Fusion of infrared and visible images by a combined guided-filter and total-variation method

Hongjun Guo, Lili Chen* and Zhenggao Pan

Laboratory of Intelligent Information Processing, Suzhou University, Suzhou 234000, Anhui, China luoxuemiwu@foxmail.com

Received: 18.01.2022

Abstract. We offer a method for fusion of infrared and visible images, which is based on combination of guided-filter and total-variation methods. For extracting and keeping gradient information of the two types of images to a maximal degree and preserving thermal targets, a guided filter is used at the first stage in order to construct two-scale image decomposition and apply it to the source images. This is done for acquiring the layers of base and details, where the details layers can retain the gradient information in advance. Then a total-variation model is applied. By constraining the fused base layer to have similar intensity distribution with the base layer of the infrared image and similar gradient information with the base layer of the visible image, one derives the fused image which can simultaneously preserve both the thermal-target and gradient information. Both subjective and objective evaluations of our experimental data indicate that the combined method suggested in this work has a superior performance to a number of known methods.

Keywords: image fusion, infrared images, guided filters, total variation

UDC: 004.932

1. Introduction

Image fusion aims to generate a new image from multi-sensor data, which can provide complementary information about a scene [1, 2]. We remind in this respect that an infrared image is associated with thermal radiation emitted from a material, which is less affected by illumination conditions [1]. In contrast, a visible image contains detailed appearance information about a scene. It is sensitive to the frequency-dependent reflection of objects, especially in a low-illumination environment. Therefore, the fusion is imperative for subsequent image processing, e.g., in intelligent surveillance, target detection and environmental perception [3].

Many fusion methods have been suggested up to now. Among them, the fusion based on multi-scale transform is one of the most popular methods. The multi-scale transform generally decomposes a source image to obtain one approximate low-pass subband and several high-pass ('detail') subbands for a subsequent fusion procedure [1, 2]. The canonical multi-scale transform includes the methods of Laplacian pyramid (LP), discrete wavelet transform (DWT), curvelet transform (CVT), non-subsampled shearlet transform et al. The multi-scale transform-based fusion methods can provide excellent performance in many fusion-related situations, although sometimes they introduce unexpected effects such as ringing artefacts and, moreover, these methods are time-consuming. Furthermore, the fused images often contain different information whenever the fusion of infrared and visible images is dealt with. Then it would not be proper to use the same fusion strategies for different image features [3]. It is known that infrared images supply a lot of target information, which often manifests itself in the form of high intensities. Meanwhile, detailed appearance information in visible images can be described by high gradients. Therefore, the fusion process should simultaneously preserve both the intensity distribution and the gradient information.

*Corresponding author

To this end, a number of fusion methods based upon a total-variation model have been suggested [3, 4]. They use different representations for different types of the source images. In particular, Ma et al. [4] have offered a fusion strategy that implies direct transfer of the intensities of infrared image and the gradients of visible image into a fused one. In order to fuse source images after their detection, the authors of the work [3] have used basically the same procedure as that applied in Ref. [4]. They have also extended the appropriate method for fusing undetected image pairs. Ma et al. [5] have suggested a multi-scale transform-based method in which a total-variation model is employed to merge the base layers obtained with a Gaussian filter. Inspired by a gradient-transfer approach adopted in Ref. [3–5], we will introduce a new method based on the total-variation model. For preserving the gradient information to a maximal degree and achieving an edge-preserving function, we will also employ a guided-image filtering. In fact, we will make a particular emphasis on preserving the gradients at each stage of our procedures.

The rest of this paper is organized as follows. Section 2 reviews a theory of guided-image filtering. Section 3 presents our fusion method based on both the total-variation model and the guided-filter approach. In Section 4 we describe the experiments performed for eight different fusion methods. Their comparison demonstrates superiority of our method. Finally, the main conclusions are drawn in Section 5.

2. Guided-image filtering

Our guided-image filter is nothing but an edge-preserving filter presented recently in Ref. [6]. A design of this filter starts from maintaining gradients. Its mathematical representation is as follows:

$$\nabla Q = a\nabla I\,, \tag{1}$$

where I is a guidance image, Q an output image ($I, Q \in \mathbb{R}^{m\times n}$), and a denotes a linear coefficient. The idea of this linear model is to ensure that the gradients of the output and guidance images are linear. On this basis, it is assumed that Q is a linear transformation of I in a local window ω_k with the radius r, which is centred at some pixel k:

$$Q_i = a_k I_i + b_k\,, \forall i \in \omega_k\,, \tag{2}$$

with (a_k, b_k) being constants in ω_k and i denoting the pixel index.

Another assumption is that Q is a result of the input image ($P \in \mathbb{R}^{m\times n}$) that subtracts an unwanted component ($N \in \mathbb{R}^{m\times n}$):

$$Q_i = P_i - N_i\,. \tag{3}$$

Here (a_k, b_k) can be estimated by minimizing the following cost function:

$$\begin{aligned} E(a_k, b_k) &= \sum\nolimits_{i\in\omega_k}\left((Q_i - P_i)^2 + \varepsilon a_k^2\right) \\ &= \sum\nolimits_{i\in\omega_k}\left((a_k I_i + b_k - P_i)^2 + \varepsilon a_k^2\right), \end{aligned} \tag{4}$$

where ε is a regularization parameter introduced to penalize too large a_k's.

