Infrared and visible image fusion using intensity transfer and phase congruency in nonsubsampled shearlet transform domain

Xin Feng, Haibo Gao, Cheng Zhang and Juanjuan Luo

Hunan International Economics University, Changsha 410205, Hunan, China

Received: 19.06.2022

Abstract. It is known in the field of infrared-and-visible image fusion that infrared targets can be not too prominent while scene-texture details can be insufficient. To solve this problem, we suggest using intensity transfer and phase congruency in the domain of nonsubsampled shearlet transform (NSST). The method first decomposes source images with the NSST to obtain low- and high-frequency subbands. Then the low-frequency subbands are fused using a fusion rule based on the intensity transfer. This enables controlling a transfer of information associated with high-intensity values, i.e. information about important targets and detailed texture. Meanwhile, the phase congruency can be extracted from the source images as an absolute measure of feature significance. This is combined with analysis of directional-vector variance of the high-frequency subband NSST coefficients. In this way the influence of noise in the infrared and low-illumination visible images on the fusion is either eliminated or notably reduced. Then the high-frequency coefficients are found. Finally, a fused image is reconstructed using inverse NSST. The experimental results reported in this work demonstrate that our method can improve efficiently the performance of fusion, since it simultaneously retains the contours and edges of infrared targets and the scene-texture details. This method reveals obvious advantages over its counterparts in terms of both subjective evaluation and quantitative metrics.

Keywords: image fusion, intensity transfer, phase congruency, nonsubsampled shearlet transform

UDC: 004.93

1. Introduction

Image fusion aims to combine the images of the same scene detected with different types of sensors into a single image for subsequent processing of the latter or relevant decision-making. Infrared-and-visible image fusion represents an important branch of the image fusion. It can be used to improve performance of different security-surveillance, target-detection and some other systems. Infrared images capture thermal radiation, which is independent of illumination. Therefore infrared imaging can work efficiently under low-light or bad-weather conditions, although its resolution remains low and noise high. Visible images capture the information associated with visible light. They can clearly reflect many details of a scene under certain conditions but visible imaging is vulnerable to influence of such natural conditions as illumination or weather. It is evident that the two types of image information are complementary and their fusion can bring more information with stronger robustness. Therefore fused images are suitable for human visual perception and have a wide range of applications [1, 2].

According to contemporary theoretical basis, standard fusion methods can be based on such approaches as multi-scale transform, sparse representation, neural networks, subspace representation, saliency detection and hybrid approach. These methods have represented research hotspots at different times. In particular, the multi-scale transform-based method has been elaborated thoroughly. It decomposes a source image through a multi-scale transform into multiple subbands, uses specific rules to fuse each subband and then constructs a fused image. A number of

specific multi-scale transform methods are known, which are associated with wavelet [3], curvelet [4], nonsubsampled contourlet [5] and nonsubsampled shearlet (NSST) [6] transformations. In particular, the NSST has been widely applied in various fields of image fusion [7, 8] due to its shift invariance, high computational efficiency and unlimited number of wavelet directions. The method involves a nonsubsampled Laplacian pyramid and several shearing filters.

In view of many advantages of the NSST, in this work we offer a fusion method implemented in the NSST domain. This method focuses on designing the fusion rules for the coefficients of low- and high-frequency subbands after NSST decomposition. The low-frequency subband concentrates the main energy of any image. In the case of infrared images, targets are mostly represented by the high-intensity pixel values. In order to retain in full the target information, we suggest an infrared-intensity transfer scheme for the low-frequency subband fusion. Regarding the high-frequency subband fusion, one must preserve detailed-texture features and so we offer a fusion rule based on phase congruency. To further avoid the influence of noise in the fusion process, a directional-vector variance is introduced into significance measure for the high-frequency coefficients. This approach employs in full the high-frequency subband NSST coefficients and the corresponding statistical characteristics.

2. Related work

2.1. Intensity transfer

An intensity-transfer scheme has originally been suggested as a fusion method by Li et al. [9]. Here the fusion is considered as a minimization problem:

$$\begin{aligned}\mathbf{x}^* &= \arg\min_{\mathbf{x}} \| \mathbf{x}-\mathbf{u} \|_2^2 + W \| \mathbf{x}-\mathbf{v} \|_2^2 \ \ or \\ &= \arg\min_{\mathbf{x}} \sum\nolimits_p ((\mathbf{x}_p-\mathbf{u}_p)^2 + w_p(\mathbf{x}_p-\mathbf{v}_p)^2),\end{aligned} \tag{1}$$

where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{mn\times 1}$ are the vector forms of respectively infrared and visible images, $\mathbf{x} \in \mathbb{R}^{mn\times 1}$ denotes the vector forms of a fused image, $\|\cdot\|_2$ the l_2-norm, $W \in \mathbb{R}^{mn\times mn}$ the diagonal weight matrix to balance the two terms, and w_p the diagonal element of W at the position p. The latter can be obtained as

