https://doi.org/ 10.480467/AFJBS.6.Si3.2024.1115-1126

Hybrid Fusion Based Target Detection Methods Visible and Images Using Guided Filters

[1]S.Kavitha, [2]Santhalakshmi M, [3]Shanthakumar M, [4]Janarthanam S

[123] Assistant Professors, Department of Computer Science, Kamban College of Arts and Science, Coimbatore.

[4]Associate Professor, School of Science and Computer Studies, CMR University, Bengaluru.

Corresponding Author : professorshanth@gmail.com

Volume:6,IssueSi3,2024

Received:14April2024

Accepted: 11May2024

doi:

10.48047/AFJBS.6.Si3.2024.1115-1126

Abstract

The combination of visible and infrared (IR) images to create an absolute, accurate, and dependable image is the most widely used technique in image processing applications. This paper proposes a method called TwoScale Decomposition (TSD) with Guided Filtering (GF) and Phase Congruency (PC) and Sum Modified Laplacian (SML) has been developed. TSD is used in TSD-PS-GF to decompose the IR and VIS pictures, respectively, and obtain the base and detail layers to use TSD-PS with Sturdy GF (TSD-PS-SGF) rather than GF to improve the image fusion's resilience. The Enhanced Preconditioned Conjugate Gradient (EPCG) method is used for each SGF iteration to optimize the conjugate routes and RLS. This method handles irregularities in the structure and achieves a high rate of convergence.

Keywords : Infrared, Visible, Two-Scale Decomposition, TSD-PS-SGF, TSD-PS-ASGF, TSD-DNN-ASGF

Introduction

Target identification is a very crucial technology for military operations that has not yet realized its complete tactical secure. It is the process of recognizing the potential military target as being a specific target. Completely reliable automated target identification can improve the lethality and survivability of the war fighter and platform. Automated target identification performs on sensor information for processing data to make the decision. The primary value added to a weapons system of automated target identification is engagement timeline reduction for targets acquisition. The rapid acquisition and servicing of targets increase lethality and survivability of the weapons platform.

Persistent Surveillance presents an earlier military application chance for providing lower technically sophisticated automated target identification. Automated target identification can provide significant enhancement to military weapons platforms over human-only performance. Also, can provide enhancements to the weapons operator or intelligence analyst for fire control, surveillance, reconnaissance, intelligence, PS and situational awareness provide completely autonomous target engagements like for missile seekers.

For satellite imagery, the actual impetus was the program of unmanned planetary satellite in the 1960s that transmitted images to the ground receiving stations; the low visual quality of the received images required the development of processing approaches to provide the images valuable. The other impetus was the Land sat program which began in 1972 and provided repeated worldwide exposure in digital plan. A third impetus is the continued growth of faster and more powerful computers, peripheral device and software that are applicable for image processing [8]. Mostly, Infrared (IR) images are used for military applications such as automatic target identification of military relevant land targets. A precise identification of target from a source IR and VIS image is a challenging process due to the fact that a target can have complicated structure and can modify shape or dimension.

The objective of this work to fuse these two types of images can join the merits of radiation details in IR images and complete texture details in VIS images. The precise, reliable and paired descriptors of the scene in fused images make these methods be widely used in different applications.

Related Works

DUQ et al., [1] was proposed for fusing IR and VIS images of different resolutions and generating high-resolution images to acquire understandable and perfect fused images. In this method, the fusion difficulty was devised as a Total Variation (TV) minimization problem. The data reliability term limits the pixel intensity resemblance of the down sampled fused image with respect to the IR image and the regularization term induces the gradient resemblance of the fused image with respect to the VIS image.

YAN et al., [2] was proposed for IR and VIS image fusion on the basis of l_2-norm. In this method, the fusion process was formulated as l_2-norm optimization problem where the primary term computed by the l_2-norm limits the fused image to have the same pixel

intensities as the IR image and the secondary term measured by l_2-norm forces the fused image to have the same gradient distribution as the VIS image. Also, two weights were introduced for optimizing the l_2-norm and acquiring the fused image effectively.

