
A fast and efficient two-scale fusion of infrared and visible images using median filter and intensity transfer

Jinjuan Wang, Xiliang Zeng, Shan Duan, Qun Zhou and Hao Peng

Hunan International Economics University, Changsha 410205, Hunan, China

Received: 25.05.2022

Abstract. We suggest a fast and efficient method for fusing infrared and visible images, which is based on a median filter and an intensity transfer. The median filter is used to construct a new two-scale image decomposition and obtain base and detail layers. Then a fused base layer is obtained by utilizing an optimization fusion rule based on intensity transfer, where the fusion is transformed into matrix operations. A fused detail layer is obtained by an improved maximum-selection rule which maximizes the details-appearance information with some image-enhancement effect. Our experimental results demonstrate that the method can provide a superior competitive fusion performance in terms of both subjective evaluation and objective metrics, when compared with seven standard state-of-the-art methods. Moreover, our method is quite suitable for real-time fusion applications.

Keywords: image fusion, infrared images, median filters, intensity transfer

UDC: 004.93

1. Introduction

Better characteristics of sensors are almost always achieved at the expense of their more complex design and implementation or higher cost [1]. Since infrared (IR) sensors capture primarily thermal-radiation information, which is not affected by illumination or disguise, they are useful for revealing and highlighting different targets and can work day and night. However, the images captured by the IR sensors are not consistent with human visual system too well. Moreover, usually they manifest low spatial resolution and lose detailed information about appearance of a scene. On the contrary, visual (VIS) images can provide high-quality appearance information captured by a VIS sensor, which is perfectly consistent with the human visual system. As a result, it would be useful to fuse the IR and VIS images into a composite one, which can provide important mutually complementary information.

Up to now, image-fusion technologies have been widely used in many fields, e.g. in medical imaging, object detection, remote sensing and surveillance. Among the well-known fusion methods, one can remind of multi-scale decomposition [2], sparse representation [3] and methods based on neural networks [4]. The mathematical models underlying these techniques are different. The fusion methods relying upon multi-scale decomposition have become prevalent in the past two decades and still remain very efficient. They decompose source images into different scales (from fine to coarse) and then combine the coefficients of the appropriate subbands using some fusion rules. The methods based on sparse representation can be derived from the idea that an image can be represented by a linear combination of sparse bases in over-complete dictionaries. Finally, the methods associated with neural networks imitate human brain's perception to deal with information.

Of course, none of the approaches mentioned above is perfect. The multi-scale decomposition-based methods depend significantly on transforms, decomposition levels and

wavelet bases. In the frame of sparse representation, constructing appropriate dictionaries which have promising representation and high fusion performance is still difficult and time-consuming. Designing suitable neural networks and tuning their parameters are complex to implement and expensive computationally, especially when it is done using recently developed deep-learning based methods.

In recent years, an edge-preserving filtering has emerged as an effective tool for image processing. It can be classified as a multi-scale decomposition. Bilateral filter [5], weighted least-squares filter [6], guided filter [7] and some others are now commonly used in edge-preserving filtering. They can decompose an image into one base layer and one or more detail layers. Notice also that some simple filters can be applied to the similar decomposition. For example, Ma et al. [8] have used a Gaussian filter to decompose source images and obtain a single base layer and a single detail layer. Li et al. [7] have suggested a simple mean filter to achieve the two-scale decomposition, while Bavirisetti et al. [9] and Naidu et al. [10] have constructed the same two-scale image decomposition, using a mean filter, and taken the difference of mean and median filtering outputs as a saliency map for the detail layer fusion. Kalaiselvi et al. [11] have employed the same median filter in order to construct a two-scale adaptive median filter for denoising. Inspired by those methodologies, we have decided to avoid the problems mentioned above. This can be achieved when using a median filter to construct a parameterless two-scale image decomposition. Then it is combined with intensity transfer and improved maximum-selection fusion rules for reducing the computational complexity.

Note that the intensity-transfer scheme [12] has originally been intended as an image-fusion method, which regards the fusion as a minimization problem and solves it by direct matrix mapping. A seemingly similar approach is the intensity-transformation function for IR images [13]. It behaves like a sigmoid function, shifts or expands the range of pixel values, and can be considered as a preprocessing of fusion.

The remainder of this article is organized as follows. In Section 2 we briefly review a theory of intensity transfer. Section 3 describes the suggested fusion method in detail and the experimental results and discussion are presented in Section 4. Finally, we draw the main conclusions in Section 5.

2. Intensity transfer

The intensity transfer [12] has been proposed for fusion of IR and VIS images in order to meet the requirements of preserving thermal objects and keeping background information. In this scheme, the fusion is formulated as a minimization problem, i.e.

