Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 3500–3506, DOI 10.12694/scpe.v25i5.3039

A HYBRID IMAGE FUSION AND DENOISING ALGORITHM BASED ON MULTI-SCALE TRANSFORMATION AND SIGNAL SPARSE REPRESENTATION

DAJUN SHENG*

Abstract. In response to the problem of denoising in image fusion, the author proposes a hybrid image fusion and denoising algorithm based on multi-scale transformation (MLT) and signal sparse representation (SRS). A hybrid model is constructed for shear transformation, and the coefficients after MLT decomposition are thresholded. Sliding window technology and translation invariance are used to form sparse representation for image fusion, and SRS algorithm is used to remove noise from the source image. The experimental results show that the algorithm reduces the contrast and spectral information distortion of the fused image, displays high-quality visual fusion effects, maintains high PSNR values under different noise levels, can provide a more complete description of the features in the image, accurately judge the focus area, maintain the structural correlation of the image, and strengthen the description of fusion edges and details in the fused image. It has been proven that the methods of multi-scale transformation and sparse signal representation can fuse and denoise images.

Key words: Multiscale transformation, Signal sparsity, Image fusion, Denoising algorithm

1. Introduction. In the real world, 20% of human perception information comes from hearing, about 70% of information comes from vision, and about 10% of information comes from taste, smell, touch, and other pathways. From the perspective of biological visual information perception, visual information such as images and videos has become the most important means for humans to perceive and recognize information [1]. In 2015, data showed that the total number of uploaded photos on social networking site Facebook reached 600 billion, with a growth rate of 500 million photos uploaded daily; The average daily video views on the video sharing website YouTube are as high as 8 billion times. With the continuous development of science and technology, the demand for visual information by humans is increasing day by day, and the amount of visual information data is rapidly increasing. The acquisition of multi-source heterogeneous visual information has brought unprecedented development opportunities to visual information processing technology [2]. However, in the process of visual information perception, many factors such as data acquisition, compression, transmission, and storage, as well as hardware device limitations and human operation errors, result in image quality problems such as data loss, noise introduction, and motion blur, which also bring huge challenges to theoretical research and engineering practice.

Natural image quality plays a crucial role in communication and visual perception. High quality images have richer content and information, providing users with a better interactive experience; Poor quality images can lose important information and even cause discomfort to users. Although improving the performance of imaging hardware can improve image quality to a certain extent, equipment costs will significantly increase. The blurring effect caused by the shaking of the shooting equipment, as well as the Gaussian, pulse, and quantization noise introduced during the shooting, storage, and compression processes, cannot be avoided by improving hardware facilities due to the degradation and distortion effects on image quality caused by the computational processing process itself or network packet loss and noise interference [3]. Therefore, utilizing computer theoretical technologies such as image processing, machine vision, and numerical analysis to analyze and process multimodal or noisy images, in order to better understand and perceive target objects, has significant theoretical and practical significance. Image denoising and fusion technology has emerged with the aim of removing or weakening image quality issues during the process of acquiring, transmitting, or storing images. Compared with hardware methods, image denoising and fusion technology has obvious characteristics such as

*College of Big Data and Artificial Intelligence, Xinyang University, Xinyang, Henan 464000, China (Corresponding author, x0376y@163.com)

