

Research Article

A Heterogeneous Image Fusion Method Based on DCT and Anisotropic Diffusion for UAVs in Future 5G IoT Scenarios

Shuai Hao ^{1,2}, Beiyi An,² Hu Wen ¹, Xu Ma,² and Keping Yu ³

¹College of Safety Science and Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

²College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

³Global Information and Telecommunication Institute, Waseda University, Tokyo 169-8050, Japan

Correspondence should be addressed to Hu Wen; wenh@xust.edu.cn and Keping Yu; keping.yu@aoni.waseda.jp

Received 26 April 2020; Revised 8 May 2020; Accepted 27 May 2020; Published 27 June 2020

Academic Editor: Di Zhang

Copyright © 2020 Shuai Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Unmanned aerial vehicles, with their inherent fine attributes, such as flexibility, mobility, and autonomy, play an increasingly important role in the Internet of Things (IoT). Airborne infrared and visible image fusion, which constitutes an important data basis for the perception layer of IoT, has been widely used in various fields such as electric power inspection, military reconnaissance, emergency rescue, and traffic management. However, traditional infrared and visible image fusion methods suffer from weak detail resolution. In order to better preserve useful information from source images and produce a more informative image for human observation or unmanned aerial vehicle vision tasks, a novel fusion method based on discrete cosine transform (DCT) and anisotropic diffusion is proposed. First, the infrared and visible images are denoised by using DCT. Second, anisotropic diffusion is applied to the denoised infrared and visible images to obtain the detail and base layers. Third, the base layers are fused by using weighted averaging, and the detail layers are fused by using the Karhunen–Loeve transform, respectively. Finally, the fused image is reconstructed through the linear superposition of the base layer and detail layer. Compared with six other typical fusion methods, the proposed approach shows better fusion performance in both objective and subjective evaluations.

1. Introduction

Internet of Things (IoT) has attracted extensive attention in academics and industry ever since it was first proposed. IoT aims at integrating various technologies, such as body domain network systems, device-to-device (D2D) communication, and unmanned aerial vehicles (UAVs) and satellite networks, to provide a wide range of services in any location by using any network. This makes it highly useful for various civil and military applications. In recent years, however, the proliferation of intelligent devices used in IoT has given rise to massive amounts of data, which brings its own set of challenges to the smooth functioning of the wireless communication network. However, the emergence of the fifth-generation (5G) wireless communication technology has provided an effective solution to this problem. Many scholars are committed to the study of key 5G characteristics such as quality-of-

service (QoS) and connectivity [1, 2]. With the development of UAV technology and 5G wireless systems, the application field of IoT has expanded further [3–5]. Due to their characteristics of dynamic deployment, convenient configuration, and high autonomy, UAVs play an extremely important role in IoT. As some wireless devices suffer from limited transmission ranges, UAVs can be used as wireless relays to improve the network connection and extend the coverage of the wireless network. Meanwhile, due to their adjustable flight altitude and mobility, UAVs can easily and efficiently collect data from the users of IoT on the ground. At present, there exists an intelligent UAV management platform, which can operate several UAVs at the same time through various terminal devices. It is capable of customizing flight routes as needed and obtaining the required user data. Intelligent transportation systems (ITS) can use UAVs for traffic monitoring and law enforcement. UAVs can also be used as base

stations in the air to improve wireless network capacity. In 5G IoT, UAV wireless communication systems will play an increasingly important role [6, 7].

The perception layer, which is the basic layer of IoT, consists of different sensors. The UAVs collect the data of IoT users by means of airborne IoT devices, including infrared (IR) cameras and visible (VI) light cameras [8]. One of the key technologies affecting the reliability of the perception layer is the accurate acquisition of multisource signals and reliable fusion of data. In recent years, heterogeneous image fusion has become an important topic regarding the perception layer of IoT. IR and visible light sensors are the two most commonly used types of sensors. IR images taken by an IR sensor are usually less affected by adverse weather conditions such as bright sunlight and smog [9]. However, IR images lack sufficient details of the scene and have a lower spatial resolution than VI images. In contrast, VI images contain more detailed scene information, but are easily affected by illumination variation. IR and VI image fusion could produce a composite image, which is more interpretable to both human and machine perception. The goal of image fusion is to combine images obtained by different types of image sensors to generate an informative image. The fused image would be more consistent with the human visual perception system than the source image individually. It can be convenient for subsequent processing or decision-making. Nowadays, the image fusion technique is widely used in such fields as military reconnaissance [10], traffic management [11], medical treatment [12], and remote sensing [13, 14].

In recent decades, a variety of IR and VI image fusion approaches have been investigated. In general, based on the different levels of image representation, fusion methods could be classified into three categories: pixel-level, feature-level, and decision-level [15]. Pixel-level fusion, conducted on raw source images, usually generates more accurate, richer, and reliable details compared with other fusion methods. Feature-level image fusion first extracts various features (including colour, shape, and edge) from the multisource information of different sensors. Subsequently, the feature information obtained from multiple sensors is analysed and processed synthetically. Although this method could reduce the amount of data and retain most of the information, some details of the image are still lost. Decision-level fusion is used to fuse the recognition results of multiple sensors to make a global optimal decision on the basis of each sensor independently, thus, completing the decision or classification. Decision-level fusion has the advantage of good real-time performance, self-adaptability, and strong anti-interference; however, the fault-tolerance ability of the decision function directly affects the fusion classification performance. In this study, we focus on the pixel-level fusion method.

The remainder of this paper is organised as follows. In Section 2, we introduce related works and the motivation behind the present work. The proposed fusion method is described in Section 3. Experimental results on public datasets are covered in Section 4. The conclusions of this study are presented in Section 5.

2. Related Works

In general, the pixel-level fusion approach can be divided into two categories: space-based fusion methods and the transform domain technique. Space-based fusion methods usually address the fusion issue via pixel grayscale or pixel gradient. Although these methods are simple and effective, they easily lose spatial details [16]. Liu et al. [17] observed that this type of method is more suitable to fusion tasks of the same type of images. The transform domain technique usually takes the transform coefficient as the feature for image fusion. The fused image is obtained by the fusion and reconstruction of the transform coefficients. The image fusion method based on multiscale transformation has been widely investigated because of its compatibility with human visual perception. In recent years, a variety of fusion methods based on multiscale transform have been proposed, such as low-pass pyramid (RP) [18], gradient pyramid (GP) [19], nonsampled contourlet transform (NSCT) [20], discrete wavelet transform (DWT) [21], and dual-tree complex wavelet transform (DTCWT) [22]. However, image fusion methods based on multiscale transform are usually complex and suffer from long processing time and energy consumption issues, which limit their application.