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

where \mathbf{u}, \mathbf{v} and $\mathbf{x} \in \mathbb{R}^{mn \times 1}$ are the vector forms of respectively IR, VIS and fused images, and $\|\cdot\|_2$ denotes the l_2 -norm. Here $W \in \mathbb{R}^{mn \times mn}$ is the diagonal weight matrix and w_p the diagonal entry of W at the position p , which is calculated as

$$w_p = |\log(S_p)|, \quad (2)$$

with S_p being the element of spatial-saliency map [14]. The latter can be obtained from the image statistics:

$$S_p = \sum_{V=0}^{255} N_V(I_p - V), \quad (3)$$

where N_V represents the total number of pixels in 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 .

To solve the optimization problem for \mathbf{x} , we assume that the derivative of $\|\mathbf{x} - \mathbf{u}\|_2^2 + W \|\mathbf{x} - \mathbf{v}\|_2^2$ in Eq. (1) is equal to zero. Then we have

$$\mathbf{x}^* = (I + W)^{-1}(\mathbf{u} + W\mathbf{v}). \quad (4)$$

where $I \in \mathbb{R}^{mn \times mn}$ is a unit matrix.

3. Fusion method

Basically, our fusion method consists of three steps: two-scale image decomposition, fusion, and reconstruction. First, the decomposition is achieved by median filtering and then the obtained base and detail layers are fused using specific fusion rules. At last, the fused image is reconstructed issuing from the fused base and detail layers.

The median filtering is a nonlinear spatial smoothing operation, in which each output pixel takes the median value of its 3-by-3 neighbourhood. This approach is more efficient than the convolution whenever one of the purposes is to preserve edges. As a consequence, the technique is also suitable when constructing a two-scale image decomposition. Compared with the mean filter [7, 9] or the Gaussian filter [8] also used in the two-scale decomposition, our approach is more advantageous for preserving gradient information in the base layer.

Assume that the two detected source IR and VIS images are denoted respectively as ir and $vi \in \mathbb{R}^{m \times n}$ and the fused image is given by $f \in \mathbb{R}^{m \times n}$. Then the base layer can be obtained by the median filtering:

$$B_n = \text{median}(I_n), \quad (5)$$

where B_n denotes the base layer, I_n the source image, $n \in \{ir, vi\}$ represents the IR (or VIS) image, and $\text{median}(\cdot)$ implies the median filtering.

After that, the detail layer can be obtained by subtracting the base layer B_n from the corresponding source image. This can be expressed as

$$D_n = I_n - B_n, \quad (6)$$

where D_n denotes the detail layer.

In this two-scale decomposition, the base layer contains the most energy of the source image, and the detail layer contains mainly the information on the appearance of details. Note that no parameters are required in the process of decomposition, whereas the classical algorithm of median filtering is robust and can be implemented at high speed [15].

In the fusion of base layers, the ideal result is transferring the object information in the source image such as contour, location and background, to the fused base layer. So the latter should simultaneously approximate the base layers of the IR and VIS images. Then we employ the intensity-transfer scheme studied in Section 2 as a fusion rule for the base layers. We arrive at the optimization problem which can be formulated as follows:

$$\min_{B_f} J(B_f) = \|\text{vec}(B_f) - \text{vec}(B_{ir})\|_2^2 + W \|\text{vec}(B_f) - \text{vec}(B_{vi})\|_2^2, \quad (7)$$

where B_f, B_{ir} and $B_{vi} \in \mathbb{R}^{m \times n}$ denote the base layers respectively of the fused, IR and VIS

images, $vec(\cdot)$ is the vec-operator applied to the matrix, which stacks the columns into a vector, and $J(\cdot)$ is the cost function of the above optimization problem.

Let us have for simplicity $\mathbf{x} = vec(B_f)$, $\mathbf{u} = vec(B_{ir})$ and $\mathbf{v} = vec(B_{vi})$. Then Eq. (7) can be substituted by

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} J(\mathbf{x}) = \|\mathbf{x} - \mathbf{u}\|_2^2 + W \|\mathbf{x} - \mathbf{v}\|_2^2. \quad (8)$$

In its form, this relation is the same as Eq. (1). Now let the derivative of $J(\mathbf{x})$ be equal to zero. Solving the equation for \mathbf{x} , we obtain Eq. (4). Then \mathbf{x}^* is reshaped to the size of B_f and denoted as \hat{B}_f .

In Eq. (4), there is a single unknown parameter, W , which is crucial in the calculation of the fused result. A spatial-saliency map based on the image statistics (i.e., Eqs. (2) and (3)) has been utilized in Ref. [12] to obtain W . However, this process involves large amount of calculations. Actually, the performance of the spatial-saliency map is very similar to image enhancement conducted in experiments. Therefore, to accelerate the fusion process with no loss of object information, we adopt a simple adaptive scheme that implies using directly the IR image as the saliency map:

$$S = I_{ir}. \quad (9)$$

The reason for this procedure is that, in the fusion process, the object information is primarily conveyed by the IR image.