DON et al., [4] was proposed by a novel hyperspectral image fusion method. In the method, the features of the panchromatic and hyperspectral images were simultaneously considered. The GF was applied for generating the spatial detail image of each hyperspectral image band efficiently.

PIA19 et al., [5] was proposed by using the deep neural network. In this technique, a Siamese CNN was applied for estimating the activity measure and automatically generating the weight map according to the saliency of each pixel for a pair of source images. The source images were decomposed into low- and high-frequency sub-bands by using the 3-level wavelet transform [5]and the fused image was acquired by restoring the wavelet images with the scaled weight maps.

ZHA20 et al., [7] was proposed based on the GAN for fusing unmatched IR and VIS images. In this model, the GAN two-player game was used for fusing IR and VIS images. By using this model, the corresponding IR image from a VIS image was generated and these two images were fused together for obtaining more information. However, the computational complexity of this model was high and also the training was difficult compared to other deep learning models.

SHO19 et al., [6] was proposed for fusing VIS and thermal images that focuses on generating fused images with high VIS similarity to normal RGB images when introducing new informative details in pedestrian regions. In this method, two types of objective functions were applied for optimization such as a similarity metric between the RGB input and the fused output for achieving natural image appearance and an auxiliary pedestrian detection error for defining relevant features of the human appearance and blending them into the output.

Limitations

From the survey a precise identification of target from a source IR and VIS image is a challenging process due to the fact that a target can have complicated structure and can modify shape or dimension. The major challenges found and should be considered when developing a system for target identification are follows:

- Targets on the edge of the field of sights.
- Targets that are invisible or in shadows.
- Targets that can be heard but not observed.
- Targets under less than ideal indirect fire illumination.
- Natural and man-made obstacles

To overcome the limitations this research work to fuse the two types of images which can join the merits of radiation details in IR images and complete texture details in VIS images. The precise, reliable and paired descriptors of the scene in fused images make these techniques be widely used in different applications mention in the proposed method.

Proposed Methodology

The proposed work consists of three stages shown in Fig.1. First work is an enhanced version of the image fusion technique called TSD-PS with Sturdy GF, Next, TSD-PS with Adaptive SGF (TSD-PS-ASGF) method is proposed for sharpening the output of SGF. Third one is TSD-Deep Neural Network with ASGF (TSD-DNN-ASGF) method is proposed that utilizes DNN instead of PC and SML schemes to construct the saliency weighted maps of base and detail layers. In this method, a weightedaveraging scheme is applied to fuse the base layers whereas DNN is applied to extract the detail layer features for preserving more information. In pre-processing step, two scale decomposition is applied for establishing them in the fused image.

TSD-PS-SGF METHOD

The proposed method has four phases such as TSD, generation of saliency of base and detail layers with PC and SML, respectively, computing weighting maps with SGF and two-scale image restoration.

TSD and Saliency Map Construction

First, the source images I_n (i,j) are decomposed by TSD method in which a median filter is used to approximate the base layer as:

$$X_n(i,j) = I_n(i,j) * \mu(i,j) \quad (1)$$

In Eq. (1), n represents the n^{th} source image and μ denotes the median filter. The range of median filter is set to 35×35 to obtain the detail layer as:

$$Y_n(i,j) = I_n(i,j) - X_n(i,j) \quad (2)$$

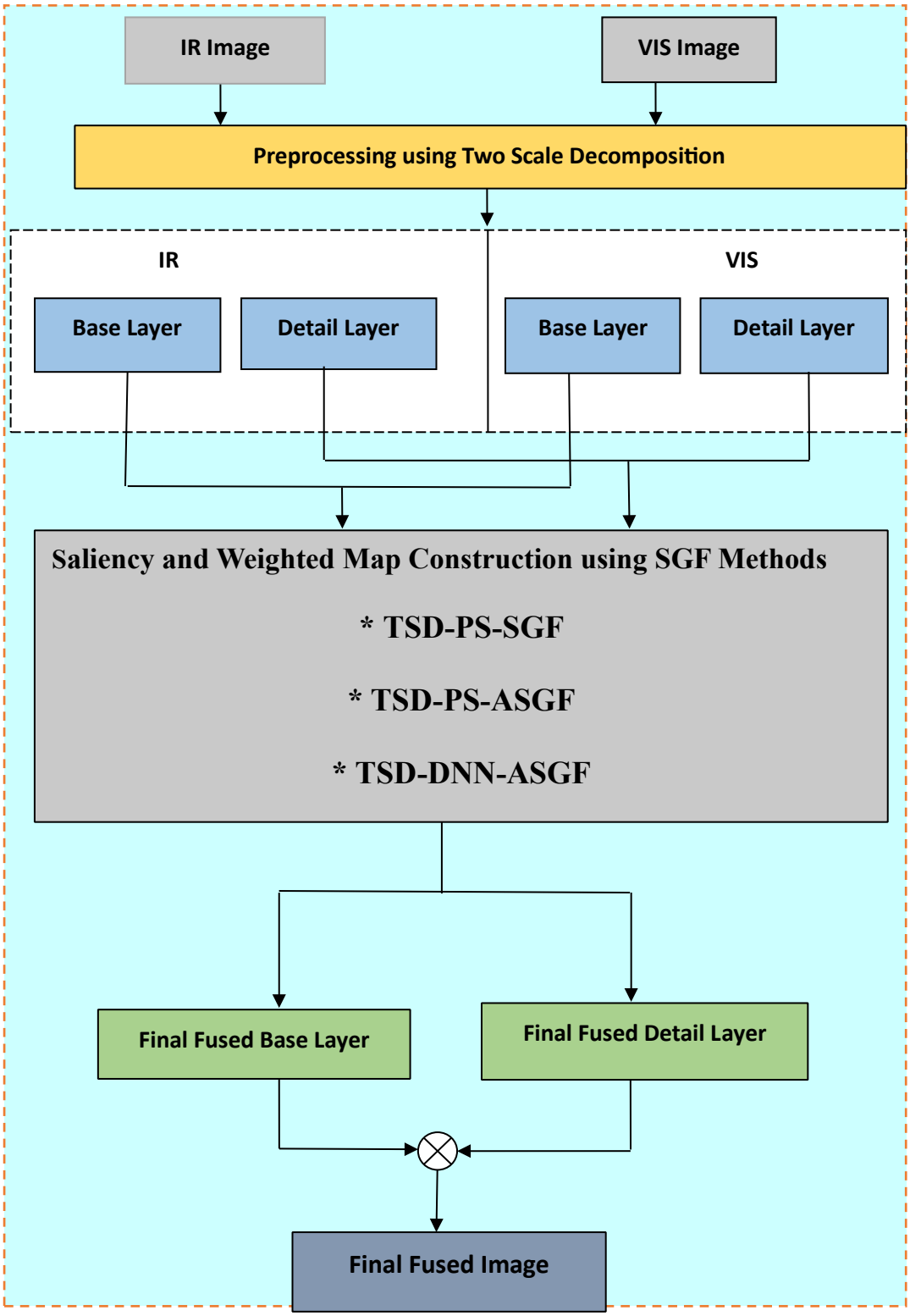


Fig .1. Proposed Fusion based Target identification Method

Afterwards, PC technique is adopted to create saliency maps of base layer. The PC is computed by extracting the phase details using log-Gabor filter as: $PC(i) = \frac{\sum_l El(i)}{\sum_l \sum_n Anl(i)+\varepsilon}$ (3)

