IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 14, 2020, accepted February 25, 2020, date of publication March 10, 2020, date of current version March 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2979760

Infrared and Visible Image Fusion Based on Gradient Transfer Optimization Model

RUIXING YU, WEIYU CHEN, AND DAMING ZHOU, (Member, IEEE)
School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China

Corresponding author: Ruixing Yu (yrxgigi@nwpu.edu.cn)

This work was supported in part by the National Nature Science Foundation of China under Grant 61101191, in part by the Aeronautical Science Foundation of China under Grant 20160153001, and in part by the Special Scientific Research Project of Shaanxi Provincial Department of Education under Grant 17JK0599.

ABSTRACT To tackle the problem of partial loss of image details in infrared and visible image fusion, a gradient transfer optimization model is proposed for the fusion of infrared and visible images. Firstly, an adaptive image decomposition method is proposed based on coupled partial differential equation, the infrared image and the visible image are decomposed into base layer and detail layer to extract the high-brightness target and the details of the two images. Based on this superior information of infrared image and visible image, the optimization model is designed to obtain the fusion image obvious target and rich details. For the proposed optimization model, Alternating Direction Method of Multipliers (ADMM) is used to decompose the original model into sub-problems that are easy to solve and iteratively optimize to obtain the optimal solution. The introduction of control parameters makes the model more flexible in different situations, and retains the thermal radiation information of the infrared image and the detailed information of the visible image to the greatest extent. The fused image visual effects and performance indicators are improved. We completed the experiment using a public data set and analyzed the experimental results. The experimental results show that the proposed method can better preserve the clear target and texture information of infrared and visible images, and the fusion results are more accurate and comprehensive. The experiment results also indicate that our method performs well and achieves comparable metric values with the state-of-the-art methods.

INDEX TERMS Image fusion, optimization model, ADMM, partial differential equation.

I. INTRODUCTION

Image fusion is an enhancement technology that plays a very important role in computer vision and image processing. It aims to generate information-rich images by combining images obtained by different types of sensors for subsequent target detection or human eye observation, recognition, classification and the like. The infrared sensor imaging system can penetrate smoke, fog, snow, and has strong anti-interference ability. It also can work at night, but the infrared image only describes the thermal radiation intensity of the object and there are not considerable texture details. The visible light sensor imaging system has high resolution, rich information on the edge and texture of the target, and strong anti-interference ability. However, the image quality is poor at night and in low visibility conditions. The fusion of infrared and visible images provides more comprehensive information for various imaging conditions, and the fusion results play a key role in aviation detection, military reconnaissance, and security monitoring [1].

According to different image fusion levels, the fusion processing can usually be divided into three levels: pixel level fusion, feature level fusion and decision level fusion. Fusion algorithms are generally divided into seven categories according to the theory used, i.e., multi-scale transformation [2], [3], sparse representation [6], neural network [7]–[9], [11], subspace [12], saliency-based methods [13]–[15], hybrid models [2], [16], [17] and other methods [19]–[22]. For the above categories, we have summarized several advanced image fusion algorithms in recent years, as shown in Table 1.

From table 1, the methods based on multi-scale transformation are the most widely studied methods in image fusion. But, inappropriate conversion methods can easily degrade performance [23], and the current fusion rules are mostly simple, so that interference information such as "halo" is introduced to the results [24]. In recent years, neural networks have been applied to the field of image fusion, such as using convolutional neural networks (CNN) to obtain image

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

TABLE 1. State-of-the-art fusion methods.

Families of fusion method	Representation methods	Fusion strategies	Advantages
multi-scale	VSM-WLS [2]	VSM-WLS is an edge-preserving filter that based on the Gaussian and rolling guidance filters.	VSM-WLS reduces halos and preserves information as specific scales.
	NSCT [3]	This paper combines salience detection with NSCT. The object region for the IR image is acquired with saliency detection and a free regions removal. And the primary fused image is generated by NSCT. The final fused image is obtained by integrating the primary fused image with the IR object region.	NSCT makes objects clearly and preserves details and has good ability to locate the object accurately.
	Wavelet [4]	Wavelet method employs explicit luminance and contrast masking models.	Wavelet method is the first to give a principled approach to image fusion from a perceptual perspective.
	MSD [5]	A hybrid-MSD with Gaussian and bilateral filters is proposed to inject the multi-scale IR spectral features into the visible image.	MSD method can capture important multi-scale IR spectral features and separate fine-scale texture details from large-scale edge features, which made the fusion results visually better.
sparse representation	CSR [6]	This is the first paper that applied CSR to multifocus image fusion and multimodal image fusion. A CSR-based image fusion framework is proposed to decompose each source image into a base layer and a detail layer.	CSR improves the detail preservation of SR and has shift invariant.
neural network	CNN [8]	The features learned by high-performing CNN are used to combine the information from visible image and infrared image. And the fusion image is generated by incorporating different ratios of different layers features of the source images and only the salient regions in the infrared image take part in the fusion.	The visible and infrared images are combined by exploiting more low-layer features of the visible image and more high-layer features of the infrared image.
	DenseFuse [9]	DenseFuse is combined by CNN layers and dense block in which the output of each layer is connected to every other layer.	(1)improves flow of information through the network, which makes network be trained easily; (2)gets more useful features from source images; (3)reduces overfitting on tasks.
subspace	RPCA [10]	RPCA takes advantage of PCA and NSCT. RPCA is employed to obtain the sparse matrix(salient information) of the source infrared and visible images. NSCT is for getting the translation invariance.	RPCA can highlight the infrared objects as well as retain the background information in visible image.
saliency	MED-GF [13]	MED-GF adopts phase congruency to obtain the saliency maps of the base and detail layers.	The image fused by MED-GF can be subsequently constructed through the base image and the detail image.
hybrid models	SIDCST [16]	SIDCST combines the advantages of the DTCWT and SF. DTCWT realizes multi-scale partition while the SF accomplishes the directional localization. The low-pass sub-bands are merged with a SR-based fusion approach while the high-pass sub-bands are fused using the 2APCNN(dual-channel adaptive pulse coupled neural network) method.	(1)perfect reconstruction; (2)approximate shift invariance; (3)good directional selectivity.
	MST-SR [17]	MST-SR is an image fusion framework by combining multi-scale transform with sparse representation.	MST-SR method improves the robustness to misregistration and overcomes the tendency of the SR-based method to smooth fine details with MST.
	STV [18]	STV integrates image fusion based on spectral total variation method and image enhancement. Basic layer fusion based on spatial intensity information effectively preserves the energy information of the source images. In addition, TV spectrum and space varying based-subband gains are applied to each fused subband decomposition to effectively enhance the edge details of the fused image.	STV solves the problem of degraded image fusion and has made significant improvements on edge details preservation and contrast of the fused image. But STV can't achieve good results when source images contain considerable noise.
other methods	GTF [20]	GTF is an infrared and visible image fusion model based on gradient transfer and total variation minimization. It formulates the fusion problem as an L1-TV minimization problem, where the data fidelity term keeps the main intensity distribution in the infrared image, and the regularization term preserves the gradient variation in the visible image.	The results of GTF look like sharpened infrared images with more appearance details. The proposed algorithm is able to fuse image pairs without pre-registration.
	FM-AO [21]	FM-AO is a feature fusion method. Alternating operators are used for extracting the multi-scale fusion features. The extracted multi-scale features are combined through the fuzzy measure-based weight strategy to form the final fusion features. The fusion result is produced through importing the final fusion features into the original infrared and visual images using the contrast enlargement strategy.	The fuzzy measure could produce the fusion features with rich spatial information, which would be useful for preserving the details and important regions of the original infrared and visual images in the final fused image. So, FM-AO is a new method to obtain an effective fusion result with clear image details.

features and reconstruct fused images, or build a new deep learning architecture for infrared and visible image fusion problems [23]. The method has good adaptability, fault tolerance and noise immunity, but designing an appropriate neural network and adjusting the corresponding parameters is still a challenging task [24]. Other techniques such as methods based on total variation [20], fuzzy theory [26] and entropy [22], which bring new horizons to the fusion of infrared and visible images. Among them, the fusion algorithms based on total variation model are gradually proposed. Total variation is a mathematical model that has achieved fruitful results in the field of image deblurring, such as Zhuo *et al.* [27] integrated a flash gradient into a maximum-a-posteriori framework to recover a sharp image by combining a blurry image and a corresponding flash image. And Je *et al.* [28] employs disparity information in image segmentation, and estimates a PSF, and restores a latent image for each region. The main difference between our method and these two methods is application background. They apply to image deblurring, while we apply to image fusion. And Son and Zhang [35] proposed a new method for decomposing the image pairs to create a new base layer of the infrared image, which can have visual appearance similar to the denoised image. This algorithm made an outstanding contribution to the denoising theory. Whereas in our algorithm, in addition to denoising and decomposing algorithms, we also propose an optimization model to achieve image fusion by minimizing the optimization equation. In 2016, Ma *et al.* [20] proposed fusion of infrared and visible images based on total variation method. The main idea of the work is to retain the intensity information of the infrared image while retaining the detailed appearance of the visible light image. The fusion result is visually a sharpened infrared image. However, the author ignores part of the information of the visible image, which may result in loss of details in the image. Zhou *et al.* [29] proposed a method called target-aware fusion, which introduced the concept of inner product to strengthen the control of gradient modulus and gradient direction. The fusion result is good and the visual information is richer, however, multiple controls are not absolutely necessary and have an adverse effect on computational efficiency.

Inspired by the above algorithms and based on the superior information of infrared image and visible image, this paper proposes to effectively utilize the difference to design an optimization model of two types of images. The brightness information retention item is designed to retain the general outline of the target brought by the thermal radiation information in the infrared image, and the texture information retention item is designed to more accurately retain the gradient information between the visible image and the infrared image. Using optimization model will avoid the cumbersome problems of designing too many fusion rules in traditional fusion method, and will not introduce unnecessary interference information to fused images. Compared with the neural network or the method based on representation, the design idea is simple and easier to understand and solve. The items in the model are independent of each other and the introduction of various control parameters makes the model more flexible in different situations. Specific effects could be exerted by changing or increasing or decreasing one item. With the flexibility, this paper presents the simplest fusion model, reducing the computational burden, speeding up the fusion, and improving the quality of the fusion.

The main contributions of this study are as follows:

(1) In order to solve the problems of distortions and artifacts in fusion images, an adaptive image decomposition method is proposed based on coupled partial differential equations(CPDE). The coupling coefficient is introduced in CPDE to protect and enhance the edge while avoiding ''block effect''. The introduction of the CPDE decomposition method also can obtain good fusion effect even in the case of image degradation.

(2) We formulate the image fusion as an optimization problem characterized by introducing a brightness information retention item and a texture information retention item and seek the optimal fused image by minimizing an optimization function. Although our method needn't design troublesome fusion rules, it overcomes the problem of missing image details and also obtains a fusion image with a significant target and detailed appearance.

(3) Our method achieves remarkable fusion performance on the ''TNO Image Fusion Dataset''. By comparing with other algorithms, our fusion results have not only the best visual effect (the fusion images look like sharpened infrared images with clear highlighted targets and abundant textures), but also superior performance index.

