

## Bounded PCA-based Multi-Sensor Image Fusion Employing Curvelet Transform Coefficients

A.K. Singh<sup>#</sup>, D. Chaudhuri<sup>\*,\*</sup>, S. Mitra<sup>#</sup>, M.P. Singh<sup>§</sup> and B.B. Chaudhuri<sup>#</sup>

<sup>#</sup>Department of Computer Science & Engineering, Techno India University, Kolkata – 700 091, India

<sup>§</sup>DRDO-Young Scientist's Laboratory- Cognitive Technologies, Chennai – 600 113, India

\*Email: debasis.chaudhuri@technoindiaeducation.com

### ABSTRACT

The fusion of thermal and visible images acts as an important device for target detection. The quality of the spectral content of the fused image improves with wavelet-based image fusion. However, compared to PCA-based fusion, most wavelet-based methods provide results with a lower spatial resolution. The outcome gets better when the two approaches are combined, but they may still be refined. Compared to wavelets, the curvelet transforms more accurately depict the edges in the image. Enhancing the edges is a smart way to improve spatial resolution and the edges are crucial for interpreting the images. The fusion technique that utilizes curvelets enables the provision of additional data in both spectral and spatial areas concurrently. In this paper, we employ an amalgamation of Curvelet Transform and a Bounded PCA (CTBPCA) method to fuse thermal and visible images. To evidence the enhanced efficiency of our proposed technique, multiple evaluation metrics and comparisons with existing image merging methods are employed. Our approach outperforms others in both qualitative and quantitative analysis, except for runtime performance. Future Enhancement-The study will be based on using the fused image for target recognition. Future work should also focus on this method's continued improvement and optimization for real-time video processing.

**Keywords:** Image fusion; Target detection; Curvelet transform; Bounded PCA

### 1. INTRODUCTION

The problems like target detection<sup>1</sup>, medical imaging<sup>2-4</sup>, remote sensing<sup>5</sup>, and surveillance<sup>6</sup> employ image fusion, to merge images from several sensors into a single, rich fused image. In a low-light or dark environment, the target is visible in the thermal image because they capture the thermal radiation of the scene. However, the thermal images have poor imaging capability, low contrast, and smaller resolution. The target, on the other hand, might be undetectable and get distorted by smoke, bad weather, and other variables, in the visible images, yet they may provide a good degree of detail. For accurate target detection and to overcome the limited imaging capabilities of thermal and visible sensors, the fusion of the thermal and visible images is useful.

The fusion techniques can be grouped into two sub-types: transform-domain fusion<sup>7</sup> and spatial-domain fusion<sup>8</sup>. Common spatial domain fusion techniques include the Select Maximum method, Averaging approach, Select Minimum method, and Principal Component Analysis method. The fused image of the spatial-domain approaches, however, exhibits spatial distortion. When the fused image is used for additional processing, like classification, the spatial distortion acts as a drawback. The development of spectral-domain fusion techniques helped to

solve this problem. Pyramid-based techniques and Wavelet-based methods<sup>9</sup> are examples of spectral-domain approaches.

In this paper, for image fusion, a combination of the fast discrete curvelet transform<sup>11</sup> and the proposed bounded PCA is applied. The objective of this study is to generate a final fused image from a pair of thermal-visible images by extracting the corresponding features using a curvelet transform on a single-channel grayscale thermal image and the color channels of the visible image, respectively, which are then combined using a bounded PCA method. The use of the curvelet transform increases the amount of information available in the spectral and spatial domains by providing us with detailed features of an image<sup>49</sup>. The weightage of each coefficient of the thermal image and its corresponding coefficient of the visible image (color channels) is determined using the bounded PCA approach for fusing the coefficients.

This paper is structured as follows. The theoretical foundation of the discrete curvelet feature extraction and bounded PCA technique is described in Section 2. The proposed discrete Curvelet Transform and Bounded PCA-Combined Image Fusion method (CTBPCA), for thermal and visible images, is introduced in Section 3. The results and analysis of our experiments are described in Section 4. The proposed method is also contrasted with other approaches for image fusion, including the wavelet approach and the anisotropic

diffusion approach in Section 4. In Section 5, we wrap up the paper with a summary and future scope of work.

**2. THEORETICAL FOUNDATION**

Feature extraction is one of the most important elements in object detection and recognition. The two most often used methods in the field of feature extraction in image fusion are the curvelet transform and PCA. In this paper, the fast discrete curvelet transform is used for feature extraction purposes.

**2.1 Fast Discrete Curvelet Transform-Based Feature Extraction**

Curvelet transform was invented by Candes and Donoho<sup>10</sup> as a result of the necessity for image analysis<sup>10</sup>. The highly redundant dictionary of the curvelet transform can offer a sparse representation of signals with edges along regular curves. The original curvelet structure was eventually revised and re-introduced as the Fast Digital Curvelet transform (FDCT)<sup>11</sup>. The goal of the second-generation curvelet transform is to make it easier to comprehend and use. The curvelet transform is also defined in the continuous and digital domains, as well as for higher dimensions. The curvelet transform is a unique member of the family of multiscale geometric transforms. The curvelet transform is conceptually represented as a multiscale pyramid with several orientations and locations at each length scale and needle-shaped parts at tiny scales. In that they have useful geometric properties, curvelets do set themselves apart from wavelets and other items. For instance, curvelets follow a parabolic scaling relation, which states that at scale  $2^j$ , each element has an envelope that is positioned along a “ridge” that is  $2^{j/2}$  in length and  $2^j$  in width.

First, a 2D Fast Fourier Transform (FFT) of the image is acquired to apply the curvelet transform. Then wedges are formed in the 2D Fourier frequency plane (like the darkened region in Fig. 1). When the Fourier plane is divided into radial and angular divisions, wedges take on a parabolic appearance. The breakdown of an image into several scales (used for band-passing the image at various scales) is accomplished by concentric circles, while the angular divisions divide the band-passed image into various angles or orientations. So, to deal with a specific wedge, we must define its scale ( $j$ ) and angle

( $\theta$ ). Let us now examine the situation in the spatial domain (Fig. 1(b)). Here, each wedge corresponds to a certain curvelet (shown by an ellipse) at a specific scale and angle. This means that if the inverse FFT of a certain wedge is performed, the curvelet coefficients for that scale and angle will be obtained. The implementation of the curvelet transform is based mostly on this notion. In a spatial Cartesian grid, curvelets with a certain size and angle are depicted in Fig. 1(b). The wrapping-based fast discrete curvelet transforms<sup>11</sup> are used in the present paper.

**2.2 Proposed Bounded PCA**

Principal Component Analysis (PCA) is a statistical method for reducing the dimensions of the data. To enhance the variance and decrease the covariance, it essentially projects data from its original space to its eigenspace, keeping the components that correspond to the biggest eigenvalues and rejecting the others. To improve the signal-to-noise ratio, PCA helps to eliminate redundant information and highlight the factors that have the most impact.

A fusion algorithm based on PCA and the Finite Ridgelet Transform (FRIT) was proposed by Miao and Wang<sup>13</sup>. The image is split into low and high-frequency components using FRIT. The low-frequency coefficients are fused with the PCA technique. Eqn. 1 is used to get the weight values.

$$\omega_A = \frac{x}{x+y}, \omega_B = \frac{y}{x+y} \tag{1}$$

here,  $(x, y)^T$  is a vector containing eigenvectors of images  $I_A$  and  $I_B$ .  $\omega_A$  and  $\omega_B$  represent the weight values of the respective images. The fused image  $I_F$  is obtained by the weighted average as given in Eqn. 2.

$$I_F = \omega_A I_A + \omega_B I_B \tag{2}$$

This PCA fusion method does, however, occasionally provide weight values that are either negative or more than 1. This happens as the eigenvectors are not normalised, which can change the image. The pixels with a negative value are trimmed to 0 for images with negative weightage and if normalised, the image appears to be inverted. The pixels with values greater than 1 are trimmed to 1 for images with weights larger than 1. This can lead to weird artifacts in the image. Now, the sum of the weight values  $\omega_A$  and  $\omega_B$  must always be 1.