$$w_p = |\log(S_p)|, \tag{2}$$

where S_p implies the value of spatial-saliency map at the position p. One can derive it from the image statistics via the relation

$$S_p = \sum\nolimits_{V=0}^{255} N_V (I_p - V). \tag{3}$$

In Eq. (3), N_V denotes the total number of pixels in the image I, of which intensity values are equal to V, while I_p is the intensity value of pixel in the image I at the position p. It is worth noting that Eq. (2) can also be represented in the form

$$W = \left|\log(S)\right|. \tag{4}$$

To solve Eq. (1), one supposes that the derivative of the objective function with respect to $\mathbf{x}$ is zero. Then one can obtain the relation

$$\mathbf{x}^* = (II + W)^{-1}(\mathbf{u} + W\mathbf{v}), \tag{5}$$

with $II \in \mathbb{R}^{mn\times mn}$ representing the unit matrix.

2.2. Phase congruency

Phase congruency is an image-feature perception model suggested by Morrone et al. [10]. It assumes that a feature should be at the maximum point of phase of the image Fourier component. Kovesi [11] has proposed a method for calculating the phase congruency, using a log Gabor filter. It allows for constructing filters with arbitrarily large bandwidth and still maintains a zero (i.e., DC) component in an even-symmetric filter. The phase congruency at the position x can be expressed as

$$PC_2(x) = \frac{\sum_o \sum_n W_o(x) \lfloor A_{no}(x)\Delta\Phi_{no}(x) - T_o \rfloor}{\sum_o \sum_n A_{no}(x) + \varepsilon}, \tag{6}$$

where o means the direction, n the scale, $W_o(x,y)$ the weight function of filtering response, $A_{no}(x,y)$ the amplitude of n-th Fourier component along the direction o, T_o the noise estimate, $\lfloor \cdot \rfloor$ some non-negative value, and ε a small number to avoid division by zero. Note that $\Delta\Phi_{no}(x,y)$ is the phase deviation [12], of which definitions are as follows:

$$\begin{aligned} A_{no}(x)\Delta\Phi_{no}(x) = {} & e_{no}(\mathrm{x})\bar{\phi}_{2o}(x) + o_{no}(x)\bar{\phi}_{1o}(x) \\ & - \left| e_{no}(\mathrm{x})\bar{\phi}_{1o}(x) + o_{no}(x)\bar{\phi}_{2o}(x) \right|, \end{aligned} \tag{7}$$

$$\bar{\phi}_{2o}(x) = \sum_n e_{no}(x) \Big/ E_o(x), \tag{8}$$

$$\bar{\phi}_{1o}(x) = \sum_n o_{no}(x) \Big/ E_o(x), \tag{9}$$

$$E_o(x) = \sqrt{\left(\sum e_{no}(x)\right)^2 + \left(\sum o_{no}(x)\right)^2}, \tag{10}$$

$$e_{no}(x) = I(x) * M_{no}^e, \tag{11}$$

$$o_{no}(x) = I(x) * M_{no}^o. \tag{12}$$

Here $I(x)$ is the input image, and M_{no}^e and M_{no}^o are even symmetric and odd symmetric wavelets, respectively.

The phase congruency represents a dimensionless quantity which is not affected by any changes in image brightness or contrast. It is suitable as an absolute measure of saliency of the image features. Meanwhile, it contains rich texture and structure information, which is consistent with human visual perception [13].

3. Our fusion method

3.1. Fusion steps

This Section introduces a framework of fusion procedures based on intensity transfer and phase consistency in the NSST domain. The basic steps are as follows.

Step 1. NSST decomposing.

Let infrared, visible and fused images be represented respectively as IR, VI and F. The multi-scale and multi-directional decomposition of NSST is performed on IR and VI, and then the NSST coefficients $\left\{L^{IR}(x,y), H_{k,\theta}^{IR}(x,y)\right\}$ and $\left\{L^{VI}(x,y), H_{k,\theta}^{VI}(x,y)\right\}$ can be obtained. Here $L_{k,\theta}^{I}(x,y), I \in \{IR, VI\}$ denote the low-frequency subband coefficients at the position (x,y), while $H_{k,\theta}^{I}(x,y), I \in \{IR, VI\}$ are the directional subband coefficients at the scale k for the direction θ.

Step 2. Fusing low-frequency and high-frequency subbands.

The NSST coefficients of the fused image are obtained according to the fusion rule for the low-frequency subbands, which is based on the intensity transfer, and the fusion rule for the high-frequency subbands, which is based on the phase consistency and the directional-vector variance.