low cost, high flexibility, and wide applicability. However, it involves the understanding, representation, and modeling of image degradation, noise characteristics, and its own characteristics, and there are many difficult problems and challenges. Image denoising and fusion technology is essentially a fundamental research in the fields of image processing and computer vision, and has received widespread attention [4]. The research in the field of image fusion and denoising began in the 1950s and 1960s. In the practical process, the multi-source images obtained through a large amount of manpower and financial resources often have a certain degree of blurriness. Due to the limited technical conditions at the time, blurred images did not have practical value. Fortunately, through the unremitting efforts of experts and scholars to accurately reconstruct real original images, there is currently a relatively mature and widely used image denoising and fusion technology. A typical successful example is in 1964, NASA's Jet Propulsion Laboratory captured images of the moon on a spacecraft using television cameras, which contained information on noise and interference [5]. Computer processing was used to remove interference and noise, correct geometric distortion and contrast loss, and greatly improve image quality. In recent years, image denoising and fusion techniques based on sparse representation theory have effectively represented and approximated the original image by constructing a dictionary using linear combinations of a few atoms, mining the relationship between representation coefficients and corresponding atoms to reveal the inherent nature of visual information, and obtaining images that conform to human visual perception characteristics. At the same time, due to its superior performance such as simple model, easy implementation, noise resistance, interpretability, and ability to process high-dimensional data, it has sparked a wave of research on image denoising and fusion technology under sparse representation frameworks in the academic and engineering fields. With the continuous development of information technology such as images and videos, image denoising and fusion technology has been widely applied in many scientific and technological fields such as military remote sensing, security monitoring, medical imaging, and consumer electronics.

It can be foreseen that with the comprehensive arrival of social media, the Internet, and the era of big data, the vigorous development of information technology mainly based on images and videos will inevitably give rise to a large number of emerging applications and new demands for image denoising and fusion technology. The large and widespread application demands in the industrial sector will also drive the continuous development of image fusion and denoising research fields. As mentioned earlier, research on image denoising and fusion is also of great significance for the development of theoretical technologies in fields such as image processing and computer vision. On the one hand, image denoising and fusion technology can serve as the underlying technical support for other high-level image processing techniques, thereby improving the efficiency, accuracy, and stability of subsequent image processing tasks; On the other hand, research on image self problems not only involves modeling, representing, and understanding the attributes of images themselves, but also involves interdisciplinary research on human visual mechanisms and psychological perception. It is a fundamental research in the fields of image processing and machine vision, and has great research value for disciplines such as machine vision, pattern recognition, and image understanding[6].

Therefore, in order to solve the noise problem in image fusion, the author proposes a hybrid image fusion and denoising algorithm based on multi-scale transformation (MLT) and signal sparse representation (SRS). The algorithm process is as follows:

Step 1: Perform shear transformation under the mixed model, thresholding the values of various coefficients after MLT decomposition;

Step 2: Utilizing sliding window technology and translation invariance to form sparse representations for image fusion;

Step 3: SRS global processing image denoising algorithm removes noise from the source image.

The experimental results show that the proposed algorithm reduces the contrast and spectral information distortion of the fused image, and has good image fusion and denoising effects.

2. A hybrid image fusion and denoising algorithm based on MLT and SRS. As shown in Figure 2.1, the fusion algorithm proposed by the author consists of the following three steps:

a) Use cartoon texture decomposition to decompose the original image into cartoon and texture parts. The cartoon part mainly includes the structural and geometric parts of the image, while the texture part mainly includes the oscillation and noise parts of the image.

b) The cartoon and texture parts of the image are fused separately. The cartoon part is fused using convolu-

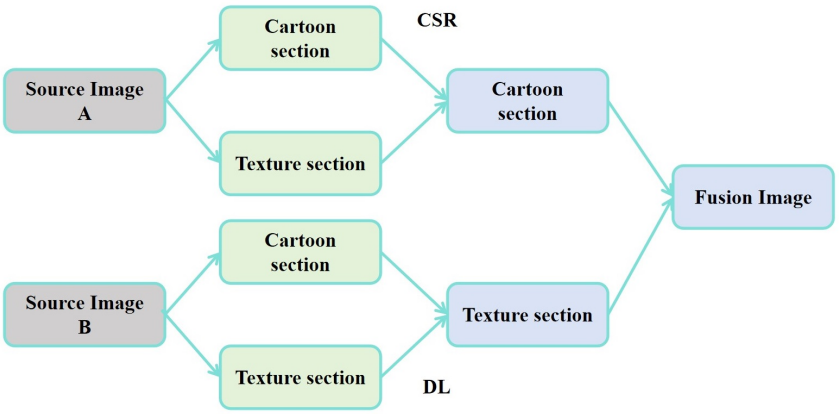


Fig. 2.1: Fusion algorithm framework

tional sparse representation method, while the texture part is fused using dictionary learning method[7].
c) The fused cartoon and texture parts are fused to obtain a fully focused image.

