[1.] Gargi J. TRIVEDI, [1.] Rajesh SANGHVI

HYBRID MODEL FOR INFRARED AND VISIBLE IMAGE FUSION

[1.] Department of Mathematics, G H Patel College of Engineering & Technology, CVM University, Vallabh Vidhyanagar-388120, INDIA

Abstract: In this work, we offer a novel hybrid approach to multimodal and multisensor image fusion employing DCT (discrete cosine transform) in the DWT (discrete wavelet transform) domain and FPDE (partial differential equations of order four). To begin, we use a technique relied upon on partial differential equation of order four to separate the source images into a base image and a detail image. Detail photos are averaged according to their primary components' weights to get the final image. The next step is to feed the foundational pictures into DWT decomposition. DCT is applied on the coefficients of the corresponding four sub-bands. Coefficients in the sub-bands may be decomposed using DCT to provide useful information. To fine-tune the retrieved features, the spatial frequency of each coefficient is determined. Discrete cosine transformation coefficients are fused using the fusion rule, which takes into account the spatial frequency value. Inverse DCT and inverse DWT are used to produce the final base picture. As a result of linearly joining the final base and detail pictures, a fused picture is produced. The suggested approach is compared to already available fusion methods. The results show that the suggested approach outperforms the state-of-the-art techniques when measured against objective metrics including mean, standard deviation, mutual information, and entropy.
Keywords: Infrared images(IR),Visual images(VI), Principal component Analysis (PCA), Discrete Wavelet Transformation (DWT), Discrete cosine Transformation (DCT), Multimodal Images, , Image fusion

1. INTRODUCTION

Over the last several years, picture fusion has risen to prominence as a topic of study [1,2]. Thanks to advancements in imaging technology and resulting price reductions, more and more individuals are able to acquire photographs. Image fusion continues to garner interest because it helps meet the growing need for more complete and trustworthy pictures. According to [3,4] When it comes to infrared and visible the image fusion approach is to merge relevant information from numerous sensors into a single informative fused picture. By extracting all complimentary features from the input images and eliminating any inconsistencies in the final fused picture, successful image fusion produced from both sensors may overcome the constraints of a single sensor and provide numerous complementary qualities. Fusion picture attempt to combine the qualities of source photos for better visual effects and give rich information to enhance decision-making. Recent studies in the area of multi-sensor fusion have focused more and more on fusing infrared and visible images. Each kind of sensor, visible and infrared, has its own set of pros and cons. Because infrared cameras are more attuned to thermal and radiative information, heat source objects seem brighter and have greater pixel values in infrared photographs. However, infrared photos lack texture details because of the sensor's limits.[5] Visible pictures provide rich information since visible image sensors catch reflected light from objects. Because of this, it's clear that combining pictures Spatial domain methods, multi-directional and multiscale decomposition methods like Discrete Cosine Transformation (DCT) [6,7], Non sub band curvelet Transformation(NSCT) [7], Discrete wavelet Transformation (DWT)[5], hybrid methods [8], and fusion schemes related to preservation of edges [10] and fusion schemes based on fuzzy approach[9] are some of the image fusion schemes that have been implemented in the literature. cause unintended distortions such artefacts, halo effects, loss of edge detail, and alterations to the original picture structure. These procedures need several iterations before they can produce a perfect picture. This article presents a novel hybrid approach to the discrete wavelet transform (DWT) based on the DCT (discrete cosine transform) and FPDE(partial differential equations of order four) [11]and First, we use a fourth-order differential equations-based approach to separate the input pictures into a base image and a detail image. The primary components of the weighted average of the detail photos are used to generate the final image. The next step is to feed the foundational pictures into DWT decomposition. DCT is applied on the coefficients of the corresponding four sub-bands. With DCT, we may get useful information from the sub band coefficients. In order to enhance the retrieved characteristics, the spatial frequency of each coefficient is determined. Last but not least, the DCT coefficients are fused using a fusion algorithm that takes into account the spatial frequency value. Inverse DCT and inverse DWT are used to produce the final base picture. Last but not least, a merged picture is produced by linearly integrating the final detail and base pictures discussed before. Our approach gets through those obstacles and yields a fused picture with more information, less fusion loss, and less noticeable artefacts. Overall, the suggested one is an efficient

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN-L 1584 – 2665

image fusion strategy due to its qualities of requiring minimal computing time, being extremely effective, and being easily implemented.