Afterward, one can obtain W as

$$S = \text{normalize}(S), \quad (10)$$

$$W = |\log(S)|, \quad (11)$$

where $\text{normalize}(\cdot)$ is the normalizing operation aimed to scale the entries of the matrix to the region $[0,1]$.

As a result of action of the median filter used in the two-scale decomposition, the coefficients of the detail layers can be either positive or negative, which usually provides redundant or complementary information according to whether their signs are the same or not. Accordingly, we improve the common maximum-selection rule. If the signs of the coefficients of the two detail layers are the same at the same position, then the result is summation of the two coefficients. The advantage of doing so is that the detail information can be ‘enhanced’. If the signs are opposite, then the result is the coefficient with larger absolute value. The relevant expression is as follows:

$$D_f(i, j) = \begin{cases} D_{ir}(i, j) + D_{vi}(i, j) & \text{if } \text{sgn}(D_{ir}(i, j)) = \text{sgn}(D_{vi}(i, j)) \\ D_{ir}(i, j) & \text{if } \text{sgn}(D_{ir}(i, j)) \neq \text{sgn}(D_{vi}(i, j)) \text{ and } |D_{ir}(i, j)| \geq |D_{vi}(i, j)|, \\ D_{vi}(i, j) & \text{otherwise} \end{cases} \quad (12)$$

where D_f represents the fused detail layer and (i, j) the pixel coordinate.

Finally, the fused image is reconstructed by simply summing the fused base and detail layers:

$$I_f = \hat{B}_f + D_f. \quad (13)$$

4. Experimental results and discussion

This section describes experimental settings, fusion results and a comparison of performances of different fusion methods. All the experiments have been carried out on a laptop characterized by Intel i5-6200U CPU and 8GB RAM.

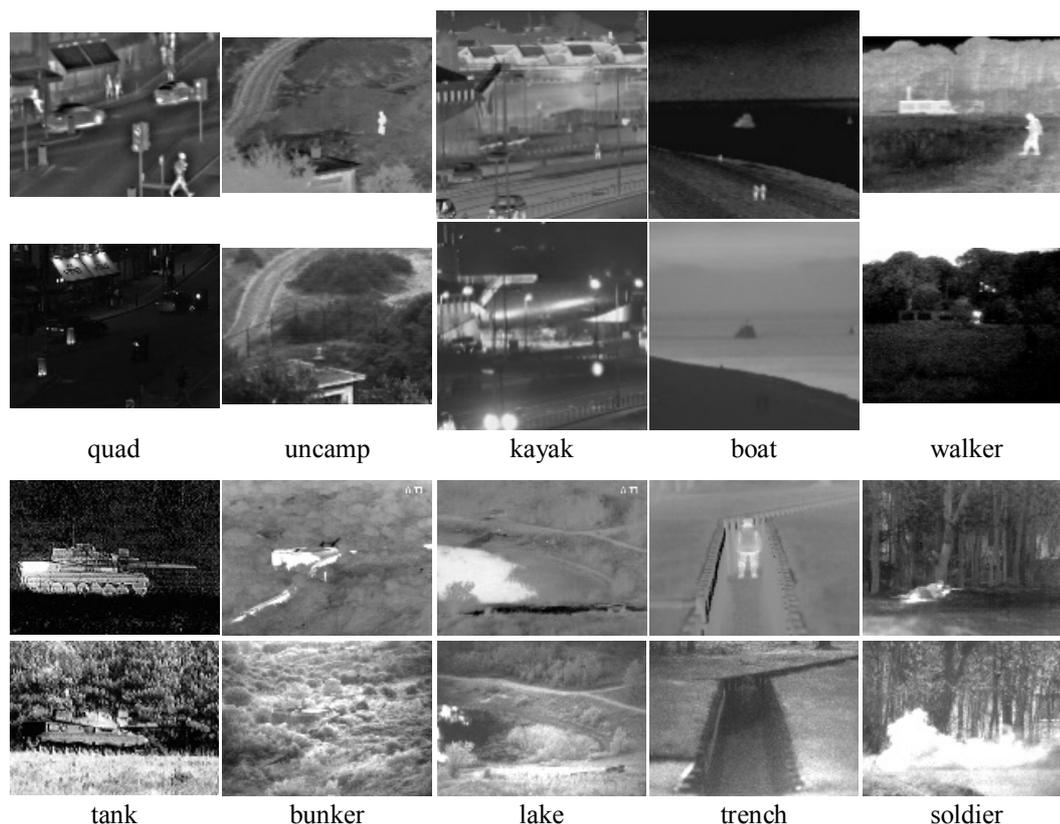


Fig. 1. Ten pairs of standard IR and VIS source images ‘quad’, ‘uncamp’, ‘kayak’, ‘boat’, ‘walker’, ‘tank’, ‘bunker’, ‘lake’, ‘trench’, ‘soldier’ (see the text): the first and second rows correspond to IR and VIS images, respectively.