The rest of this article is arranged as follows. In part 2, we present an optimization model for this algorithm and give a detailed solution for the model. In part 3, we show the experimental results and comparisons with other algorithms. In part 4, we summarize the proposed algorithm.

II. THE PROPOSED FUSION METHOD

In order to obtain more comprehensive information in the fused image, the proposed algorithm expresses the image fusion as an optimization problem. By iterative optimization of the designed objective function, a fused image is obtained which retains both the infrared image thermal radiation information and the visible image detailed appearance information. The framework of our proposed algorithm is shown in Fig.1. In this section, we comprehensively describe the proposed fusion algorithm.

Fig.2 depicts a framework of the algorithm we proposed herein and related algorithms. The following algorithms are similar to ours: GTF [20], ALF [30], TAD-PGF [29]. Inspired by these articles, we proposed the algorithm of this paper. In order to clearly show the difference between our method and the above algorithms, we drew the following flow chart. The blue box in the figure represents the algorithm proposed in this paper. The red box is the GTF algorithm proposed by Ma *et al.* [20], and the green box is the ALF algorithm

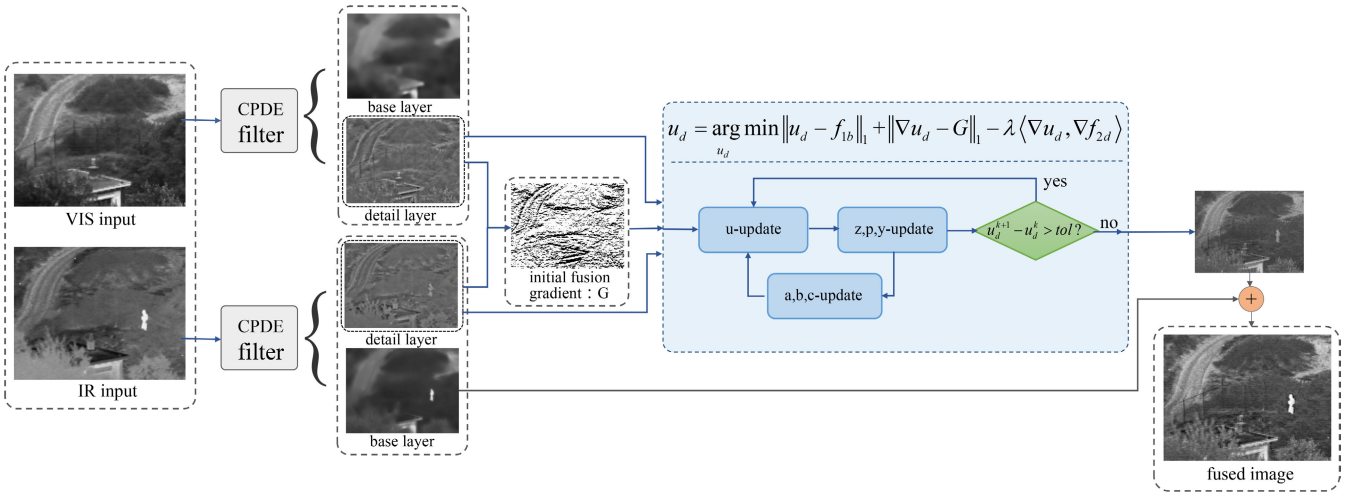

FIGURE 1. Framework of the proposed method.

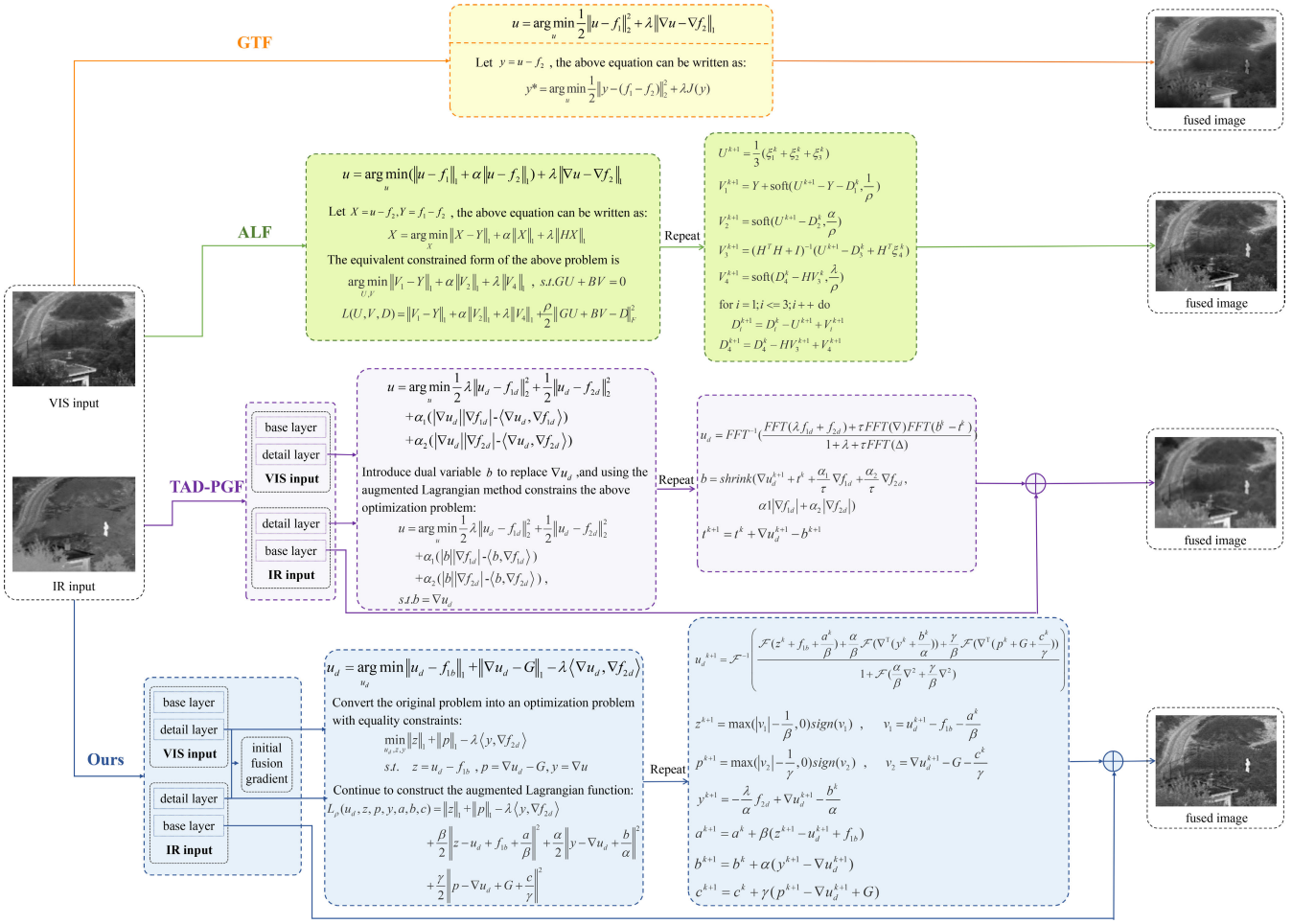

FIGURE 2. The framework of the algorithm we proposed and related algorithms.

proposed by Guo *et al.* [30], and the purple box is the TAD-PGF algorithm proposed by Zhou *et al.* [29].

For readability to readers, the notation table of this article is listed in Table 2.

A. IMAGE DECOMPOSITION BASED ON COUPLED PARTIAL DIFFERENTIAL EQUATIONS

Due to the difference of image feature information between the infrared and the visible image, we can extract these

TABLE 2. The notation table of this article.

f	the original input image
f_b	the base layer of the original image
f_d	the detail layer of the original image
$c(\cdot)$	the diffusion coefficient
g	the diffusion coefficient in the proposed decomposition algorithm
f_{xx}	the second derivative in the x direction of the input image
f_{xy}	the partial derivative in the xy direction of the input image
f_{yy}	the second derivative in the y direction of the input image
f_1	the input infrared image
f_2	the input visible image
u_d	the initial fused image
u	the final fused image
f_{1b}	the base layer of the original infrared image
f_{2b}	the base layer of the original visible image
f_{1d}	the detail layer of the original infrared image
f_{1d}	the detail layer of the original visible image
G	initial fusion gradient
∇	gradient
$\langle\cdot\rangle$	inner product

significant information and analyze them separately, that is, separate the infrared target from the background, and extract the detail information in the visible image. This processing method can highlight thermal infrared targets while preserving image details. This method can be achieved by using the method of decomposing images. In this way, the original image is decomposed into a base layer and a detail layer, and then analyzed separately. For example, the fusion strategy of the base layer and the detail layer are formulated separately, and then the obtained base fusion layer and the detail fusion layer are merged, so as to utilize the source image information as much as possible. In the past work, the decomposition method based on anisotropic diffusion played an important role, such as the traditional anisotropic diffusion algorithm:

$$SPDE(f) = div\left[c(|\nabla f|)\nabla f\right] \tag{1}$$

$$f_b = SPDE(f) \tag{2}$$

$$f_d = f - f_b \tag{3}$$

where $c(|\nabla f|) = \dfrac{1}{1+\dfrac{|\nabla f|^2}{k}}$. The above algorithm effectively preserves edge information while removing unnecessary noise. Anisotropic diffusion is a second-order partial differential equation, which essentially evolves the image toward the segmentation plane. Although it maintains a good effect on the edge of the image, it also causes the image to have a ''block effect'', the image looks like it covers a large mosaic, which adversely affects the subsequent process of the image.