2. METHODS

Our article makes use of a number of theoretical frameworks, some of which are briefly introduced below.

Partial Differential Equations

In the framework of PDE (partial differential equations), all the images are treated as continuous objects. In our case, a rethinking and reworking of the strategy has led to a proper and novel execution. Second-order PDEs and fourth order PDEs are two types of partial differential equations. Anisotropic diffusion [12] is the smoothing of isotropic areas while keeping edges intact using second order partial differential equations. However, this approach has the downside of producing blocky effects, which makes for unreliable object recognition in computer vision. By compensating between smoothing and edge preservation, we may prevent these aesthetically unpleasant consequences and herald for partial differential equations of order four. Instead of using step pictures to mimic the observed images, FPDE uses planner images. Planner photos have clean contours and a realistic appearance, resulting in minimal artefacts. Decomposing a picture using the FPDE technique yields the basic and detail images, which are then examined in depth [13]. The FPDE technique is first used on the source images to get the base images. The source photos are subtracted from the appropriate base images to get the detail images.

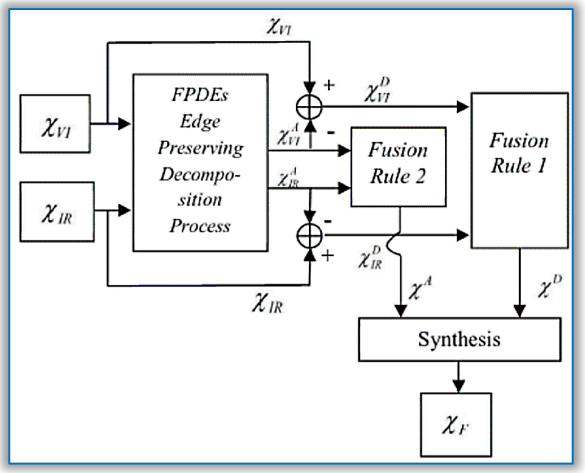

Figure 1. Block diagram of proposed FPDE method taken from [10]

Principle component analysis (PCA)

Traditionally, principal component analysis (PCA) included projecting the data from its original space to its eigenspace, where the variance was maximized and the covariance was minimized, in order to find patterns in the data.[14-15] Figure 2 depicts the process flow of traditional PCA.

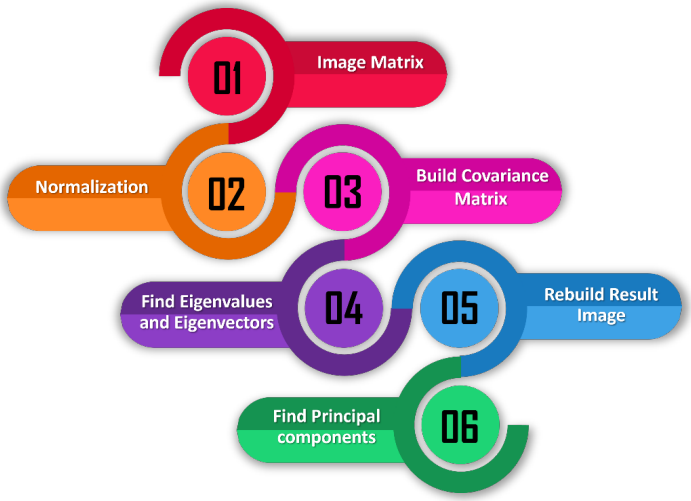

Figure 2. Process flow of Principle component analysis (PCA)

Discrete wavelet Transformation

The discrete two-dimensional wavelet transform may be computed by sequentially applying low pass and high pass filters over each row and column of the input image, followed by subsampling the results. This modification involves discretely sampling the wavelets, which results in poor spatial resolution at low frequency but excellent resolution at high ones. Image fusion allows for the production of multi-scale coefficients by fusing coefficients at the same level[16-17]. The equation for converting 1-D wavelet decomposition to 2-D is just the tensor product of the 1-D complements of the 2-D scaling and wavelet functions.

$$\left.\begin{aligned} \mathrm{Img}(m,n) &= (mn)^{\frac{-1}{2}} \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} \mathrm{Img}(x,y)\ \emptyset_{m,n}(x,y) \\ \mathrm{Img}(m,n) &= (mn)^{\frac{-1}{2}} \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} \mathrm{Img}(x,y)\ \psi^{c}_{m,n}(x,y) \end{aligned}\right\} \quad (1)$$