Due to the aforementioned reasons, many researchers implemented image fusion methods by using discrete cosine transform (DCT). In [23], the authors pointed out that image fusion methods based on DCT were efficient due to their fast speed and low complexity. Cao et al. [24] proposed a multifocus image fusion algorithm based on spatial frequency in the DCT domain. The experimental results showed that it could improve the quality of the output image visually. Amin-Naji and Aghagolzadeh et al. [25] employed the correlation coefficient in the DCT domain for multifocus image fusion, proving that the proposed method could improve image quality and stability in noisy images. In order to provide better visual effects, Jin et al. [26] proposed a heterogeneous image fusion method by combining DSWT, DCT, and LSF. Jin et al. proved that the proposed method was superior to the conventional multiscale method. Although image fusion methods based on DCT have achieved superior performance, the fused results show undesirable side effects such as blocking artifacts [27]. While performing DCT, the image is usually required to be divided into small blocks prior, which causes discontinuities between adjacent blocks in the image. In order to address this problem, several filtering techniques have been proposed, such as weighted least square filter [28], bilateral filter [29], and anisotropic diffusion filter [30]. Xie and Wang [31] pointed out that the anisotropic diffusion processing of images could retain the image edge contour information. Compared with other fusion methods based on filtering, image fusion based on anisotropic diffusion could retain more edge profiles. It preferably suppresses noise and obtains better visual evaluation. However, most of the proposed methods for anisotropic diffusion models are based on the diffusion equation itself, ignoring the image's own feature information, which may lead to loss or blurring of image details (textures, weak edges, etc.). Inspired by the above research, a heterogeneous image fusion method

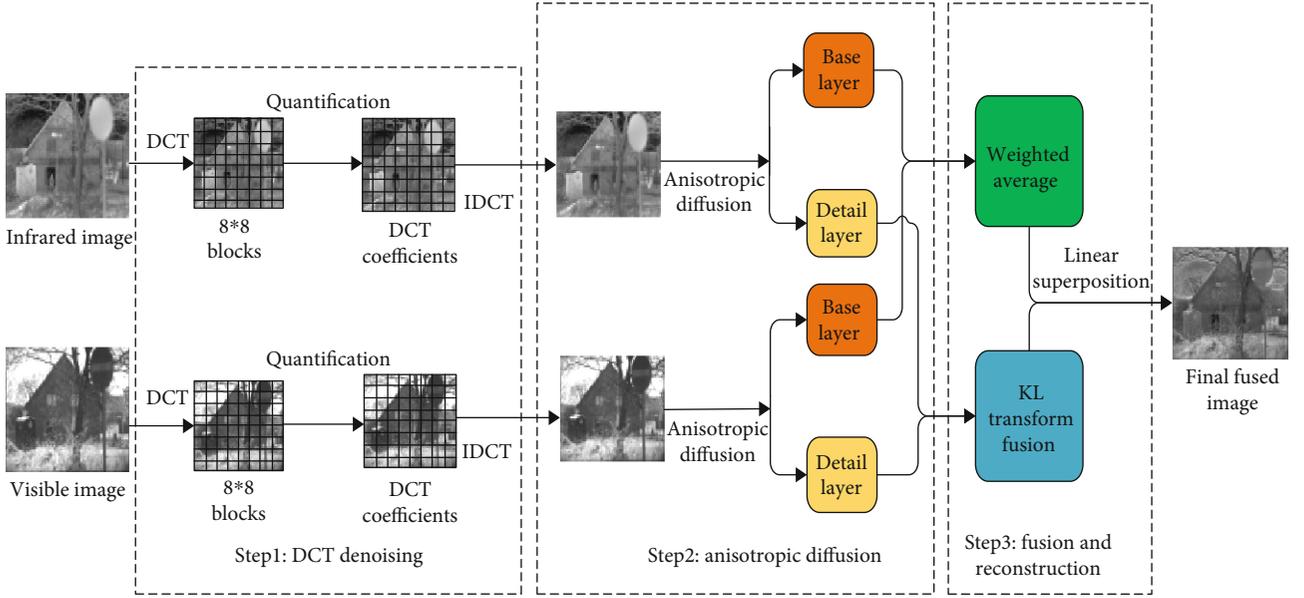


FIGURE 1: Flowchart of the proposed method framework.

based on DCT and anisotropic diffusion is proposed. The advantages of the proposed method mainly lie in the following three aspects:

- (1) Due to the use of DCT transform, the fusion algorithm proposed in this paper shows good denoising ability
- (2) The final fusion images show satisfactory detail resolution
- (3) The proposed algorithm is easy to implement, and the real-time performance is very good, which is suitable for real-time requirements

3. Proposed Fusion Method

In this section, the operation mechanism of the proposed algorithm is described in detail. The proposed image fusion framework can be divided into three components, as shown in Figure 1. In the first step, in order to eliminate the noise in the original images, DCT and inverse discrete cosine transform are performed on the IR and VI images, respectively. In the second step, anisotropic diffusion is adopted to decompose IR and VI images to obtain the detail and base layers. In the third step, base layers are fused by using the weighted averaging, and detail layers are fused by using the Karhunen–Loeve transformation. Finally, the fused base and detail layers are linearly superimposed to obtain the final fusion result.

3.1. DCT. As an effective transform tool, DCT can transform the image information from the time domain to the frequency domain so as to effectively reduce the spatial redundancy of the image. In this study, the 2D DCT of an $N \times N$

image block $f(i, j)$ is defined as follows [32]:

$$F(u, v) = \frac{2}{N} c(u)c(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \cos\left(\frac{(2i+1)\pi u}{2N}\right) \times \cos\left(\frac{(2j+1)\pi v}{2N}\right), \text{ where } u, v = 0, 1, \dots, N-1, \quad (1)$$

where $f(i, j)$ denotes the (i, j) -th image pixel value in the spatial domain and $F(u, v)$ denotes the (i, j) -th DCT coefficient in the frequency domain; $c(k)$ is a multiplication factor, defined as follows:

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}}, & \text{if } u = 0, \\ 1 & \text{if } u \neq 0. \end{cases} \quad (2)$$

Similarly, the 2D inverse discrete cosine transform (IDCT) is defined as:

$$f(i, j) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} c(u)c(v)F(u, v) \cos\left(\frac{(2i+1)\pi u}{2N}\right) \times \cos\left(\frac{(2j+1)\pi v}{2N}\right), \text{ where } i, j = 0, 1, \dots, N-1. \quad (3)$$

When performing the DCT transform, most of the image information is concentrated on the DC coefficient and low-frequency spectrum nearby. Therefore, the coefficient close to 0 is deleted, and the coefficient containing the main information of the image is reserved for inverse transformation. The influence of noise can be effectively removed without

causing image distortion. In this study, if the coefficient obtained by the DCT transformation is less than 0.1, we set the coefficient as 0. The results of denoising by DCT are shown in Figure 2.