These points will be described in a more detail in Subsections 3.2 and 3.3.

Step 3. NSST reconstructing.

The fused image F is formed in a straightforward manner, using the inverse NSST transform.

3.2. Low-frequency subband fusion rule based on intensity transfer

The main task of infrared-and-visible image fusion is highlighting the targets in images, e.g., persons or weapons. The low-frequency subband obtained by the multi-scale decomposition concentrates the main energy and reflects the basic contours of image features. Therefore the strategy of the low-frequency subband fusion is more essential than that of the high-frequency subbands. In particular, infrared images associated with thermal-radiation imaging can better reflect the contours, the position and the other information about the targets. Any target manifests some smooth textures different from the surrounding environment, which is represented mainly as smooth high-intensity pixel values.

The target information in the visible images cannot be ignored. Even in its low-frequency subband, it still contains rich texture details that must be transferred to the fused image. The most direct scheme is to transfer the pixel intensity with the texture details. It is worthwhile that the target information in the infrared images should be taken into account when transferring the intensity of texture-detail features and prevent interference. In the present work, we improve the intensity-transfer scheme in the low-frequency subband fusion. Namely, we modify the optimization problem as follows:

$$\min_{L^F} J(L^F) = \| vec(L^F) - vec(L^{IR}) \|_2^2 + W \| vec(L^F) - vec(L^{VI}) \|_2^2 , \tag{13}$$

where $vec(\cdot)$ represents the operation of stacking the matrix by columns as vectors. For simplicity, let we have $\mathbf{x} = vec(L^F)$, $\mathbf{u} = vec(L^{IR})$ and $\mathbf{v} = vec(L^{VI})$. Then Eq. (13) can be rewritten as

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} J(\mathbf{x}) = \| \mathbf{x} - \mathbf{u} \|_2^2 + W \| \mathbf{x} - \mathbf{v} \|_2^2 . \tag{14}$$

Eq. (14) has the same form as Eq. (1). In the same way, assuming that the derivative of $J(\mathbf{x})$ is zero, one can obtain $\mathbf{x}$, as shown in Eq. (5). Then it will be reshaped to the shape of the original subband.

In Eq. (5), there is only one unknown parameter, W, which is crucial in the intensity transfer. In Ref. [9], acquisition of spatial-saliency map has been based on image-information statistics, namely Eqs. (2) and (3). However, a lot of calculations are involved in this acquisition. In fact, we have found experimentally that the effect of spatial-saliency map is similar to the effect of enhancing the infrared image. Therefore, in order to accelerate the fusion process and reduce both the noise itself and the influence of that noise on the intensity transfer, we apply a bilateral filter to the infrared image to obtain a new saliency map:

$$S = filter(IR) . \tag{15}$$

After S is normalized, the weight matrix W can be obtained with Eq. (4).

3.3. High-frequency subband fusion rule based on phase congruency and directional-vector variance

After NSST decomposition, the subbands of high-frequency details contain mainly the edge, texture and target-contour information. Since such information should be retained at all costs, its identification is of primary importance. Here a traditional way is comparing the absolute values of

the high-frequency coefficients and taking the larger ones as fused high-frequency coefficients. However, this method has the two problems. The first one is that the edges, the texture and the other features in the source image are usually continuous and smooth, although these features can become discrete and too rough in the high-frequency coefficients, as a result of traditional processing. The other problem is that the existence of noise can perturb selection of the coefficients. As a matter of fact, the noise is inevitable in the infrared images and the low-illumination visible images, which usually corresponds to relatively large absolute values lying inside the high-frequency subbands.

The solution of the first problem adopted in this work is employing phase congruency to extract the edges, the textures and the other features directly from the source images. This can be expressed as

$$P^I(x,y) = PC_2(I(x,y)) \tag{16}$$

where $P^I, I \in \{IR, VI\}$ is the result of phase congruency, of which shape is the same as that of the source image.

The element range of the phase-consistency result is [0,1], where 1 and 0 correspond respectively to significant and insignificant features. This parameter is suitable for comparing the feature significance of different pixels and then selecting the pixels or the high-frequency coefficients with richer information.

To cope with the second problem, we identify the noise and eliminate or reduce its influence on the fusion. Since we use the NSST decomposition, it would be a good choice to take full advantage of the characteristics of NSST coefficients. Because the edges, the texture and the other features of the image reveal directionality, the amplitudes of the high-frequency coefficients along some directions (and the corresponding adjacent directions) will be large. On the other hand, the noise has a different behaviour and, in general, it is non-directional. It can also have a large amplitude but the distribution of each directional coefficient will remain relatively uniform. Therefore, a feature can be distinguished from a noise by judging about the distribution patterns of the high-frequency coefficients at each scale.