The relevant solution is as follows:

$$a_k = \frac{\frac{1}{|\omega|}\sum_{i\in\omega_k} I_i P_i - \mu_k \overline{P}_k}{\sigma_k^2 + \varepsilon}\,, \tag{5}$$

$$b_k = \overline{P}_k - a_k \mu_k\,. \tag{6}$$

Here μ_k and σ_k^2 are respectively the mean and the variance of I in ω_k, and $\overline{P}_k$ represents the mean of P in ω_k. Then Q can be calculated according to Eq. (2). Actually, another popular

strategy [6] reads as

$$Q_i = \overline{a}_i I_i + \overline{b}_i , \tag{7}$$

where $\overline{a}_i = \frac{1}{|\omega|}\sum_{k\in\omega_i} a_k$ and $\overline{b}_i = \frac{1}{|\omega|}\sum_{k\in\omega_i} b_k$. In summary, the entire filtering process can be denoted as $G_{r,\varepsilon}(P, I)$, where the parameters are the same as above.

3. Fusion method

In Refs. [3, 4], a minimization based upon the total variation is used directly to solve the problem of finding an optimal fused image. However, the information concerning tiny details and edges can be ignored in the process of iterative calculations, thus reducing the performance of fusion. The appropriate examples will be illustrated in Section 4.2. To avoid unwanted side-effects, we use a two-scale image decomposition to preserve the gradient information in advance.

Similar to a traditional wavelet-based image fusion, our method is divided into four steps: (1) performing two-scale decomposition on the source images to get base and details layers, (2) fusing the base layers using the total-variation model, (3) fusing the details layers using a traditional highly efficient fusion rule of maximum selection, and (4) reconstructing the fused image with the base and details layers being combined together.

3.1. Two-scale decomposition

The guided-image filter is employed to implement the two-scale decomposition within our fusion method. First, the infrared and visible images are decomposed into two-scale representation. Let us denote the infrared and visible images respectively as $ir, vi \in \mathbb{R}^{m\times n}$, while the fused image is given by $f \in \mathbb{R}^{m\times n}$. The base layer can be obtained as follows:

$$B_n = G_{r,\varepsilon}(I_n, I_n) , \tag{8}$$

where I_n is the source image and B_n the base layer. Note that the parameters $G_{r,\varepsilon}$ that appear when filtering the input and guidance images are identical. In this case the guided filter behaves as a gradient-preserving filter. It can avoid efficiently a gradient reversal, in contrast to bilateral filtering [6]. The details layer D_n can be defined as

$$D_n = I_n - B_n . \tag{9}$$

3.2. Fusion of base and details layers, and reconstruction

Note that the base layer contains an overwhelming majority of energy of the image. In especial case of the infrared image, it depicts mainly thermal-radiation information of the target under interest, of which natural intuition-based notion is the pixels with high intensities. Therefore, preserving this thermal-radiation information in the base layer is a top priority in all our efforts. This prompts us to formulate the relationship among the base layers of the fused and infrared images as an empirical error:

$$E_1(B_f) = \frac{1}{p}\left\| B_f - B_{ir} \right\|_p^p , \tag{10}$$

where B_f and B_{ir} are the base layers of respectively f and ir, and $\|\cdot\|_p$ stands for the l^p norm. Note that $E_1(B_f)$ should be as small as possible.

On the other hand, the base layer of the visible image is still rich in detailed information. This concerns especially the information on contours and edges of background and, probably, some

appearance information of targets. This kind of information plays an important role in targets location and environmental perception. An explicit way for fusing this details-related information is to make the fused base layer have similar gradient information with the visible base layer. As a consequence, we choose the following formula to relate the details layers of the fused and visible images:

$$E_2(f) = \frac{1}{q}\left\|\nabla B_f - \nabla B_{vi}\right\|_q^q. \tag{11}$$

Combing Eqs. (10) and (11), one can form an objective function:

$$\begin{aligned} E(f) &= E_1(f) + E_2(f) \\ &= \frac{1}{p}\left\|B_f - B_{ir}\right\|_p^p + \lambda\frac{1}{q}\left\|\nabla B_f - \nabla B_{vi}\right\|_q^q, \end{aligned} \tag{12}$$

where λ is the parameter introduced for controlling a trade-off between the two terms. The first term in the r. h. s. of Eq. (12) requires the fused base layer to have similar pixel intensities with the infrared base layer, and the second requires that the base layers have similar gradients. This can be regarded as a minimization problem of Eq. (12).