3.2. Anisotropic Diffusion. In computer vision, anisotropic diffusion is widely used to reduce noise while preserving image details. The model of anisotropic diffusion of an image I can be represented as follows [33]:

$$It = c(x, y, t)\Delta I + \nabla c \cdot \nabla I, \quad (4)$$

where $c(x, y, t)$ is rate of diffusion, ∇ and Δ are used to represent the gradient operator and the Laplacian operator, respectively, and t is time. Then, the equation (4) can be discretized as:

$$I_{i,j}^{t+1} = I_{i,j}^t + \lambda \left[c_N \cdot \Omega_N I_{i,j}^t + c_S \cdot \Omega_S I_{i,j}^t + c_E \cdot \Omega_E I_{i,j}^t + c_W \cdot \Omega_W I_{i,j}^t \right]. \quad (5)$$

In (5), $I_{i,j}^{t+1}$ is an image with coarse resolution at $t+1$ scale. λ denotes a stability constant satisfying $0 \leq \lambda \leq 1/4$. The local image gradients along the north, south, east, and west directions are represented as:

$$\begin{cases} \Omega_N I_{i,j} = I_{i-1,j} - I_{i,j}, \\ \Omega_S I_{i,j} = I_{i+1,j} - I_{i,j}, \\ \Omega_E I_{i,j} = I_{i,j+1} - I_{i,j}, \\ \Omega_W I_{i,j} = I_{i,j-1} - I_{i,j}. \end{cases} \quad (6)$$

Similarly, the conduction coefficients along the four directions of north, south, east, and west can be defined as:

$$\begin{cases} c_{N_{i,j}}^t = g\left(\left\|(\nabla I)_{i+1/2,j}^t\right\|\right) = g\left(\left|\Omega_N I_{i,j}^t\right|\right), \\ c_{S_{i,j}}^t = g\left(\left\|(\nabla I)_{i-1/2,j}^t\right\|\right) = g\left(\left|\Omega_S I_{i,j}^t\right|\right), \\ c_{E_{i,j}}^t = g\left(\left\|(\nabla I)_{i,j+1/2}^t\right\|\right) = g\left(\left|\Omega_E I_{i,j}^t\right|\right), \\ c_{W_{i,j}}^t = g\left(\left\|(\nabla I)_{i,j-1/2}^t\right\|\right) = g\left(\left|\Omega_W I_{i,j}^t\right|\right). \end{cases} \quad (7)$$

In (7), $g(\cdot)$ is a decreasing function. In order to maintain smoothing and edge preservation, in this paper, we choose $g(\cdot)$ as:

$$g(\nabla I) = \begin{cases} \left[1 - \frac{(\|\nabla I\|)^2}{T^2}\right]^2, & (\|\nabla I\|)^2 \leq T^2, \\ 0, & (\|\nabla I\|)^2 > T^2, \end{cases} \quad (8)$$

where T is an edge magnitude parameter.

Let $I_{IR}(x, y)$ and $I_{VI}(x, y)$ be IR and VI images, respectively, which have been coregistered. Anisotropic diffusion for an image I is denoted as $AD(I)$. The base layers $I_{IR}^B(x, y)$

and $I_{VI}^B(x, y)$ are obtained after performing anisotropic diffusion processes on $I_{IR}(x, y)$ and $I_{VI}(x, y)$, respectively, which are represented by:

$$\begin{cases} I_{IR}^B(x, y) = AD(I_{IR}(x, y)), \\ I_{VI}^B(x, y) = AD(I_{VI}(x, y)). \end{cases} \quad (9)$$

Then, the detail layers $I_{IR}^D(x, y)$ and $I_{VI}^D(x, y)$ are obtained by subtracting the base layers from the respective source images, which are shown in (10).

$$\begin{cases} I_{IR}^D(x, y) = I_{IR}(x, y) - I_{IR}^B(x, y), \\ I_{VI}^D(x, y) = I_{VI}(x, y) - I_{VI}^B(x, y). \end{cases} \quad (10)$$

The results obtained by the anisotropic diffusion of IR image and VI image are shown in Figure 3.

3.3. Construction and Fusion

3.3.1. Base Layer Fusion. The weighted average is adopted to fuse the base layers of IR and VI images. The fused base layer I_F^B is calculated by:

$$I_F^B = \omega_1 I_{IR}^B(x, y) + \omega_2 I_{VI}^B(x, y), \quad (11)$$

where ω_1 and ω_2 are normalized weighted coefficients, set to 0.5 in this paper.

3.3.2. Detail Layer Fusion. KL-transform can make the new sample set approximate to the original sample set distribution with minimum mean square error, and it eliminates the correlation between the original features. We use KL-transform to fuse the detail layers of IR and VI images. Let $I_{IR}^D(x, y)$ and $I_{VI}^D(x, y)$ be detail layers of the IR and VI images, respectively. The fused process based on KL-transform is described as follows.

Step 1. Arrange $I_{IR}^D(x, y)$ and $I_{VI}^D(x, y)$ as column vectors of a matrix, denoted as X .

Step 2. Calculate the autocorrelation matrix C of X .

Step 3. Calculate the eigenvalues λ_1 matrix λ_2 , and eigenvectors $\mu_1 = [\mu_1(1) \mu_1(2)]^T$ and $\mu_2 = [\mu_2(1) \mu_2(2)]^T$ of C .