According to the above algorithm, the high-frequency information in the source image needs to be extracted first[8,9,10]. Based on the MLT algorithm to obtain the required data, and according to the composite wavelet and affine system theory, when the dimension n=2, the affine system for composite expansion is defined as follows:

$$A_{AS(\psi)} = \{\psi_{j,l,k}(x) = |det(A)|^{\frac{j}{T}}\psi(S^l A^j - k)\}_j, l \in Z, k \in Z \tag{2.1}$$

In the formula, A and S are both 2×2 non singular matrices, $\Psi \in l^2(R^2)$ is a composite wavelet, $|detS| = 1$. Let A be the parabolic scaling matrix, and S represent the shear matrix for $\forall_a > 0, s \in R$. Among them, $\widehat{\psi_1} \in C^\infty(R)$ is a wavelet.

$$supp\widehat{\psi_1} \subset [-\frac{1}{2}, \frac{1}{16}] \cup [-\frac{1}{16}, \frac{1}{2}] \tag{2.2}$$

And $supp\widehat{\psi_1} \subset [-1, 1]$, therefore $\widehat{\psi}(0) \in C^\infty(R)$ and $\widehat{\psi}(0) \subset [-\frac{1}{2}, \frac{1}{2}]^2$ assume: $\sum_{j \geqslant 0} |\widehat{\psi_1}(2^{-2j}\omega)|^2 = 1, |\omega| \geqslant \frac{1}{8}$, and for $\forall_j \geqslant 0$, $\sum\limits_{L=-2^j}^{2^j-1} = |\widehat{\psi_2}(2^j\omega - l)|^2 = 1, |\omega| \leqslant 1$. From this, it can be concluded that:

$$\sum_{j \geqslant 0} \sum_{L=-2^j}^{s^j-1} |\widehat{\psi}(0)(\xi A_0^{-j} S_0^{-1})|^2 = \sum_{j \geqslant 0} \sum_{L=-2^j}^{2^j-1} |\widehat{\psi_1}(2^{-2j}\xi_1)|^2 |\widehat{\psi_2}(2^{-2j}\frac{\xi_2}{\xi_1} - 1)|^2 = 1 \tag{2.3}$$

For $\xi = R^2$, where x^D is the indicator function of D, $\xi \subset [-\frac{1}{8}, \frac{1}{8}]^2, \widehat{\psi} \subset [-\frac{1}{8}, \frac{1}{8}]^2$ and $\widehat{\psi} = 1$, set $\{\psi(x-k) : k \in Z^2\}$ is a framework of $L^2([-\frac{1}{16}, \frac{1}{16}]^2)^V$, one attribute of ψ^d, d=0.1, $G(\xi) = (\xi_1, \xi_2)$, changes continuously along the straight line $\xi_2 = \pm\xi_1$, and is used to establish a shear transformation hybrid model.

$F \in R$ is a real value sample signal, SRS is a linear combination of dictionary based prototype signals, forming sparse representation theory based on $D \in R^{n \times m}$ dictionary, among them, there are m prototype signals, and in dictionary D, there is a linear combination of prototype signals indicating $\forall x \in f, \exists s \in R^T$, such as $x \approx Ds$, where s is the sparse coefficient in D. It is usually assumed that the dictionary follows a restricted isometric attribute and is redundant, which solves the problem of reconstructing signals using optimization problems to find the non-zero component with the smallest s:

$$min_s||s||_0 subto||D_s - x|| < \varepsilon \tag{2.4}$$

Column based composite system testing solves the problem of the number of non-zero coefficients in sparse matrices, where sparse representation globally processes images, depending on the local information of the source image[11]. The general source image is divided into small blocks with a fixed dictionary D, and the sparse representation of the image is fused using sliding window technology and translation invariance. The source image I is divided into j small blocks of size $n \times n$, represented in dictionary order as vectors $v.V^j$, which can be represented as: $v = \sum_{t=1}^{T} S^j(t)d_t$, where j is the number of image blocks, dt is the prototype from D and $D = [d_{l...}d_{t...}d_T]$, it contains T prototype vectors, and $S^j = [s^1(1)..., s^j(t), ..., s^j(T)]$ is a sparse representation, therefore, the image block of I is used to reconstruct a matrix v, v=DS, where S is a sparse matrix.