The coefficients at the position (x, y) along each direction at the scale j can be extracted to form a composite directional vector [14]:

$$DV_j^I(x,y) = [H_{j,1}^I(x,y), H_{j,2}^I(x,y), \cdots, H_{j,\theta}^I(x,y), \cdots]. \tag{17}$$

Its variance is given by

$$DVAR_j^I(x,y) = \frac{1}{N_\theta} \sum\nolimits_\theta (H_{j,\theta}^I(x,y) - \bar{V}_j^I(x,y))^2, \tag{18}$$

where $\bar{V}_j^I(x,y) = \frac{1}{N_\theta} \sum\nolimits_\theta H_{j,\theta}^I(x,y)$ and N_θ denotes the directional quantity. If the directional-vector variance of a given pixel is large, this pixel belongs to some texture, edge or other feature. Otherwise, it is simply a noise.

In summary, a combination of the phase congruency and the directional-vector variance, PD, can overcome many disadvantages of the traditional fusion rule. This can be expressed as

$$PD_j^I(x,y) = P^I(x,y) \cdot DVAR_j^I(x,y), \tag{19}$$

$$H_{j,\theta}^F(x,y) = \begin{cases} H_{j,\theta}^{IR}(x,y), \left|PD_j^{IR}(x,y)\right| \geq \left|PD_j^{VI}(x,y)\right| \\ H_{j,\theta}^{VI}(x,y), otherwise \end{cases}. \tag{20}$$

It is worth noting that the same evaluation indices $PD_j^{IR}(x,y)$ and $PD_j^{VI}(x,y)$ are used in Eq. (20) for each directional subband at the scale j. In this way one can reduce the amount of calculations needed and, moreover, eliminate the consistency problem which sometimes occurs after the inverse NSST.

4. Experimental results and their analysis

To verify the efficiency of our fusion method, five advanced fusion methods have been used for comparison. These are an original NSST method [15], a gradient-transfer fusion (GTF) [16], a latent low-rank representation (LatLRR) [17], an infrared-feature extraction (IFE) [18], and a fusion based on generative adversarial network (FusionGAN) [19].

To facilitate correct and unbiased comparison, our method and the traditional NSST method involve the same NSST parameters: 3-level decomposition, "9-7" wavelet basis for the multi-scale decomposition and "23-45" wavelet basis for the multi-directional decomposition. Moreover, the numbers of corresponding directions are 8, 8, and 16 in the both cases. Notice that the NSST-based methods use simple but efficient rules [20], an average scheme for the low-frequency coefficients and a maximum-selection scheme for the high-frequency coefficients. In order to obtain easily reproducible data, we have fixed the parameters of the other fusion methods at the default values described in the relevant original works [16–19].

(a) (b) (c) (d)

Fig. 1. Infrared (upper row) and visible (lower row) images of standard image pairs 'quad' (a), 'kayak' (b), 'soldier' (c) and 'bunker' (d).

Four groups of infrared and visible images have been used for our fusion-performance comparison. These are the standard pairs 'quad', 'kayak', 'soldier' and 'bunker' (see Fig. 1). The former two pairs come from the page http://www.imagefusion.org and the last two can be found in the TNO dataset (https://figureshare.com/articles/TNO_Image_Fusion_Dataset/1008029).

The fused images are shown in Fig. 2. It is evident that all the methods under comparison, including our method, highlight successfully the infrared targets and retain the necessary background information. Nonetheless, different methods impose different effects. In the first group of 'quad' images, the brightness of background and target and the contrast obtained by the NSST-based method are low, the background information is blurred, and the contours of target are not clear enough. A similar situation also occurs with the images fused by the GTF and FusionGAN

methods. Loosely speaking, the image quality obtained by the LatLRR and IFE methods is significantly improved but there are still some deficiencies in the target-contours clarity and the background details, if compared with the fused image due to our method. This can be clearly observed at least on the billboard, the remote pedestrians, the vehicles and the trash bins.

NSST GTF LatLRR

IFE FusionGAN Our method

(a)

NSST GTF LatLRR

IFE FusionGAN Our method

(b)

