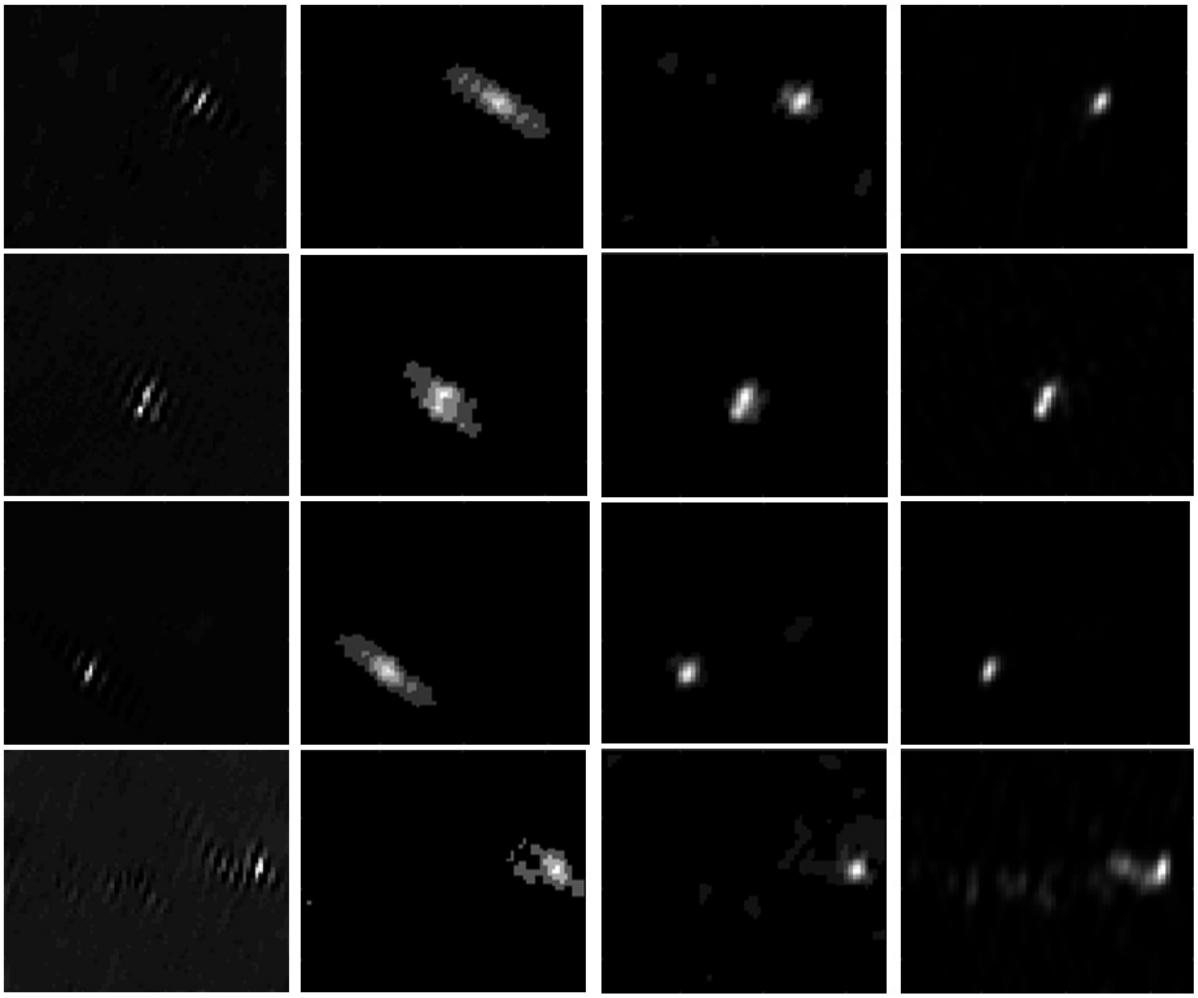


FIGURE 8. Reconstruction results. Left to right: TV, proposed without TV, proposed method with TV, and the original image. Top to bottom: image G, H, I, and J.

4. **Conclusions.** In this paper, we perform an image reconstruction that has incomplete frequency information. We modify the random raw patches method by removing bicubic interpolation and iterative back projection and adding TV and HMA. The proposed method is composed of the training and reconstruction process. In the training, we use natural images and their dirty images to obtain the low-high resolution dictionaries. Then in the reconstruction, we use the compressive sensing method to reconstruct astronomical dirty images. For the result compared to TV on 10 test images, the proposed method has a higher SSIM. Meanwhile, compared to without TV, only seven have a higher SSIM, but visually ten images are of better quality. It should be noted that the proposed combination of TV and random raw patches is designed to reconstruct an image from the incomplete frequency with a sampling pattern in Figure 2 and no noise environment. In the future, we plan to adopt this proposed approach in the case of measurements with different sampling patterns and noisy environments. For example, we can find cases of incomplete frequency measurement with the noisy environment in astronomical imaging, magnetic resonance imaging, medical imaging, molecular imaging, and others. Applying several denoising methods as in [22] can also be a research opportunity to overcome noise problems in reconstruction.