In order to overcome the ''block effect'' brought by the second-order partial differential equation, and take advantage of its good edge-preserving features, we proposed an adaptive decomposition strategy based on Coupled Partial Differential Equations(CPDE), it contains second- and fourth-order partial differential equations. Unlike the second-order partial differential equations, the fourth-order partial differential equations approach the original image by segmented slopes. The area of the grayscale gradation does not transform the image into blocks with different gray values as in the second-order equation. The fourth-order partial differential equations smooth it into a region of grayscale gradation, thereby eliminating the ''block effect''. So, it can avoid the above problems well, the generated image looks more natural, and the artifacts are reduced accordingly. In order to make full use of the advantages of fourth-order model and second-order model, a decomposition method proposed in this paper is as follow:

$$\begin{aligned} CPDE(f) &= \frac{\partial f}{\partial t} \\ &= -g\left[\begin{array}{l}\left(\dfrac{f_{xx}}{|D^2 f|}\right)_{xx} + \left(\dfrac{f_{xy}}{|D^2 f|}\right)_{yx} \\ +\left(\dfrac{f_{yx}}{|D^2 f|}\right)_{xy} + \left(\dfrac{f_{yy}}{|D^2 f|}\right)_{yy}\end{array}\right] \\ &\quad +(1-g)div\left[c\left(|\nabla f|\right)\nabla f\right] \end{aligned} \tag{4}$$

$$f_b^m = CPDE(f^{m-1}), \quad (m = 1 \ldots N) \tag{5}$$

$$f_d^m = f^{m-1} - f_b^m, \quad (m = 1 \ldots N) \tag{6}$$

where g is the coupling coefficient, m is the number of decomposition layers, and f_b^0 is the input image f when $m = 1$. The corresponding second derivative is calculated as follows:

$$(f_{xx})_{i,j}^n = f_{i+1,j}^n - 2f_{i,j}^n + f_{i-1,j}^n \tag{7}$$

$$(f_{yy})_{i,j}^n = f_{i,j+1}^n - 2f_{i,j}^n + f_{i,j-1}^n \tag{8}$$

$$(f_{xy})_{i,j}^n = \frac{f_{i+1,j+1}^n - f_{i+1,j-1}^n - f_{i-1,j+1}^n + f_{i-1,j-1}^n}{4} \tag{9}$$

$$(\left|D^2 f\right|)_{i,j} = \sqrt{(f_{xx})_{i,j}^2 + (f_{xy})_{i,j}^2 + (f_{yx})_{i,j}^2 + (f_{yy})_{i,j}^2} \tag{10}$$

The above decomposition method includes an adaptive decomposition process based on detecting image edges, the coupling coefficient g is selected as the edge detection function $g = \dfrac{1}{1+k|\nabla G_\sigma * f|^2}$, G_σ represents a Gaussian function with variance σ. Gaussian filtering is performed on the image to smooth the effect of noise while protecting edge. At the edge of the image, the smaller g is, the larger $1-g$ is, that is, the coefficient of the fourth-order term is smaller while the coefficient of the second-order term is larger, then, the second-order term plays a major role to protect and enhance the edge. Conversely, in the flat region of the image, the fourth-order term plays a major role, thereby avoiding ''block effect''. The decomposition algorithm proposed in this paper is an edge-preserving algorithm. Pixels with large changes in the gradient are regarded as edges, at this time, the second-order term plays a leading role and performs edge-preserving. Therefore, although the filtering operation is performed on the image first, the edge of the image is still maintained. The effect of the decomposition algorithm is shown in Fig.5. Fig.6 shows the edge retention ability of the decomposition algorithm, while Fig.7 shows the impact of the decomposition algorithm on the fused image. And performance index comparison of the fused image edge strength is shown in Table 3.

TABLE 3. Performance indicators comparisons of EN on figures.

Alg	*Camp*	*Kayak*	*Duine*	*Boat*	*Lake*	*Street*	*Kap*	*Road*
DenseFuse	7.06	**6.93**	**6.81**	5.45	6.99	6.53	6.75	7.08
MDLatLRR	6.25	5.86	5.77	4.88	6.56	5.99	6.54	6.98
VSM-WLS	6.60	6.11	6.11	5.41	7.00	6.72	6.96	7.26
SAF	7.03	6.91	5.97	6.31	7.20	6.71	7.13	7.31
CSR	7.08	6.05	6.22	6.37	**7.31**	6.71	6.51	6.93
FPDE	6.30	6.31	6.27	6.34	7.20	6.67	7.20	7.37
GTF	7.02	6.29	6.34	5.91	7.19	5.42	**7.28**	7.12
Ours	7.08	6.34	6.05	6.39	6.99	6.74	6.94	7.39

B. DESIGN OF FUSION IMAGE OPTIMIZATION MODEL

In this paper, we assume that the given infrared image is completely registered with the visible image. The infrared images, visible images and fused images are both grayscale images of the size $m \times n$, with gray-scale values ranging from 0 to 255.

Inspired by Ma *et al.* [20] which aims to make the fused image look like a sharp infrared image with a clearly prominent target, we present a new fusion optimization model. The design idea of our model is to make the fusion image not only retains the intensity information of the infrared target, but also preserves the rich texture information of the visible image. Such a design can make the fusion result full use of the context information of the target to better accomplish the subsequent target detection and recognition tasks. Due to the obvious thermal radiation target of the infrared image, we expect the infrared image to control the overall contrast and target brightness of the fused image; at the same time, in order to reduce the loss of texture details, the texture of the fused image is obtained from the infrared and visible images. According to this idea, the proposed optimization model includes a contrast information retention item and a texture information retention item. The contrast information retention item constrains the fused image to have a pixel intensity similar to that of the infrared image. The texture information retention item controls the gradient distribution in the original infrared and visible images to be transferred into the fused image. The algorithm model in this paper is as follows:

$$\begin{cases} u_d = \arg\min_{u_d} \|u_d - f_{1b}\|_1 + \|\nabla u_d - G\|_1 - \lambda\langle \nabla u_d, \nabla f_{2d}\rangle \\ u = u_d + f_{1b} \end{cases} \tag{11}$$

1) DESIGN OF CONTRAST INFORMATION RETENTION ITEM

By observing a large number of images, the brightness information of the target is obvious in the infrared image, while that of the visible image is very weak and even difficult to observe the target in some images. In other words, infrared targets are easier to be detected from the background than visible light targets. Therefore, when retaining the overall contrast of the fusion image, we only consider constraining the base layer of infrared image. Then the contrast constraint can be expressed as follows:

$$\xi_1 = \|u_d - f_{1b}\|_1 \quad or \quad \xi_1 = \frac{1}{2}\|u_d - f_{1b}\|_2^2 \tag{12}$$

In the previous work, both L1 and L2 norm constraints appeared. But as pointed out in Reference [30], when the difference between the unknown fused image and the constrained target is Gaussian, it is best to use the L2 norm. The L1 norm specification will be more appropriate when it is a Laplace or pulse distribution. In our problem, most pixels are expected to have similar strength to the source image, with only a small portion being used to provide gradient information, in which case the L1 norm specification is better. Thus, we determine the luminance and contrast information retention term as equation (13):

$$\xi_1 = \|u_d - f_{1b}\|_1 \tag{13}$$

2) DESIGN OF TEXTURE INFORMATION RETENTION ITEM

Our goal is to include more detail information than the source image. In this paper, we describe the detail information using texture features, while the texture features are often represented by the gradient of the image. Since in most cases, for a source image with a complicated scene, the detail information is mainly concentrated in the visible image, whereas the infrared image also has some texture, if only consider the texture information of the visible image, the details may be lost. Therefore, in order to reduce the loss of details and get as much detail information as possible for the fused image, we constrain the gradient of the visible and the infrared image at the same time in the design of the texture information retention item. We take the gradient of the infrared and visible image into account by integrating the gradient information of the above two to obtain an initial fusion gradient map. We control the gradient information and introduce the weight coefficient to control the proportion of the two in different situations:

$$G = \omega_1 \nabla f_1 + \omega_2 \nabla f_2 \tag{14}$$

ω_1, ω_2 is the weight coefficients of the infrared and visible light images, respectively. The weighting factors in this paper are defined as follows:

$$\omega_1 = \frac{|\nabla f_1|}{|\nabla f_1| + |\nabla f_2|}, \quad \omega_2 = 1 - \omega_1 \tag{15}$$

Using the initial fusion gradient field obtained above as the target gradient [31], the gradient of the fused image u_d is close to the target gradient:

$$\xi_2 = \|\nabla u_d - G\|_1 \tag{16}$$

Then, we have basically realized the control of the gradient information of the fused image, but for the gradient information we only consider the gradient magnitude as the constraint. Gradient amplitude constraints may corrupt the pixel intensity distribution, which may result in the loss of small amounts of information or distortion of the original brightness value of the infrared target. Each pixel has its gradient amplitude and gradient direction. In the constraints above, we control the amplitude of the gradient. Next, we add the gradient direction to obtain more accurate constraint information. Inspired by [29], we use the inner product to

further control the direction of the gradient to obtain a more accurate fusion result, which is expressed as follows:

$$\xi_2 = \lambda \langle \nabla u_d, \nabla f_{2d} \rangle \tag{17}$$

where $\langle \nabla u_d, \nabla f_{2d} \rangle$ represents the inner product of the fused image and detail layer of the visible image, and it is used to constraints the gradient of the fused image. The inner product here can be expanded to:$\langle \nabla u_d, \nabla f_{2d} \rangle = |\nabla u_d| \, |\nabla f_{2d}| \cos(\theta_{\nabla u_d} - \theta_{\nabla f_{2d}})$, where $|\nabla u_d|$ is the gradient magnitude of the fused image, $|\nabla f_{2d}|$ is the gradient magnitude of the visible image detail layer, and $(\theta_{\nabla u_d} - \theta_{\nabla f_{2d}})$ is the difference in the gradient direction of the two. That is, when the gradient information of the iterated fused image and the detail layer of the visible image is closer, the inner product will be larger. And even when the gradient information of the fused image obtained in this iteration is consistent with the gradient information of the detail layer of the visible image, this item is 1, and then the fused image obtains complete gradient information.

From the above, the texture information retention item can be written as:

$$\xi = \|\nabla u_d - G\|_1 - \lambda \langle \nabla u_d, \nabla f_{2d} \rangle \tag{18}$$

With the contrast information retention term constraining the pixel intensity, the texture information retention term constrains the gradient magnitude and direction. We get the final optimization function:

$$u_d = \arg\min_{u_d} \|u_d - f_{1b}\|_1 + \|\nabla u_d - G\|_1 - \lambda \langle \nabla u_d, \nabla f_{2d} \rangle \tag{19}$$

where λ is the control parameter and controls the balance between the two. Then we obtained the fused image. In order to further enhance the brightness of the target to facilitate later processing, we add the base layer of the infrared image to the fused image:

$$u = u_d + f_{1b} \tag{20}$$

As a result, we end up expressing the fusion problem as a minimization problem in order to obtain a fused image with significant targets and details. This model was inspired by the fusion models of Ma *et al.* [20] and Zhou *et al.* [29], and their models both achieved excellent results. The method of Ma *et al.* is based on the total variation denoising model, which directly constrains brightness and gradient of the image. The fusion result is visually a sharpened infrared image. But the author ignores part of the information of the visible image, which may result in loss of details in the image. Compared to the model proposed by Ma *et al.*, we first decompose the image before constructing the model, and the constraints on the information increase, so that more accurate control of important information such as targets and textures is achieved. The method of Zhou *et al.* decomposed the image into base layer and detail layer, and introduced the concept of inner product to strengthen the control of gradient modulus and gradient direction. The fusion result has good visual effects, however, multiple controls are not absolutely necessary. Compared to the model proposed by Zhou *et al.*, we applied the CPDE to decompose the image into base layer and detail layer, and at the same time the CPDE is also an edge-preserving filter for image denoising. So, our method is a joint denoising fusion algorithm. In the design of the model, the algorithm in this paper is more streamlined, reducing unnecessary constraints and reducing the computational burden. The model of Zhao *et al.* [18] also provided ideas for the proposed algorithm, after decomposing the source image, they enhanced the edges of the detail layer image and then fused, so they achieved excellent fusion results. In contrast to their method, the decomposition algorithm in this paper is an edge-preserving filter, so even though there is no enhanced edge, a good fusion result is obtained.

The detailed solution of the algorithm is given next.

C. OPTIMIZATION ITERATIVE METHOD BASED ON ALTERNATING DIRECTION MULTIPLIER METHOD

Alternating Direction Method of Multipliers (ADMM) is a computational framework for solving optimization problems [30]. The objective function of the original problem model is equivalently decomposed into several solvable sub-problems, and then solve each sub-problem. Finally, the solution to the sub-problem offers the global solution of the original problem. In the algorithm of this paper, we use ADMM to tackle the issue.