In equation (1), c represents the horizontal detail coefficient, v represents the vertical detail coefficient, and d represents the diagonal detail coefficient. where (m,n) denotes the image size, ψ stands for the wavelet functions, φ is the scaling function, and the functions supplied by

$$\left.\begin{aligned} \emptyset_{m,n}(x,y) &= 2^{i/2}\ \emptyset\left(2^{j}x-m\,,2^{j}y-n\right) \\ \varphi^{c}_{m,n}(x,y)\ 2^{i/2}\ &\psi(2^{j}x-m\,,2^{j}y-n) \end{aligned}\right\} \quad (2)$$

If we write "j" for each degree of breakdown, we get "j" levels of breakdown.

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN-L 1584 – 2665

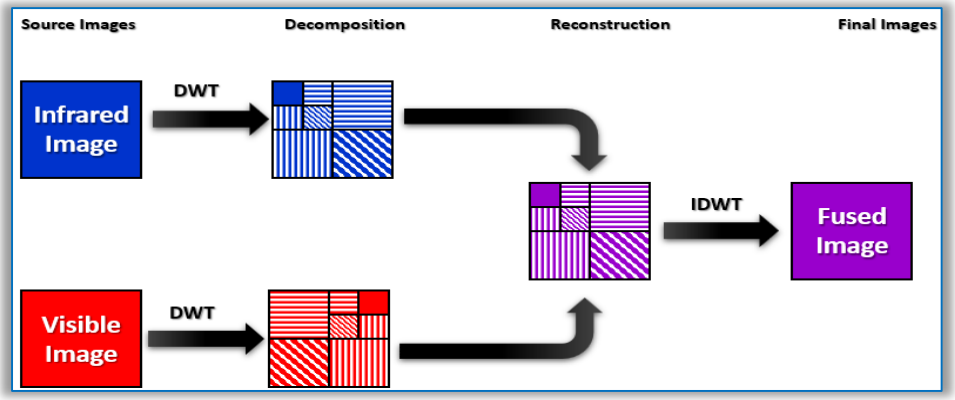

Figure 3. Block diagram of Discrete wavelet Transformation (DWT)

Discrete Cosine Transformation (DCT)

The discrete cosine transform (DCT) is a useful mathematical tool for digital image processing. The majority of the DCT's large coefficients are found in the low-frequency range; hence, it is well-known that it has great energy compactness qualities [18]. Specifically, [19] defines the discrete cosine transform (,) 1 2 X k k of a 1 2 N xN picture or 2D signal:

$$\mathrm{Img}(p,q) = \alpha(p)\,\alpha(q)\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}\mathrm{Img}(x,y)\cos\left(\frac{\pi\,(2x+1)p}{2m}\right)\cos\left(\frac{\pi\,(2y+1)q}{2n}\right)$$

$p \in [0, m-1]$ and $q \in [0, n-1]$ is a discrete frequency variable (x, y) pixel index

$$\alpha(p) = \begin{cases} \frac{1}{\sqrt[2]{m}} & ,p = 0 \\ \sqrt[2]{\frac{2}{m}} & ,1 \le p \le m-1 \end{cases} \quad \text{and} \quad \alpha(q) = \begin{cases} \frac{1}{\sqrt[2]{n}} & ,q = 0 \\ \sqrt[2]{\frac{2}{n}} & ,1 \le q \le m-1 \end{cases}$$

The DCT's foundational functions go by the name "base functions." Thus, the DCT coefficients $\mathrm{Img}(p,q)$ may be thought of as the weights used for each basis function. This picture shows the 64 basic functions that apply to matrices of size 8 by 8.

The frequencies go up the screen from top to bottom in two - dimensional image or signal of size n by n and across the screen from left to right. Based on its position at the top left, the constant-valued basis function is commonly referred to as the DC (Discrete Cosine) basis function, and the DCT coefficient $\mathrm{Img}(0,0)$ is often referred to as the DC coefficient.