Now we are to solve this optimization problem. Considering computational complexity and feasibility, we set the parameters of the norm in Eq. (12) to unity, i.e. $p = 1$ and $q = 1$. Meanwhile, let we have

$$B_y = B_f - B_{vi}. \tag{13}$$

Then the above-mentioned problem can be reduced to

$$B_{y^*} = \arg\min_{B_y}\left\|B_y - (B_{ir} - B_{vi})\right\|_1 + \left\|\nabla B_y\right\|_1. \tag{14}$$

An iteratively reweighted-norm total-variation algorithm [7] can be adopted to solve it. As a result, the fused image can be obtained:

$$B_{f^*} = B_{y^*} + B_{vi}. \tag{15}$$

In the two-scale decomposition, much of the detailed information about the edges and textures is maintained in the details layers. The fusion of the details layers should ensure that the information about details is retained and can later be transferred into the final fused image. Hereafter, we adopt a simple and efficient fusion rule of maximum selection in order to fuse the details layers. It can be expressed as

$$D_f = \max(D_{ir}, D_{vi}). \tag{16}$$

Then the fused image can be reconstructed via

$$f = B_{f^*} + D_f. \tag{17}$$

4. Experimental results and discussion

4.1. Datasets

We have verified the performance of our fusion method using six image pairs, which are referred to as ‘quad’, ‘kayak’, ‘walker’, ‘tank’, ‘bunker’ and ‘lake’ (see Fig. 1). They have been downloaded from the Liu’s homepage [8] and the TNO image-fusion dataset [9]. We have compared our method with eight typical fusion methods corresponding to the state-of-the-art in the field. These are LP [10] and DWT [11] methods, a dual-tree complex wavelet transform (DTCWT) [12], a CVT method [13], a multi-resolution singular-value decomposition (MSVD) [14], a fusion method based on guided