Due to the impact of noise on image fusion, in order to improve the effectiveness of image fusion, threshold processing is performed on the coefficients after MLT decomposition. The threshold is defined as:

$$\widehat{T}(\widehat{\sigma_I}) = \widehat{\sigma_n}/\widehat{\sigma_I} \tag{2.5}$$

Among them, σ_I and σ_n is the standard differentiation and noise of the image source, respectively, assuming that source image I and source noise n are independent of each other, and the noise model of image z is represented as $x = I+n$. Therefore, the average noise and signal source calculations are as follows: $\sigma_x^2 = \sigma_I^2+\sigma_n^2$. Among them, σ_x^2 is the variance of the observed signal, σ_n^2 is the noise density of the source image, and the output noise power of the source image σ_I^2 is: $\widehat{\sigma_I} = \sqrt{max((\widehat{\sigma_s^2} - \widehat{\sigma_s^2}), 0)}$.

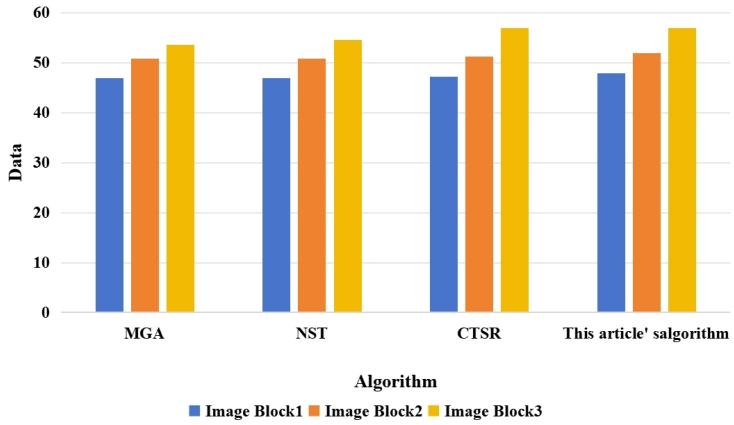
The image fusion algorithm based on MLT and SRS proposed by the author can prevent and reduce the contrast and spectral information distortion of the fused image. When some noise is detected in the source image, a given threshold is applied for filtering. The steps are as follows:

1. Perform MLT decomposition on two source images $\langle I_A, I_B\rangle$, and apply MLT to obtain their low pass band $\langle L_A, L_B\rangle$ and high pass band $\langle H_A, H_B\rangle$.
2. Perform threshold processing on low-pass and high pass using the threshold obtained from equation (5) to remove unnecessary coefficients from the decomposition.
3. Perform low-pass fusion by applying sliding window technology to $\langle I_A, I_B\rangle$, dividing image I into image blocks of size $\sqrt{n} \times \sqrt{n}$, and dividing them by step size pixels from top left to bottom right, due to the presence of $\{P_A^i\}_{i=1}^T$ and $\{P_B^i\}_{i=1}^T$ in L_A and L_B respectively, rearrange $\langle P_A^i, P_B^i\rangle$ into column vectors and rearrange $\langle \widehat{V_A^i}, \widehat{V_B^i}\rangle$ in each iteration.
4. Perform high-throughput fusion and filter using the threshold rule of formula (5) to ensure that the fused image contains the source image.
5. Perform image reconstruction and perform corresponding inverse MLT on L_F and H_F to reconstruct the final fused image I_F.