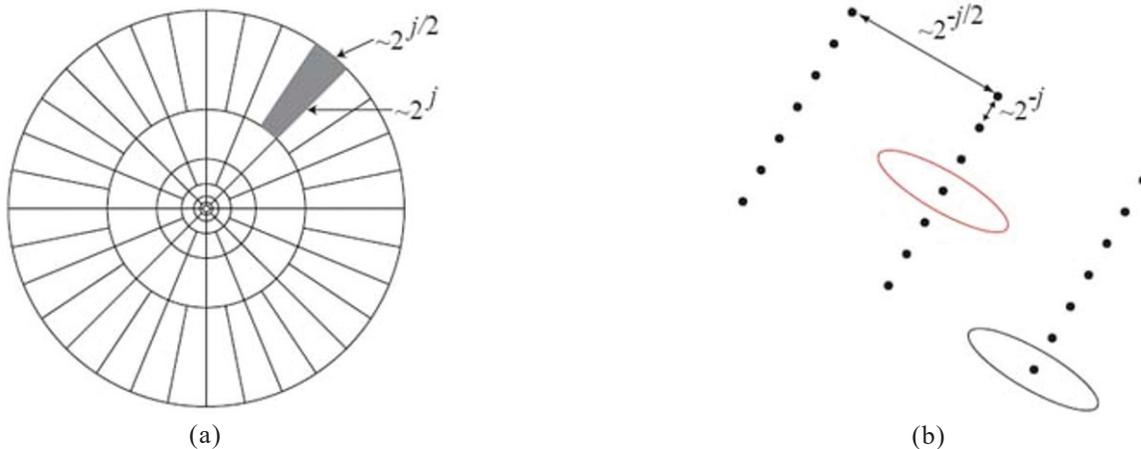


Figure 1. Curvelets in fourier frequency as (a) and spatial domain (b) (Courtesy: Candes<sup>11</sup> et al.)

$$\text{(Proof: As } \omega_A = \frac{x}{x+y} \& \omega_B = \frac{y}{x+y} \therefore \omega_A + \omega_B = \frac{x+y}{x+y} = 1)$$

To acquire weight values between  $[0,1]$ , the method must be modified keeping the sum of the weight values fixed at 1. In this paper, we used a bounded PCA approach to resolve this issue. A covariance matrix  $C$  is first generated from the images  $I_A$  and  $I_B$  as:

$$C = \begin{bmatrix} \sigma_A^2 & \sigma_{AB} \\ \sigma_{AB} & \sigma_B^2 \end{bmatrix} \quad (3)$$

In Eqn. (3),  $\sigma_A^2$  denotes the variance of image  $I_A$ ,  $\sigma_B^2$  denotes the variance of image  $I_B$  and  $\sigma_{AB}$  denotes the covariance of images  $I_A$  and  $I_B$ . Let  $D$  be the matrix containing the eigenvalues of  $C$  and  $V$  be the matrix with the corresponding eigenvectors of  $C$ .  $D$  is a 2-element matrix where each element is an eigenvalue of  $C$ .  $V$  is a  $2 \times 2$  matrix, where each column is a  $2 \times 1$  eigenvector of each eigenvalue in  $D$ . We take the eigenvector corresponding to the maximum eigenvalue as the principal eigenvector.  $x$  and  $y$  denote the 2 elements of the principal eigenvector, while  $\omega_A$  and  $\omega_B$  represent the weight values of the respective images  $I_A$  and  $I_B$ .  $\omega_A$  and  $\omega_B$  are calculated using Eqn. 1. For weight values going beyond the range  $[0,1]$ , we use 0 as the demarcation point to calculate new weights.  $\omega_A$  becomes  $\omega'_A$  and  $\omega_B$  becomes  $\omega'_B$ , where,  $\omega'_A + \omega'_B = 1$ .

If  $\omega_A < 0$  and  $\omega_B > 1$  and, the new weight values are found according to Eqn. 4.

$$\omega'_A = \frac{|\omega_A|}{\omega_B - \omega_A}, \omega'_B = 1 - \omega'_A \quad (4)$$

Else, if  $\omega_B < 0$  and  $\omega_A > 1$ , the new weight values are found by Eqn. 5.

$$\omega'_B = \frac{|\omega_B|}{\omega_A - \omega_B}, \omega'_A = 1 - \omega'_B \quad (5)$$

$$\text{Else, if } \omega_A \text{ and } \omega_B \text{ are bound to } [0,1], \omega'_A = \omega_A, \omega'_B = \omega_B \quad (6)$$

Since  $\omega_A + \omega_B = 1$ , it is evident that  $\omega_A < 0$  and  $\omega_B < 0$  will not occur concurrently. The fusion is then carried out with the new weight values by Eqn. (2).

For the sake of simplicity, we used the bounded PCA approach and the PCA approach proposed by Miao and Wang on a single pair of thermal-visible images and applied the fusion logic directly to the images themselves. We have substituted the bounded PCA approach as fusion logic for the specific curvelet coefficients later in this study.

Figure 2(a) and Fig. 2(b) are the original thermal and visible images, respectively and it has been noted that the scene coverage of the area thermal sensor is more than that of the visible sensor (R-band). By using the thermal image as the reference image, Fig. 2(c) represents the registered visible image. The thermal image is multiplied by its calculated PCA weight value of  $-0.312716$ , where pixels with values lower than 0 are clipped to 0 and the resultant image is completely dark (see Fig. 2(d)). Figure 2(e) shows the inverted thermal image after being multiplied by the appropriate PCA weight value and then normalized. Figure 2(f) depicts the visible image multiplied by its appropriate PCA weight value of  $1.312716$ , making the image look whiter as the pixels with values larger than 1 are clipped to 1. The normalized visible image is displayed in Fig. 2(g) after being multiplied by the corresponding PCA weight value. A thermal image that has