For the original problem (19), we first convert it into an optimization problem with equality constraints:

$$\begin{aligned} &\min_{u,z,y} \|z\|_1 + \|p\|_1 - \lambda \langle y, \nabla f_{2d} \rangle \\ &s.t. z = u_d - f_{1b} \quad , p = \nabla u_d - G, y = \nabla u_d \end{aligned} \tag{21}$$

After obtaining the above ADMM form, we continue to construct the augmented Lagrangian function:

$$\begin{aligned} L_\rho(u_d, z, p, y, a, b, c) = \; & \|z\|_1 + \|p\|_1 - \lambda \langle y, \nabla f_{2d} \rangle \\ & + \frac{\beta}{2} \left\| z - u_d + f_{1b} + \frac{a}{\beta} \right\|^2 \\ & + \frac{\alpha}{2} \left\| y - \nabla u_d + \frac{b}{\alpha} \right\|^2 \\ & + \frac{\gamma}{2} \left\| p - \nabla u_d + G + \frac{c}{\gamma} \right\|^2 \end{aligned} \tag{22}$$

At this point, the solution to the original problem can be obtained. The equation (21) is decomposed into three sub-problems and iterated continuously. In the process of solving each sub-problem, the remaining quantities are regarded as fixed values.

- u-update

$$\begin{aligned} \frac{\partial L_\rho(u_d, z, p, y, a, b, c)}{\partial u_d} = \; & -\beta(z^k - u_d^k + f_{1b} + \frac{a^k}{\beta}) \\ & -\alpha \nabla^T (y^k - \nabla u_d^k + \frac{b^k}{\alpha}) \\ & -\gamma \nabla^T (p^k - \nabla u_d^k + G + \frac{c^k}{\gamma}) \\ = \; & 0 \end{aligned} \tag{23}$$

Perform a Fourier transform on equation (23):

$$\begin{aligned}&\mathcal{F}(z^k + f_{1b} + \frac{a^k}{\beta}) + \frac{\alpha}{\beta}\mathcal{F}(\nabla^T(y^k + \frac{b^k}{\alpha})) \\ &+\frac{\gamma}{\beta}\mathcal{F}(\nabla^T(p^k + G + \frac{c^k}{\gamma})) \\ &= \mathcal{F}(u_d^k + \frac{\alpha}{\beta}\nabla^2 u_d^k + \frac{\gamma}{\beta}\nabla^2 u_d^k)\end{aligned} \tag{24}$$

Thus, performing an inverse Fourier transform to get u_d^{k+1} :

$$u_d^{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(A) + \frac{\alpha}{\beta}\mathcal{F}(B) + \frac{\gamma}{\beta}\mathcal{F}(C)}{1 + \mathcal{F}(D)}\right)$$

where,

$$\begin{aligned}A &= z^k + f_{1b} + \frac{a^k}{\beta} \\ B &= \nabla^T(y^k + \frac{b^k}{\alpha}) \\ C &= \nabla^T(p^k + G + \frac{c^k}{\gamma}) \\ D &= \frac{\alpha}{\beta}\nabla^2 + \frac{\gamma}{\beta}\nabla^2\end{aligned} \tag{25}$$

- z-update
  If $z > 0$,

$$\begin{aligned}&\frac{\partial L_\rho(u_d, z, p, y, a, b, c)}{\partial z} \\ &= 1 + \beta(z^k - u_d^{k+1} + f_{1b} + \frac{a^k}{\beta}) \\ &= 0 \\ z^{k+1} &= -\frac{1}{\beta} + u_d^{k+1} - f_{1b} - \frac{a^k}{\beta} = v_1 - \frac{1}{\beta}, \\ v_1 &= u_d^{k+1} - f_{1b} - \frac{a^k}{\beta}\end{aligned} \tag{26}$$

If $z < 0$,

$$\begin{aligned}&\frac{\partial L_\rho(u_d, z, p, y, a, b, c)}{\partial z} \\ &= -1 + \beta(z^k - u_d^{k+1} + f_{1b} + \frac{a^k}{\beta}) \\ &= 0 \\ z^{k+1} &= \frac{1}{\beta} + u_d^{k+1} - f_{1b} - \frac{a^k}{\beta} = v_1 + \frac{1}{\beta}, \\ v_1 &= u_d^{k+1} - f_{1b} - \frac{a^k}{\beta}\end{aligned} \tag{27}$$

Then we can get z^{k+1} :

$$\begin{aligned}z^{k+1} &= \max(|v_1| - \frac{1}{\beta}, 0)sign(v_1) \\ v_1 &= u_d^{k+1} - f_{1b} - \frac{a^k}{\beta}\end{aligned} \tag{28}$$

- p-update
  The update method of p is the same as z:

$$\begin{aligned}p^{k+1} &= \max(|v_2| - \frac{1}{\gamma}, 0)sign(v_2) \\ v_2 &= \nabla u_d^{k+1} - G - \frac{c^k}{\gamma}\end{aligned} \tag{29}$$

- y-update

$$\begin{aligned}&\frac{\partial L_\rho(u_d, z, p, y, a, b, c)}{\partial y} \\ &= \lambda\nabla f_{2d} + \alpha(y^k - \nabla u_d^{k+1} + \frac{b^k}{\alpha}) \qquad (30) \\ &= 0 \\ y^{k+1} &= -\frac{\lambda}{\alpha}f_{2d} + \nabla u_d^{k+1} - \frac{b^k}{\alpha} \qquad (31)\end{aligned}$$

- a-update

$$a^{k+1} = a^k + \beta(z^{k+1} - u_d^{k+1} + f_{1b}) \tag{32}$$

- b-update

$$b^{k+1} = b^k + \alpha(y^{k+1} - \nabla u_d^{k+1}) \tag{33}$$

- c-update

$$c^{k+1} = c^k + \gamma(p^{k+1} - \nabla u_d^{k+1} + G) \tag{34}$$

This is the solution to the original optimization problem we have worked out. By iterating the above sub-problems to convergence, the optimal solution of the original problem, that is, the detail layer fused image u_d, can be obtained. And the final fusion image can be obtained by $u = u_d + f_{1b}$.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In the above sections, we have introduced the model of the algorithm and its specific implementation process. This section will be devoted to the analysis of the advantages and disadvantages of the proposed method by comparing with other advanced algorithms.

A. EXPERIMENTAL ENVIRONMENT AND PARAMETER SELECTION

This experiment is implemented by MATLAB R2018a. The computer is configured with a main frequency of 3.30GHz, Intel Core CPU, 4GB of memory, and 64-bit win7 operating system.

There is a control parameter λ in equation (19). During the experiment, we found that the fusion effect will change according to the parameter λ. Fig.3 shows the change of performance indicators when different parameters are selected. As shown in the figure, when the values of λ are 0.2, 0.4, and 0.6, respectively, the values of the performance indicators gradually increase, but are lower than the values of $\lambda = 0.8$. When $\lambda = 0.8$, each performance indicator shows an optimal trend. When λ is 1.0 and 1.2, the values change slightly and tend to be stable. At the same time, we observed the fused images obtained through different parameters, a part

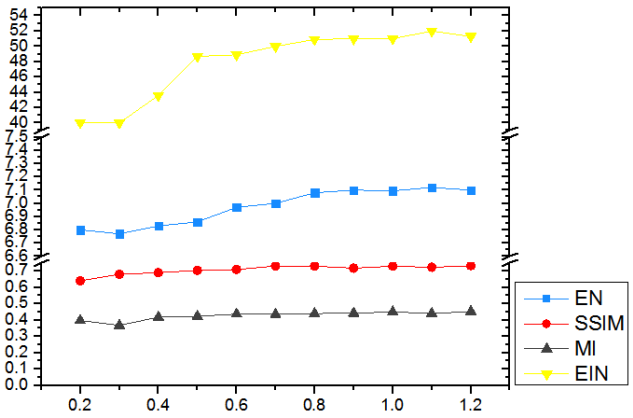

FIGURE 3. Performance indicators with different control parameter. The horizontal axis indicates the value of different parameters.

of fusion results is shown in Fig.4. When we set $\lambda = 0.2$, the fusion result is severely distorted, which only reflects the approximate outline of the fused image, but the target edges are blurred, and a lot of image details are lost, the texture of the fused image cannot be presented. When the value of λ is gradually increased, it can be seen from fig.4(b) that the fusion result is gradually clearer, the texture retained by the image is gradually increased, the contrast is also improved. As shown in fig.4(c), the fused image has clear targets and textures when $\lambda = 0.8$, and the visual effect of the image is more in line with the observation of the human eye. As λ continues to increase, the fused image gradient is more obvious, as shown in fig.4(d), the fused image is visually like an over-sharpened image. After compared all the values of λ, we found $\lambda = 0.8$ is the best. The fused image not only has good visual effect, but also has well performance indicators. So $\lambda = 0.8$ is finally selected, and it kept unchanged for all the images to be fused.

B. EXPERIMENTAL RESULTS AND COMPARISON

We use a common data set to test the proposed algorithm. The dataset used in the experiments is the TNO Image Fusion Dataset, which contains registered infrared-visible image pairs in various conditions. The performance of the algorithm is tested by direct observation and comparison with seven excellent fusion algorithms, they are DenseFuse [9], MDLatLRR [32], VSM-WLS [2], SAF [33], CSR [6], FPDE [19], GTF [20].

1) DECOMPOSITION ALGORITHM RESULTS

- Visual quality
  Using the above decomposition method, we decomposed the infrared image and the visible image into the base layer and the detail layer to perform the next fusion operation more effectively. The decomposed images are shown in Fig.5. We obtained the base layer f_b

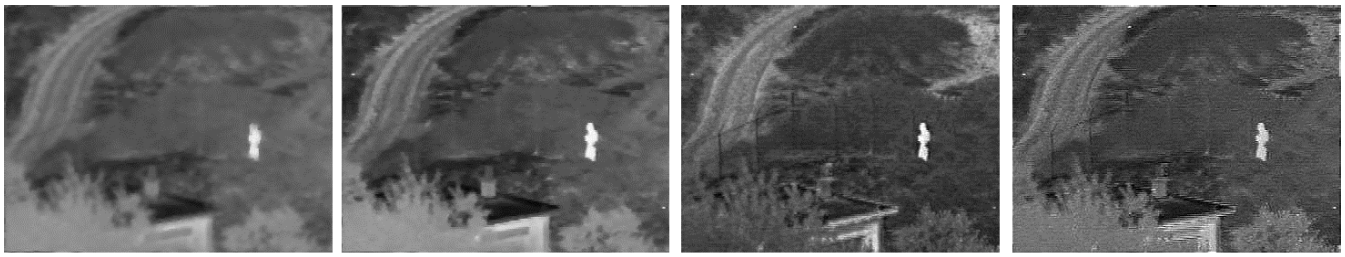

FIGURE 4. From (a) to (d) are the fusion results with different control parameter.