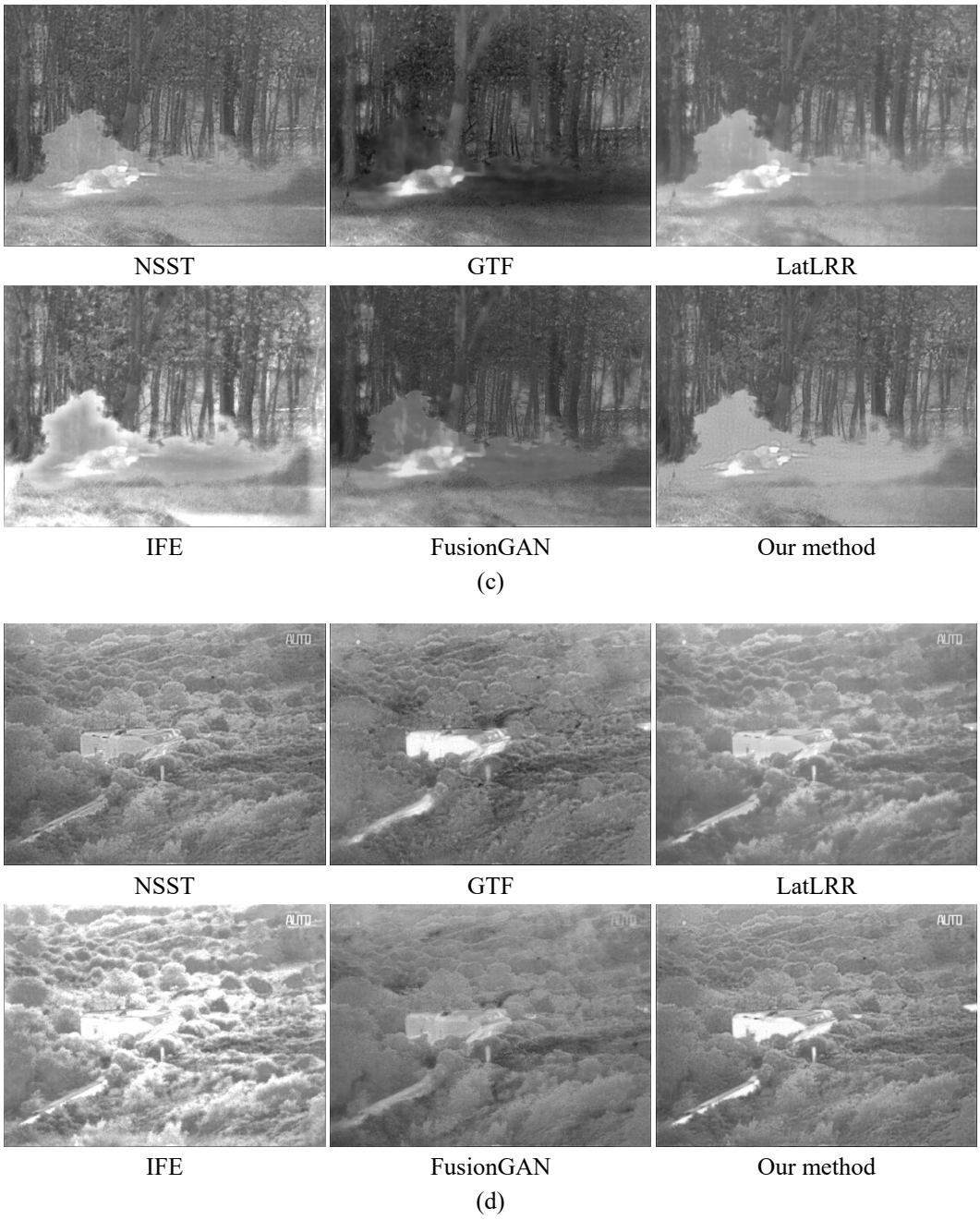

Fig. 2. Fused images obtained using different methods indicated in the legend. Panels a, b, c and d correspond respectively to the image pairs ‘quad’, ‘kayak’, ‘soldier’ and ‘bunker’ (see Fig. 1).

To facilitate a more detailed analysis, the information on the remote pedestrian in the fused ‘quad’ images is extracted and enlarged in Fig. 3. It is seen from Fig. 3 that the fused image obtained using our method has obvious advantages in both target highlighting and edge preserving. The results of comparison of the fusion effects for the other three image pairs are basically similar to those mentioned above for the first image pair. In a word, our method reveals outstanding performance with respect to preservation of details and its fusion quality is the best.