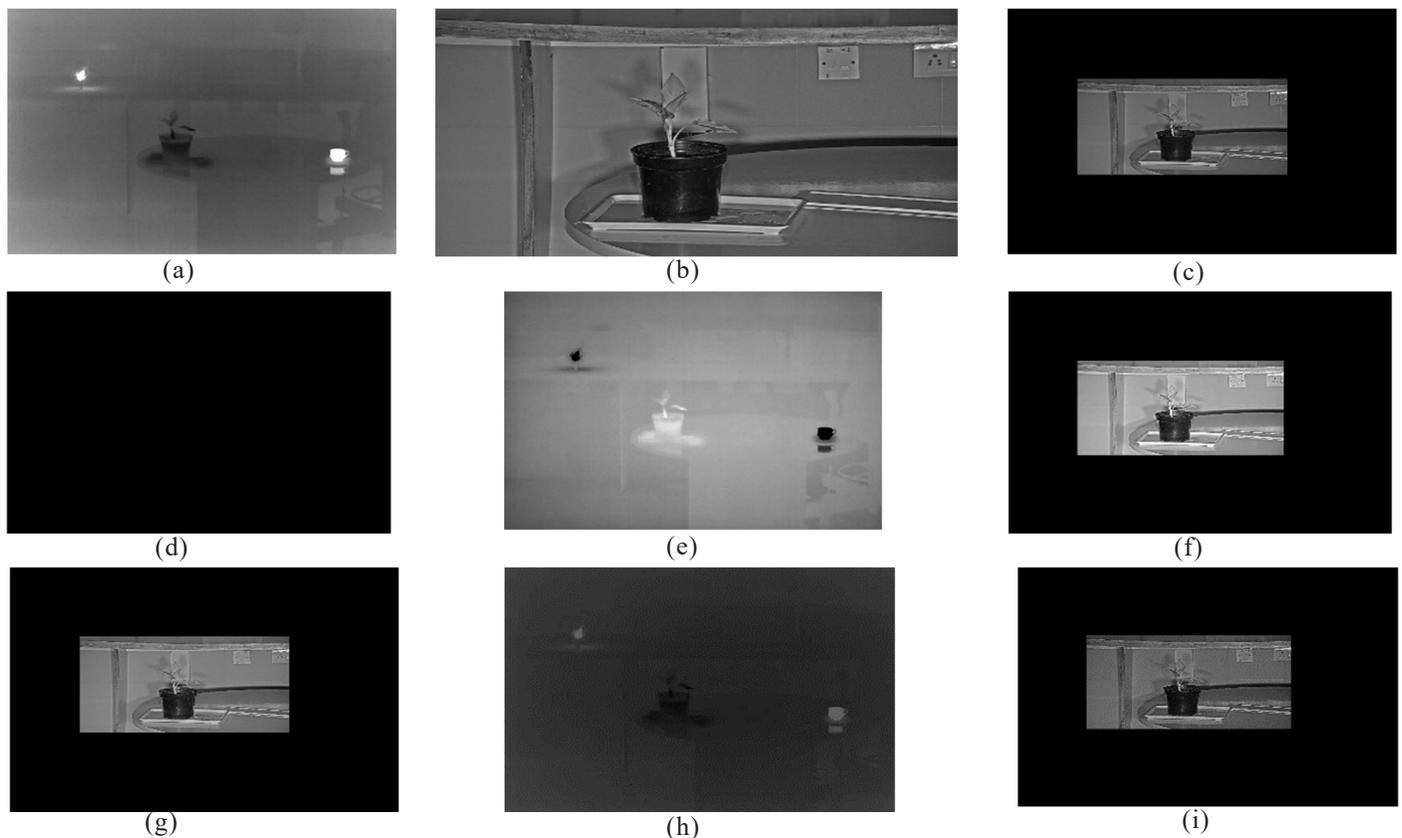


Figure 2. Effect of PCA weight values and bounded PCA weight values on both visual and thermal images.

been multiplied by the relevant bounded PCA weight value of 0.19239 is clip-free as shown in Fig. 2(h). The results are better than Fig. 2(d) and 2(e). There is absolutely no clipping in Fig. 2(i), which displays the visible image multiplied by its matching bounded PCA weight value of 0.80761. The outcome is superior to Fig. 2(g). This makes the recommended bounded PCA technique effective and eliminates the need for any extra computing for normalisation.

### 3. METHODOLOGY

Many industries, including the military, remote-sensing, security, surveillance, etc., have made extensive use of the fusion of infrared (IR) and visible video sequences. Infrared and visible video sequences can be combined to produce good IR target characteristics and a clear visible backdrop since infrared sensors can recognise thermal targets and visible sensors can offer details of the scene. For thermal and visual images, the suggested discrete curvelet transform and bounded PCA-combined image fusion (CTBPCA) approach is described below. The diagram depicting the methodology being put forth can be observed in Fig. 3.

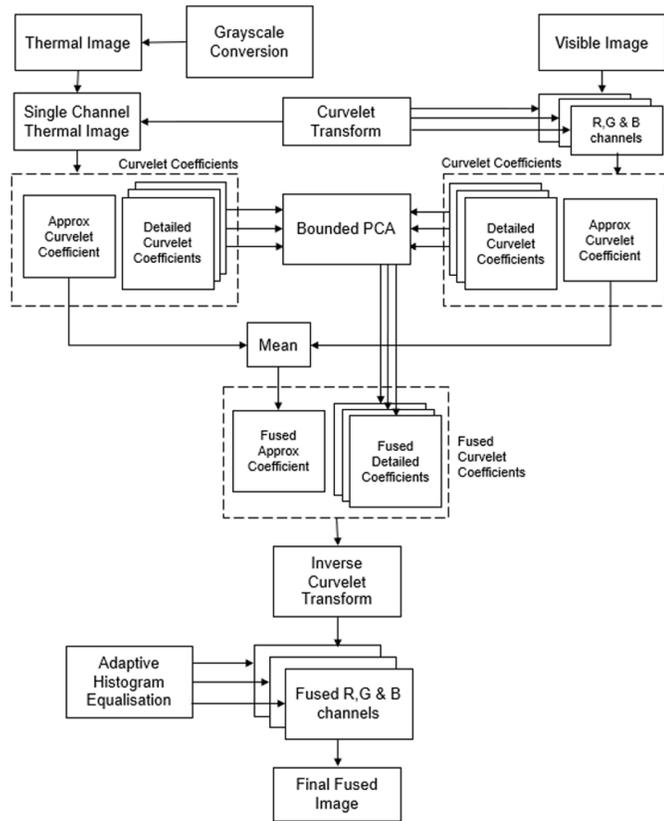


Figure 3. The flowchart of the proposed method.

#### Algorithm: CTBPCA

- **Step 1:** At first, the red, green, and blue color channels are extracted from the visible image, and the corresponding thermal image is transformed into a single-channel grayscale image.
- **Step 2:** The thermal image and each color channel of the visible image of the provided image pair are then subjected to the curvelet transforms to produce the approximation coefficient and detailed coefficients from all images.

- **Step 3:** Next, the mean of the approximate coefficients from each color channel and the single-channel grayscale thermal image are used as the approximate coefficients of each fused color channel. The detailed coefficients are then combined using the bounded PCA approach (Eqn. (2)).
- **Step 4:** Then, we apply adaptive histogram equalisation to create an image with a fused color image for each channel using the inverse curvelet transform on the combined coefficients.
- **Step 5:** Ultimately, the fused and enhanced channels are joined to provide the final fused image.

### 4. EXPERIMENTAL RESULTS AND DISCUSSION

A total of 24 thermal-visible image pairs are used in our investigations. This section presents the results of the experiment. All investigations were conducted on a mobile Dell Precision 3561 workstation with an i7-11800H CPU. All other algorithms' source codes, aside from the proposed approach and a wavelet-based image fusion technique, were taken from the code library provided by VIFB. For the deep learning-based methods in our investigation, we exploited the pre-trained models that their authors provided rather than retraining them.

#### 4.1 Baseline Algorithms

The VIFB benchmark includes 20 recently released visible-infrared image fusion methods. Many of these algorithms were developed to combine grayscale images. The creators of VIFB have modified them to fuse color images by merging each RGB channel with the corresponding infrared image. Their code library was used to conduct a comparative analysis of our suggested approach. We used wavelet-based fusion together with 16 of the 20 methods in our comparisons. For further details on these algorithms, see Table 1.

#### 4.2 Dataset

Although research into the combination of thermal and visual images has been ongoing for some time, the industry has not yet developed a dataset that is widely acknowledged and used in the open. While OTB<sup>27-28</sup>, and VOT<sup>29</sup> are well-known benchmarks, the visual tracking community has created and regularly uses many more. This makes an objective comparison challenging as various image pairings are frequently used in the research on the fusion of visual and thermal images. A couple of the visible and infrared image fusion datasets that are already available are the OSU Color- Thermal Database<sup>a</sup> (Davis and Sharma<sup>30</sup>), TNO Image fusion dataset<sup>b</sup> (Toet<sup>31</sup>), and VLIRVDIF<sup>c</sup> (Ellmauthaler<sup>32</sup>).

The fusion tracking dataset and various other images from the internet<sup>d</sup> were used by the authors of VIFB to obtain 21 image pairings for their dataset<sup>30, 33, 34</sup>. These image pairs were used in this paper.