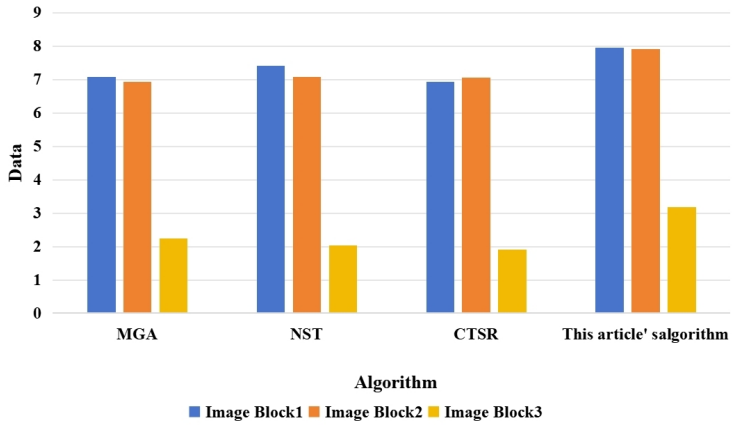
3. Experiments and Results. The experiment randomly assigns values from dictionary D with size 128×512 from image blocks in the training dataset, and then performs sparse encoding to obtain the sparse matrix of the signal. Estimated 160000 training data and 16×16 patches, randomly sampled into images, with a dictionary size set to 128. The experiment uses three commonly used metrics to evaluate the quality of fused images, namely mutual information (MI), standard deviation (SD), and entropy. The proposed algorithm is compared and analyzed with MGA, NST, and CTSR algorithms. As shown in Figure 3.1(a)-(c), the proposed algorithm achieved the best results in SD, MI, and entropy metrics, outperforming MGA, NST, and CTSR algorithms. The proposed algorithm displayed high-quality visual fusion images [12,13,14].

Experimental analysis of different noise levels, that is different standard deviations The application threshold of Equation 2.1 after MLT is used to verify the effect of removing noise and obtaining high-quality denoised images. The proposed algorithm is compared with MGA, NST, and CTSR algorithms, as shown in Figure 3.2. Figure 3.2 shows adding different levels of noise to different images σ the PSNR value shows that the proposed algorithm has higher PSNR values than MGA, NST, and CTSR algorithms in all cases of noise levels[15].

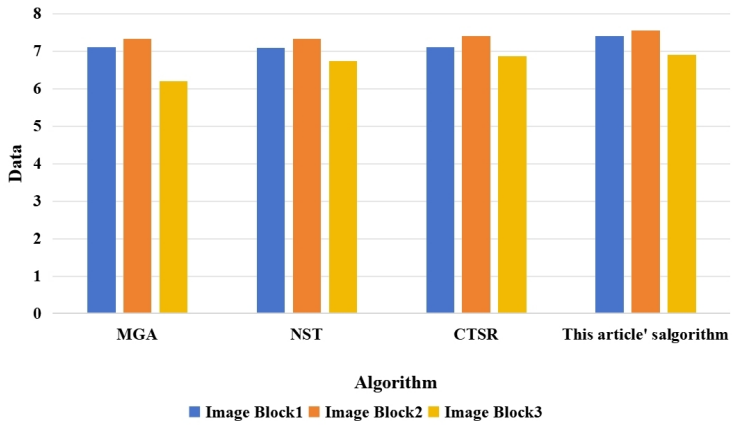
From Figure 3.3, it can be seen that the algorithm proposed by the author has the highest values in desk, Pepsi, and book. Although the comprehensive evaluation criteria in the images of flower, lab, and plane did not reach the maximum value, it can be observed that algorithms such as MGA all exhibit varying degrees of blurring of focus area judgment, block effects, and distortion of fusion boundaries during fusion. The