Several widely used image pairs of standard IR and VIS images have been employed to conduct our experiments. They are referred to as ‘quad’, ‘uncamp’, ‘kayak’, ‘boat’, ‘walker’, ‘tank’, ‘bunker’, ‘lake’, ‘trench’ and ‘soldier’ (see Fig. 1). All of these images have been obtained from the website imagefusion.org or the TNO image fusion dataset [16].

In order to demonstrate the efficiency of our fusion method, seven reference methods which provide the state-of-the-art results have been used for comparison. These are a dual-tree complex wavelet transform (DTCWT) [17], a curvelet transform (CVT) [18], a multi-resolution singular-value decomposition (MSVD) [19], a guided filter-based fusion (GFF) [7], a Laplacian pyramid with sparse representation (LP-SR) [3], a gradient transfer fusion (GTF) [20], and an intensity transfer and direct matrix mapping (IT-DMM) [12]. These methods have been implemented in Matlab. The codes of the first six methods are publicly available at the Ma’s homepage [21], while the last method has been coded by us according to the lines reported in the work [12]. The parameters involved in the reference methods have been set according to the original works [3, 7, 12, 17–20].

First we evaluate the fusion performance in a subjective manner. Due to the space constraints, we report only on the fused images obtained for the image pairs ‘quad’, ‘uncamp’ and ‘soldier’ (see Fig. 2, Fig. 3 and Fig. 4, respectively). Further on, the fused images obtained by the seven reference methods DTCWT, CVT, MSVD, GFF, LP-SR, GTF and IT-DMM are compared with that obtained by our method.

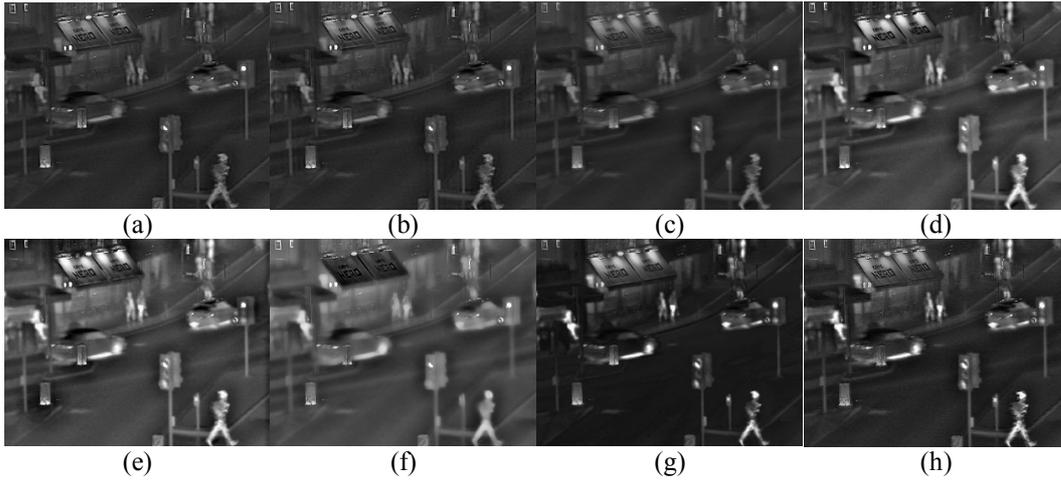


Fig. 2. Fusion results for a 'quad' image pair obtained by the DTCWT (a), CVT (b), MSVD (c), GFF (d), LP-SR (e), GTF (f) and IT-DMM (g) reference methods and our method (h).

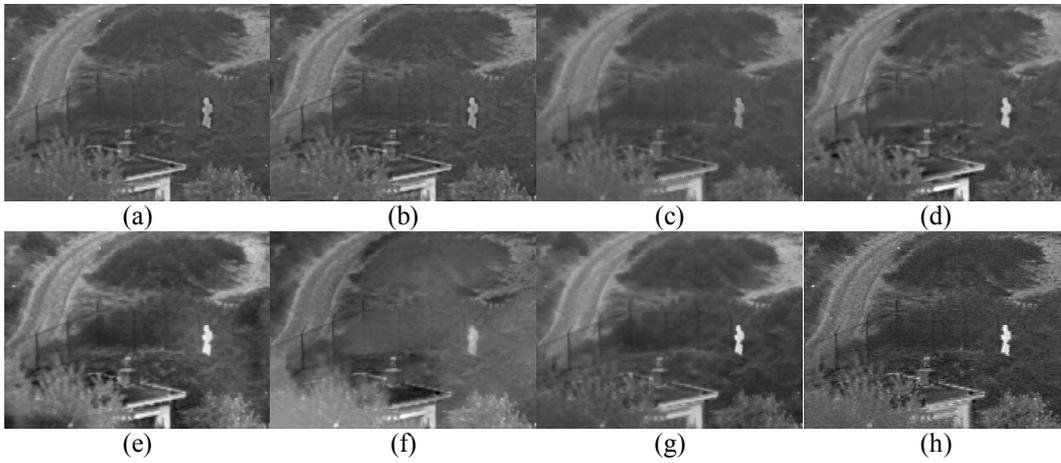


Fig. 3. Fusion results for an 'uncamp' image pair obtained by the DTCWT (a), CVT (b), MSVD (c), GFF (d), LP-SR (e), GTF (f) and IT-DMM (g) reference methods and our method (h).