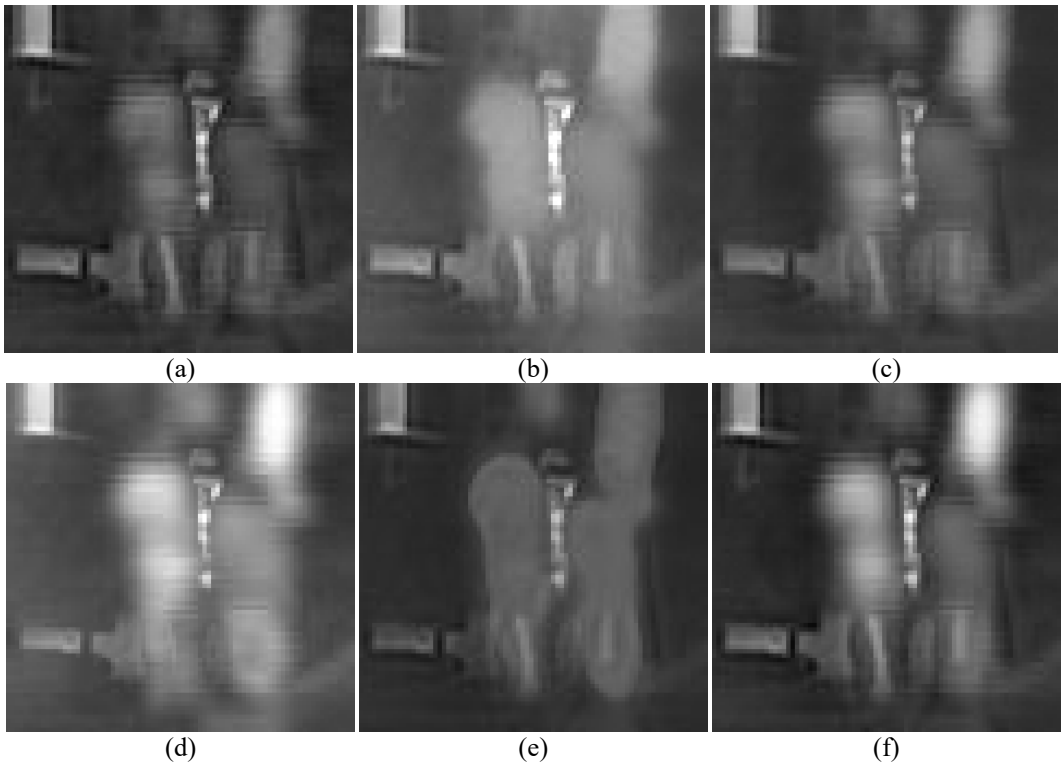

Fig. 3. Data of local enlargement of a remote pedestrian seen in the fused 'quad' images, as obtained using the NSST (a), GTF (b), LatLRR (c), IFE (d) and FusionGAN (e) methods and our method (f).

To verify the resources of our method further on, we use such canonical quantitative metrics of the fused images as a standard deviation (SD), a spatial frequency (SF), a structural similarity (SSIM) [21], a normalized mutual information (NMI) [22] and a feature mutual information (FMI) [23]. The relevant results are gathered in Table 1. The SD and SF can directly evaluate fused images and reflect the appropriate amount of detailed information. The SSIM quantizes the amount of information transferred from a source image to a fused one. In this work, we first derive the metric values corresponding to infrared, visible and fused images and then average them to obtain the resulting value. Finally, the NMI and the FMI also represent useful evaluation indicators for the fusion results.

Note also that each of the evaluation metrics mentioned above assesses the image quality from a different perspective. Therefore a large single-metrics value would not be enough to indicate superiority of a given fusion method. Moreover, since all of the fused images inspected by us are about the same scene and include the same target features and textures, the metrics values for different methods should be close to each other. Then multiple metrics must be simultaneously considered when evaluating the fusion results. In Table 1, both optimal and suboptimal values for every metrics are marked in bold. It is obvious that our method is superior with respect to the other methods in terms of the objective evaluation and in the viewpoint of the basics of statistics.

Let us compare in detail the fusion methods, using all the metrics and taking the first image pair as example. It is seen from Table 1 that the optimal values for the SF, SSIM and FMI metrics correspond to our method. Regarding the SD, our method obtains suboptimal value, though it is close enough to the optimal one. Finally, the performance of our method proves to be mediocre from the viewpoint of the NMI. Note that the similar conclusions can also be drawn in the case of fusion of the rest of three image pairs. Summing up, our analysis demonstrates superiority of our method in terms of the overall fusion quality.

Table 1. Results of quantitative evaluation of the fused ‘quad’, ‘kayak’, ‘soldier’ and ‘bunker’ images obtained by different methods.

Fused image	Metrics	NSST	GTF	LatLRR	IFE	FusionGAN	Our method
‘quad’	SD	0.091	0.117	0.109	**0.146**	0.082	**0.125**
	SF	**0.048**	0.037	0.037	0.044	0.031	**0.048**
	SSIM	0.668	0.618	**0.671**	0.649	0.653	**0.674**
	NMI	0.323	**0.429**	0.323	**0.953**	0.370	0.377
	FMI	**0.921**	0.917	0.918	0.920	0.907	**0.921**
‘kayak’	SD	0.129	**0.164**	0.143	**0.182**	0.125	0.148
	SF	0.043	0.036	0.034	**0.051**	0.030	**0.043**
	SSIM	0.749	0.720	**0.766**	0.697	0.732	**0.751**
	NMI	0.303	0.307	0.331	**0.400**	0.314	**0.387**
	FMI	0.884	0.877	0.877	**0.888**	0.869	**0.886**
‘soldier’	SD	0.142	0.121	0.143	**0.183**	0.130	**0.147**
	SF	**0.054**	0.046	0.033	**0.050**	0.046	0.048
	SSIM	**0.688**	0.628	0.709	0.665	0.652	**0.685**
	NMI	**0.269**	0.230	0.263	0.228	0.240	**0.342**
	FMI	**0.923**	0.901	0.897	0.892	0.898	**0.922**
‘bunker’	SD	0.111	**0.121**	0.112	**0.156**	0.097	0.118
	SF	**0.056**	0.051	0.035	**0.060**	0.042	0.050
	SSIM	0.655	0.629	**0.689**	0.595	0.655	**0.655**
	NMI	0.193	0.161	**0.217**	0.211	0.168	**0.215**
	FMI	**0.927**	0.917	0.894	0.883	0.899	**0.922**