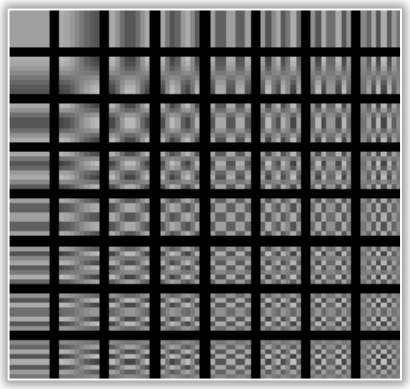

Figure 4. The 64 Basis Functions of an 8-by-8 Matrix

3. PROPOSED METHOD

Extraction of base layer and detailed layer from Images using FPDE

Coregistered source images $\{\mathrm{Img}_m(x,y)\}_{m=1}^{2}$ of size a × b passthrough FPDE process for edge preservation, smoothing and decomposition.

$$B_m(x,y) = \mathrm{FPDE}\{\mathrm{Img}_m(x,y)\}_{m=1}^{2}$$
$$D_m(x,y) = \mathrm{Img}_m(x,y) - B_m(x,y)$$

Here $B_m(x,y)$ represents and $D_m(x,y)$ are represents m[th] base layer and detailed layer of source images. Detail and base images are fused using distinct criteria.

Detail Layer Fusion

First, take into account the $D_m(x,y)$ for corresponding $\mathrm{Img}_m(x,y)$.

— calculating the covariance matrix C and the column vectors of the input pictures $D_m(x,y)$.

— Eigen vectors are computed by applying processing to the relevant diagonal values of the matrix.

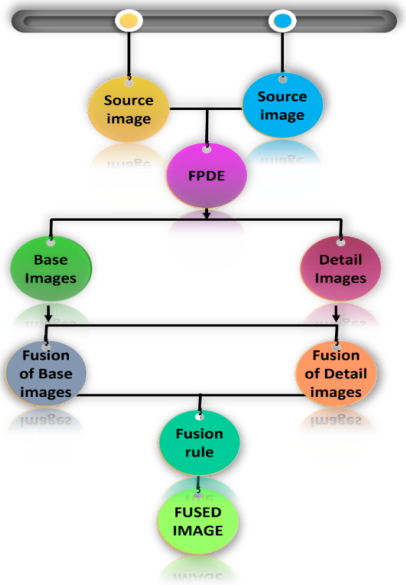

Figure 5: Process flow of our Hybrid method

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN-L 1584 – 2665

— principal components P_m are derived from the column vector with the greatest Eigenvalue (divide each element by the Eigen vector's mean).
— the input photos are multiplied by the normalized Eigen vector values, and then the resulting images are added.
— utilize the resulting fused detail picture as the base layer for your detail texture.

$$D(x,y) = \sum_{m=1}^{2} P_m \, D_m(x,y)$$

Base Layer Fusion:

— Step I: Utilizing DWT to decompose the $B_m(x,y)$ to high and low sub bands.
— Step II: the DCT is applied to each sub and to extract the most important information and store them in relatively few coefficients.
— Step III: Spatial frequency calculated using below equation

$$SF_{ij} = \sqrt{\left[\sum_{i=1}^{4}\sum_{j=2}^{4} DCT(i,j) - DCT(i,j-1)\right]^2 + \left[\sum_{i=2}^{4}\sum_{j=1}^{4} DCT(i,j) - DCT(i-1,j)\right]^2}$$

— Step IV: The average fusion rule has been establishing the merged sub-image of base layer fusion coefficient.

$$BFC_{ij} = \frac{C_{1-ij} + C_{2-ij}}{2} \text{ from } SF_{ij}$$

— Step V: Apply IDCT on (BFC_{ij}) coefficients and reconstructs the fused sub-bands of DWT.
— Step VI: Inverse DWT is then utilized to rebuild the final base picture $B(x,y)$ from the fused sub-bands.

Fused Image Reconstruction:

$$F(x,y) = B(x,y) + D(x,y)$$

4. EXPERIMENTS AND RESULTS

Experiments in simulation were run using MATLABR2021a on a PC equipped with an Intel i7 processor running at 144 Hz and 16 GB of RAM. To evaluate the efficacy of the suggested strategy, a number of tests have been conducted. All the pictures are grayscale. The source photos, available at [21-23] multisensory image fusion respectively. Hybrid method is compared with PCA [23], DCT [22], DWT [23], FPDE [10].Quantitative and qualitative comparisons of the fusion techniques are made to highlight their respective benefits and drawbacks. To objectively evaluate the techniques, we employ the performance assessment measures outlined in this section.