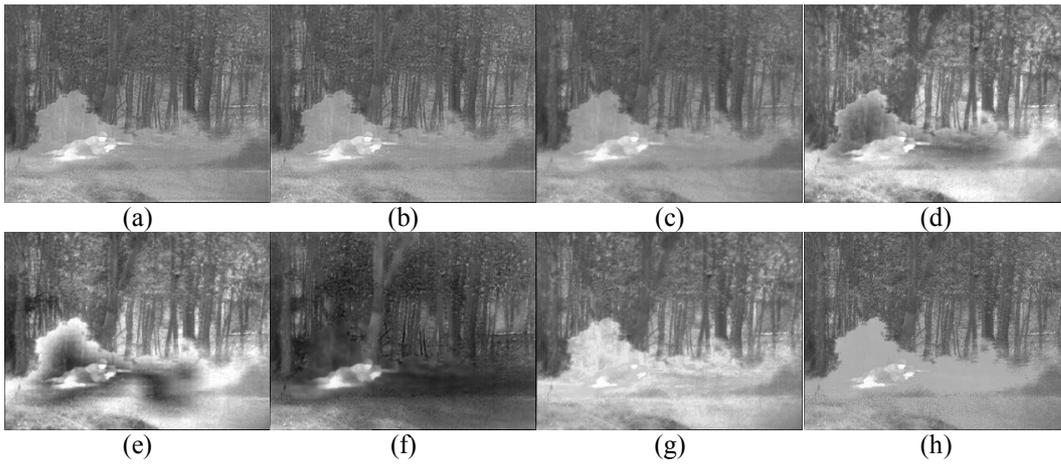


Fig. 4. Fusion results for a 'soldier' image pair obtained by the DTCWT (a), CVT (b), MSVD (c), GFF (d), LP-SR (e), GTF (f) and IT-DMM (g) reference methods and our method (h).

One can see from Fig. 2 that, perhaps, the poorest fusion effect is provided by the GTF method (Fig. 2f). Here the detailed information about pedestrians, cars, store and so on is smoothed out, and only the outline information of the object is retained. Similar situations take place in Fig. 2c, Fig. 2d, and Fig. 2g, although the degree of smoothing-out effect is not so serious. On the whole, the patterns observed in Fig. 2a and Fig. 2b look dark and the contrast is low. This implies that the object information is not prominent and conducive to subsequent processing (e.g., target detection). Concerning the last two images (Fig. 2e and Fig. 2h), it looks like the corresponding methods are roughly characterized by the same (and high) performance. A more careful observation testifies that Fig. 2h can reveal more detailed information than Fig. 2e.

To make further comparison of the visual appearance, a number of important local regions of the fused images corresponding to the ‘quad’ source pair have been extracted and magnified in Fig. 5. Now it becomes evident that our fusion method achieves the best visual effect.

Similar conclusions can also be drawn in case of the fused images displayed in Fig. 3 and Fig. 4. In other words, the fused images obtained by our method contain clearer object information with higher contrast and more detailed information.