In addition, we have conducted experiments with the multiplier k involved in the weight matrix:

$$W = k \cdot \left|\log(S)\right|, k = 1, 2, \cdots . \tag{21}$$

Fig. 4 shows the appropriate partial fusion results derived by our method for the case of ‘bunker’ image pair as an example. As seen from Fig. 4, the information about details in the image increases with increasing k, and the same is true of the overall brightness. Table 2 lists the results of quantitative evaluation of these fused images. It becomes evident that all of the SD, SF, NMI and FMI metrics improve with increasing k, although to different degrees. On the other hand, the SSIM then decreases slightly but maintains basically the same value. Hence, a better fusion effect can be achieved by setting the appropriate k parameter.

(a) $k = 2$ (b) $k = 4$ (c) $k = 6$ (d) $k = 8$

Fig. 4. Fused ‘bunker’ images obtained by our method at different *k* values (see the text).

Table 2. Quantitative evaluation of the fused images shown in Fig. 4 obtained by our method.

Metric	$k = 1$	$k = 2$	$k = 4$	$k = 6$	$k = 8$
SD	0.117	0.131	0.147	0.156	0.161
SF	0.050	0.051	0.052	0.052	0.052
SSIM	0.654	0.658	0.654	0.651	0.648
NMI	0.214	0.296	0.389	0.437	0.465
FMI	0.922	0.924	0.927	0.928	0.928

5. Conclusion

In the present work we have suggested a method for the infrared-and-visible image fusion, which is based on the intensity transfer and the phase congruency. The intensity transfer considers the low-frequency subband NSST fusion as a minimization problem. To solve this problem, the information about high-intensity values in the two fused images can be retained in full by controlling a balance between the salient target information involved in the infrared image and the detail-texture information in the visible image. Since the phase congruency is not affected by either image brightness or its contrast, it can be used as a measure of significance of any image feature and, therefore, as a measure for selecting the high-frequency NSST-subband coefficients. Considering a need in eliminating or reducing the influence of noise, we have also introduced the directional-vector variance as a factor for selecting the high-frequency subband coefficients.

In our experimental studies, six up-to-date fusion methods have been employed to fuse four standard image pairs. The corresponding data has been compared with each other. The results have testified that our method can preserve successfully the infrared-target and texture-detail information. According to both the subjective evaluation and the objective quantitative metrics, this method is superior to the well-known NSST, GTF, LatLRR, IFE and FusionGAN methods.

Acknowledgments

This work was supported by the Key scientific research project “Research on image data recognition method based on in-depth learning” of Hunan Education Department (Hunan Education Tong [2021] 352, No. 21A0585).

References

1. Ma J, Ma Y and Li C, 2019. Infrared and visible image fusion methods and applications: A survey. Inform. Fusion. **45**: 153–178.
2. Jin X, Jiang Q, Yao S, Zhou D, Nie R, Hai J and He K, 2017. A survey of infrared and visual image fusion methods. Infrared Phys. Techn. **85**: 478–501.
3. Zhang Z and Blum R S, 1999. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proc. IEEE. **87**: 1315–1326.
4. Starck J L, Candès E J and Donoho D L, 2002. The curvelet transform for image denoising. IEEE Transact. Image Processing. **11**: 670–684.
5. Peng H, Li B, Yang Q and Wang J, 2021. Multi-focus image fusion approach based on CNP systems in NSCT domain. Comp.Vis.Image Underst., **210**: 103228.
6. Easley G, Labate D and Lim W Q, 2008. Sparse directional image representations using the discrete shearlet transform. Appl. Comp. Harm. Analysis. **25**: 25–46.
7. Wang L, Dou J, Qin P, Lin S, Gao Y, Wang R and Zhang J, 2021. Multimodal medical image