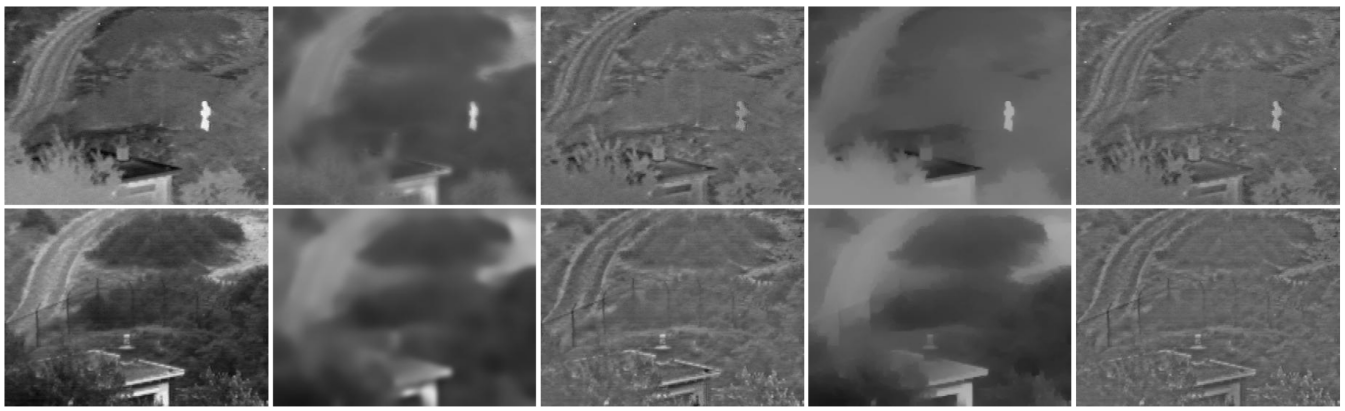

FIGURE 5. Results of decomposition.

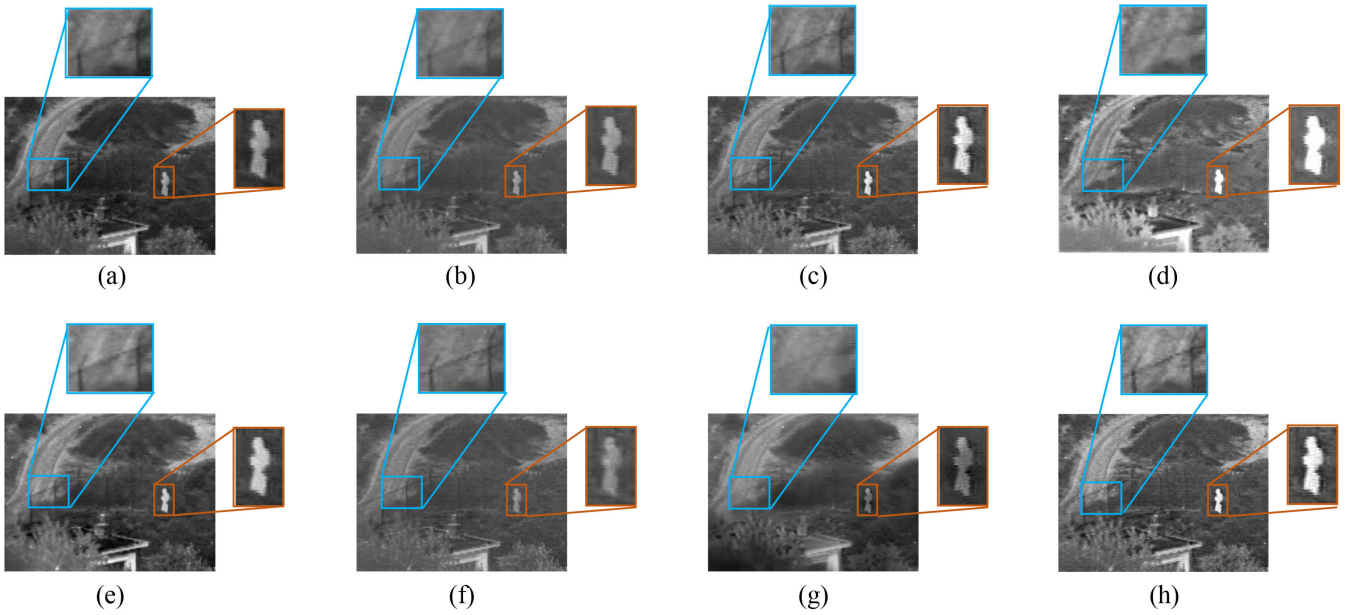

FIGURE 6. Partial enlarged images. (a) DenseFuse, (b) MDLatLRR, (c) VSM-WLS, (d) SAF, (e) CSR, (f) FPDE, (g) GTF, (h) ours.

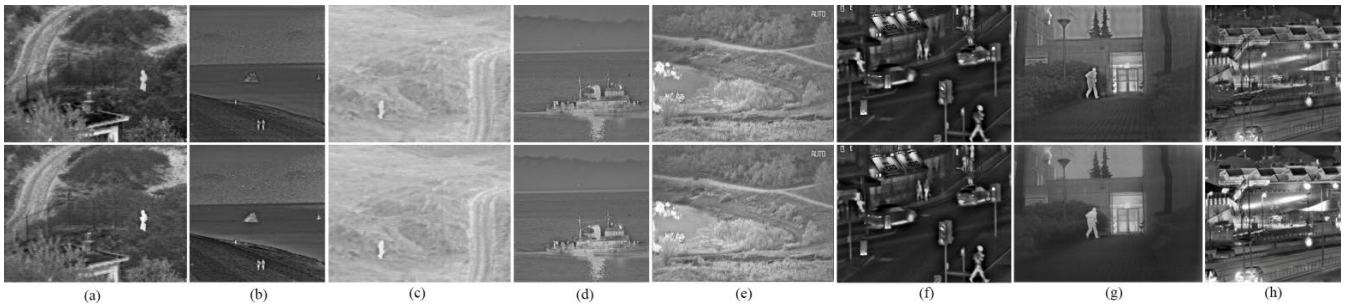

FIGURE 7. From top to bottom are the undecomposed fused images and the fused images after decomposition, respectively.

through CPDE, and then use $f_d = f - f_b$ to obtain the detail layer f_d. In order to facilitate observation, we have increased the gray level of the detail layer image by 0.2. In Fig.5, (a) is source image, from (b) to (e) are the results of decomposition. From top to bottom are infrared images and visible images, respectively. (b) and (c) are the images obtained by the method proposed in this paper, (d) and (e) are the results obtained by the second-order partial differential equation decomposition. (b) and (d) are base layers, (c) and (e) are detail layers. It can be seen from (b) and (d) that the image transition in (b) is more natural, whereas the partial region in (d) has stratification.

It can be seen from the Fig.5 that the base layer images obtained by the two algorithms are significantly different. The images obtained by the algorithm in this paper do not appear ''black shadow'' and the transition is natural, whereas the partial region in (d) has stratification.

Figure 6 shows the edge retention capability of the algorithm in this paper. Figures (a) to (h) are the fusion images obtained by other advanced algorithms and our algorithm. We have intercepted and enlarged important information in the picture (such as the edges of the character and some texture information). The figure shows that the fusion image (d) of the SAF algorithm, the target person has the highest brightness and the most obvious edges, but the texture of the fence and road is not clear. It can be seen that the algorithm makes full use of infrared information. Image (b),(f) and (g) processed by MDLatLRR, GTF and FPDE respectively have a low brightness, and the edges of the characters are not obvious. From the images (c) and (h) we can see, the image information obtained by VSM-WLS and the algorithm proposed has high brightness and clear edges, and the texture of fences and roads is more obvious. It can be seen that although the algorithm in this paper performs filtering on the image, it still retains the edge information of the image.

Fig.7. shows the effect of the decomposition algorithm on the fusion results. The first row is the fusion result obtained without decomposing and recombining, and the second row is the image obtained by decomposing

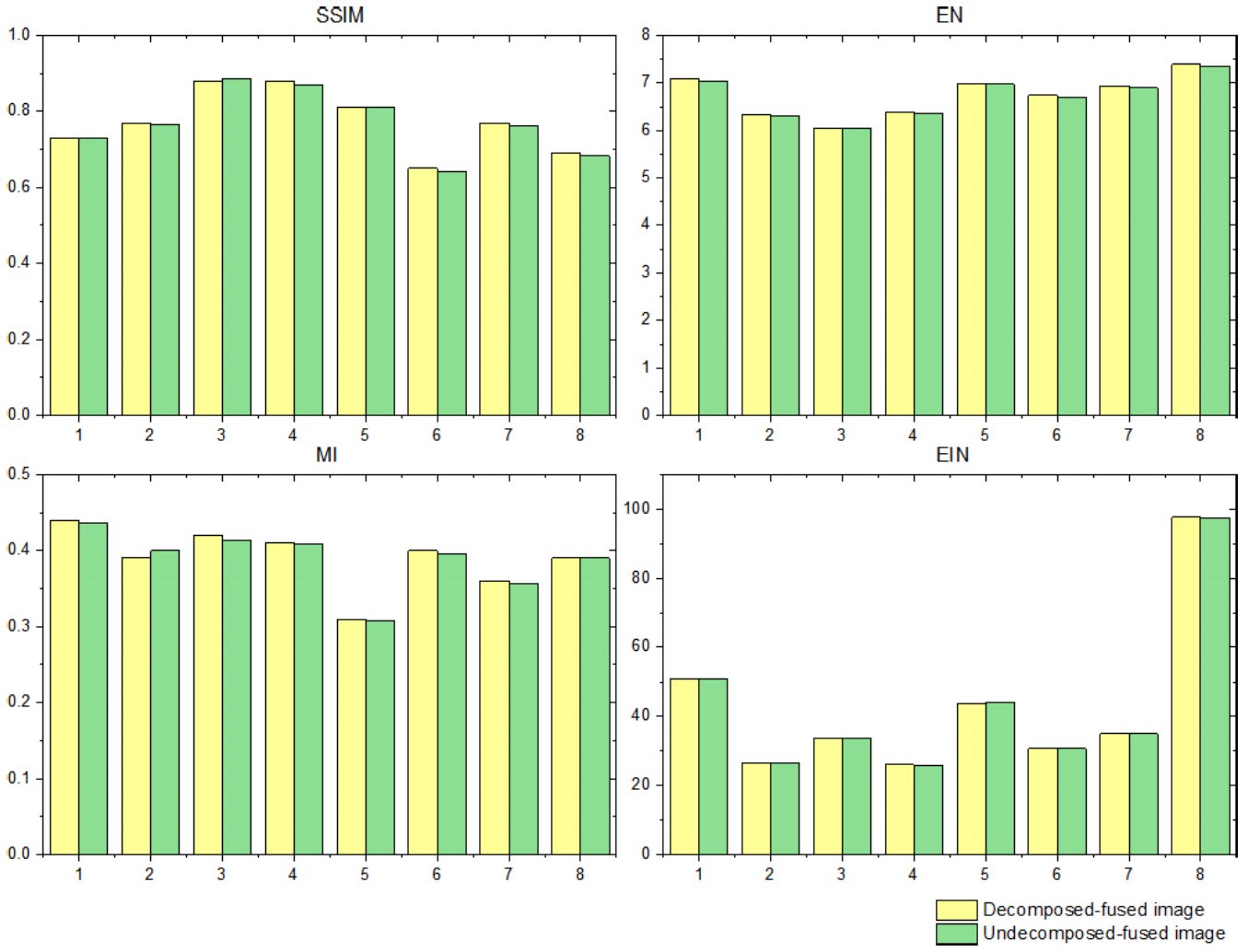

FIGURE 8. Quantitative comparisons of four evaluation indices on the eight image pairs of Nato-camp, Kayak, Duine, Steamboat, Lake, Street, Kaptein, Road (from left to right).

and recombining, and from (a) to (h) are the results on the eight image pairs of Nato-camp, Kayak, Duine, Steamboat, Lake, Street, Kaptein, Road. As can be seen from the figure, the fusion result after the decomposition operation is more natural in image "Nato-camp", and there is no "black shadow". The results of the fusion after the decomposition of image "Kayak" and "Steamboat" are clearer and brighter.