These images show a variety of places and working circumstances. Resolutions for the images in the collection range from 320×240 to 630×460, or 512×184 to 452×332 for the VIFB dataset. Additionally, we have experimented with the

**Table 1. Image fusion techniques used for visible and infrared images**

Method	Full form of the methods	Category
ADF (Bavirisetti and Dhuli <sup>35</sup> , 2016)	Application Development Framework	
CBF (Shreyamsha <sup>36</sup> , 2015)	Cross Bilateral Filter	
GFCE (Zhou <sup>39</sup> <i>et al.</i> , 2016)	Guided Filter-based Context Enhancement	
GFF (Li <i>et al.</i> , 2013)	Gated Fully Fusion	Multi-scale
MGFF (Bavirisetti <sup>44</sup> <i>et al.</i> , 2019)	Multi-scale Guided Fast Fusion	
MSVD (Naidu <sup>46</sup> <i>et al.</i> , 2011)	Multi-resolution Singular Value Decomposition	
Wavelet (Chipman <sup>9</sup> <i>et al.</i> , 1995)	Wavelet Transform	
CNN (Liu <sup>37</sup> <i>et al.</i> , 2018)	Convolution Neural Network	DL-based
FPDE (Bavirisetti <sup>38</sup> <i>et al.</i> , 2017)	Fourth-order Partial Differential Eqn.s	Subspace-based
GTF (Ma <sup>41</sup> <i>et al.</i> , 2016)	Gradient Transfer Fusion	Other
IFEVIP (Zhang <sup>42</sup> <i>et al.</i> , 2017)	Infrared Feature Extraction and Visual Information Preservation	
LatLRR (Li and Wu <sup>43</sup> , 2022)	Latent Low-Rank representation	Saliency-based
TIF (Bavirisetti and Dhuli <sup>47</sup> , 2016)	Temporal Image Fusion	
MST SR (Liu <sup>45</sup> <i>et al.</i> , 2015)	Multi-Scale Transformation and Sparse Representation	
NSCT SR (Liu <sup>45</sup> <i>et al.</i> , 2015)	Non-Subsampled Contourlet Transform and Sparse Representation	Hybrid
RP SR (Liu <sup>45</sup> <i>et al.</i> , 2015)	Ration-of-low-pass Pyramid and Sparse Representation	
VSMWS (Ma <sup>48</sup> <i>et al.</i> , 2017)	Visual Saliency Map and Weighted Least Square	



**Figure 4.** The visible and infrared test set. 24 pairs of visible and infrared images are included in the collection. RGB images are shown in the first, third, and fifth rows, while the corresponding infrared images are shown in the second, fourth, and sixth rows.

AXIS Q8741-LE Bi-spectral PTZ Network Camera<sup>®</sup> to fuse three visual-thermal image pairs (1<sup>st</sup> and 2<sup>nd</sup>-row col 6<sup>th</sup>, 3<sup>rd</sup> and 4<sup>th</sup>-row col 5<sup>th</sup> and 5<sup>th</sup> and 6<sup>th</sup>-row col 2<sup>nd</sup>). These three images each have a resolution of 800×600. To ensure that the image

fusion can be completed properly, each pair of visible and infrared images has been pre-registered. In total, we employed 24 image pairings for our research. Figure 4 displays a few illustrations of the images in the dataset.

### 4.3 Methods for Image Fusion Quality Assessment

A particular image fusion technique may be evaluated in three different ways: visually, quantitatively, and in terms of runtime. In visible interpretation, the quality of the fused image is assessed by contrasting it with the input images and using various optical metrics, such as spatial features, geometric patterns, color, and so forth. Quantitative assessment evaluates the quality of the fused image by comparing the spectral and spatial similarities between the fused image and the raw input images using a set of predetermined quality criteria. Runtime is the amount of time it will take for an algorithm to complete its execution.

#### 4.3.1 Visible Interpretation

One approach to evaluating the fused image is visible interpretation, sometimes referred to as qualitative analysis. Visible analysis has the advantage of being a simple, mild, and honest way to assess the quality of a fused image, but it also largely depends on the observers' background and the viewing context. Applying qualitative assessment approaches that are founded on the human visual system is essential for effectively analysing the quality of fused images. If the observer is aware

of the ground truth, (s)he may judge the quality of the fused image more accurately. Numerical models are not used in this method.

#### 4.3.2 Quantitative Evaluation

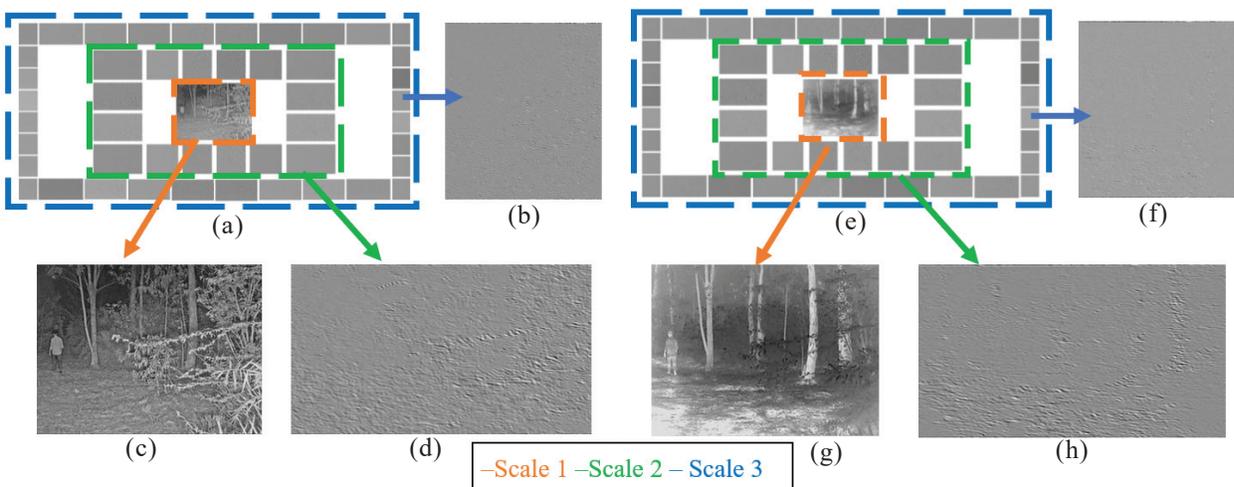
The objective analysis method of quantitative evaluation, which is based on mathematical modeling, is well known. It determines the quality of the fused image by comparing the spectral and spatial similarities between the fused image and the raw input images using a set of pre-defined quality indicators. Many evaluation techniques have been proposed for thermal-visible image fusion. To compare performance thoroughly and objectively, we employed 11 assessment metrics. The assessment measures that have been applied are listed in Table 2 along with the groups to which they belong. As can be seen, the adopted metrics offer information for all four categories as proposed by Liu, *et al.* (2012).

#### 4.3.3 Runtime Performance

Performance time is determined by how long it takes the image fusion algorithm to run. Performance time becomes a key consideration when evaluating an algorithm for real-time

**Table 2.** Following are the evaluation metrics implemented in this paper. A high value is denoted by "+," while a low value is denoted by "-", both of which imply strong performance. (Data from VIFB (Zhang<sup>16</sup> *et al.*))

Category	Name	Meaning	+/-
Information theory-based	CE (Bulanon, <sup>17</sup> <i>et al.</i> )	Cross entropy	-
	EN (Roberts, <sup>19</sup> <i>et al.</i> )	Entropy	+
	MI (Qu, <sup>21</sup> <i>et al.</i> )	Mutual Information	+
	PSNR (Jagalingam & Hegde <sup>23</sup> )	Peak signal-to-noise ratio	+
Structural similarity-based	RMSE (Jagalingam & Hegde <sup>23</sup> )	Root mean squared error	-
	AG (Cui, <sup>18</sup> <i>et al.</i> )	Average gradient	+
	EI (Rajalingam, <sup>20</sup> <i>et al.</i> )	Edge intensity	+
Image feature-based	SD (Rao <sup>22</sup> )	Standard deviation	+
	SF (Eskicioglu & Fisher <sup>24</sup> )	Spatial frequency	+
	Q <sub>CB</sub> (Chen & Blum <sup>25</sup> )	Chen-Blum metric	+
Human perception inspired	Q <sub>CV</sub> (Chen & Varshney <sup>26</sup> )	Chen-Varshney metric	-



**Figure 5.** Curvelet coefficients in different scale factors of visible image red channel and corresponding grayscale thermal image.

use cases and is influenced by both the algorithm and the platform on which it is run.

#### 4.4 Results of the Proposed Method

Figure 5(a) shows the curvelet coefficients of the red channel of the visible image, separated into three scales:

- **Scale 1:** We get the approximation coefficient (marked in orange)
- **Scale 2:** (marked in green) is divided into 16 angular divisions. Each of the 16 divisions has a corresponding detail coefficient.
- **Scale 3:** (marked in blue) is made up of 32 angular divisions. Each of these divisions provides a unique detail coefficient.