Step 4. Calculate the uncorrelated coefficients K_1 and K_2 corresponding to the largest eigenvalue λ_{\max} , which is defined as:

$$\lambda_{\max} = \max(\lambda_1, \lambda_2). \quad (12)$$

The eigenvector corresponding to λ_{\max} is denoted as μ_{\max} . And, K_1 and K_2 are given by:

$$\begin{cases} K_1 = \lambda_{\max}(1) / \sum_i \lambda_{\max}(i), \\ K_2 = \lambda_{\max}(2) / \sum_i \lambda_{\max}(i). \end{cases} \quad (13)$$

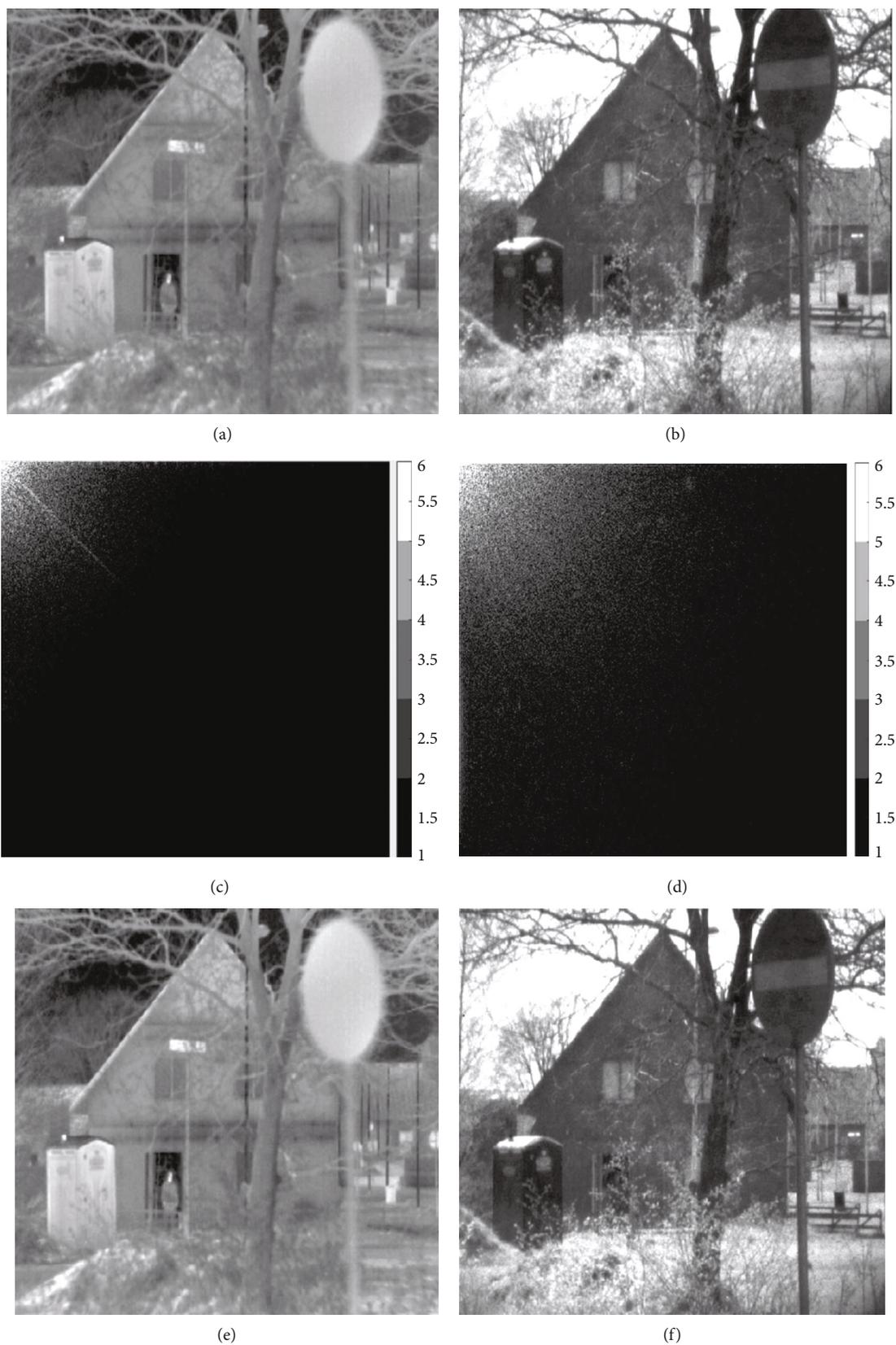


FIGURE 2: Results of denoising by DCT. (a) IR image. (b) VI image. (c) DCT coefficient matrix of the IR image. (d) DCT coefficient matrix of the VI image. (e) IR image after IDCT. (f) VI image after IDCT.