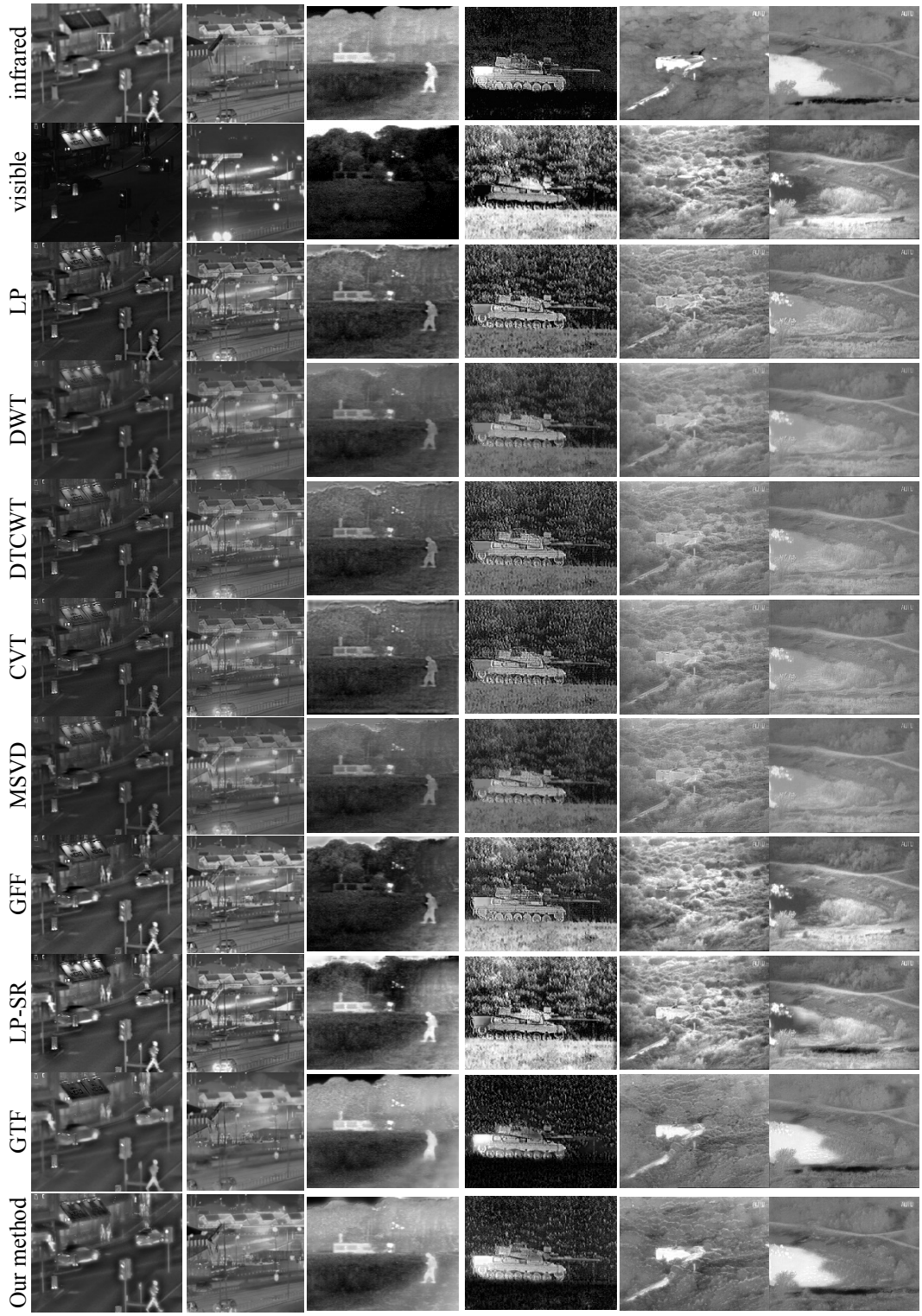

Fig. 1. Source and fused images: the image pairs termed as ‘quad’, ‘kayak’, ‘walker’, ‘tank’, ‘bunker’ and ‘lake’ follow from left to right, whereas the images following from top to bottom correspond to infrared source image, visible source image, and the images fused by LP, DWT, DTCWT, CVT, MSVD, GFF, LP-SR and GTF methods and our fusion method.

filtering (GFF) [15], a Laplacian-pyramid sparse representation (LP-SR) [16], and a gradient-transfer fusion (GTF) [17]. All of these methods are implemented in Matlab. They are publicly available at the Ma's homepage [17]. Regarding our sets of parameters, we follow the corresponding original papers.

In the frame of our fusion method, the radius r and the regularization parameter ε of the guided filter are set to 8 and 0.4^2, respectively (see Ref. [6]), while the parameter λ is equal to 4. Note that the recent GTF method has been proven to retain the largest (or almost the largest) information amount of the source images. Then a comparison of our method with the GFT is the most important.

The fusion results are evaluated using both subjective and objective techniques. The objective metrics are given by standard deviation (SD), correlation coefficient (CC), peak signal-to-noise ratio (PSNR), information entropy (IE), mutual information (MI), and structural-similarity index measure (SSIM) [18]. We define these metrics below.