In Eq. (3), l is the orientation, $A_{nl}(i)$ denotes the sum of amplitudes, E_l (i) denotes the local energy and ε refers to the small constant coefficient for preventing the zero denominator, therefore:

$$E_l(i) = \sqrt{F(i)^2 + H(i)^2} \quad (4)$$

$$F(i) = \sum_n s_{nl}(i), H(i) = \sum_l h_{nl}(i) \quad (5)$$

$$A_{nl}(i) = \sum_l \sqrt{s_{nl}(i)^2 + h_{nl}(i)^2} \quad (6)$$

In Eq. (5), s_{nl} (i) and h_{nl} (i) are outcome of the even symmetric and odd filters, correspondingly. A large SML is used to create the saliency maps of detail layer. Consider the location of pixel is a, thus SML at pixel a is described like:

$$SML(a) = \sum_{k=\omega a} ML(k)^2 \quad (7)$$

In this algorithm, the pixel a is used as the centroid to construct a rectangular window ω_a. If k=(i,j), then d_l is the n^{th} detail layer and step is the distance between pixels.

Weighted Map Construction

Initially, the weighted maps are generated by evaluating the saliency maps. Then, the GF is applied to enhance the spatial stability of weighted maps in which the parameters are optimized based on the IRL Swith EPCG algorithm. The weighted maps of detail layers in VIS and IR images are P_{VIS}^d and P_{IR}^d determined by comparing the saliency maps of this layer and Similarly, the weighted maps of base layers in VIS and IR images are also computed as:

$$P_{VIS}{}^b = \begin{cases} 1, & S_{VISb} > S_{IRb} \\ 0, & Otherwise \end{cases}; P_{IR}{}^b = 1 - P_{VIS}{}^b \quad (8)$$

Normally, the weighted maps are not properly associated with object margins which may generate artifacts in the fused image. Hence, SGF is carried out on each weight map P_{VIS}^d,P_{VIS}^b,P_{IR}^b and P_{IR}^d with equivalent source image I_n(i,j). Therefore, the resultant weighted maps of base and detail layers of both VIS and IR images are as:

$$W_{VIS}{}^b = G(P_{VIS}{}^b, X_{VIS}, r_1, \varepsilon_1), W_{IR}{}^b = G(P_{IR}{}^b, X_{IR}, r_1, \varepsilon_1) \quad (9)$$

$$W_{VIS}{}^d = G(P_{VIS}{}^d, Y_{VIS}, r_2, \varepsilon_2), W_{IR}{}^d = G(P_{IR}{}^d, Y_{IR}, r_2, \varepsilon_2) \quad (10)$$

Thus, the EPCG is used for determining the residual and the corresponding weighting maps until the convergence is satisfied.

Adaptive Sturdy Guided Filtering

Given a source image$I_n(i,j)$, the weighted maps of base (P^b) and detail layers (P^d) for IR and VIS images are applied to the ASGF. Therefore, the resultant weighted maps are:

$$W_{VIS}{}^b = G(P_{VIS}{}^b, X_{VIS}, r_1, \varepsilon_1), W_{IR}{}^b = G(P_{IR}{}^b, X_{IR}, r_1, \varepsilon_1) \quad (11)$$

$$W_{VIS}{}^d = G(P_{VIS}{}^d, Y_{VIS}, r_2, \varepsilon_2), W_{IR}{}^d = G(P_{IR}{}^d, Y_{IR}, r_2, \varepsilon_2) \quad (12)$$

Here X_{VIS} and Y_{VIS} are base and detail layers of VIS images, respectively. In the same manner, X_{IR} and Y_{IR} are base and detail layers of IR images, respectively.

The weight $w^{ASGF}(a,z)$ is computed as a reproduction of weighted maps of base and details layers for VIS and IR images:

$$w^{ASGF}(a,z) = \xi_z{}' + ((P_{VIS}{}^b + P_{IR}{}^b) \cdot (P_{VIS}{}^d + P_{IR}{}^d)) \quad (13)$$

$$\text{Where } P_{VIS}{}^d = \begin{cases} 1, & S_{VISd} > S_{IRd}; \ P_{IRd} = 1 - P_{VISd} \\ 0, & Otherwise \end{cases} \quad (14)$$

$$P_{VISb} = \begin{cases} 1, & S_{VISb} > S_{IRb}; \quad P_{IR}{}^b = 1 - P_{VIS}{}^b \\ 0, & Otherwise \end{cases} \quad (15)$$