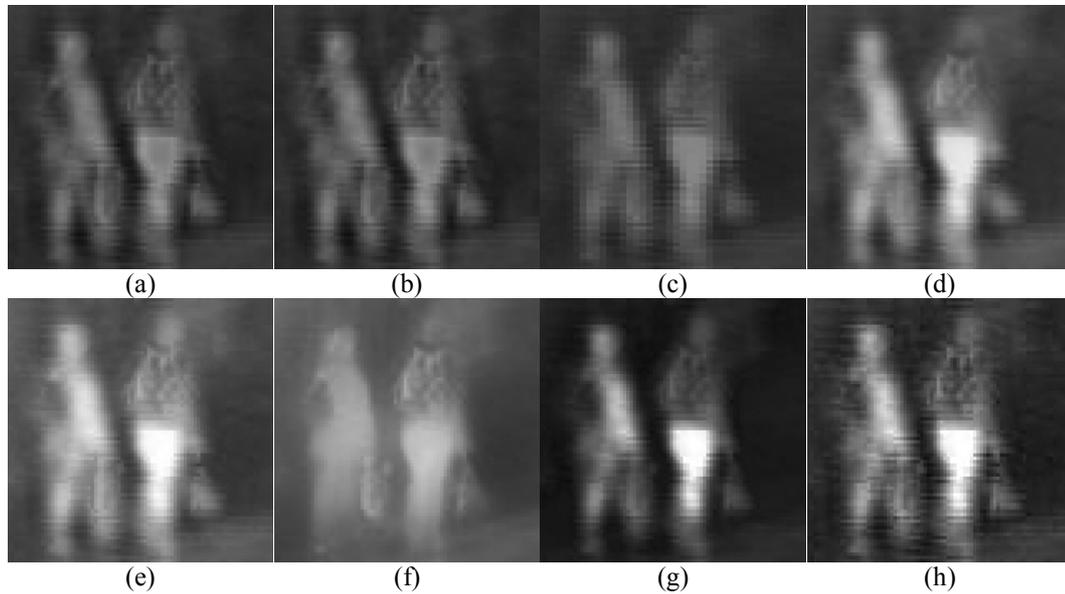


Fig. 5. Magnified local regions of the fused images of a ‘quad’ pair obtained by the DTCWT (a), CVT (b), MSVD (c), GFF (d), LP-SR (e), GTF (f) and IT-DMM (g) reference methods and our method (h).

Still there is no standard fusion metric that can measure the fusion performance alone, and no specific metric has been definitely proven to be better than the others. Hence, we will use several metrics such as spatial frequency (SF) [22], information entropy (IE) [23] and universe image-quality index (UIQI) [24] to evaluate the fusion performance in an objective manner. The metrics adopted by us are defined as follows:

$$\begin{cases} RF = \left(\frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^{n-1} (F(i, j+1) - F(i, j))^2 \right)^{\frac{1}{2}} \\ CF = \left(\frac{1}{(m-1)n} \sum_{i=1}^{m-1} \sum_{j=1}^n (F(i+1, j) - F(i, j))^2 \right)^{\frac{1}{2}}, \\ SF(F) = (RF^2 + CF^2)^{\frac{1}{2}} \end{cases}, \quad (14)$$

$$IE(F) = -\sum_{i=0}^{L-1} p(i) \log_2 p(i), \quad (15)$$

$$UIQI(A, F) = \frac{\sigma_{AF}}{\sigma_A \sigma_F} \cdot \frac{2\bar{A}\bar{F}}{\bar{A}^2 + \bar{F}^2} \cdot \frac{2\sigma_A \sigma_F}{\sigma_A^2 + \sigma_F^2}, \quad (16)$$

where A and F are respectively the IR and fused images, L represents the gray level, $p(i)$ the histogram of intensity i , \bar{A} and \bar{F} denote respectively the mean values of A and F , and σ_A , σ_F and σ_{AF} imply the variances of A and F and covariance of A and F , respectively.

In addition, we have also used the running time as a metric that measures the computational complexity of the fusion methods. To avoid the influence of running environment on the running time, we have calculated the running time by taking the average of 10 executions of each method in seconds.

The results of fusion evaluation are illustrated in Table 1. Here only the results concerned with the fused images ‘quad’, ‘uncamp’ and ‘soldier’ are presented for brevity. The best value in each row of Table 1 is marked in bold.

Table 1. Comparison of different fusion metrics for the fused images ‘quad’, ‘uncamp’ and ‘soldier’, as obtained by the DTCWT, CVT, MSVD, GFF, LP-SR, GTF and IT-DMM reference methods and our method.

	Metric	DTCWT	CVT	MSVD	GFF	LP-SR	GTF	IT-DMM	Our method
‘quad’ 632 × 496	SF	0.0474	0.0474	0.0391	0.0477	0.0506	0.0370	0.0435	0.0703
	IE	5.9963	6.0041	5.9350	6.5308	6.8285	6.6205	5.6679	6.3289
	UIQI	0.6325	0.5999	0.6106	0.6787	0.6739	0.6048	0.2048	0.6337
	Time, s	0.4462	1.6652	0.6840	0.5531	0.1932	7.6043	0.4237	0.1393
‘uncamp’ 360 × 270	SF	0.0408	0.0414	0.0336	0.0384	0.0476	0.0333	0.0342	0.0677
	IE	6.3632	6.3910	6.2495	6.3778	7.0171	6.6779	6.7134	6.6913
	UIQI	0.4054	0.4154	0.4599	0.4689	0.4263	0.2323	0.2080	0.4700
	Time, s	0.1392	0.5686	0.2088	0.1208	0.0549	1.2873	0.1309	0.0497
‘soldier’ 768 × 576	SF	0.0526	0.0526	0.0425	0.0548	0.0565	0.0463	0.0423	0.0653
	IE	7.0105	7.0116	6.9441	7.3549	7.5273	6.7850	7.2972	6.9236
	UIQI	0.1992	0.2149	0.3342	0.0543	0.2427	0.2254	0.1227	0.2037
	Time, s	0.6395	2.3300	0.9347	0.8030	0.2371	8.8869	0.6034	0.2117

Regarding the ‘quad’ image group, our method provides the two best values in terms of the SF and the running time, although it is not the best according to the IE and UIQI metrics (see Table 1). Nonetheless, the scores of our method are very close to the best ones. This indicates that the method proposed in this work provides at least a quite competitive performance. Similar conclusions can also be drawn from the data referred to the ‘uncamp’ and ‘soldier’ image groups.