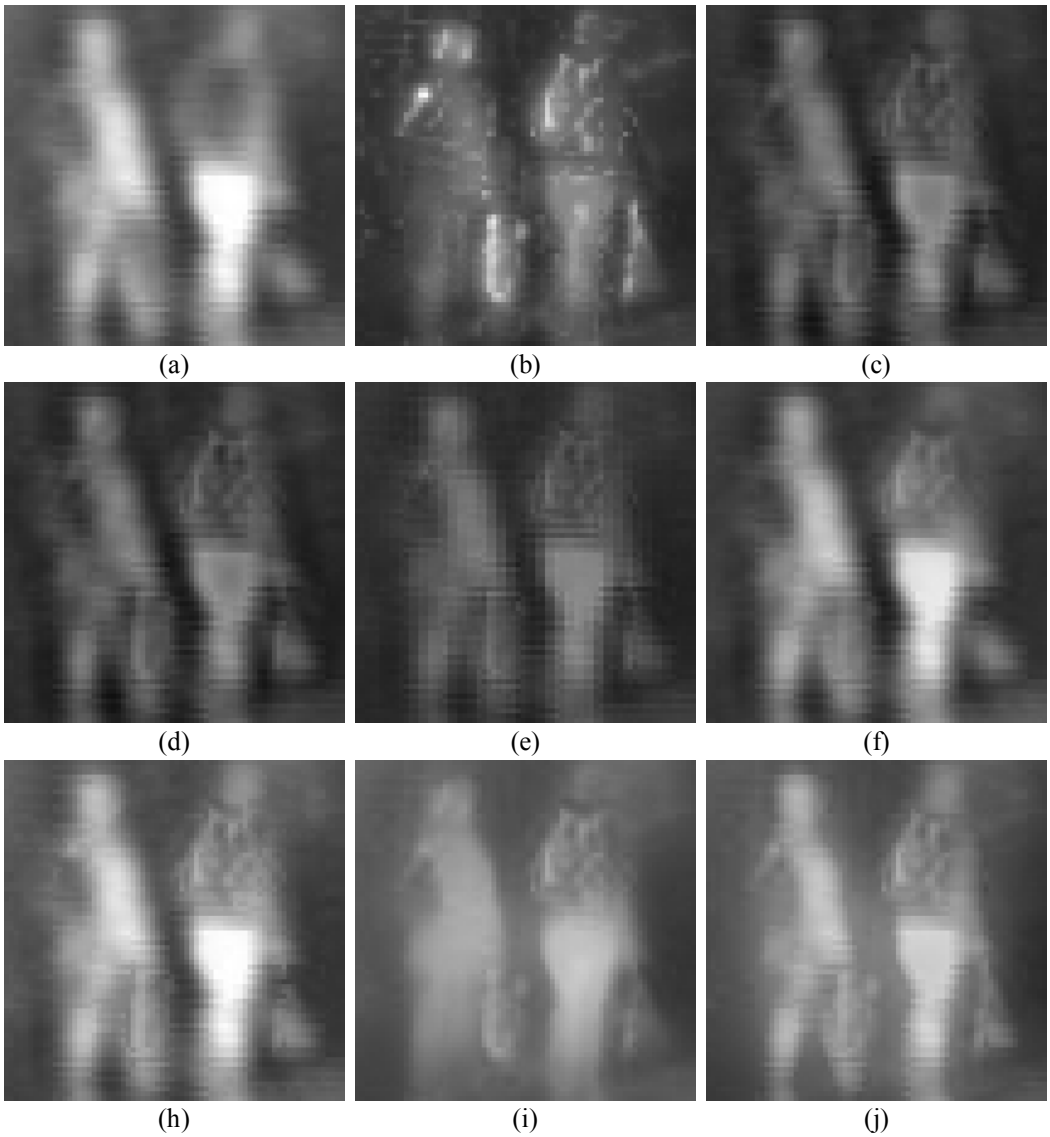

Fig. 2 Magnified regions of the 'quad' image (see Fig. 1), as extracted from (a) infrared and (b) visible images, and the fused images obtained using (c) DWT, (d) DTCWT, (e) CVT, (f) MSVD, (g) GFF, (h) LP-SR, (i) GTF and (j) our fusion method.

(1) the SD measures overall contrast of the fused image F:

$$SD(F)=\left[\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(F(i,j)-u)\right]^{\frac{1}{2}},u=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}F(i,j); \tag{18}$$

(2) the CC quantizes correlation between the fused image F and the infrared image I:

$$CC(I,F)=\frac{\sum_m\sum_n(F(m,n)-\mu_F)(I(m,n)-\mu_I)}{\sqrt{\sum_m\sum_n(F(m,n)-\mu_F)^2\cdot\sum_m\sum_n(I(m,n)-\mu_I)^2}}, \tag{19}$$

where μ implies the mean value;

(3) the PSNR is defined by the relation

$$PSNR(I,F)=10\log_{10}\frac{255^2mn}{\sum_{i=1}^{m}\sum_{j=1}^{n}(F(i,j)-I(i,j))^2}; \tag{20}$$

(4) the IE reflects the average amount of information contained in the fused image:

$$IE(F)=-\sum_{i=0}^{L-1}p(i)\log_2 p(i), \tag{21}$$

where L is the total number of gray levels and $p(i)$ the normalized image histogram;

(5) the MI measures similarity of the intensities of the images:

$$MI(I,F)=\sum_{i,f}p_{I,F}(i,f)\log\frac{p_{I,F}(i,f)}{p_I(i)p_F(f)}, \tag{22}$$

where $p_{I,F}$ is the jointly normalized histogram of the images I and F, p_I and p_F are the normalized histograms of I and F, and i and f represent the pixel values of I and F, respectively;

(6) the SSIM for the infrared (I) and fused (F) images is defined as

$$SSIM(I,F)=\frac{(2u_Iu_F+C_1)(2\sigma_{IF}+C_2)}{(u_I^2+u_F^2+C_1)(\sigma_I^2+\sigma_F^2+C_2)}, \tag{23}$$

where u_I and u_F denote the mean value of I and F. σ_I, σ_F are the standard deviation of I and F, respectively, and σ_{IF} is the covariance of I and F, and C_1 and C_2 constants that serve to avoid any instability in the relation given by Eq. (23).

Of course, the larger the fusion metrics, the better the fusion quality is.

4.2. Subjective evaluation

To compare visual appearances of the fused images intuitively, we show all the source and fused images in Fig. 1. For each column in Fig. 1, the first two images (i.e., the images in the first and second rows) are the original infrared and visible images, respectively. The rest are the fused images obtained by different methods (see the legend). One can conclude from Fig. 1 that, roughly speaking, all the nine methods reveal comparable (and quite good) performances. For example, the information about all pedestrians, cars, traffic lights and trashcans in the initial ‘quad’ images is transferred fairly well into the fused images.

To facilitate further observation of the fusion effects, a specific region of the ‘quad’ images is extracted and magnified in Fig. 2. The location of this region is shown in the infrared ‘quad’ image (see Fig. 1). One can see from Fig. 2 that the figures of two pedestrians in Fig. 2j can be more clearly distinguished, and their edges are clearer than those in the other images. The pedestrians in the rest of images have different degrees of blur, which is probably accompanied by low intensities. This is not good for targets detection. In particular, one can observe that the image of Fig. 2i obtained by the GTF method has lost a lot of detailed information. Similar conclusions can also be inferred with the other image groups, although we omit the corresponding discussion for conciseness.