Quality Evaluation metrics

— **The Mean, or Average Intensity,** of a Combined Image is Defined as

$$AI = \frac{1}{N}\sum_{i=1}^{N} F_i$$

— **Entropy** is used to measure the amount of spatial Information in a signal. Where the total number of grayscale value in the histogram of image is denoted by n, and p_n shows corresponding normalization Histogram of grayscale values. The grater the value of E, the more info the picture contains.

$$E = -\sum_{n=0}^{n-1} p_n \log_2 p_n$$

— **Mutual Information** calculates the quantity of data transferred between two images., The Kullback-Leibler criterion shown below equation, which calculates the MI among two randomly selected variables, where $p_X(x)$ indicate the marginal histograms of X (inputted image) and $p_F(f)$ shows for F (fused image), also $p_{(X,F)}(x,f)$ is the joint histogram.

$$MI_{(X,F)} = \sum_{x,f} p_{(X,F)}(x,f)\log\frac{p_{(X,F)}(x,f)}{p_X(x)p_F(f)}$$

MI is computed using source images IR and VR separately. The results are summed to give the final MI score as in below equation. A higher MI score indicates that more information is transmitted from the X (source images) to F.

$$MI = MI_{IR,F} + MI_{VI,F}$$

— **Standard deviation** indicates the contrast ratio of the image. It is mathematically defined as:

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN-L 1584 – 2665

$$SD = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(F(i,j)-\mu)^2}$$

where μ is the mean over the entire fused image and $\mathbf{F(i,j)}$ is the corresponding pixel in the image. High SD attracts human attention.

Table 2. Performance Comparison of Multi sensor images the Four Fusion Methods.

Methods	DWT	PCA	DCT	FPDE	Hybrid Method
"Solider with Jeep."					
Mean	4.10	3.512	2.35	2.41	4.30
Cross Entropy	1.12	1.49	1.37	1.44	2.06
Entropy	7.078	6.98	6.14	6.24	6.84
Mutual information	1.64	1.56	1.48	1.50	1.65
"Nato camp."					
Mean	3.30	3.47	4.58	5.20	5.89
Standard deviation	5.94	6.33	6.44	6.37	6.89
Entropy	6.32	6.40	6.19	6.36	6.84
Mutual information	1.50	1.59	1.29	1.43	2.03
" Traffic Signal"					
Mean	4.37	4.42	4.51	5.01	5.65
Standard deviation	7.83	8.05	8.00	7.86	8.32
Entropy	6.33	6.35	6.19	6.42	6.74
Mutual information	1.34	1.55	**2.62**	1.48	2.59