(a) MI



(b) SD



(c) Entropy

Fig. 3.1: Performance evaluation of the different image fusion algorithms

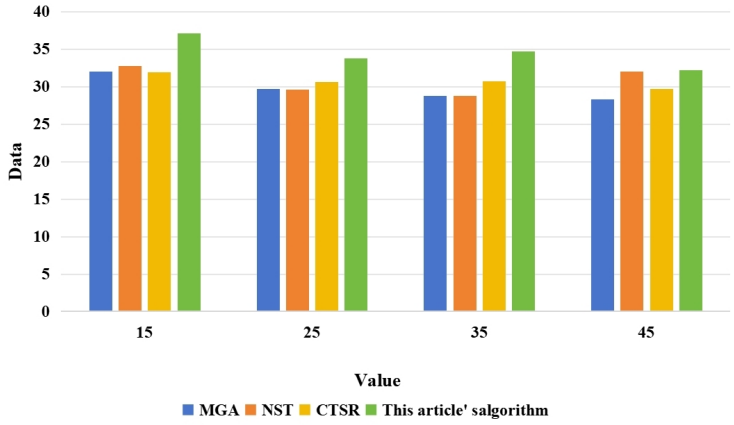


Fig. 3.2: The performance evaluation at different noise levels

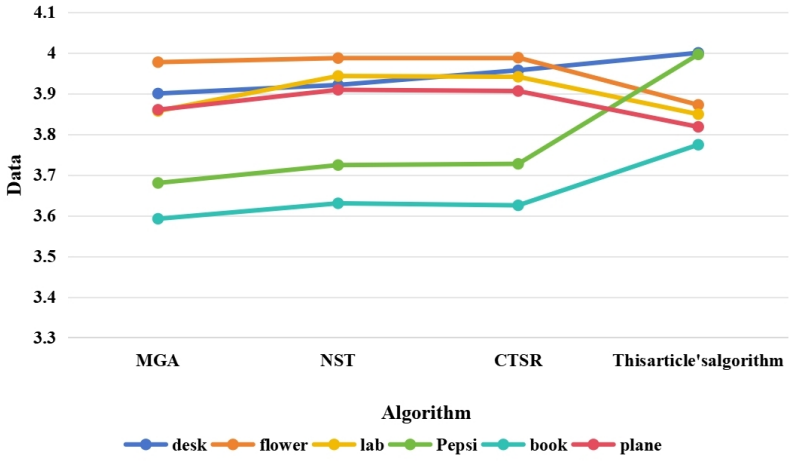


Fig. 3.3: Comparison of algorithm evaluation

emergence of these fusion results is due to the inability of the image decomposition algorithm used to describe all features in the image, inaccurate judgment of the focus area of the source image, and neglect of the structural correlation of the image Using filtering algorithms to fuse images (as filtering algorithms can ensure that images have relatively smooth edges during fusion, but it is also because of the use of filtering algorithms that the description of fusion boundaries is not accurate enough, such as distortion and Gibbs effect) [16].

In summary, compared to the other three algorithms, the algorithm proposed by the author can provide a more complete description of the features in the image, accurately judge the focus area, maintain the structural correlation of the image, and strengthen the fusion of edge description and detail information in the fused image by fusing images [17,18,19,20].

4. Conclusion. Through the above simulation experiments and analysis, it can be concluded that the proposed hybrid image fusion and denoising algorithm based on MLT and SRS has good applicability. Under appropriate threshold conditions, it can obtain high-quality fused images and achieve the effect of removing noise from the source image, reducing the contrast and spectral information distortion of the fused image. The comparison of this algorithm with MGA, NST, and CTSR algorithms shows that the algorithm can display

high-quality visual fusion effects and maintain high PSNR values under different noise levels. However, there are also cases where there is more noise when dealing with different efficient transformation domains and less appropriate thresholds. Therefore, in the next step of research, the focus should be on improving the model transformation algorithm to address these situations, so as to better grasp the geometric shape changes in the image and graphics fusion process.

REFERENCES

[1] Hu, Z. W. H. (2021). An efficient fusion algorithm based on hybrid multiscale decomposition for infrared-visible and multi-type images. Infrared physics and technology, 112(1),158.

[2] Li, G., Lin, Y., & Qu, X. (2021). An infrared and visible image fusion method based on multi-scale transformation and norm optimization. Information Fusion, 71(2),145-148.

[3] Tu, P., Huang, C., & Zhu, J. (2022). A hand gesture recognition algorithm based on multi-scale hybrid features. Journal of Physics: Conference Series, 2218(1), 012038.