4.3. Quantitative evaluation

The data of objective comparisons of the nine fusion methods is gathered in Table 1, where the highest values corresponding to the best results are indicated in bold for each metric and each image group. For convenient observation of the data, we plot them in Fig. 3 as line charts categorized by different metrics. As seen from Table 1 and Fig. 3, our combined method may not

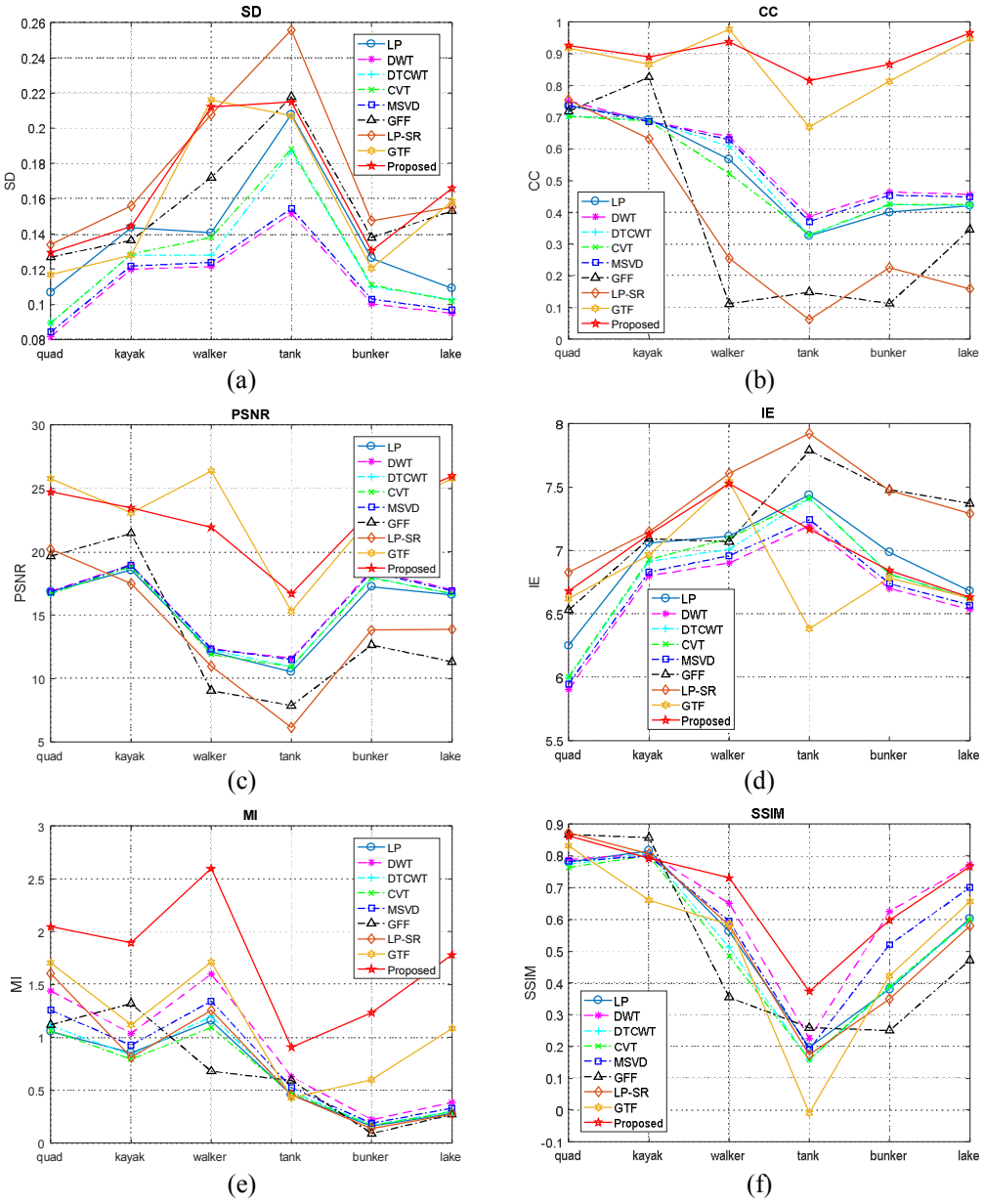

Fig. 3. Line charts of the fusion results with respect to the metrics (a) SD, (b) CC, (c) PSNR, (d) IE, (e) MI and (f) SSIM. Data concerns the image pairs 'quad', 'kayak', 'walker', 'tank', 'bunker' and 'lake'. Abbreviations in the legend correspond to different fusion methods, where our combined method is denoted as 'Proposed'.

provide the best result for each metric, although it does achieve the upper-class performance. More specifically, from the viewpoint of average of the metric data for the six image pairs, our method is superior to the other methods with respect to the CC, PSNR, MI and SSIM metrics. Concerning

Table 1. Objective data of image fusion for different pairs of images, as obtained by different methods.