For still better comparison and more precise conclusions, the scores of fusion performances corresponding to all 10 fused images are plotted as line charts in terms of each fusion metric (see Fig. 6). The appropriate data confirms the above conclusions.

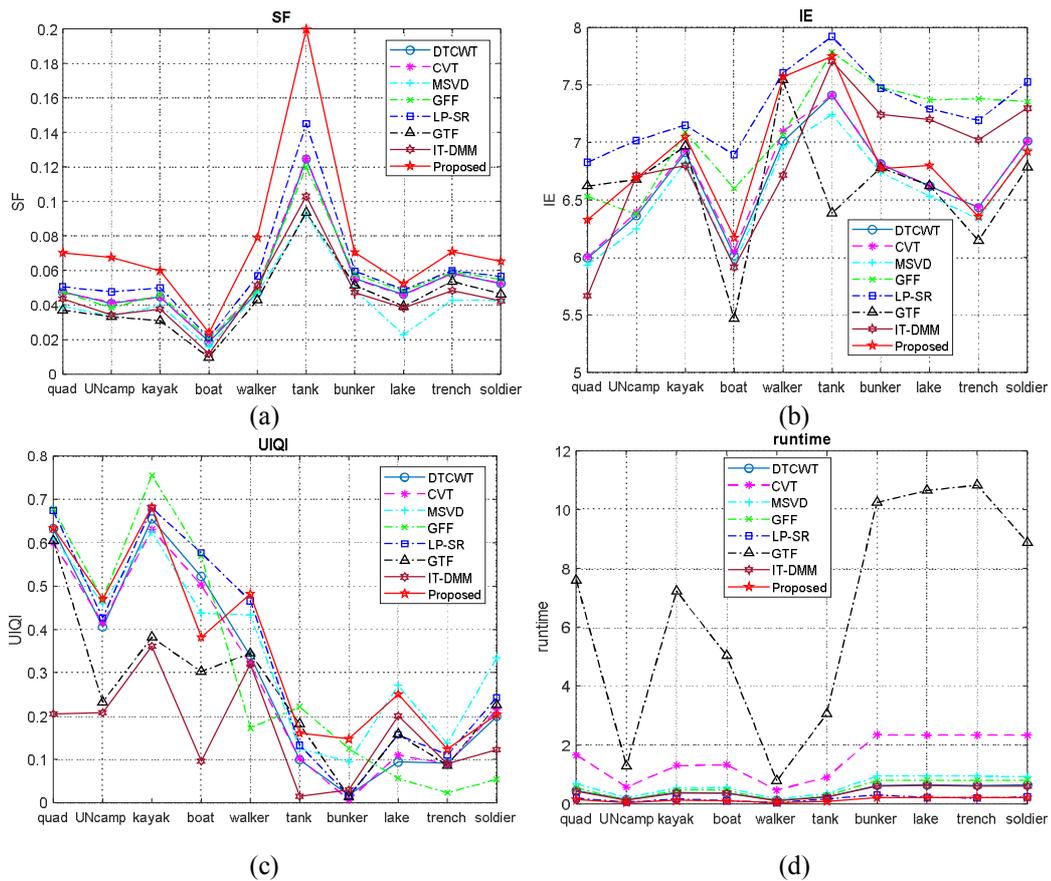


Fig. 6. Quantitative comparisons of SF (a), IE (b), UIQI (c) and running time (d) for the eight fusion methods applied to 10 groups of images.

In summary, both the subjective and objective evaluations make it evident that the method proposed in this work can quickly and efficiently solve the problems of fusion. It demonstrates superior or, at least, quite competitive performance, if compared to the common state-of-the-art methods. As a consequence, our method can be applied for real-time fusion.

5. Conclusion

We suggest a new two-scale method for fusing IR and VIS images, which is based upon the median filter and the intensity-transfer approach. Inspired by the method of two-scale image decomposition developed in the previous works, we use the median filter to implement a new two-scale decomposition, which proves to be simple, efficient and parameterless. Afterwards, the fused base layer is obtained by utilizing a novel optimization fusion rule based on the intensity transfer. This enables us to approximate simultaneously the IR and VIS images. Then the fusion is transformed into matrix operations. The fusion of the detail layers adopts an improved maximum-selection rule to maximize the information about details, with some image-enhancement effect.