Here, $S_{VIS}{}^b$ and $S_{VIS}{}^d$ are saliency maps of base and detail layer for VIS images. Likewise, $S_{IR}{}^b$ and $S_{IR}{}^d$ are the saliency maps of base and detail layer for IR images. $\xi_z{}'$is the offset which is defined by the naive offset selection scheme as:

$$\xi_z{}' = \begin{cases} Max(\omega_{i,j}) - I, & if\ \Delta' > 0 \\ Min(\omega_{i,j}) - I, & if\ \Delta' < 0 \\ 0, & if\ \Delta' = 0 \end{cases} \quad (16)$$

$$\text{Where } \Delta' = I - \mu_{i,j}$$

Here, Δ' is the intensity variance between pixel z and the mean ($\mu_{i,j}$) of the guided image I in Gaussian window$\omega_{i,j}$.

The proposed TSD-DNN-ASGF method is briefly explained. Assume there are N preregistered source images($I_n, n \in \{1, \dots, N\}$). Then, TSD method is applied to decompose the source images. For each source imageI_n, the base layers (X_n) and the detail layers (Y_n) are obtained by TSD. Then, the base layers are fused

by the weighted-averaging method and the detail information is reconstructed by the DNN. Finally, the fused image F is reconstructed by combining the fused base $\bar{X}$ and detail layers$\bar{Y}$.

For the detail layerY_1, Y_2, ... , Y_n, a DNN i.e., VGG-19 network is applied for extracting the deep features. After that, the weighted maps are obtained by a multi-layer fusion method. At last, the fused detail layer is reconstructed by these weighted maps and the detail layer.

Assume the detail layer Y_n and $\phi_n{}^{x,m}$ is the feature maps of n^{th} detail layer is extracted by the x^{th} layer where m ($m \in \{1,2, \dots, M\}$, $M = 64 \times 2^{i-1}$) denotes the channel number of x^{th} layer.

$$\phi_n{}^{x,m} = \Phi_x(Y_n) \tag{17}$$

After the deep features $\phi_n{}^{x,m}$ are obtained, the activity level map $C_n{}^x$ is computed by l_1-norm and block-based average operator. The l_1-norm of $\phi_n{}^{x,1:M}(i,j)$ is the activity level measure of the source detail layer. Therefore, the primary activity level map $C_n{}^x$ is obtained as:

$$C_n{}^x(i,j) = \|\phi_n{}^{x,1:M}(i,j)\|_1 \tag{18}$$

Then, the block-based average operator is used for computing the final activity level map $C_n{}^x(i, j)$ for providing the fusion method more robust to misregistration.

$$C_{nx}(i,j) = \sum\_\_\_\_\_ r\beta=-r \sum r\theta(=2-rr+C1nx)2(i+\beta,j+\theta) \tag{19}$$

consider $r = 1$. Then, the key weighted maps $W_n{}^x$ are determined by the softmax operator as:

$$^x(i,j) = \_\_\_\_\_ NC_n{}^x(Cinx,j()i,j) \tag{20}$$

$$W_n \quad \sum n=1$$

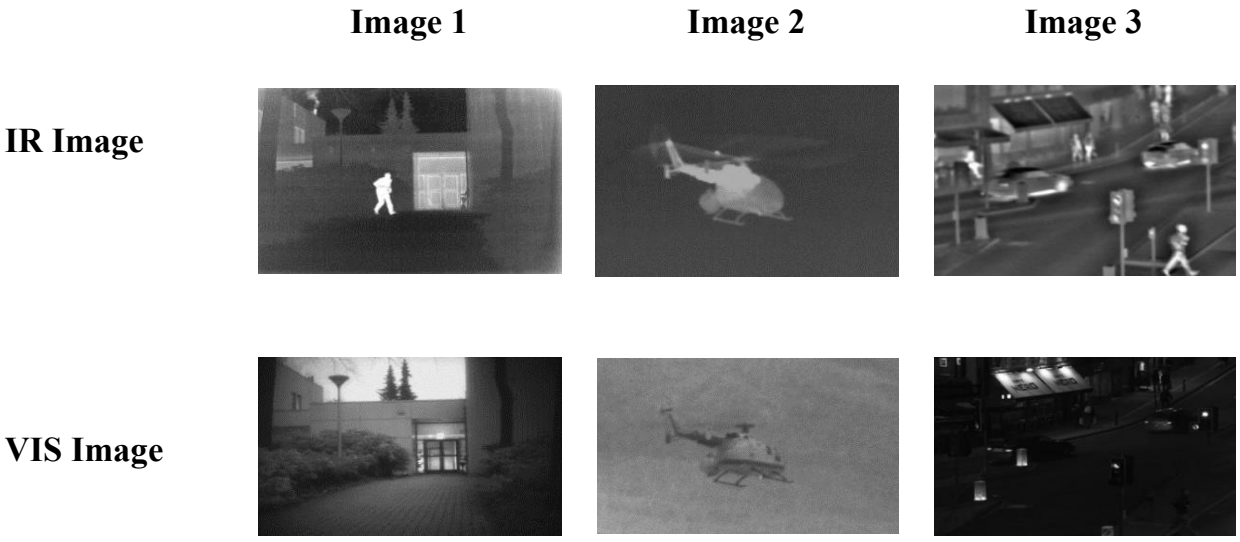
The key weighted map values $W_n{}^x(i, j)$ are in the range of[0,1]. The pooling operator in VGG-network is a type of sub-sampling method. Each time this operator can resize the feature maps to ^{1} times of the original size where s $_s$ denotes the stride of the pooling operator which is equal to 2. After all the weighted maps are obtained, these are applied to ASGF to improve the smoothness and sharpness of the fused detail layer efficiently.

Experimental Results

The existing and proposed methods are evaluated in terms of information theory-based metric (Mutual Information (MI)), a Visual Information Fidelity (VIF), an image feature-based metric(Q^{AB}_{F}), an quality of weight fusion metric (Q^W), an image edge-based metric (Q^E) and an image structure-based metric(Q^{SSIM}) to ensure the effectiveness of the proposed methods.

Table 1 Numerical Analysis of Proposed Methods for TNO Image Fusion Datasets

Metrics/Methods	NSCT-PS-GF	TSD-PS-GF	TSD-PSSGF	TSD-PSASGF	TSD-DNNASGF
MI	2.14	2.43	2.51	2.62	2.83
Q^{AB}_{F}	0.544	0.568	0.592	0.615	0.637
Q_{SSIM}	0.628	0.662	0.666	0.681	0.702
VIF	0.4740	0.4956	0.5108	0.5397	0.5516
Q_W	0.8306	0.8351	0.8474	0.8596	0.8759
Q_E	0.1006	0.1886	0.2051	0.2284	0.2499