Figure 5(b) shows the detail curvelet coefficient for the red channel of the visible image at scale 3 at the 4<sup>th</sup> angular division. The approximate curvelet coefficient of the visible image's red channel at scale 1 is seen in Fig. 5(c). Figure 5(d) shows the detail curvelet coefficient for the red channel of the visible image at scale 2 at the 2<sup>nd</sup> angular division. Figure 5(e)

demonstrates the curvelet coefficients of a thermal image in grayscale separated into three scales:

- The approximation coefficient is provided by scale 1 (marked in orange)
- Each of the 16 angular divisions that make up Scale 2 (marked in green) contains a corresponding detail coefficient
- Each of the 32 angular divisions making up Scale 3 (marked in blue) provides a unique detail coefficient.

Detail curvelet coefficient of a corresponding thermal image in grayscale at scale 3 at angular division 4 is shown in Fig. 5(f). Figure 5(g) displays the approximate curvelet coefficient of a corresponding thermal image in grayscale, at Scale 1.

Detail curvelet coefficient of a corresponding thermal image in grayscale at the 2<sup>nd</sup> angular division of Scale 2 is shown in Fig. 5(h).

The five key steps of the recommended method for fusing two thermal-visible images are: curvelet transform, mean of approximation curvelet coefficient, bounded PCA on detailed

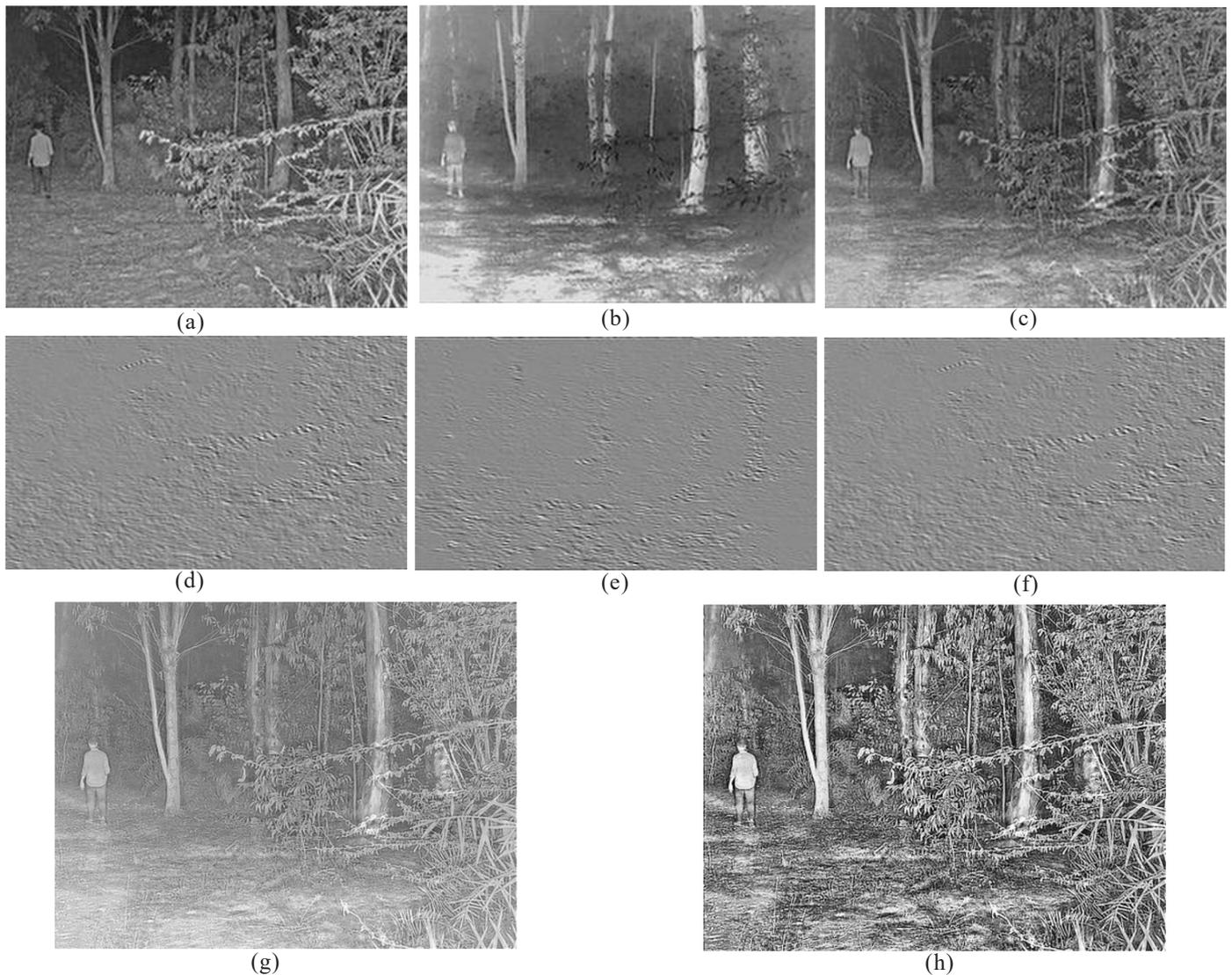


Figure 6. Stepwise results of the proposed algorithm for the red channel of the visible image and corresponding thermal image.



(a)



(b)



(c)



(d)



(e)



(f)



(g)

Figure 7. Channel-wise fusion images with the final fused image.

coefficients, inverse curvelet transform, and adaptive histogram equalisation. The stepwise results of the proposed fusion method are shown in Fig. 6. The original images considered for this purpose are shown in Fig. 4. The visual image is in Fig. 4 (row 1 column 6) and its corresponding thermal image is in Fig. 4 (row 2 column 6).

For the fusion process of the curvelet coefficients, averaging method is used for approximate coefficients while detail coefficients are fused by the proposed bounded PCA method.

The approximate curvelet coefficient of the red channel of the visible image  $[A_{vis}]$  is shown in Fig. 6(a). Figure 6(b) shows the approximate curvelet coefficient of the corresponding thermal image  $[A_{therm}]$ . The approximate coefficient of the fused image, calculated by averaging method  $\left[A_{fused} = \frac{A_{vis} + A_{therm}}{2}\right]$ , is shown in Fig. 6(c). The detailed curvelet coefficients of the red channel of the visible image  $[D_{22_{vis}}]$ , thermal image  $[D_{22_{therm}}]$  and the fused image calculated by the bounded PCA method