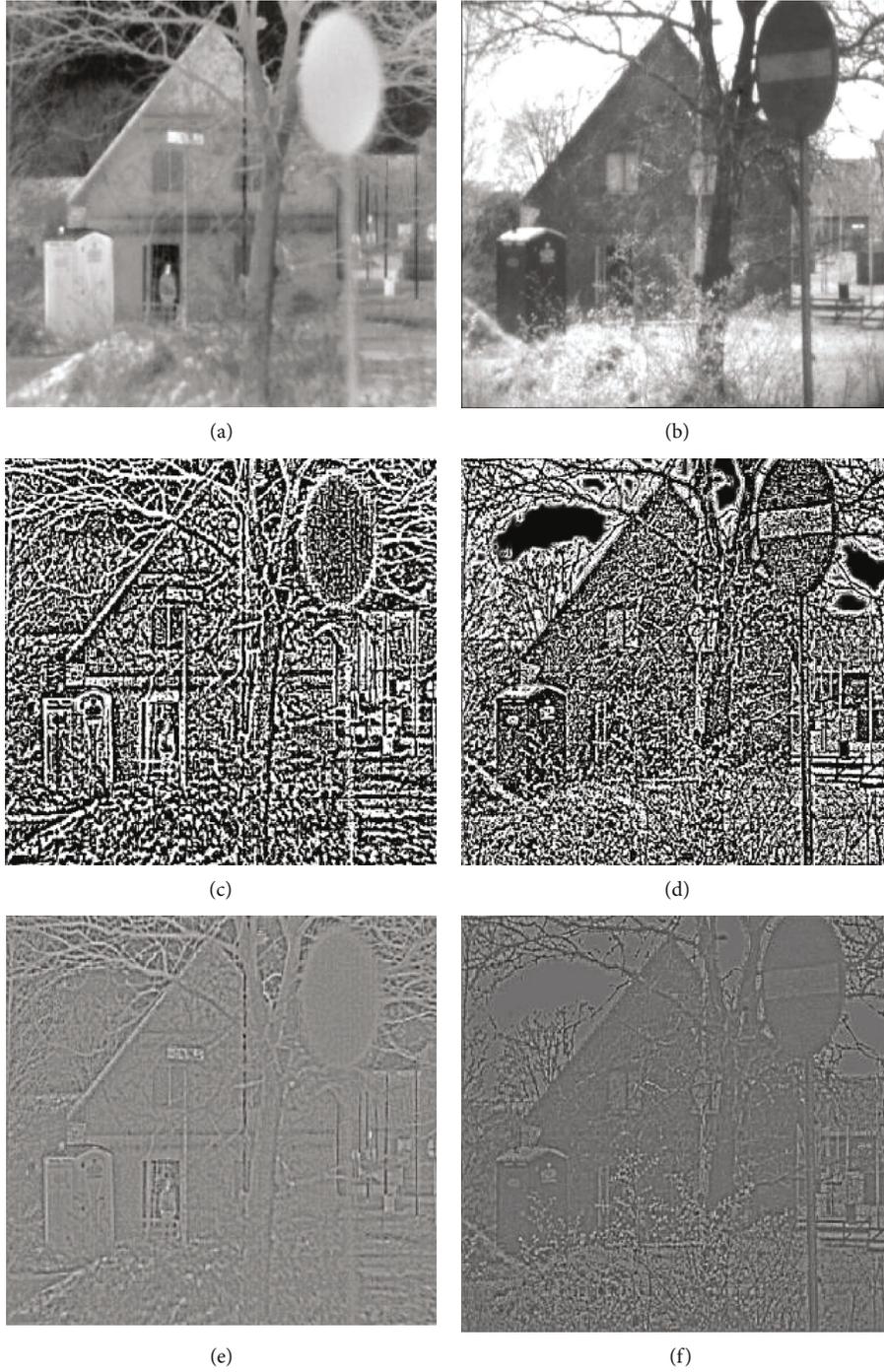


FIGURE 3: (a) IR image. (b) VI image. (c) Detail layer of the IR image. (d) Detail layer of the VI image. (e) Base layer of the IR image. (f) Base layer of the VI image.

Step 5. The fused detail layer I_F^D is calculated using:

$$I_F = I_F^B + I_F^D. \quad (15)$$

$$I_F^D = K_1 I_{IR}^D(x, y) + K_2 I_{VI}^D(x, y). \quad (14)$$

3.3.3. Reconstruction. In this study, the final fused image is obtained by the linear superposition of I_F^B and I_F^D , as shown in equation (14).

4. Experiment

In this section, the coregistered IR and VI images from the TNO image fusion dataset are used to evaluate our algorithm. This database provides a large number of coregistered infrared and visible images. All the experiments are implemented

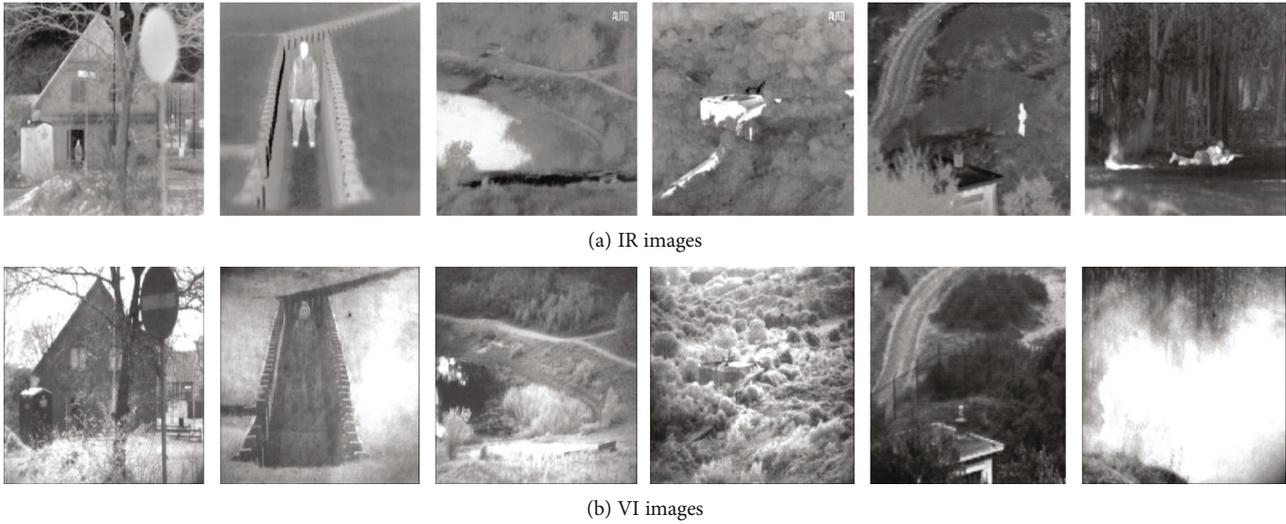


FIGURE 4: Source images used in the experiment.

in MATLAB 2019a on 1.8 GHz Intel(R) Core(TM) i7 CPU with 8 GB RAM. The proposed method is compared with six typical fusion methods: multiscale singular value representation (MSVD) [34], discrete harmonic wavelet transform (DCHWT) [35], discrete wavelet transform (DWT) [21], two-scale fusion (TS) [36], dual-tree complex wavelet transform (DTCWT) [22], and curvelet transform (CVT) [37]. In order to verify the advantages of the proposed method, the experimental verification is divided into two parts. Subjective evaluation results of the fused image are shown in the first part. In the second part, we compare the objective evaluation results of the proposed algorithm with six comparison algorithms. The six pairs of coregistered source images used in this experiment are depicted in Figure 4.

4.1. Subjective Evaluation. The fusion results obtained by the proposed method and six compared methods are shown in Figure 5. The fusion experimental results from pair 1 to pair 6 are represented from top to bottom, respectively.

In order to show a better comparison, the details in the fused images are highlighted with red boxes. As can be seen from Figure 5, our method preserves more detail information and contains less artificial noise in the red window. The image details in fused images obtained by MSVD, DWT, and TS are blurred, which are clearly seen from the first three pairs of the experiment. Compared with the above three fusion methods, the fusion results based on DCHWT preserve more detail information, while showing obvious VI artifacts in them (clearly visible in the last pair of the experiment). The fused images obtained by DTCWT, CVT, and the proposed method could preserve more detail information. Compared with DTCWT and CVT fusion methods, the fusion results of the proposed algorithm look more natural. In the next section, several different objective quality metrics are evaluated to demonstrate the advantages of the proposed method.