NSCT-PS-GF

TSD-PS-GF

TSD-PS-SGF

TSD-PS-ASGF

TSD-DNN-ASGF

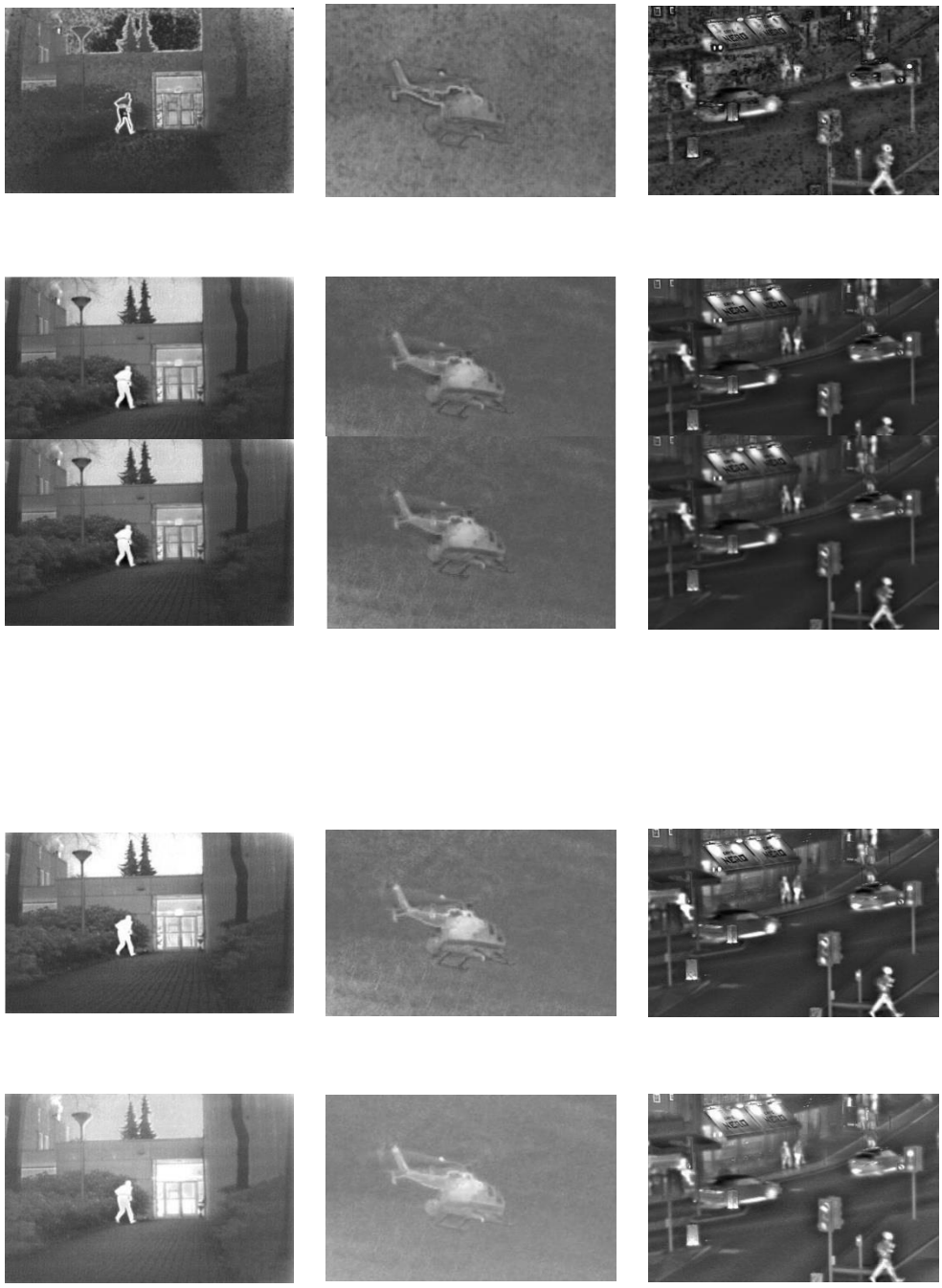


Figure 2 Fused Image Results of Proposed Method for TNO Image Fusion Dataset

Image 1 **Image 2** **Image 3**

IR Image

VIS Image

NSCT -PS- GF

TSD- PS-GF

TSD- PS- SGF

TSD-

PS-

ASGF

TSDDNN-

ASGF

Figure 3 Fused Image Results Proposed Method for TRICLOBS Image Dataset

CONCLUSION

The enhanced image fusion technique for achieving effective military target identification is proposed that integrates the decomposition and filtering methods to enhance the fused image visual quality by preserving the most essential details of the source images. This research focuses on enhancing the visual quality of the fused image by sharpening and preserving more edge details with the minimum computational complexity. The key objective of this enhanced fusion is to use an adaptive robust filtering and deep learning methods in decomposing the source images, constructing the weighted maps to fuse them and get the final fused image.