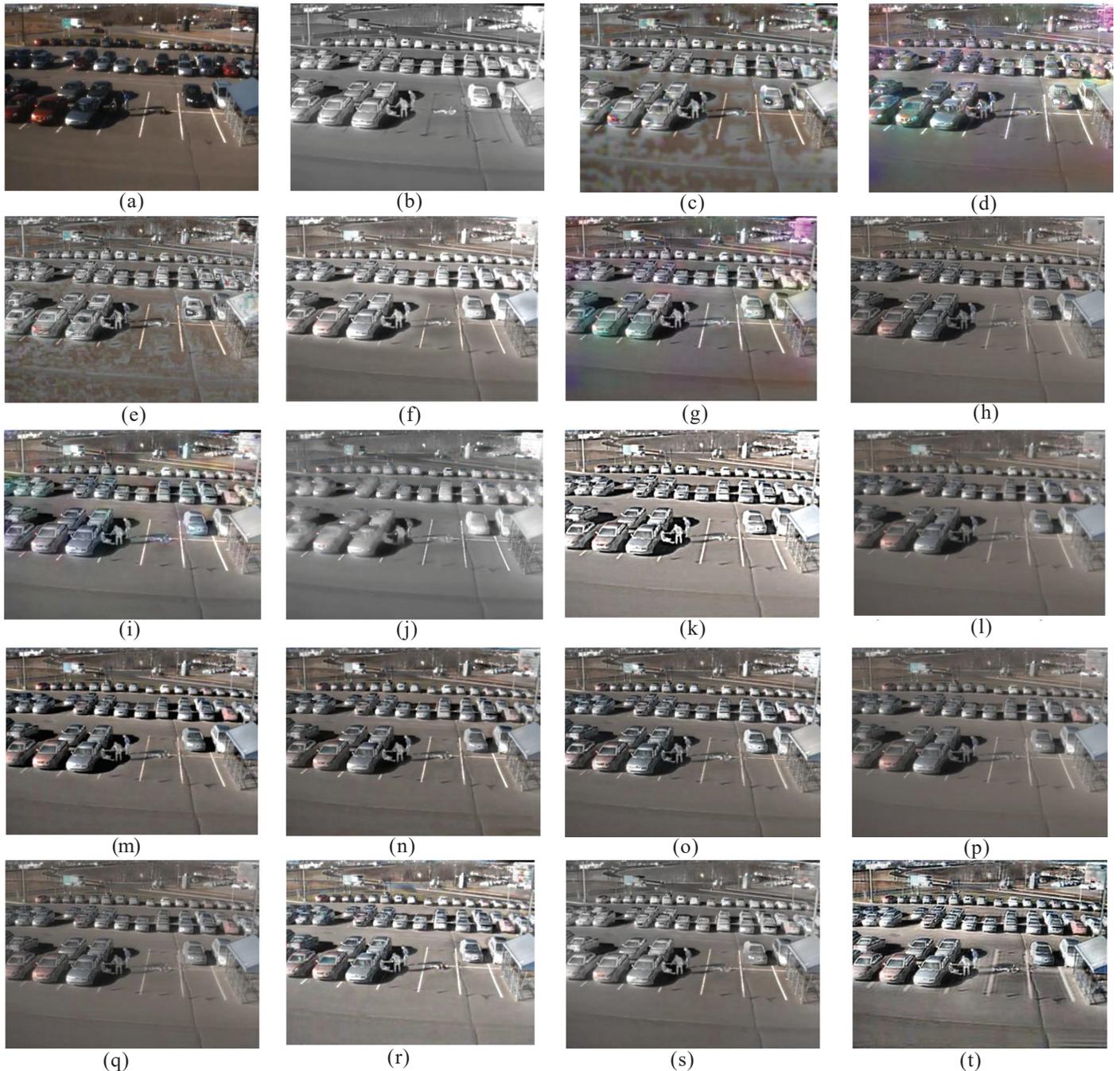


Figure 8. A qualitative assessment of 18 techniques using an image pair from the dataset in Figure 4: (a) Original visual image, (b) Original thermal image, (c) NSCT SR output image, (d) RP SR output image, (e) CBF output image, (f) IFEVIP output image, (g) MST SR output image, (h) ADF output image, (i) CNN output image, (j) GTF output image, (k) LatLRR output image, (l) MSVD output image, (m) MGFF output image, (n) TIF output image, (o) VSMWLS output image, (p) Wavelet output image, (q) FPDE output image, (r) GFCE output image, (s) GFF output image, and (t) Proposed output image.

$[D_{22_{fused}} = \text{boundedPCA}(D_{22_{vis}}, D_{22_{therm}})]$  at scale 2 and angle 2 are shown in Fig. 6(d), 6(e) and 6(f), respectively.

Figure 6(g) shows the fused image red channel obtained by applying inverse curvelet transform on the fused curvelet coefficients and finally, the fused red channel obtained by further applying adaptive histogram equalization is shown in Fig. 6(h).

We also took into consideration the thermal and visual images displayed from the VIFB dataset in Fig. 4 (row 1 column 5) and (row 2 column 5), respectively. Here, we have combined a thermal image with images from the red, green, and blue channels. The end product of the proposed algorithm's fusion of the red channel and thermal image is depicted in Fig. 7(b), which shows the results of adaptive histogram equalization applied to the red channel fused image in Fig. 7(a). Similar results can be seen in Figs. 7(c) and 7(d) and in Figs. 7(e) and 7(f), which show the fused green channel image after applying the inverse curvelet transform and adaptive histogram equalisation, and the fused blue channel image after applying the inverse curvelet transform and adaptive histogram equalisation, respectively. Figure 7(g) displays the final integrated fused image, which is the final product of the proposed algorithm. As can be observed, the finished product provides more information than both the thermal and visual images.

#### 4.5 Qualitative Performance Comparison

In this part, we have contrasted the proposed method's qualitative performance with that of alternative approaches. Figure 8 compares the proposed method's qualitative performance on a pair of images to that of 18 previous fusion approaches. In this image pair, there are several people visible

in the infrared image but invisible in the visual image because they are obscured by a car's shadow. These people are virtually always discernible in fused images, as can be shown in Fig. 8. Multiple algorithms produce composite images with flaws. These techniques comprise the NSCT SR, RP SR, CBF, IFEVIP, and MST SR approaches. Moreover, the combined visuals generated by utilizing various techniques such as ADF, CNN, GTF, LatLRR, and MSVD do not sufficiently conserve the intricate particulars presented in the observable imagery. According to Fig. 8, the fused images, created by the MGFF, TIF, VSMWLS, wavelet approach, FPDE, GFCE, and GFF as well as the proposed CTBPCA method (Fig. 8(t)), are more accurate for human sensibility and retain more data.

#### 4.6 Quantitative Performance Comparison

Table 3 displays the average values for 11 assessment measures on 24 visual-thermal image pairs from the dataset in Fig. 4 for all methods. The suggested CTBPCA technique, with its four best values, offers the best overall quantitative performance, as can be seen. The CNN approach yields the second-best performance with 2 best and 2 second-best results. CBF, GTF, MGFF, MSVD, TIF, or VSMWLS did not produce any best, second-best, or third-best values. Table 3, however, makes it abundantly clear that there isn't a single fusion technique that can outperform competing ones in all or the majority of evaluation metrics.

#### 4.7 Runtime Comparison

The runtimes for each algorithm are listed in Table 4. The runtimes of various image fusion algorithms vary greatly. This holds even for techniques falling under the same category. Multi-scale approaches like ADF, CBF, and MSVD exist,

**Table 3. On 24 image pairings, the average assessment metrics for all procedures. The top three values for each statistic are indicated by the colors Green, blue, and red, respectively.**

Method	AG	CE	EI	EN	MI	PSNR	$Q_{cb}$	$Q_{cv}$	RMSE	SF	SD
ADF	4.975	1.59	49.969	6.817	2.304	<b>59.594</b>	0.476	716.95	<b>0.098</b>	15.328	36.26
CBF	7.48	1.075	77.256	7.307	2.559	58.814	0.528	1426.48	0.119	21.305	47.78
CNN	6.252	<b>0.967</b>	64.167	7.315	<b>2.888</b>	<b>59.101</b>	<b>0.613</b>	<b>463.62</b>	0.112	19.821	<b>59.563</b>
Proposed(CTBPCA)	<b>11.055</b>	1.381	<b>113.217</b>	<b>7.617</b>	1.125	57.166	0.461	889.3	0.134	<b>31.56</b>	53.912
FPDE	4.934	1.425	49.79	6.798	2.263	<b>59.259</b>	0.463	717.99	<b>0.098</b>	14.351	36.037
GFCE	<b>7.784</b>	2.12	<b>79.928</b>	7.276	1.876	56.384	0.526	813.95	0.162	<b>23.15</b>	52.682
GFF	5.848	1.402	59.912	7.219	2.646	58.243	<b>0.603</b>	822.37	0.108	18.524	49.023
GTF	4.692	1.439	47.082	6.586	1.978	58.017	0.398	1948.54	0.113	15.795	36.325
IFEVIP	5.641	1.406	57.885	6.956	2.291	57.423	0.457	<b>539.6</b>	0.131	17.290	48.513
LatLRR	<b>10.114</b>	1.713	<b>103.634</b>	6.915	1.659	56.399	0.484	652.17	0.161	<b>31.339</b>	<b>57.119</b>
MGFF	6.251	1.448	64.261	7.124	1.75	58.377	0.526	626.31	0.105	18.892	44.638
MST SR	6.305	<b>1.062</b>	64.835	7.367	<b>2.676</b>	58.077	<b>0.621</b>	<b>490.1</b>	0.113	19.826	<b>56.933</b>
MSVD	3.97	1.746	39.768	6.729	1.954	58.574	0.417	758.98	0.1	13.723	35.080
NSCT SR	6.898	<b>1.031</b>	71.419	<b>7.37</b>	<b>3.039</b>	57.626	0.596	1316.83	0.126	20.428	51.276
RP SR	6.814	1.123	69.248	<b>7.376</b>	2.249	57.921	0.585	812.05	0.117	22.235	55.56
TIF	5.954	1.559	61.478	7.086	1.749	58.389	0.531	562.73	0.104	18.596	43.053
VSMWLS	6.026	1.696	60.842	7.032	2	58.364	0.484	697.18	0.105	18.612	46.021
Wavelet	3.579	1.835	36.377	6.702	2.045	58.608	0.403	716.32	<b>0.099</b>	11.272	34.689