4.2. Objective Evaluation. In order to verify the advantages of the proposed method, cross entropy (CE) [38], mutual information (MI) [39], the average gradient (AG) [39], relative

standard deviation (RSD) [39], mean gradient (MG) [39], and running time are used as objective evaluation metrics. CE represents the cross entropy between the fused image and the source image. The smaller the cross entropy, the smaller the difference between the images. MI represents the calculation of mutual information between the fused image and the source image. The larger the value of MI, the higher the similarity between the two images. Calculating the average gradient of the image involves calculating the definition of the image, which reflects the expressive ability of the image to the detail contrast. RSD represents the relative standard deviation of the source image and the fused image, which reflects the degree of deviation from the true value. The smaller the relative standard deviation, the higher the fusion accuracy. The mean gradient represents the definition of the fused image. It refers to the clarity of each detail shadow and its boundary on the image. The objective evaluation results of the 6 pairs of the experiment in the “Subjective Evaluation” section are shown in Tables 1–3, and the best results are highlighted in bold.

The comparison results of the objective evaluation indexes of the first two pairs of experiments are given in Table 1. As seen from Table 1, the proposed method outperforms other fusion methods in terms of all metrics except for CE. Although the CE value of the proposed method is not the minimum value, it is very close to the minimum.

The comparison results of objective evaluation indexes of the third and fourth pairs of the experiment are shown in Table 2. As can be seen from Table 2, the fusion result of the proposed algorithm in this paper contains more information. The proposed method outperforms other methods as regards AG, RSD, MG, and running time. The CE values of the proposed method in the third pair of the experiment are very close to the best values produced by MSVD. In the fourth pair of the experiment, all fusion quality indexes of image fusion generated by the proposed method are the best.

The comparison results of the objective evaluation indexes of the last two pairs of the experiment are given in Table 3. As can be seen from Table 3, MI, RSD, MG, and

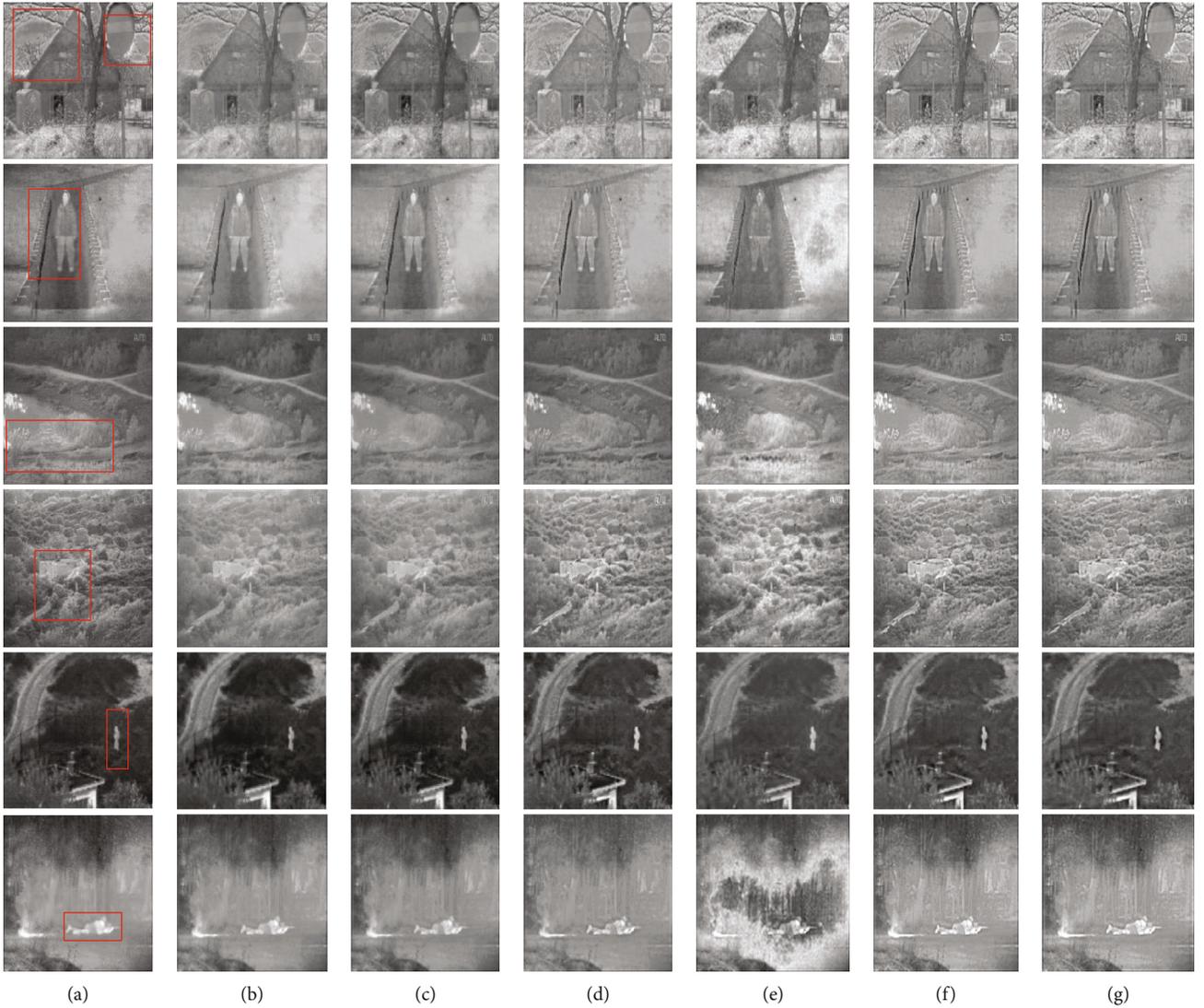


FIGURE 5: Comparison of the fusion results of various methods. (a) Proposed method; (b) MSVD; (c) DWT; (d) TS; (e) DCHWT; (f) DTCWT; (g) CVT.

TABLE 1: Objective evaluation results of the first and the second pair of fusion images.