- Performance indicators

  In this text, we measured some objective parameters to evaluate the fusion results. We quantitatively evaluated the performances of different fusion methods using four metrics, i.e., entropy(EN), structural similarity index measure (SSIM) [34], mutual information (MI) [34], and edge intensity(EIN). EN reflects the amount of information contained in the image. The larger the entropy value in the image, the richer the image information contains. SSIM is used to measure the level of similarity between images. The higher the value of the structural similarity, the higher the similarity between the fused image and the corresponding source image reveals. The higher the EIN, the sharper the edges of the image. Fig.8 shows a histogram of the results of this experiment, in which the yellow column represents the experimental parameters of the fused image after using the decomposition algorithm, and the green one is the parameters of the image obtained by directly fusion. It can be seen from the histogram that under the premise of using the optimization model of this paper, the performance indicators of the fused image obtained by using the decomposition algorithm is slightly better than that of the directly fused image.

2) FUSION RESULTS

- Visual quality

  Some samples of the images obtained from the experiment are shown in Fig.9, and from left to right are the results on the eight image pairs of Nato-camp, Kayak, Duine, Steamboat, Lake, Street, Kaptein, Road.

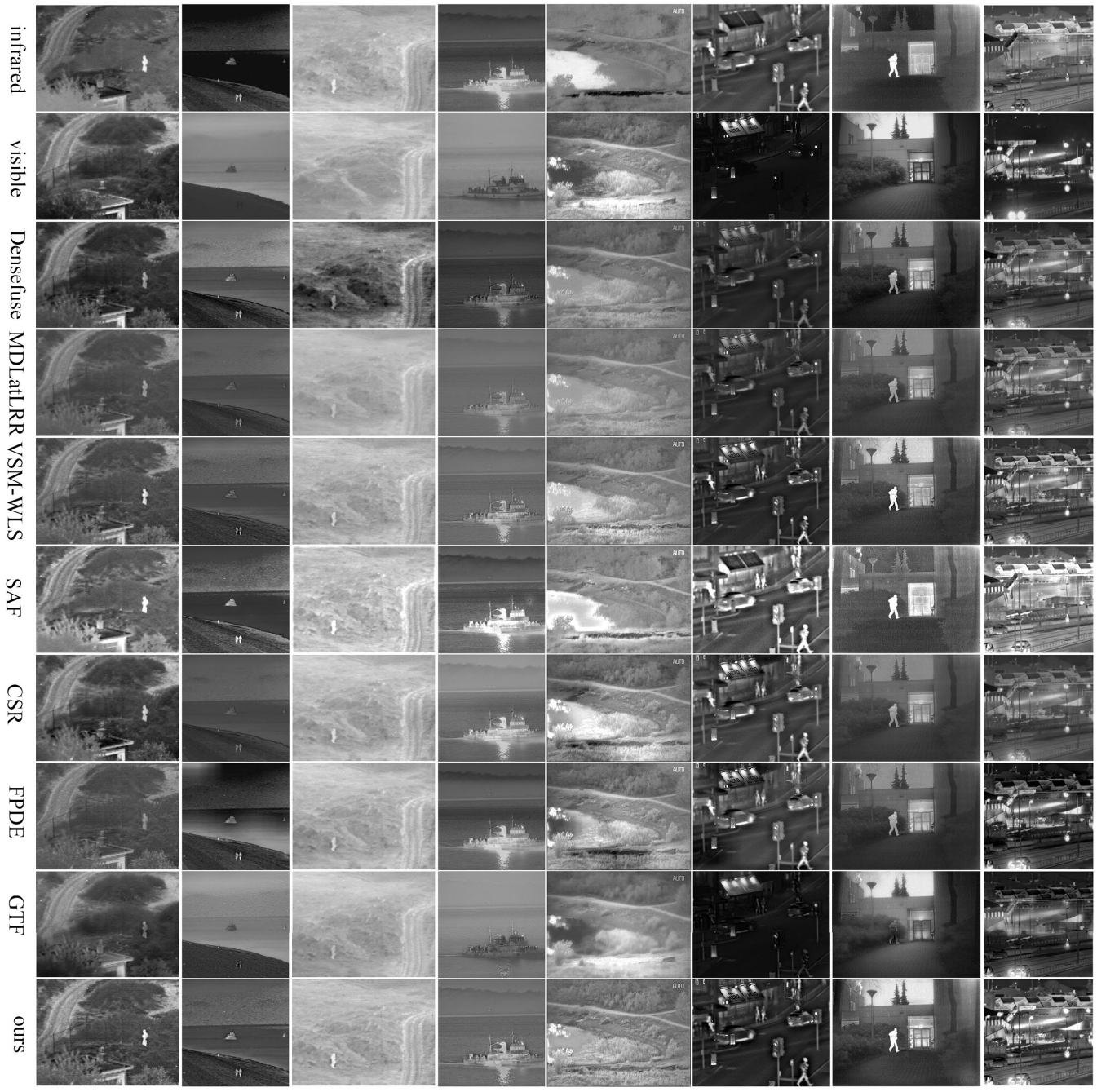

FIGURE 9. Qualitative fusion results. (a) DenseFuse, (b) MDLatLRR, (c) VSM-WLS, (d) SAF, (e) CSR, (f) FPDE, (g) GTF, (h) ours.

We visually observed each set of test images, and the algorithms involved showed their excellent performance. In general, the details of the fusion image based on the Densefuse algorithm are clear, but the overall tone are heavier, especially the ''Duine''. The images obtained by the VSM-WLS and SAF algorithm have large image contrast, the thermal radiation targets have high brightness; but ''black shadows'' appear in some areas in images fused using SAF methods, for example, in image ''Nato-camp'', the transition at the fence is not natural, and this phenomenon also appears in the image ''Lake''. The details of the fused images obtained by the FPDE method are not much lost, but the infrared targets are not significant. The fused images based on the CSR method has better visual effects, with obvious targets and clear textures. The GTF results tend to preserve the thermal radiation distribution of the infrared images, so the infrared target brightness is suitable, but the texture information needs to be further enriched, for example, the image ''Lake'' misses texture information.

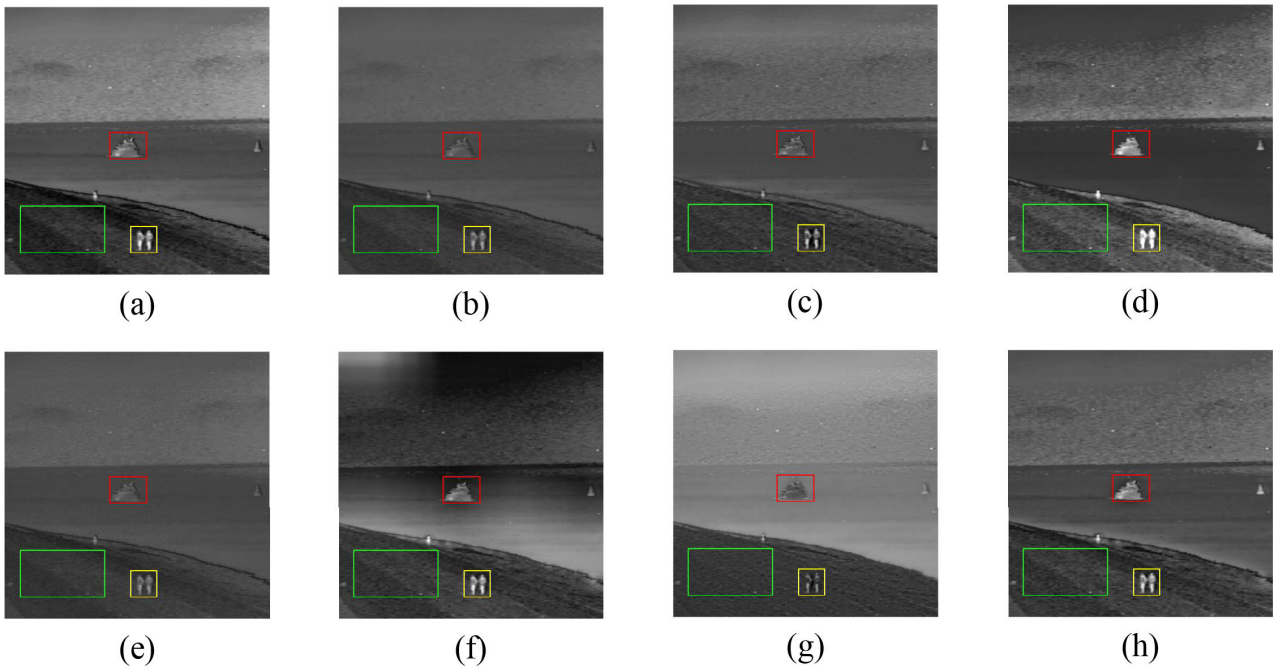

FIGURE 10. Results on "Kayak" images. (a) DenseFuse, (b) MDLatLRR, (c) VSM-WLS, (d) SAF, (e) CSR, (f) FPDE, (g) GTF, (h) ours.

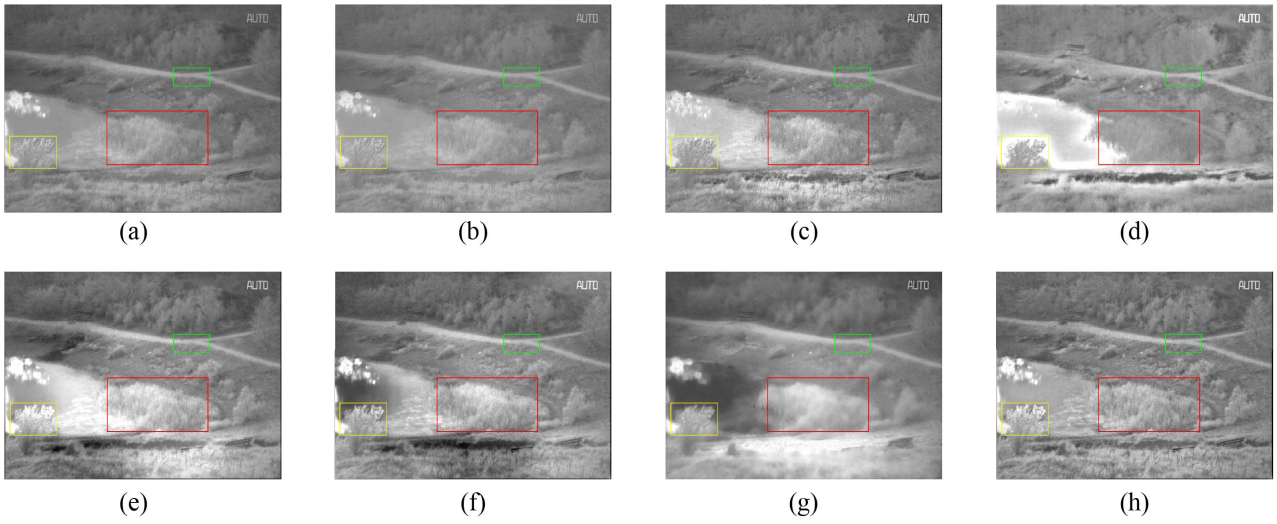

FIGURE 11. Results on "Lake" images. (a) DenseFuse, (b) MDLatLRR, (c) VSM-WLS, (d) SAF, (e) CSR, (f) FPDE, (g) GTF, (h) ours.