Image pair	Metric	LP	DWT	DTCWT	CVT	MSVD	GFF	LP-SR	GTF	Our method
'quad'	SD	0.1068	0.0816	0.0896	0.0895	0.0843	0.1268	**0.1438**	0.1168	0.1293
	CC	0.7354	0.7505	0.7041	0.7027	0.7369	0.7192	0.7549	0.9179	**0.9264**
	PSNR	16.8605	16.8823	16.6966	16.6901	16.8356	19.6416	20.2037	**25.7695**	24.7179
	IE	6.2486	5.9035	5.9963	6.0041	5.9474	6.5308	**6.8285**	6.6205	6.6791
	MI	1.0567	1.4381	1.1119	1.0638	1.2569	1.1210	1.6057	1.7065	**2.0464**
	SSIM	0.7813	0.7890	0.7725	0.7637	0.7828	0.8676	**0.8722**	0.8324	0.8630
'kayak'	SD	0.1435	0.1199	0.1281	0.1284	0.1218	0.1363	**0.1560**	0.1279	0.1444
	CC	0.6924	0.6868	0.6904	0.6879	0.6863	0.8262	0.6321	0.8661	**0.8893**
	PSNR	18.5532	18.9828	18.8703	18.8342	18.9457	21.4558	17.5008	23.0821	**23.4614**
	IE	7.0598	6.7989	6.9125	6.9277	6.8302	7.0926	**7.1498**	6.9670	7.1281
	MI	0.8528	1.0376	0.8316	0.7934	0.9260	1.3207	0.8142	1.1196	**1.8933**
	SSIM	0.8166	0.8064	0.8105	0.8009	0.7995	**0.8573**	0.8055	0.6608	0.7915
'walker'	SD	0.1406	0.1215	0.1278	0.1381	0.1238	0.1721	0.2079	**0.2162**	0.2123
	CC	0.5668	0.6385	0.6078	0.5222	0.6290	0.1119	0.2549	**0.9763**	0.9372
	PSNR	12.1325	12.3824	12.2688	11.9130	12.3495	9.0668	10.9668	**26.3891**	21.9325
	IE	7.1111	6.9016	7.0098	7.0967	6.9576	7.0685	**7.6058**	7.5417	7.5275
	MI	1.1558	1.5988	1.1999	1.0929	1.3424	0.6798	1.2581	1.7105	**2.5964**
	SSIM	0.5626	0.6507	0.5122	0.4849	0.5953	0.3547	0.5816	0.5846	**0.7306**
'tank'	SD	0.2078	0.1516	0.1868	0.1882	0.1544	**0.2178**	0.2557	0.2070	0.2148
	CC	0.3261	0.3872	0.3274	0.3295	0.3709	0.1482	0.0629	0.6701	**0.8157**
	PSNR	10.5287	11.5988	10.9380	10.9265	11.5072	7.8642	6.1425	15.3354	**16.6872**
	IE	7.4373	7.1933	7.4104	7.4118	7.2426	7.7879	**7.9218**	6.3849	7.1670
	MI	0.4579	0.6300	0.4870	0.4716	0.5253	0.5927	0.4582	0.4256	**0.9084**
	SSIM	0.1974	0.2261	0.1582	0.1589	0.1912	0.2596	0.1734	-0.0090	**0.3740**
'bunker'	SD	0.1262	0.1003	0.1102	0.1110	0.1030	0.1379	**0.1474**	0.1206	0.1304
	CC	0.4003	0.4643	0.4248	0.4257	0.4539	0.1124	0.2260	0.8138	**0.8662**
	PSNR	17.2464	18.5553	17.9664	17.9451	18.4013	12.6440	13.8197	22.7737	**23.2483**
	IE	6.9858	6.6989	6.8118	6.8129	6.7388	**7.4778**	7.4723	6.7806	6.8423
	MI	0.1610	0.2227	0.1688	0.1692	0.1876	0.0900	0.1369	0.6009	**1.2340**
	SSIM	0.3800	**0.6253**	0.3865	0.3902	0.5208	0.2504	0.3487	0.4230	0.5980
'lake'	SD	0.1091	0.0949	0.1027	0.1021	0.0968	0.1530	0.1553	0.1586	**0.1660**
	CC	0.4218	0.4568	0.4211	0.4245	0.4485	0.3455	0.1593	0.9466	**0.9653**
	PSNR	16.5844	17.0016	16.6715	16.7019	16.9265	11.3373	13.8871	25.7726	**25.9680**
	IE	6.6811	6.5352	6.6210	6.6208	6.5693	**7.3712**	7.2908	6.6236	6.6313
	MI	0.2817	0.3851	0.2992	0.2958	0.3315	0.2709	0.2767	1.0828	**1.7805**
	SSIM	0.6031	**0.7722**	0.5964	0.5998	0.7007	0.4718	0.5792	0.6563	0.7662

the SD and the IE, the results of our method are only somewhat worse than those of the LP-SR method. Of course, none of the metrics can be treated as absolutely better than the others, so that we utilize multiple metrics to judge about the fusion performance. It is also worth mentioning that, in every metric, the results of our method are basically better than those of the recent GTF method.

Summing up the subjective and objective evaluations, one can conclude that our fusion method keeps efficiently the thermal-radiation information and, moreover, preserves the transferred gradient information. In addition, it achieves a superior performance if compared to the eight common fusion methods, which reflect the state-of-the-art in the field.