[4] Wei, B., Feng, X., & Wang, W. (2021). 3m: a multi-scale and multi-directional method for multi-focus image fusion. IEEE Access, PP(99), 1-1.

[5] Fallah, M., & Azadbakht, M. (2021). Fusion of thermal infrared and visible images based on multi-scale transform and sparse representation. Journal of Geospatial Information Technology, 8(3), 39-59.

[6] Qu, S., Liu, X., & Liang, S. (2021). Multi-scale superpixels dimension reduction hyperspectral image classification algorithm based on low rank sparse representation joint hierarchical recursive filtering. Sensors (Basel, Switzerland), 21(11),63-65.

[7] Wang, X., Su, Y., Zhang, H., & Zou, C. (2021). A new hybrid image encryption algorithm based on gray code transformation and snake-like diffusion. The Visual Computer, 85(7),1-22.

[8] Wang, J., & Gao, Y. (2021). Suspect multifocus image fusion based on sparse denoising autoencoder neural network for police multimodal big data analysis. Scientific Programming,54(7),12-15.

[9] Gao, X., Mou, J., Li, B., Banerjee, S., Sun, B., & Taylor, T. (2023). Multi-image hybrid encryption algorithm based on pixel substitution and gene theory. Fractals,69(3),52-56.

[10] Chen, J., Zhu, Z., Hu, H., Qiu, L., Zheng, Z., & Dong, L. (2023). A novel adaptive group sparse representation model based on infrared image denoising for remote sensing application. Applied Sciences,74(1),63-65.

[11] LIU Yong-sheng, CAI Shi-yang, CHEN Yi-xin, XU Zhi-bo. (2023). Point cloud denoising algorithm based on hybrid filtering and improved bilateral filtering. Journal of Northeastern University(Natural Science), 44(5), 682-688.

[12] Li, Y. Q. X. (2021). An infrared and visible image fusion method based on multi-scale transformation and norm optimization. Information Fusion, 71(1),65.

[13] Wang, C., Wu, Y., Yu, Y., & Zhao, J. Q. (2022). Joint patch clustering-based adaptive dictionary and sparse representation for multi-modality image fusion. Machine Vision and Applications, 33(5), 1-16.

[14] Wang, H., Li, Y., Ding, S., Pan, X., Gao, Z., & Wan, S., et al. (2022). Adaptive denoising for magnetic resonance image based on nonlocal structural similarity and low-rank sparse representation. Cluster Computing, 26(5), 2933-2946.

[15] Prateek, G. V., Ju, Y. E., & Nehorai, A. (2021). Sparsity-assisted signal denoising and pattern recognition in time-series data. Circuits Systems and Signal Processing,124(9), 1-50.

[16] Rebollo-Neira, L., & Inacio, A. (2023). Enhancing sparse representation of color images by cross channel transformation. PloS one, 18(1), e0279917.

[17] Miao, Y., Zakharov, Y. V., Sun, H., Li, J., & Wang, J. (2021). Underwater acoustic signal classification based on sparse time–frequency representation and deep learning. IEEE Journal of Oceanic Engineering,36(7),52-56.

[18] Zhang, J., Chen, J., Yu, H., Yang, D., & Xing, M. (2021). Learning an sar image despeckling model via weighted sparse representation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14,(7) 7148-7158.

[19] An, F. P., Ma, X. M., & Bai, L. (2022). Image fusion algorithm based on unsupervised deep learning-optimized sparse representation. Biomedical signal processing and control(Jan. Pt.B), 85(3),71.

[20] Liu, Z., Wang, L., Feng, Y., Qian, Z., & Chen, X. (2021). A recognition method for time-frequency overlapped waveform-agile radar signals based on matrix transformation and multi-scale center point detection. Applied Acoustics, 175(4), 107855.

Edited by: Zhigao Zheng
Special issue on: Graph Powered Big Aerospace Data Processing
Received: Dec 16, 2023
Accepted: Jan 9, 2024