The last row is the algorithm proposed in this paper. It can be seen by observing the images "Nato-camp", "Street" and "Road", the contour of the object in the image is clear, and image "Kayak" and "Lake" show that the infrared target is significant; the design goal is basically realized. The experiment achieved satisfactory result.

In general, the images obtained by DenseFuse, VSM-WLS, CSR and ours have better visual effects, retaining important information of infrared and visible images. However, DenseFuse and CSR need to build a network and train it, and VSM-WLS needs to design a fusion strategy. It can be seen that the advantages of the algorithm in this paper are simple design and easy to understand.

Figures 10 and 11 show two representative fusion examples in our experiments. For better comparison, some of the focus areas in the source image and the fused image are labeled with a rectangular box. Fig. 10 shows the fusion of the different methods for the "Kayak" source image. The scenes described contain a wealth of content, including beaches, pedestrians, ships, clouds and so on. Infrared images mainly capture the thermal radiation information of pedestrians and ships, whereas visible images extract most of the spatial details of beaches and clouds. In this example, the ideal fusion result is

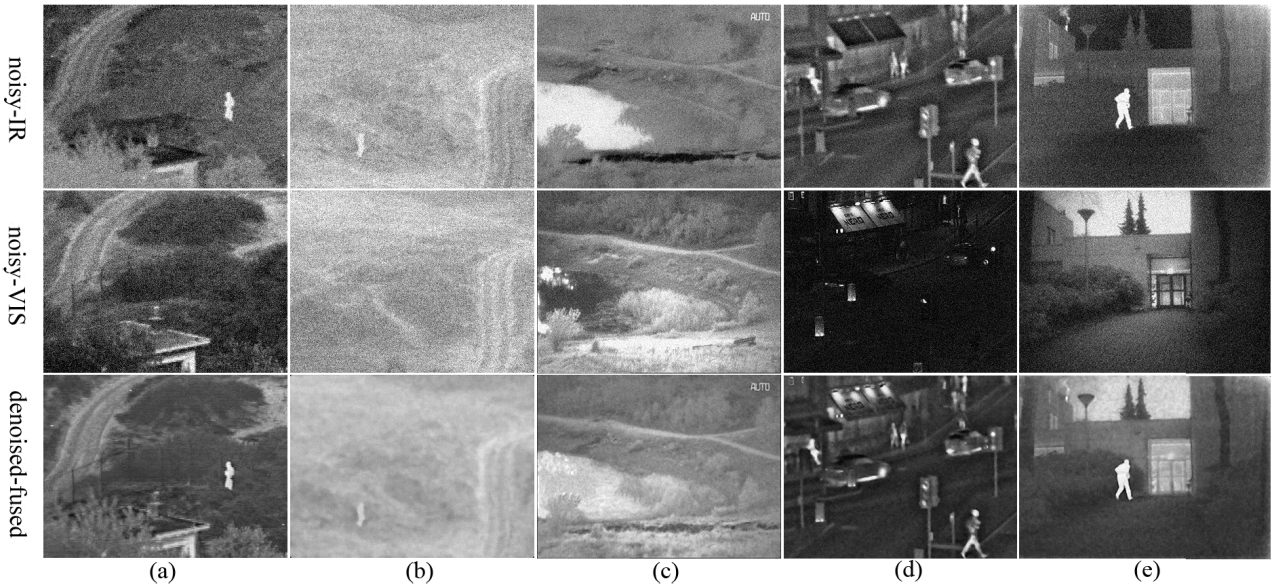

FIGURE 12. Fusion results on noisy images. (a) Nato-camp, (b) Duine, (c) Lake, (d) Street, (e) Kaptein.

that the characters are prominent and the background information is obvious. The fused images based on SAF and FPDE have clear texture details, amd the texture of clouds and beaches is obvious. However, the high contrast of these two methods is also not friendly to visual observation, and appear visual artifacts (such as the sky and the sea in Fig.10(d)(f)). The fused image based on MDLatLRR and GTF methods have relatively low performance in the extraction of thermal radiation information from infrared images, resulting in low brightness for pedestrians and ship targets (pedestrians in Fig.10(b)(g)). The fusion of the DenseFuse method have high quality with clear texture and infrared targets. It can be seen from Fig.10(h) that the proposed method achieves higher performance in detail extraction (such as the stripes on the beach, clouds in the sky and coastline), and the heat radiation target has a higher brightness (such as pedestrians and ships). Fig. 11 shows the fusion results of the different methods for the ''Lake'' source image. The scenes describe information such as lakes, trees, grassland and so on. The SAF- and GTF-based methods are not well completed in this example, the texture of the grassland area in the fused images is not clear(see the fence in Fig.11(d)), and in addition, some of the background areas in the fusion result appear black, which also introduces an unnatural visual experience.The DenseFuse-,SVM-WLS-based and proposed methods have achieved a good visual experience. The lake is clear and the background information is rich. The method proposed in this paper is superior to retain the details of the visible image, and the texture is clearer than other methods (such as the trees and grassland in Fig.11(h)).

Finally, we add Gaussian noise to the source image to verify the stability of our algorithm when the image environment is poor. The experimental results are shown in Fig.12.

It can be seen from the Fig.12 that despite adding noise to the source images, the fused images still obtain clearer target and texture. It shows that our algorithm can complete the fusion work even when the image quality is poor.

- Performance indicators

Observed assessment is more subjective, we also measured some objective parameters to evaluate the fusion results. We quantitatively evaluated the performances of different fusion methods using four metrics mentioned above, that is, EN, SSIM [34], MI [34], and EIN. According to the above four indicators, the image with higher value is better in our test. The experimental results are shown in Table 3 to Table 6. The red numbers indicate that our method performs best; the blue numbers indicate that our result is second only to the first, and the black bold numbers indicate that the relevant method is optimal. It can be seen that the indicators measured by the algorithm proposed in this paper are not optimal for each item, but the overall effect is good and the performance is relatively stable. Among them, the EIN and SSIM of the fused image are better, it shows that the algorithm in this paper retains more source image information and has stronger edge retention ability. It is worth mentioning that DenseFuse also shows extraordinary results, but its framework is cumbersome and the calculation process is complicated. It is not as easy to calculate and understand as the method of this paper.

At the same time, we used the image fusion data set ''Nato-camp'' as a test sequence to verify the versatility of the method, which contains 32 pairs of infrared and visible images. The quantitative results of the test are shown in Fig.13, where our method is represented by a

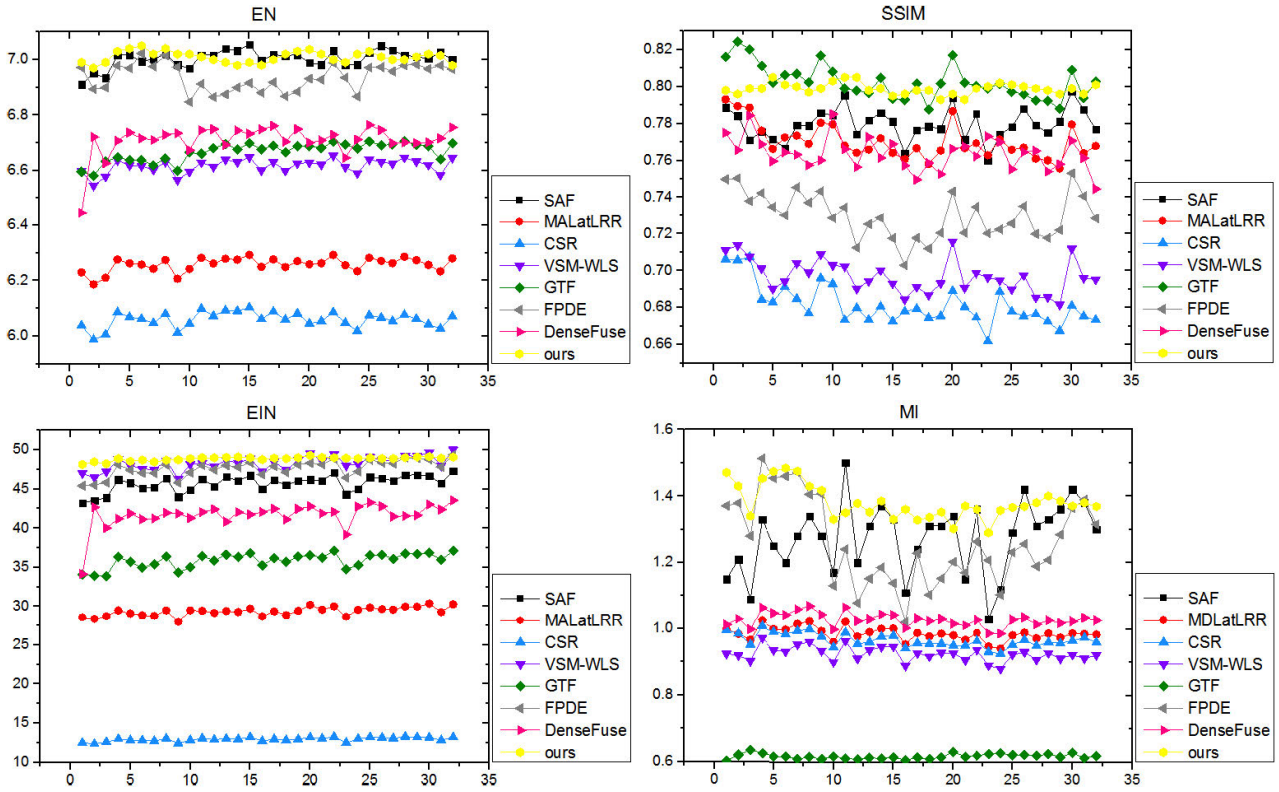

FIGURE 13. Quantitative comparisons of four evaluation indices on the "Nato-camp" sequence. The eight state-of-the-art methods were used for comparison.

TABLE 4. Performance indicators comparisons of SSIM on figures.