fusion based on nonsubsampled shearlet transform and convolutional sparse representation. Multim. Tools and Appl. **80**: 36401–36421.
8. Chu T, Tan Y, Liu Q and Bai B, 2020. Novel fusion method for SAR and optical images based on non-subsampled shearlet transform. Intern. J. Rem. Sens. **41**: 4590–4604.
9. Li Z and Yan H, 2019. Infrared and visible image fusion via intensity transfer and direct matrix mapping. Infrared Phys. Techn. **102**: 103030.
10. Morrone M C and Owens R A, 1987. Feature detection from local energy. Patt. Recog. Lett. **6**: 303–313.
11. Kovesi P, 1999. Image features from phase congruency. Videre: J. Comp. Vis. Res. **1**: 1–26.
12. Zhan K, Li Q, Teng J, Wang M and Shi J, 2015. Multifocus image fusion using phase congruency. J. Electron. Imag. **24**: 033014.
13. Liu Z, Feng Y, Chen H and Jiao L, 2017. A fusion algorithm for infrared and visible based on guided filtering and phase congruency in NSST domain. Opt. Las. Engin. **97**: 71–77.
14. Zhang Q and Guo B, 2009. Multifocus image fusion using the nonsubsampled contourlet transform. Signal Proc. **89**: 1334–1346.
15. Defa H, Hailiang S and Weijin J, 2018. Infrared and visible image fusion based on shiftable complex directional pyramid transform and SUSAN edge detector. Ukr. J. Phys. Opt. **19**: 199–210.
16. Ma J, Chen C, Li C and Huang J, 2016. Infrared and visible image fusion via gradient transfer and total variation minimization. Inform. Fusion. **31**: 100–109.
17. Li H and Wu X J, 2018. Infrared and visible image fusion using latent low-rank representation. Preprint arXiv:1804.08992.
18. Zhang Y, Zhang L, Bai X and Zhang L, 2017. Infrared and visual image fusion through infrared feature extraction and visual information preservation. Infrared Phys. Techn. **83**: 227–237.
19. Ma J, Yu W, Liang P, Li C and Jiang J, 2019. FusionGAN: A generative adversarial network for infrared and visible image fusion. Inform. Fusion. **48**: 11–26.
20. Wang W W, Shui P L and Song G X. 2003. Multifocus image fusion in wavelet domain. Proc. 2003 IEEE International Conference on Machine Learning and Cybernetics. **5**: 2887–2890.
21. Wang Z, Bovik A C, Sheikh H R and Simoncelli E P, 2004. Image quality assessment: from error visibility to structural similarity. IEEE Transact. Image Proc. **13**: 600–612.
22. Liu Z, Blasch E, Xue Z, Zhao J, Laganiere R and Wu W, 2011. Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative study. IEEE Trans. Pattern Analysis Machine Intellig. **34**: 94–109.
23. Haghighat M B A, Aghagolzadeh A and Seyedarabi H, 2011. A non-reference image fusion metric based on mutual information of image features. Comp. Electr. Eng. **37**: 744–756.

Xin Feng, Haibo Gao, Cheng Zhang and Juanjuan Luo. 2022. Infrared and visible image fusion using intensity transfer and phase congruency in nonsubsampled shearlet transform domain. Ukr.J.Phys.Opt. **23**: 215 – 227. doi: 10.3116/16091833/23/4/215/2022

Анотація. *У галузі злиття інфрачервоних і видимих зображень відомо, що інфрачервоні об'єкти можуть бути не надто помітними, а деталі текстури сцени – недостатніми. Щоб вирішити цю проблему, ми пропонуємо використовувати передавання інтенсивності та конгруентність фази в області несубдискретизованого шерлетового перетворення (NSST). Метод спочатку розкладає вихідні зображення за допомогою NSST для отримання низько- та високочастотних піддіапазонів. Потім низькочастотні піддіапазони зливають*

за допомогою правила злиття, заснованого на передаванні інтенсивності. Це дає змогу контролювати передавання інформації, пов'язаної зі значеннями високої інтенсивності, тобто інформації про важливі об'єкти та детальну текстуру. Водночас, конгруентність фази можна одержати з вихідних зображень як абсолютну міру значущості ознаки. Це поєднують з аналізом відхилення напрямного вектора для коефіцієнтів NSST високочастотних піддіапазонів. У такий спосіб усувають або значно зменшують вплив шумів в інфрачервоних зображеннях або видимих зображеннях зі слабким освітленням на дані злиття. Потім знаходять високочастотні коефіцієнти. Нарешті, злите зображення реконструюють за допомогою оберненого NSST. Експериментальні результати, наведені в цій роботі, демонструють, що наш метод може значно поліпшити ефективність злиття завдяки одночасному збереженню контурів і країв інфрачервоних цілей і деталей текстури сцени. Цей метод засвідчує очевидні переваги над аналогами з точки зору і суб'єктивних оцінок, і кількісних показників.