References

[1] Du, Q., Xu, H., Ma, Y., Huang, J., & Fan, F. (2018). Fusing infrared and visible images of different resolutions via total variation model. Sensors, 18(11), 3827.

[2] Yan, H., & Li, Z. (2019). Novel model for infrared and visible image fusion based on l_2 norm. OSA Continuum, 2(11), 3076-3090.

[3] Zhang, Y., Wei, W., & Yuan, Y. (2019). Multi-focus image fusion with alternating guided filtering. Signal, Image and Video Processing, 13(4), 727-735.

[4] Dong, W., Xiao, S., & Qu, J. (2018). Fusion of hyperspectral and panchromatic images with guided filter. Signal, Image and Video Processing, 12(7), 1369-1376.

[5] Piao, J., Chen, Y., & Shin, H. (2019). A New Deep Learning Based MultiSpectral Image Fusion Method. Entropy, 21(6), 570.

[6] Shopovska, I., Jovanov, L., & Philips, W. (2019). Deep visible and thermal image fusion for enhanced pedestrian visibility. Sensors, 19(17), 3727.

[7] Zhao, Y., Fu, G., Wang, H., & Zhang, S. (2020). The Fusion of Unmatched Infrared and Visible Images Based on Generative Adversarial Networks. Mathematical Problems in Engineering, 2020.

[8] Feng, Y., Lu, H., Bai, J., Cao, L., & Yin, H. (2020). Fully convolutional network-based infrared and visible image fusion. Multimedia Tools and Applications, 1-14.

[9] Ge, Y., & Jing, G. (2019). Infrared and visible image fusion using multiresolution convolution neural network. In Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing (pp. 1-5).

[10] Han, X., Lv, T., Song, X., Nie, T., Liang, H., He, B., &Kuijper, A. (2019). An adaptive two- scale image fusion of visible and infrared images. IEEE Access, 7, 56341-56352.

[11] Jinju, J., Santhi, N., Ramar, K., & Bama, B. S. (2019). Spatial frequency discrete wavelet transform image fusion technique for remote sensing applications. Engineering Science and Technology, an International Journal, 22(3), 715-726.

[12] Ma, J., Ma, Y., & Li, C. (2019). Infrared and visible image fusion methods and applications: A survey. Information Fusion, 45, 153-178.

[13] Ma, J., Liang, P., Yu, W., Chen, C., Guo, X., Wu, J., & Jiang, J. (2020). Infrared and visible image fusion via detail preserving adversarial learning. Information Fusion, 54, 85-98.

[14] Mishra, A., Mahapatra, S., & Banerjee, S. (2017). Modified Frei-Chen operator-based infrared and visible sensor image fusion for real-time applications. IEEE Sensors Journal, 17(14), 4639-4646.

[15] Piao, J., Chen, Y., & Shin, H. (2019). A New Deep Learning Based MultiSpectral Image Fusion Method. Entropy, 21(6), 570.

[16] Precilla, A. C., George, J., Kannan, S. R., &Prabhu, R. (2018). Modified PCA based image fusion using feature matching. International Journal of Pure and Applied Mathematics, 119(15), 477-483.

[17] Shao, Z., &Cai, J. (2018). Remote sensing image fusion with deep convolutional neural network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(5), 1656-1669.

[18] Shopovska, I., Jovanov, L., & Philips, W. (2019). Deep visible and thermal image fusion for enhanced pedestrian visibility. Sensors, 19(17), 3727.

[19] Zhang, Y., Wei, W., & Yuan, Y. (2019). Multi-focus image fusion with alternating guided filtering. Signal, Image and Video Processing, 13(4), 727-735.

[20] Zhao, Y., Fu, G., Wang, H., & Zhang, S. (2020). The Fusion of Unmatched Infrared and Visible Images Based on Generative Adversarial Networks. Mathematical Problems in Engineering, 2020.