**Table 4.** Execution times for all algorithms (in sec.). The shortest and longest average times taken are shown in green and red, respectively

Method	Avg. time (s)	Total time (s)	Method	Avg. time (s)	Total time (s)
ADF	0.85938	20.625	LatLRR	<b>241.9932</b>	5807.836
CBF	14.7503	354.008	MGFF	0.98625	23.67
CNN	29.5178	708.428	MSTSR	0.84771	20.345
Proposed	1.06	25.4407	MSVD	0.81396	19.535
FPDE	2.3215	55.715	NSCTSR	105.2959	2527.101
GFCE	1.8518	44.443	RPSR	0.89125	21.39
GFF	0.33971	8.153	TIF	<b>0.12171</b>	2.921
GTF	4.8922	117.413	VSMWLS	3.1725	76.139
IFEVIP	0.13762	3.303	Wavelet	0.25397	6.0953

although CBF has a runtime that is roughly 17 times longer than ADF's and about 18 times longer than MSVD's. The quickest method for fusing images was by using TIF.

### 5. Summary and Future Scope of the Study

The usefulness of thermal imaging in many applications has increased significantly, yet there are still some situations in which they fall short. Combining the unique characteristics of thermal images with visual sensor images is more appropriate to deal with such a circumstance. In this regard, the current paper suggested a quick discrete curvelet transform for feature extraction and a bounded PCA methodology for the fusing of a thermal-visible image's detailed coefficients. The viability of the suggested technique was evaluated by conducting tests on thermal-visible image pairs acquired from both our images shot with AXIS Q8741-LE Bi-spectral PTZ Network Camera and the publicly available VIFB dataset. The suggested methodology and the already used image fusion techniques have been compared, along with several evaluation indicators. It has been found that the suggested strategy produces superior results concerning various measurements.

Our future study will be based on using the fused image for target recognition. Future work should also focus on this method's continued improvement and optimization for real-time video processing.

### REFERENCES

- Ma, J.; Ma, Y. & Li, C. Infrared and visible image fusion methods and applications: A survey. *Inf. Fusion*, 2019, **45**, 153-178. doi: 10.1016/j.inffus.2018.02.004.
- Wang, Z.; Cui, Z. & Zhu, Y. Multi-modal medical image fusion by Laplacian pyramid and adaptive sparse representation. *Comput. Biol. Med.*, 2020, **123**, 103823. doi: 10.1016/j.combiomed.2020.103823.
- Polinati, S. & Dhuli, R. A review on multi-model medical image fusion. *In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019, 0554-0558. doi: 10.1109/ICCSP.2019.8697906.
- James, A.P. & Dasarathy, B.V. Medical image fusion: A survey of the state of the art. *Inf. Fusion*, 2014, **19**, 4-19. doi: 10.1016/j.inffus.2013.12.002.
- Jinju, J.; Santhi, N.; Ramar, K. & Sathya Bama, B. Spatial frequency discrete wavelet transform image fusion technique for remote sensing applications. *Eng. Sci. Technol., An Int. J.*, 2019, **22**(3), 715-726. doi: 10.1016/j.jestch.2019.01.004.
- Paramanandham, N. & Rajendiran, K. Infrared and visible image fusion using discrete cosine transform and swarm intelligence for surveillance applications, *Infrared Physics & Technol.*, 2018, **88**, 13-22. doi: 10.1016/j.infrared.2017.11.006.
- Sathaki, T. *Image fusion: Algorithms and applications*, Academic Press, 2008.
- Wang, Z.J.; Ziou, D.; Armenakis, C.; Li, D.R. & Li, Q.Q. A comparative analysis of image fusion Methods. *IEEE Trans. Geoscie. Remote Sensing*, 2005, **6**, 1391-1402. doi: 10.1109/TGRS.2005.846874
- Chipman, L.J.; Orr, T.M. & Graham, L.N. Wavelets and image fusion. *In Proceedings International Conference on Image Processing*, 1995, **3**, 248-251. doi: 10.1109/ICIP.1995.537627.
- Candes, E.J. & Donoho, D.L. *Curvelets—A surprisingly effective non-adaptive representation for objects with Edges*. Vanderbilt University Press, Nashville, TN, 2000.
- Candes, E.J.; Demanet, L.; Donoho, D.L. & Ying, L. Fast discrete curvelet transform. *SIAM Multiscale Modeling and Simulations*, 2006, **5**(3), 695-1043. doi: 10.1137/05064182X.
- Zheng, Y.; Hou, X.; Bian, T. & Qin, Z. Effective image fusion rules of multi-scale image decomposition. 5<sup>th</sup> International Symposium on Image and Signal Processing and Analysis, 2007, 362-366. doi: 10.1109/ISPA.2007.4383720.
- Miao, Q. & Wang, B. A novel image fusion algorithm using FRIT AND PCA. *In 10<sup>th</sup> International Conference on Information Fusion*, 2007, 1-5. doi: 10.1109/ICIF.2007.4408067.
- Shi, W.; Zhu, C.; Tian, Y. & Nichol, J. Wavelet-based image fusion and quality assessment. *Int. J. Appl. Earth Observation and Geoinf.*, 2005, **6**, 241-251. doi: 10.1016/j.jag.2004.10.010.
- Liu, Z.; Blasch, E.; Xue, Z.; Zhao, J.; Laganiere, R. & Wu,