Method	Pair 1						Pair 2					
	CE	MI	AG	RSD	MG	Time	CE	MI	AG	RSD	MG	Time
Proposed	0.3214	0.1573	4.8756	0.2504	5.7140	0.7371	0.4494	0.3736	4.4364	0.0234	5.5580	0.7742
MSVD	0.5391	0.1384	2.8604	0.5931	3.4728	2.8533	0.8568	0.2637	3.0245	0.1651	3.9973	2.3541
DWT	0.4817	0.1342	3.4970	0.3698	4.2660	32.3347	0.4454	0.2662	2.9412	0.1340	3.9847	21.9847
TS	0.5264	0.1028	3.5827	0.4832	3.9747	1.9536	0.9366	0.2065	2.8596	0.3254	3.3589	0.9680
DCHWT	0.5510	0.0985	4.4494	0.4986	5.0204	5.4881	0.3876	0.2673	3.7730	0.4195	4.6367	5.4581
DTCWT	0.2203	0.1201	4.9425	0.3500	5.6020	0.7506	0.4819	0.2120	4.0197	0.2122	4.9407	0.8207
CVT	0.2456	0.1186	4.9845	0.3488	5.6325	1.5946	0.4815	0.2123	4.0254	0.2123	4.9427	1.3500

running time values of the proposed method are clearly better than that of the other six compared methods. The CE values and AG values of the proposed method in the last two pairs of the experiment are close to the best values separately.

As can be seen from Tables 1–3, MI, RSD, MG, and running time for the proposed method are better than those obtained for other fusion methods. Although the CE value and the AG values of the proposed method are not the best, they are very close to the best. The proposed fusion method

TABLE 2: Objective evaluation results of the third and the fourth pair of fusion images.

Method	Pair 3						Pair 4					
	CE	MI	AG	RSD	MG	Time	CE	MI	AG	RSD	MG	Time
Proposed	0.6179	0.3137	3.6422	0.3473	4.3254	0.7891	0.4129	0.1269	5.2188	0.1213	6.4463	0.907382
MSVD	0.4948	0.3561	1.9869	0.3458	2.3983	2.6334	0.8562	0.1167	3.4208	0.3490	4.5526	2.147643
DWT	1.0674	0.3396	1.9595	0.4984	2.3480	18.1276	0.6749	0.1217	2.8319	0.3167	3.6561	23.632156
TS	1.1629	0.2586	2.4866	0.5446	2.8115	0.861984	0.6039	0.0981	4.1365	0.2013	4.8300	1.765118
DCHWT	0.6763	0.3179	3.2024	0.4503	3.7068	7.161516	0.8383	0.1458	4.6607	0.3401	5.6594	6.912164
DTCWT	0.9942	0.2898	3.4753	0.4383	4.0355	0.8003	0.4626	0.1077	5.2010	0.1486	6.2683	1.0240
CVT	0.8881	0.2885	3.4959	0.4386	4.0508	1.6419	0.4835	0.1098	5.1896	0.1471	6.2543	1.8738

TABLE 3: Objective evaluation results of the fifth and the sixth pair of fusion images.

Method	Pair 5						Pair 6					
	CE	MI	AG	RSD	MG	Time	CE	MI	AG	RSD	RSD	Time
Proposed	0.3292	0.2950	2.5466	0.4469	3.2846	0.752436	1.2609	0.4962	3.3075	0.4887	4.1645	0.865061
MSVD	0.9859	0.2949	3.0297	0.3864	3.8383	3.382514	1.5970	0.4498	2.4004	0.5577	3.1041	3.234588
DWT	0.9207	0.2743	3.4283	0.4230	4.5347	8.195448	1.2014	0.4476	2.2759	0.5613	3.0004	11.950211
TS	0.4923	0.2458	3.3674	0.5271	4.0203	1.811981	4.0630	0.3963	2.2917	0.6697	2.6333	1.063052
DCHWT	0.3756	0.2173	2.6355	0.5991	3.1812	1.047889	2.3590	0.2674	3.9069	0.4493	4.7613	6.561848
DTCWT	0.2432	0.2138	2.5781	0.6611	3.1841	0.8231	2.2847	0.4043	3.8098	0.5344	4.7439	1.0223
CVT	0.2399	0.2276	2.6379	0.6494	3.2569	0.8582	2.0645	0.4057	3.8285	0.5334	4.7663	2.3507

in this paper can guarantee high fusion accuracy and a satisfactory real-time performance, which is suitable for real-time requirements.

5. Conclusions

Data fusion, which is a key technology in IoT, can effectively reduce the amount of data in the communication network and thus reduce the energy loss. In this article, we focused on the fusion of the IR image and the VI image and proposed a novel heterogeneous fusion approach. Considering that DCT has shown good denoising ability and anisotropic diffusion has shown satisfactory detail resolution, we fused the two algorithms through effective fusion strategies. Experimental results show that the proposed method can achieve better performance compared with other six state-of-the-art fusion approaches as regards both subjective and objective indexes.

However, IR and VI images in this experiment are all cor-registered. The actual multisource data may be unregistered. In the future, we will focus on unregistered image fusion.

Data Availability

Experimental data can be found at https://figshare.com/articles/TN_Image_Fusion_Dataset/1008029.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the China Postdoctoral Science Foundation under Grant 2019M653874XB, the National Natural Science Foundation of China under Grant 51804250, the Scientific Research Program of Shaanxi Provincial Department of Education under Grant 18JK0512, the Natural Science Basic Research Program of Shaanxi under Grant 2020JQ-757, and the Doctoral Starting Fund of Xi'an University of Science and Technology under Grants 2017QDJ026 and 2018QDJ017.

References

- [1] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.
- [2] X. Li, M. Liu, C. Deng, P. T. Mathiopoulos, Z. Ding, and Y. Liu, "Full-duplex cooperative NOMA relaying systems with I/Q imbalance and imperfect SIC," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 17–20, 2020.
- [3] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things Communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, 2017.
- [4] F. A. Turjman and S. Alturjman, "5G/IoT-enabled UAVs for multimedia delivery in industry-oriented applications," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 8627–8648, 2020.
- [5] X. Li, J. Li, and L. Li, "Performance analysis of impaired SWIPT NOMA relaying networks over imperfect weibull