Acknowledgment. This work has been partially supported by F5 Indonesia, and Indonesia Research & Higher Education Ministry.

REFERENCES

[1] H. Nyquist, Certain topics in telegraph transmission theory, *Transaction of AIEE*, vol.47, pp.617-644, 1928.

[2] C. E. Shannon, Communication in the presence of noise, *Proc. of Institute of Radio Engineers*, vol.37, no.1, 1949.

[3] D. L. Donoho, Compressed sensing, *IEEE Trans. Information Theory*, vol.52, no.4, pp.1289-1306, 2006.

[4] E. Candes, Compressive sampling, *Proc. of the International Congress of Mathematicians*, pp.1-20, 2006.

[5] J. Jiang and Y. Chen, Adaptive greedy algorithms based on parameter domain decomposition and reconstruction for the reduced basis method, *International Journal for Numerical Methods in Engineering*, vol.121, no.23, pp.5426-5445, 2020.

[6] Y. Li, J. Zhang, G. Sun and D. Lu, The sparsity adaptive reconstruction algorithm based on simulated annealing for compressed sensing, *Journal of Electrical and Computer Engineering*, vol.19, pp.1-8, 2019.

[7] D. Puneeth and M. Kulkarni, Data aggregation using compressive sensing for energy efficient routing strategy, *Procedia Computer Science*, vol.171, pp.2242-2251, 2020.

[8] T. T. Nguyen, J. Idier, C. Soussen and E. H. Djermoune, Non-negative orthogonal greedy algorithms, *IEEE Trans. Signal Processing*, vol.67, no.21, pp.5643-5658, 2019.

[9] A. Purica, B. Boyadjis, B. P. Popescu, F. Dufaux and C. Bergeron, A convex optimization framework for video quality and resolution enhancement from multiple descriptions, *IEEE Trans. Image Processing*, vol.28, no.4, pp.1661-1674, 2019.

[10] Y. Yang, W. Ma, Y. Zheng, J. F. Cai and W. Xu, Fast single image reflection suppression via convex optimization, *Proc. of IEEE Conf. Comput. Vis. Pattern Recog.*, pp.8141-8149, 2019.

[11] K. Usman, H. Gunawan and A. B. Suksmono, Sparse signal reconstruction using weight point algorithm, *J. ICT Res. Appl.*, vol.12, no.1, pp.35-53, 2018.

[12] J. C. Yang, J. Wright, T. Huang and Y. Ma, Image super-resolution as sparse representation of raw image patches, *Proc. of IEEE Conf. Comput. Vis. Pattern Recog.*, pp.1-8, 2008.

[13] L. I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation noise removal algorithm, *Physica D*, vol.60, pp.259-268, 1992.

[14] K. Akiyama, S. Ikeda, M. Pleau, V. L. Fish, F. Tazaki, K. Kuramochi, A. Broderick, J. Dexter, M. Moscibrodzka, M. Gowanlock, M. Honma and S. S. Doeleman, Super-resolution full polarimetric imaging for radio interferometry with sparse modeling, *The Astronomical Journal*, vol.153, no.4, pp.1-10, 2017.

[15] T. Obuchi, S. Ikeda, K. Akiyama and Y. Kabashima, Accelerating cross-validation with total variation and its application to super-resolution imaging, *PLoS ONE*, 2017.

[16] E. J. Candes, J. Romberg and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Information Theory*, vol.52, no.2, pp.489-509, 2006.

[17] H. Lee, A. Battle, R. Raina and A. Y. Ng, Efficient sparse coding algorithms, *Advances in Neural Information Processing Systems*, 2007.

[18] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of Royal Statistical Society, Series B*, vol.58, no.1, 1996.

[19] *https://web.njit.edu/∼gary/728/Lecture6.html*, Accessed on 07/25/2019.

[20] *http://www.bu.edu/blazars/VLBAproject.html*, Accessed on 10/11/2020.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image qualifty assessment: From error visibility to structural similarity, *IEEE Trans. Image Processing*, vol.13, no.4, 2004.

[22] J. Si, W. Sun and Y. Cheng, Image denoising using low rank matrix completion via bilinear generalized approximate message passing, *International Journal of Innovative Computing, Information and Control*, vol.16, no.5, pp.1547-1558, 2020.