5. Conclusions

In this work we propose a novel method for fusing infrared and visible images, which is based on the combined approach of guided filtering and the total-variation model. Inspired by the known gradient-transfer approach, we preserve the gradients prior to each fusion stage. In the first step of the two-stage decomposition, the guided filter is employed to preserve the gradient information in the base layer of infrared image to a maximal degree. Then fusion of the base layers takes place. After that the total-variation minimization makes the fused base layer have similar gradient information to the base layer of visible image, while we approximate the base layer of infrared image. After that, fusion of the details layers adopts the classic maximum-selection rule to maximize the detailed information of fused image.

In our experiments, eight common methods corresponding to the state-of-the-art in the field of image fusion (namely, the methods of LP, DWT, DTCWT, CVT, MSVD, GFF, LP-SR and GFF) are used. The appropriate results are compared with those of our method both in terms of subjective and objective techniques. The objective measures include the well-known SD, CC, PSNR, IE, MI and SSIM metrics. The results obtained by us demonstrate that our method can achieve superior performance in terms of both human visual perception and the objective fusion-quality metrics.

Acknowledgements

This work was supported by the Key Project of Supporting Excellent Youth in Colleges and Universities of Anhui Province (No. gxyqZD2019080), the Horizontal Project at Suzhou University of China (No. 2020xhx093) and the Key Research and Technology Development Projects of Anhui Province (No. 202004a06020045).

References

1. Li S, Kang X, Fang L, Hu J and Yin H, 2017. Pixel-level image fusion: a survey of the state of the art. Information Fusion. **33**: 100–112.
2. Jiayi Ma, Ma Yong and Chang Li. 2019. Infrared and visible image fusion methods and applications: a survey. Information Fusion. **45**: 153–178.
3. Ma J, Chen C, Li C and Huang J, 2016. Infrared and visible image fusion via gradient transfer and total variation minimization. Information Fusion. **31**: 100–109.
4. Ma Y, Chen J, Chen C, Fan F and Ma J, 2016. Infrared and visible image fusion using total variation model. Neurocomputing. **202**: 12–19.
5. Ma T, Ma J, Fang B, Hu F, Quan S and Du H, 2018. Multi-scale decomposition based fusion of infrared and visible image via total variation and saliency analysis. Infrared Phys. Technol. **92**: 154–162.
6. Kaiming He, Jian Sun and Xiaoou Tang, 2012. Guided image filtering. IEEE Transactions on

Pattern Analysis and Machine Intelligence. **35**: 1397–1409.
7. Rodríguez P and Wohlberg B, 2009. Efficient minimization method for a generalized total variation functional. IEEE Transactions on Image Processing. **18**: 322–332.
8. http://www.escience.cn/people/liuyu1
9. https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029
10. Burt P and Adelson E, 1983. The Laplacian pyramid as a compact image code. IEEE Transactions on Communications. **31**: 532–540.
11. Pajares G and De La Cruz J M, 2004. A wavelet-based image fusion tutorial. Pattern Recognition. **37**: 1855–1872.
12. Lewis J J, O'Callaghan R J, Nikolov S G, Bull D R and Canagarajah N, 2007. Pixel-and region-based image fusion with complex wavelets. Information Fusion. 8: 119–130.
13. Nencini F, Garzelli A, Baronti S and Alparone L, 2007. Remote sensing image fusion using the curvelet transform. Information Fusion. **8**: 143–156.
14. Naidu V P S, 2011. Image fusion technique using multi-resolution singular value decomposition. Defense Sci. J. **61**: 479–484.
15. Li S, Kang X and Hu J, 2013. Image fusion with guided filtering. IEEE Transactions on Image Processing. **22**: 2864–2875.
16. Liu Y, Liu S and Wang Z, 2015. A general framework for image fusion based on multi-scale transform and sparse representation. Information Fusion. **24**: 147–164.
17. https://github.com/jiayi-ma
18. Wang Z, Bovik A C, Sheikh H R and Simoncelli E P, 2004. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing. **13**: 600–612.

Hongjun Guo, Lili Chen and Zhenggao Pan. 2022. Fusion of infrared and visible images by a combined guided-filter and total-variation method. Ukr.J.Phys.Opt. **23**: 57 – 67.
doi: 10.3116/16091833/23/2/57/2022

Анотація. *Запропоновано метод злиття інфрачервоних і видимих зображень, який базується на підходах керованої фільтрації та повної варіації. Для здобуття та максимального збереження градієнтної інформації двох типів зображень та збереження теплових цілей на першому етапі використовують керований фільтр, щоби побудувати двомасштабну декомпозицію зображення та застосувати її до вихідних зображень. Це роблять для одержання базового шару і шарів деталей, де шари деталей можуть заздалегідь зберігати інформацію про градієнти. Потім застосовують модель повної варіації. Обмежуючи злитий базовий шар так, щоб він мав подібний розподіл інтенсивності до базового шару інфрачервоного зображення та подібну інформацію про градієнт до базового шару видимого зображення, можна одержати злите зображення, яке одночасно зберігає інформацію і про термічні цілі, і про градієнти. І суб'єктивне, і об'єктивне оцінювання наших експериментальних даних свідчать про те, що комбінований метод, запропонований у цій роботі, виявляє вищу ефективність порівняно з низкою відомих методів.*