- channels,” *IEEE Systems Journal*, vol. 14, no. 1, pp. 669–672, 2020.
- [6] X. Li, Q. Wang, H. Peng et al., “A unified framework for HS-UAV NOMA Networks: performance analysis and location optimization,” *IEEE Access*, vol. 8, pp. 13329–13340, 2020.
 - [7] T. Tang, T. Hong, H. Hong, S. Ji, S. Mumtaz, and M. Cheriet, “An improved UAV-PHD filter-based trajectory tracking algorithm for multi-UAVs in future 5G IoT scenarios,” *Electronics*, vol. 8, no. 10, pp. 1188–1203, 2019.
 - [8] H. Li, S. Liu, Q. Duan, and W. Li, “Application of multi-sensor image fusion of internet of things in image processing,” *IEEE Access*, vol. 6, pp. 50776–50787, 2018.
 - [9] H. Shakhatreh, A. Sawalmeh, A. Al-Fuqaha et al., “Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges,” *IEEE Access*, vol. 7, pp. 48572–48634, 2019.
 - [10] L. Meng, C. Liao, Z. Wang, and Z. Shen, “Development and military applications of multi-source image fusion technology,” *Aerospace Electronic Warfare*, vol. 27, no. 3, pp. 17–19, 2011.
 - [11] D. Fan and Y. Cai, “Research on the fusion of dual-source traffic image based on transform domain and edge detection,” *Infrared Technology*, vol. 39, no. 9, pp. 740–745, 2015.
 - [12] G. Luo, S. Dong, K. Wang, W. Zuo, S. Cao, and H. Zhang, “Multi-views fusion CNN for left ventricular volumes estimation on cardiac MR images,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 1924–1934, 2018.
 - [13] S. Li, H. Yin, and L. Fang, “Remote sensing image fusion via sparse representations over learned dictionaries,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 9, pp. 4779–4789, 2013.
 - [14] Z. Wu, Y. Huang, and K. Zhang, “Remote sensing image fusion method based on PCA and curvelet transform,” *Journal of the Indian Society of Remote Sensing*, vol. 46, no. 5, pp. 687–695, 2018.
 - [15] J. Ma, C. Chen, C. Li, and J. Huang, “Infrared and visible image fusion via gradient transfer and total variation minimization,” *Information Fusion*, vol. 31, pp. 100–109, 2016.
 - [16] N. Aishwarya and C. B. Thangammal, “Visible and infrared image fusion using DTCWT and adaptive combined clustered dictionary,” *Infrared Physics & Technology*, vol. 93, pp. 300–309, 2018.
 - [17] Y. Liu, X. Chen, R. Ward, and Z. J. Wang, “Image fusion with convolutional sparse representation,” *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.
 - [18] A. Toet, “Image fusion by a ratio of low-pass pyramid,” *Pattern Recognition Letters*, vol. 9, no. 4, pp. 245–253, 1989.
 - [19] M. Li, Y. Dong, and X. Wang, “Image fusion algorithm based on gradient pyramid and its performance evaluation,” *Applied Mechanics and Materials*, vol. 525, pp. 715–718, 2014.
 - [20] J. Liu, J. Zhang, and Y. Du, “A fusion method of multi-spectral image and panchromatic image based on NSCT transform and adaptive gamma correction,” in *2018 3rd International Conference on Information Systems Engineering (ICISE)*, pp. 10–15, Shanghai, China, May 2018.
 - [21] H. Kaur and J. Rani, “Image fusion on digital images using Laplacian pyramid with DWT,” in *2015 Third International Conference on Image Information Processing (ICIIP)*, pp. 393–398, Wagnaghat, India, December 2015.
 - [22] A. Kushwaha, A. Khare, O. Prakash, J. I. Song, and M. Jeon, “3D medical image fusion using dual tree complex wavelet transform,” in *2015 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pp. 251–256, Changshu, China, October 2015.
 - [23] M. Amin-Naji and A. Aghagolzadeh, “Multi-focus image fusion using VOL and EOL in DCT domain,” in *Proceedings of the International Conference on New Research Achievements in Electrical and Computer Engineering (ICNRAECE’16)*, pp. 728–733, Tehran, Iran, May 2016.
 - [24] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang, and Y. Zhang, “Multi-Focus image fusion based on spatial frequency in discrete cosine transform domain,” *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 220–224, 2015.
 - [25] M. Amin-Naji and A. Aghagolzadeh, “Multi-focus image fusion in DCT domain based on correlation coefficient,” in *2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, pp. 632–639, Tehran, Iran, November 2015.
 - [26] X. Jin, Q. Jiang, S. Yao et al., “Infrared and visual image fusion method based on discrete cosine transform and local spatial frequency in discrete stationary wavelet transform domain,” *Infrared Physics & Technology*, vol. 88, pp. 1–12, 2018.
 - [27] Y. A. V. Phamila and R. Amutha, “Discrete cosine transform based fusion of multi-focus images for visual sensor networks,” *Signal Processing*, vol. 95, pp. 161–170, 2014.
 - [28] Y. Jiang and M. Wang, “Image fusion using multiscale edge-preserving decomposition based on weighted least squares filter,” *IET Image Processing*, vol. 8, no. 3, pp. 183–190, 2014.
 - [29] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, “Digital photography with flash and no-flash image pairs,” *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 664–672, 2004.
 - [30] M. Alrefaya, “Adaptive speckle reducing anisotropic diffusion filter for positron emission tomography images based on anatomical prior,” in *2018 4th International Conference on Computer and Technology Applications (ICCTA)*, pp. 194–201, Istanbul, Turkey, May 2018.
 - [31] M. Xie and Z. Wang, “Edge-directed enhancing based anisotropic diffusion denoising,” *Acta Electronica Sinica*, vol. 34, no. 1, pp. 59–64, 2006.
 - [32] E. Vakaimalar, K. Mala, and S. R. Babu, “Multifocus image fusion scheme based on discrete cosine transform and spatial frequency,” *Multimedia Tools and Applications*, vol. 78, no. 13, pp. 17573–17587, 2019.
 - [33] D. P. Bavirisetti and R. Dhuli, “Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform,” *IEEE Sensors Journal*, vol. 16, no. 1, pp. 203–209, 2016.
 - [34] B. Liu, W. J. Liu, and Y. P. Wei, “Construction of the six channel multi-scale singular valuedecomposition and its application in multi-focus image fusion,” *Systems Engineering & Electronics*, vol. 37, no. 9, pp. 2191–2197, 2015.
 - [35] B. K. Shreyamsha Kumar, “Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform,” *Signal, Image and Video Processing*, vol. 7, pp. 1125–1143, 2013.
 - [36] D. P. Bavirisetti and R. Dhuli, “Two-scale image fusion of visible and infrared images using saliency detection,” *Infrared Physics & Technology*, vol. 76, pp. 52–64, 2016.
 - [37] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, “Remote sensing image fusion using the curvelet transform,” *Information Fusion*, vol. 8, no. 2, pp. 143–156, 2007.

- [38] S. Rajkumar and S. Kavitha, "Redundancy discrete wavelet transform and contourlet transform for multimodality medical image fusion with quantitative analysis," in *2010 3rd International Conference on Emerging Trends in Engineering and Technology*, pp. 134–139, Goa, India, November 2010.
- [39] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: a survey," *Information Fusion*, vol. 45, no. 2, pp. 153–178, 2019.