Alg	*Camp*	*Kayak*	*Duine*	*Boat*	*Lake*	*Street*	*Kap*	*Road*
DenseFuse	**0.77**	0.76	0.89	0.65	**0.81**	0.60	0.75	0.65
MDLatLRR	0.76	0.73	0.84	0.84	0.79	0.50	0.76	0.58
VSM-WLS	0.72	0.71	0.79	0.73	0.79	0.46	0.69	0.48
SAF	0.75	0.60	0.80	0.57	0.80	0.32	0.63	0.48
CSR	0.76	0.76	0.80	0.72	0.80	0.35	0.76	0.60
FPDE	0.76	0.56	0.80	0.62	0.79	0.38	0.70	0.55
GTF	0.70	0.69	0.83	0.83	0.74	0.59	0.70	0.55
Ours	0.73	0.77	0.88	0.88	0.81	0.65	0.77	0.69

TABLE 5. Performance indicators comparisons of EIN on figures.

Alg	*Camp*	*Kayak*	*Duine*	*Boat*	*Lake*	*Street*	*Kap*	*Road*
DenseFuse	42.24	25.65	**48.36**	26.00	36.80	30.50	29.42	89.99
MDLatLRR	28.76	12.25	18.74	12.30	23.90	23.07	22.92	59.60
VSM-WLS	47.05	21.46	32.56	23.48	42.67	36.31	**40.92**	92.42
SAF	44.47	24.68	25.67	21.63	42.65	36.42	36.59	89.53
CSR	47.47	20.18	29.26	21.09	41.07	**37.89**	20.92	52.67
FPDE	36.14	23.06	29.38	21.39	41.93	37.61	37.74	92.08
GTF	33.20	18.57	24.52	17.39	18.64	20.10	30.53	68.03
Ours	50.91	26.64	33.56	26.03	43.70	30.73	35.02	97.65

TABLE 6. Performance indicators comparisons of MI on figures.

Alg	*Camp*	*Kayak*	*Duine*	*Boat*	*Lake*	*Street*	*Kap*	*Road*
DenseFuse	1.04	**2.64**	0.96	0.60	1.45	0.74	1.78	1.36
MDLatLRR	0.95	1.24	0.91	0.62	1.47	0.41	1.16	1.17
VSM-WLS	0.94	1.40	0.88	0.53	1.46	0.37	1.09	1.09
SAF	1.36	1.38	0.78	0.97	1.16	0.41	1.68	1.28
CSR	1.41	0.91	0.93	0.96	1.46	0.34	1.15	1.23
FPDE	1.01	0.88	1.18	1.08	1.38	0.35	1.12	**1.41**
GTF	**1.53**	1.39	**1.34**	**1.60**	1.53	**0.84**	1.90	1.34
Ours	0.96	1.43	0.74	1.48	1.55	0.83	1.92	1.20

yellow-brown star. The results show that our method is optimal for EN and SSIM. In most cases MI performs well. This shows that our method is generally effective, not only retains important target information in the infrared image, but also retains rich details of the visible image.

IV. CONCLUSION

In this paper, an algorithm for infrared and visible image fusion based on gradient transfer optimization model is proposed. Firstly, a decomposition algorithm based on the coupled partial differential equation is used to decompose the source images, and then we build an optimization model and solve it to obtain a fused image with significant target contour and rich details information. Since the control parameter between the brightness contrast information holding item and the texture information holding item is introduced, the fused image with different effects can be obtained by increasing or decreasing the parameter to control the proportion of the infrared and visible images, and it greatly expands the application scenario of fused images. Experimental results show that the proposed method gives a performance comparable with state-of-the-art fusion approaches, compared with the other eight fusion methods, the visual effects and performance parameters show that the proposed algorithm not only has good visual effects, but also has superior performance indicators.

REFERENCES

[1] Q. Luo, *Study on the Theory of Multi-Sensor Image Fusion and Its Applications*. Wuxi, China: Jiangnan Univ., 2010.

[2] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infr. Phys. Technol.*, vol. 82, pp. 8–17, May 2017.
[3] F. Meng, M. Song, B. Guo, R. Shi, and D. Shan, "Image fusion based on object region detection and non-subsampled contourlet transform," *Comput. Electr. Eng.*, vol. 62, pp. 375–383, Sep. 2016.
[4] P. Hill, M. E. Al-Mualla, and D. Bull, "Perceptual image fusion using wavelets," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1076–1088, Mar. 2017.
[5] Z. Zhou, B. Wang, S. Li, and M. Dong, "Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters," *Inf. Fusion*, vol. 30, pp. 15–26, Jul. 2016.
[6] Y. Liu, X. Chen, R. K. Ward, and Z. Jane Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1882–1886, Dec. 2016.
[7] T. Xiang, L. Yan, and R. Gao, "A fusion algorithm for infrared and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain," *Infr. Phys. Technol.*, vol. 69, pp. 53–61, Mar. 2015.
[8] Y. Liu, X. Chen, J. Cheng, H. Peng, and Z. Wang, "Infrared and visible image fusion with convolutional neural networks," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 16, no. 3, May 2018, Art. no. 1850018.
[9] H. Li and X.-J. Wu, "DenseFuse: A fusion approach to infrared and visible images," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2614–2623, May 2019.
[10] Z. Fu, X. Wang, J. Xu, N. Zhou, and Y. Zhao, "Infrared and visible images fusion based on RPCA and NSCT," *Infr. Phys. Technol.*, vol. 77, pp. 114–123, Jul. 2016.
[11] W. Kong, L. Zhang, and Y. Lei, "Novel fusion method for visible light and infrared images based on NSST–SF–PCNN," *Infr. Phys. Technol.*, vol. 65, pp. 103–112, Jul. 2014.
[12] W. Kong, Y. Lei, and H. Zhao, "Adaptive fusion method of visible light and infrared images based on non-subsampled shearlet transform and fast non-negative matrix factorization," *Infr. Phys. Technol.*, vol. 67, pp. 161–172, Nov. 2014.
[13] W. Gan, X. Wu, W. Wu, X. Yang, C. Ren, X. He, and K. Liu, "Infrared and visible image fusion with the use of multi-scale edge-preserving decomposition and guided image filter," *Infr. Phys. Technol.*, vol. 72, pp. 37–51, Sep. 2015.
[14] X. Zhang, Y. Ma, F. Fan, Y. Zhang, and J. Huang, "Infrared and visible image fusion via saliency analysis and local edge-preserving multi-scale decomposition," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 34, no. 8, pp. 1400–1410, Aug. 2017.
[15] J. Zhao, Y. Chen, H. Feng, Z. Xu, and Q. Li, "Infrared image enhancement through saliency feature analysis based on multi-scale decomposition," *Infr. Phys. Technol.*, vol. 62, pp. 86–93, Jan. 2014.
[16] M. Yin, P. Duan, W. Liu, and X. Liang, "A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation," *Neurocomputing*, vol. 226, pp. 182–191, Feb. 2017.
[17] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Inf. Fusion*, vol. 24, pp. 147–164, Jul. 2015.
[18] W. Zhao, H. Lu, and D. Wang, "Multisensor image fusion and enhancement in spectral total variation domain," *IEEE Trans. Multimedia*, vol. 20, no. 4, pp. 866–879, Apr. 2018.
[19] D. P. Bavirisetti, G. Xiao, and G. Liu, "Multi-sensor image fusion based on fourth order partial differential equations," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Jul. 2017, pp. 1–9.
[20] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Inf. Fusion*, vol. 31, pp. 100–109, Sep. 2016.
[21] X. Bai, "Infrared and visual image fusion through fuzzy measure and alternating operators," *Sensors*, vol. 15, no. 7, pp. 17149–17167, Jul. 2015.
[22] J. Zhao, G. Cui, X. Gong, Y. Zang, S. Tao, and D. Wang, "Fusion of visible and infrared images using global entropy and gradient constrained regularization," *Infr. Phys. Technol.*, vol. 81, pp. 201–209, Mar. 2017.
[23] Q. Zhang, Y. Liu, R. S. Blum, J. Han, and D. Tao, "Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review," *Inf. Fusion*, vol. 40, pp. 57–75, Mar. 2018.
[24] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: A survey," *Inf. Fusion*, vol. 45, pp. 153–178, Jan. 2019.
[25] L. Chang, X. Feng, R. Zhang, H. Huang, W. Wang, and C. Xu, "Image decomposition fusion method based on sparse representation and neural network," *Appl. Opt.*, vol. 56, no. 28, p. 7969, Oct. 2017.
[26] S. Rajkumar and P. C. Mouli, "Infrared and visible image fusion using entropy and neuro-fuzzy concepts," in *Proc. 48th Crit. Infrastruct. Annu. Conv.*, 2014, pp. 93–100.
[27] S. Zhuo, D. Guo, and T. Sim, "Robust flash deblurring," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2440–2447.
[28] C. Je, H. S. Jeon, C.-H. Son, and H.-M. Park, "Disparity-based space-variant image deblurring," *Signal Process., Image Commun.*, vol. 28, no. 7, pp. 792–808, Aug. 2013.
[29] Y. Zhou, K. Gao, Z. Dou, Z. Hua, and H. Wang, "Target-aware fusion of infrared and visible images," *IEEE Access*, vol. 6, pp. 79039–79049, 2018.
[30] H. Guo, Y. Ma, X. Mei, and J. Ma, "Infrared and visible image fusion based on total variation and augmented lagrangian," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 34, no. 11, pp. 1961–1968, Nov. 2017.
[31] F. Fang, *Remote Sensing Image Fusion Based on Variational Methods*. Shanghai, China: East China Normal Univ., 2013.
[32] H. Li, X.-J. Wu, and J. Kittler, "MDLatLRR: A novel decomposition method for infrared and visible image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 4733–4746, Feb. 2020.
[33] W. Li, Y. Xie, H. Zhou, Y. Han, and K. Zhan, "Structure-aware image fusion," *Optik*, vol. 172, pp. 1–11, Nov. 2018.
[34] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Comput. Electr. Eng.*, vol. 37, no. 5, pp. 744–756, Sep. 2011.
[35] C.-H. Son and X.-P. Zhang, "Layer-based approach for image pair fusion," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2866–2881, Jun. 2016.

RUIXING YU was born in Xi'an, China, in 1978. She received the M.S. and Ph.D. degrees in navigation, guidance, and control from Northwestern Polytechnical University (NWPU), Xi'an, in 2003 and 2006, respectively. She was sponsored by the Chinese Scholarship Council to work for one year with the Multimedia Research Group, University of Alberta. She is currently an Associate Professor with School of Astronautics, Northwestern Polytechnical University. Her research interests include video tracking, image recognition, and feature detection and analysis.

WEIYU CHEN received the B.E. degree in measurement-control technology and instrumentation from the North China Institute of Aerospace Engineering, China, in 2017, and the M.S. degree from the School of Astronautics, Northwestern Polytechnical University. Her research interests include image processing and deep learning.

DAMING ZHOU (Member, IEEE) was born in Xi'an, China, in 1989. He received the M.S. degree in navigation, guidance, and control from Northwestern Polytechnical University (NPU), Xi'an, China, in 2013, and the Ph.D. degree in electrical engineering from the University of Bourgogne Franche-Comte, UTBM, Belfort, France.

He is currently a Full Professor with the School of Astronautics, Northwestern Polytechnical University. His research interests include energy management of fuel cell hybrid systems, and power system stability and control.