In our experiments, we have compared seven common and state-of-the-art fusion approaches (DTCWT, CVT, MSVD, GFF, LP-SR, GTF and IT-DMM) with our method. The results of fusion have been compared both subjectively and objectively on a number of standard IR-and-VIS image pairs. The objective metrics include the SF, the IE, the UIQI and the running time. The results obtained by us indicate that our method can, in principle, achieve a superior performance in terms

of both the human visual perception and the objective metrics. Furthermore, our method is computationally efficient and parameterless (or adaptive), which makes it quite suitable for real-time fusion applications.

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. Ma J, Ma Y and Li C, 2019. Infrared and visible image fusion methods and applications: a survey. *Information Fusion*. **45**: 153–178.
3. 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.
4. Ma J, Yu W, Liang P, Li C and Jiang J, 2019. FusionGAN: a generative adversarial network for infrared and visible image fusion. *Information Fusion*. **48**: 11–26.
5. Zhou Z, Wang B, Li S and Dong M, 2016. Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters. *Information Fusion*. **30**: 15–26.
6. Jiang W, Yang X, Wu W, Liu K, Ahmad A, Sangaiah A K and Jeon G, 2018. Medical images fusion by using weighted least squares filter and sparse representation. *Computers & Electr. Eng.* **67**: 252–266.
7. Li S, Kang X and Hu J, 2013. Image fusion with guided filtering. *IEEE Trans. Image Proc.* **22**: 2864–2875.
8. 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.
9. Bavirisetti D P and Dhuli R, 2016. Two-scale image fusion of visible and infrared images using saliency detection. *Infrared Phys. & Technol.* **76**: 52–64.
10. Naidu A R, Bhavana D, Revanth P, Gopi G, Kishore M P and Venkatesh K S, 2020. Fusion of visible and infrared images via saliency detection using two-scale image decomposition. *Int. J. Speech Technol.* **23**: 815–824.
11. Kalaiselvi T, Selvi S K and Abirami B, 2020. Two scale adaptive median filter for denoising MRI images. *Conference: Computational Methods, Communication Techniques and Informatics*. 155–159.
12. Li Z and Yan H, 2019. Infrared and visible image fusion via intensity transfer and direct matrix mapping. *Infrared Phys. & Technol.* **102**: 103030.
13. Jiao L C and Tang Q X, 2013. Fusion for visual context enhancement using intensity transformation function of infrared images. *Electron. Lett.* **49**: 751–752.
14. Zhai Y and Shah M. Visual attention detection in video sequences using spatiotemporal cues. In: *Proc. 14th ACM International Conference on Multimedia*, October 2006, pp. 815–824.
15. Huang T, Yang G and Tang G, 1979. A fast two-dimensional median filtering algorithm. *IEEE Trans. Acoustics, Speech, and Signal Proc.* **27**: 13–18.
16. https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029
17. 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.
18. Nencini F, Garzelli A, Baronti S and Alparone L, 2007. Remote sensing image fusion using the curvelet transform. *Information Fusion*. **8**: 143–156.

-
19. Naidu V P S, 2011. Image fusion technique using multi-resolution singular value decomposition. Defence Sci. J. **61**: 479.
 20. 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.
 21. <https://github.com/jiayi-ma>
 22. Eskicioglu A M and Fisher P S, 1995. Image quality measures and their performance. IEEE Trans. Commun. **43**: 2959–2965.
 23. Xu W, Li M and Wang X, 2017. Information fusion based on information entropy in fuzzy multi-source incomplete information system. Int. J. Fuzzy Syst. **19**: 1200–1216.
 24. Wang Z and Bovik A C, 2002. A universal image quality index. IEEE Signal Proc. Lett. **9**: 81–84.

Jinjuan Wang, Xiliang Zeng, Shan Duan, Qun Zhou and Hao Peng. 2022. A fast and efficient two-scale fusion of infrared and visible images using median filter and intensity transfer. Ukr.J.Phys.Opt. **23**: 155 – 165. doi: 10.3116/16091833/23/3/155/2022

***Анотація.** Запропоновано швидкий та ефективний метод злиття інфрачервоних і видимих зображень, який базується на медіанному фільтрі та передачі інтенсивності. Медіанний фільтр використовують для побудови нової двомасштабної декомпозиції зображення та отримання базового шару та шарів деталей. Потім злитий базовий шар одержують за допомогою оптимізаційного правила злиття, заснованого на передачі інтенсивності, де злиття зводиться до матричних операцій. Об'єднаний шар деталей одержують за допомогою вдосконаленого правила максимального вибору, яке максимізує інформацію про зовнішній вигляд деталей із деяким ефектом покращення зображення. Наші експериментальні результати демонструють, що такий метод може забезпечити вищу продуктивність злиття з точки зору і суб'єктивних оцінок, і об'єктивних показників у порівнянні з сімома стандартними сучасними методами. Крім того, запропонований метод цілком підходить для злиття зображень у реальному часі.*