Figure 6 : "Solider with Jeep." – Visual Quality comparison

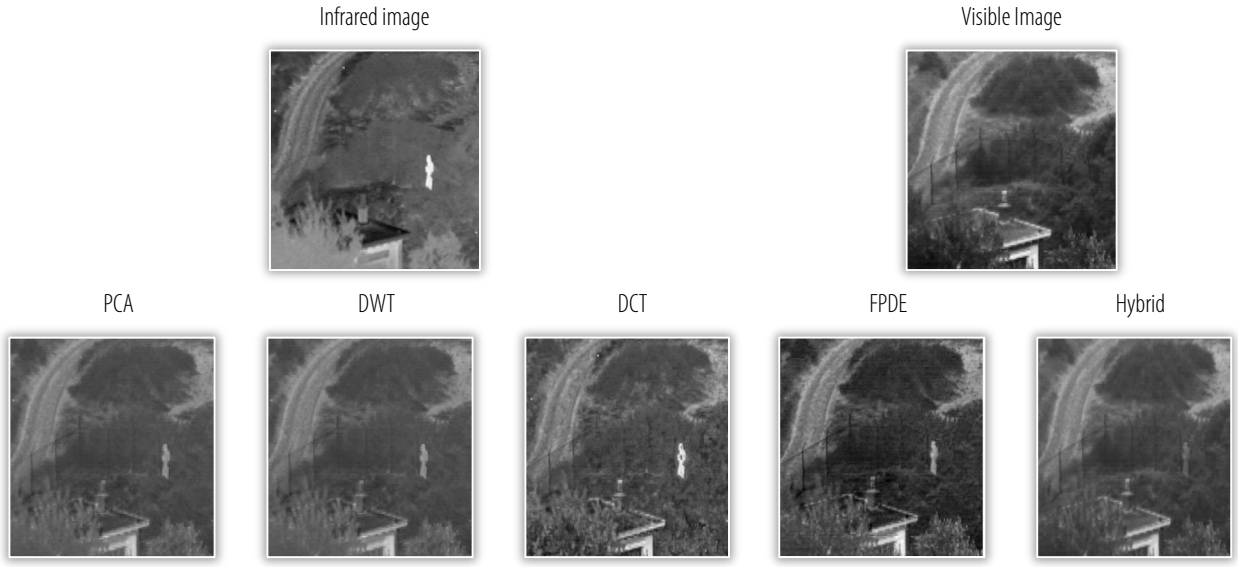

Figure 7 : "Nato camp." – Visual Quality comparison

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN-L 1584 – 2665

Figure 8: "Traffic Signal." – Visual Quality comparison

5. SUBJECTIVE ANALYSIS AND DISCUSSION:

From the first table, it is obvious that the fused picture retains the high level of detail found in the individual source photographs. for the measures of quality. The suggested approach yields a fused picture with improved contrast, information, clarity, and symmetry as compared to the alternatives.

Figure 5 seems to indicate that the created method yields better results in terms of discernibility, contrast, and brightness in the combined picture. Therefore, our technique can save the edges and the most valuable data while achieving minimal fusion artefacts and loss. Our solution not only removes block effects and artefacts from the combined picture, but it also recovers the original textures.

Excellent values for the suggested method's entropy (6.84), mutual information (2.59), and as well as the mean (4.30), and standard deviation (8.32), indicate that the fused picture is of high quality.

Our strategy is compared to four others (shown in Figures 4–7). When generating photographs, PCA seems to lose too much information, which results in blurry pictures that are difficult to make out. There is a loss of texture and sharpness in some areas of the three combined DWT pictures. Images fused by FPDE, such as the figure and sky in Figure 6 and Figure 7, have more artificial noise and muddy saliency traits. Also Fused pictures are often darker and lack some conspicuous elements. Images created with the Hybrid approach include more infrared characteristics and an excessive amount of brightness, as shown in Figure 7. This makes for an artificial viewing experience. Our fusion approach retains more information about the photos' details and textures than any of the other four methods, and the images are sufficiently clear for human vision.

Here, we compare our approach with the aforementioned five techniques and their respective assessment indicators. In Table 1, the best results of the five indicators are highlighted to indicate the objective assessment results of five metrics obtained by the comparative techniques and suggested approach for the three groups of source pictures. The results show that our approach performs better than the other approaches across almost four different parameters. Our approach yields fused pictures with less artefacts and noise and more detail and texture characteristics. The visual impression of the combined photographs is enhanced. In a nutshell, the suggested technique enhances human visual perception by better extracting and preserving detail information and texture elements from source pictures and fusing with the optimal scale. Furthermore, assessment criteria are used to check the suitability of the fusion findings for further observation and detection.

6. CONCLUSION

Feature extraction, fusion, and reconstruction are often proposed as three distinct steps by most state-of-the-art approaches. Every stage is analyzed independently. Our suggested strategy, on the other hand, integrates and optimizes all three stages simultaneously. That allows for the establishment of more robust links between procedures, hence increasing the method's stability. To further improve fusion quality and information transfer from input photos, the suggested approach is trained using an unique loss function. Because of this, the resulting merged picture has excellent quantitative and qualitative qualities. The

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN-L 1584 – 2665

effectiveness of the suggested strategy is shown by contrasting it with both deep learning-based approaches and state-of-the-art approaches.

References:

[1] G. J. Trivedi and R. Sanghvi, "Medical Image Fusion Using CNN with Automated Pooling," Indian Journal Of Science And Technology, vol. 15, no. 42, pp. 2267–2274, Nov. 2022
[2] Y. Chen, L. Cheng, H. Wu, F. Mo, and Z. Chen, "Infrared and visible image fusion based on iterative differential thermal information filter," Optics and Lasers in Engineering, vol. 148, p. 106776, Jan. 2022
[3] Y. Zeng, X. Wang, H. Zhao, Y. Jin, G. A. Giannopoulos, and Y. Li, "Image Fusion Methods in High-speed Railway Scenes: A Survey," High-speed Railway, Jan. 2023, Published
[4] W. Tang, F. He, Y. Liu, and Y. Duan, "MATR: Multimodal Medical Image Fusion via Multiscale Adaptive Transformer," IEEE Transactions on Image Processing, vol. 31, pp. 5134–5149, 2022
[5] L. Tang, J. Yuan, H. Zhang, X. Jiang, and J. Ma, "PIAFusion: A progressive infrared and visible image fusion network based on illumination aware," Information Fusion, vol. 83–84, pp. 79–92, Jul. 2022
[6] Rajinikanth, V., Satapathy, S.C., Dey, N.and Vijayarajan, R., DWT-PCA Image Fusion Technique to Improve Segmentation Accuracy in Brain Tumor Analysis.Electromagnetics and Telecommunications,471.2018
[7] X. Jin et al., "Infrared and visual image fusion method based on discrete cosine transform and local spatial frequency in discrete stationary wavelet transform domain," Infrared Physics & Technology, vol. 88, pp. 1–12, Jan. 2018
[8] K. Vanitha, D. Satyanarayana, and M. N. G. Prasad, "Multi-modal Medical Image Fusion Algorithm Based on Spatial Frequency Motivated PA-PCNN in the NSST Domain," Current Medical Imaging Formerly Current Medical Imaging Reviews, vol. 17, no. 5, pp. 634–643, 2021
[9] K. Bhalla, D. Koundal, S. Bhatia, M. Khalid Imam Rahmani, and M. Tahir, "Fusion of Infrared and Visible Images Using Fuzzy Based Siamese Convolutional Network," Computers, Materials & Continua, vol. 70, no. 3, pp. 5503–5518, 2022
[10] Bavirisetti, D.P., Xiao, G. and Liu, G. , Multi-sensor image fusion based on fourth order partial differential equations. 2017 20th International Conference on Information Fusion (Fusion), 2017
[11] Bavirisetti, D. P., & Dhuli, R. (2016). Fusion of Infrared and Visible Sensor Images Based on Anisotropic Diffusion and Karhunen-Loeve Transform. IEEE Sensors Journal, 16(1), 203–209
[12] Hao, S., He, T., An, B., Ma, X., Wen, H. and Wang, F., 2022. VDFEFuse: A novel fusion approach to infrared and visible images. Infrared Physics & Technology, 121, p.104048.
[13] Qin, X., Ban, Y., Wu, P., Yang, B., Liu, S., Yin, L., Liu, M. and Zheng, W., 2022. Improved Image Fusion Method Based on Sparse Decomposition. Electronics, 11(15), p.2321.
[14] Singh, S., Singh, H., Gehlot, A., kaur, J. and Gagandeep , IR and visible image fusion using DWT and bilateral filter. Microsystem Technologies. 2022
[15] Naidu, V.P.S. and Elias, B., 2013. A novel image fusion technique using DCT based Laplacian pyramid. International Journal of Inventive Engineering and Sciences (IJIES), pp.2319-9598.
[16] Jiang, H., Maharjan, P., Li, Z. and York, G., 2022, October. DCT-Based Residual Network for NIR Image Colorization. In 2022 IEEE International Conference on Image Processing (ICIP) (pp. 2926-2930). IEEE.
[17] Budhiraja, S.; Agrawal, S.; Sohi, B. S. Performance Analysis of Multi-Scale Transforms for Saliency-Based Infrared and Visible Image Fusion. Proceedings of International Conference on Data Science and Applications, 2021, pp.801–809
[18] Im, C.-G., Son, D.-M., Kwon, H.-J., & Lee, S.-H.. Tone Image Classification and Weighted Learning for Visible and NIR Image Fusion. Entropy, 24(10), 1435, 2022
[19] H. Jiang, P. Maharjan, Z. Li and G. York, "DCT-Based Residual Network for NIR Image Colorization," 2022 IEEE International Conference on Image Processing (ICIP), 2022, pp. 2926-2930
[20] Meher, B., Agrawal, S., Panda, R., Dora, L. and Abraham, A., 2022. Visible and infrared image fusion using an efficient adaptive transition region extraction technique. Engineering Science and Technology, an International Journal, 29, p.101037.
[21] Toet, A., The TNO Multiband Image Data Collection. Data in Brief, [online] 15, pp.249–251.2017
[22] Datset, https://github.com/xingchenzhang/VIFB
[23] Zhang, X., Ye, P., & Xiao, G., VIFB: A Visible and Infrared Image Fusion Benchmark. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.468-478.2022
[24] Gonzalez, R.C., Woods, R.E. and Eddins, S.L. , Digital image processing using MATLAB. Knoxville: Gatesmark Publishing.2020

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN–L 1584 – 2665