- W. Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study. *IEEE Transact. Pattern Analysis and Machine Intelligence*, 2012, **34**, 94–109.  
doi: 10.1109/TPAMI.2011.109.
16. Zhang, X.; Ye, p. & Xiao, G. VIFB: A visible and infrared image fusion benchmark. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, 468-478.  
doi: 10.1109/CVPRW50498.2020.00060.
  17. Bulanon, D.M.; Burks, T. & Alchanatis, V. Image fusion of visible and thermal images for fruit detection. *Biosyst. Eng.*, 2009, **103**(1), 12–22.  
doi: 10.1016/j.biosystemseng.2009.02.009.
  18. Cui, G., Feng, H., Xu, Z., Li, Q. & Chen, Y. Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition. *Optics Commun.*, 2015, **341**, 199 – 209.  
doi: 10.1016/j.optcom.2014.12.032.
  19. Roberts, J.W.; Aardt, J.V. & Ahmed, F. Assessment of image fusion procedures using entropy, image quality, and multispectral classification. *Jo. Appl. Remote Sensing*, 2008, **2**(1), 1-28.  
doi: 10.1117/1.2945910.
  20. Rajalingama, B.; Priya, R. & Bhavani, R. Hybrid multimodal medical image fusion using combination of transform techniques for disease analysis. In *International Conference on Pervasive Computing Advances and Applications – PerCAA 2019*, Procedia Computer Science., 2019, **152**, 150–157.  
doi: 10.1016/j.procs.2019.05.037.
  21. Qu, G.; Zhang, D. & Yan, P. Information measure for performance of image fusion. *Electron. Lett.*, 2002, **38**(7), 313–315.  
doi: 10.1049/el:20020212.
  22. Rao, Y.J. In-fibre bragg grating sensors. *Measure. Sci. Technol.*, 1997, **8**(4), 355.
  23. Jagalingam, P. & Hegde, A.V. A review of quality metrics for fused image. *Aquatic Procedia*, 2015, **4**, 133–142.  
doi: 10.1016/j.aqpro.2015.02.019.
  24. Eskicioglu, A.M. & Fisher, P.S. Image quality measures and their performance. *IEEE Transact. Commun.*, 1995, **43**(12), 2959–2965.  
doi: 10.1109/26.477498.
  25. Chen, Y. & Blum, R.S. A new automated quality assessment algorithm for image fusion. *Image and Vision Comput.*, 2009, **27**(10), 1421–1432.  
doi: 10.1016/j.imavis.2007.12.002.
  26. Chen, H. & Varshney, P.K. A human perception inspired quality metric for image fusion based on regional information. *Inf. Fusion*, 2007, **8**(2), 193–207.  
doi: 10.1016/j.inffus.2005.10.001.
  27. Wu, Y.; Lim, J. & Yang, M.-H. Online object tracking: A benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, 2411–2418.  
doi: 10.1109/CVPR.2013.312.
  28. Wu, Y.; Lim, J. & Yang, M.-H. Object tracking benchmark. *IEEE Transact. Pattern Analysis and Machine Intell.*, 2015, **37**(9), 1834–1848.  
doi: 10.1109/TPAMI.2014.2388226.
  29. Kristan, M.; Matas, J.; Leonardis, A.; Vojir, T.; Pflugfelder, R.; Fernandez, G.; Nebhay, G.; Porikli, F. & Cehovin, L. A novel performance evaluation methodology for single target trackers. *Pattern Analysis and Machine Intell.*, 2016, **38**(11), 2137–2155.
  30. Davis, J.W. & Sharma, V. Background-subtraction using contour-based fusion of thermal and visible imagery. *Comput. Vision and Image Understanding*, 2007, **106**(2-3), 162–182.  
doi: 10.1016/j.cviu.2006.06.010.
  31. Toet, A. TNO image fusion dataset. *Figshare. Data*, 2014.
  32. Ellmauthaler, A.; Pagliari, C.L.; Silva, E.A. da; Gois J.N. & Neves, S.R. A visible-light and infrared video database for performance evaluation of video/image fusion methods. *Multidimensional Syst. Signal Process.*, 2019, **30**(1), 119–143.  
doi: 10.1007/s11045-017-0548-y.
  33. O’Conaire, C.; O’Connor, N.E.; Cooke, E. & Smeaton, A.F. Comparison of fusion methods for thermo-visual surveillance tracking. In *9<sup>th</sup> International Conference on Information Fusion*. IEEE, 2006, 1–7.  
doi: 10.1109/ICIF.2006.301618.
  34. Li, C.; Liang, X.; Lu, Y.; Zhao, N. & Tang, J. RGB-T object tracking: Benchmark and baseline. *Pattern Recognition*, 2019, **96**, 106977.  
doi: 10.1016/j.patcog.2019.106977.
  35. Bavirisetti, D.P. & Dhuli, R. Fusion of infrared and visible sensor images based on anisotropic diffusion and karhunen-loeve transform. *IEEE Sensors J.*, 2016, **16**(1), 203–209.  
doi: 10.1109/JSEN.2015.2478655.
  36. Shreyamsha Kumar, B.K. Image fusion based on pixel significance using cross bilateral filter. *Signal, Image and Video Process.*, 2015, **9**(5), 1193–1204.  
doi: 10.1007/s11760-013-0556-9.
  37. Liu, Y.; Chen, X.; Cheng, J.; Peng, H. & Wang, Z. Infrared and visible image fusion with convolutional neural networks. *Int. J. Wavelets, Multiresolution and Inf. Process.*, 2018, **16**(3), 1850018.  
doi: 10.1142/S0219691318500182.
  38. Bavirisetti, D.P.; Xiao, G. & Liu, G. Multi-sensor image fusion based on fourth-order partial differential. In *20<sup>th</sup> International Conference on Information Fusion*. IEEE, 2017, 1–9.  
doi: 10.23919/ICIF.2017.8009719.
  39. Zhou, Z.; Dong, M.; Xie, X. & Gao, Z. Fusion of infrared and visible images for night-vision context enhancement. *Applied Optics*, 2016, **55**(23), 6480–6490.  
doi: 10.1364/AO.55.006480.
  40. Li, S.; Kang, X. & Hu, J. Image fusion with guided filtering. *IEEE Transact. Image Process.*, 2013, **22**(7), 2864–2875.  
doi: 10.1109/TIP.2013.2244222.
  41. Ma, J.; Chen, C.; Li, C. & Huang, J. Infrared and visible

- image fusion via gradient transfer and total variation minimization. *Inf. Fusion*, 2016, **31**, 100–109.  
doi: 10.1016/j.inffus.2016.02.001.
42. Zhang, Y.; Zhang, L.; Bai, X. & Zhang, L. Infrared and visual image fusion through infrared feature extraction and visual information preservation. *Infrared Physics Technol.*, 2017, **83**, 227 – 237.  
doi: 10.1016/j.infrared.2017.05.007.
  43. Li, H. & Wu, X. Infrared and visible image fusion using latent low-rank representation.  
doi: 10.48550/arXiv.1804.08992, 2022.
  44. Bavirisetti, D.P.; Xiao, G.; Zhao, J.; Dhuli, R. & Liu, G. Multi-scale guided image and video fusion: A fast and efficient approach. *Circuits, Syst. Signal Process.*, 2019, **38**(12), 5576–5605.  
doi: 10.1007/s00034-019-01131-z.
  45. Liu, Y.; Liu, S. & Wang, Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion*, 2015, **24**, 147–164.  
doi: 10.1016/j.inffus.2014.09.004.
  46. Naidu, V. Image fusion technique using multi-resolution singular value decomposition. *Defence Sci. J.*, 2011, **61**(5), 479–484.  
doi: 10.14429/dsj.61.705
  47. Bavirisetti, D.P. & Dhuli, R. Two-scale image fusion of visible and infrared images using saliency detection. *Infrared Physics Technol.*, 2016, **76**, 52–64.  
doi: 10.1016/j.infrared.2016.01.009.
  48. Ma, J.; Zhou, Z.; Wang, B. & Zong, H. Infrared and visible image fusion based on visual saliency map and weighted least square optimization. *Infrared Physics Technol.*, 2017, **82**, 8–17.  
doi: 10.1016/j.infrared.2017.02.005.
  49. Kaur, H.; Koundal, D. & Kadyan, V. Image fusion techniques: A survey. *Archives of Computational Methods Eng., Springer*, 2021, **28**, 4425–4447.  
doi: 10.1007/s11831-021-09540-7.

## CONTRIBUTORS

**Mr A.K. Singh** obtained BTech in Electronics and Communication Engineering from the Heritage Institute of Technology, Kolkata, West Bengal, India. He is working at Techno India University, Kolkata, as a Scientific Programmer Scientist. His research interests include: Computer vision, image processing, machine learning, and remote-sensing.

In the current study he reviewed the pertinent literature and implementing the project, and testing the suggested procedure.

**Dr Prof D. Chaudhuri** obtained PhD in image processing and pattern recognition. He is a Professor of computer science and engineering at Techno India University in Kolkata. His research interest include: Span image processing, pattern recognition, computer vision, remote sensing, machine learning, deep learning, and target detection from satellite, SAR, thermal, and MMW imageries.

In the current study he assisted the project scientists and developed the entire algorithm, which led to the current research.

**Mr S. Mitra** obtained MTech in Computer Science and Technology from Techno India University, Kolkata and working as a Scientific Programmer Scientist. His current research interest include: Computer vision, image processing, machine learning, and remote sensing.

He has contributed to the current study by doing the literature review and analysing the experiments.

**Mr M.P. Singh** obtained his MTech from IIT Kharagpur in Computer Science and Engineering and working as a Director of DRDO Young Scientist’s Laboratory-Cognitive Technologies at Chennai. His area of interest include: Pattern recognition, satellite image processing, and cognitive Surveillance.

In the current study he worked on the algorithm’s overall architecture.

**Dr B.B. Chaudhuri** obtained his PhD degree from Indian Institute of Technology, Kanpur. He did his post-doctorate work during 1981-1982 at Queen’s University, U.K, through Leverhulme Overseas Fellowship. He is working as a Pro-Vice-Chancellor (academic) of Techno India University in West Bengal, India. He received the Om Prakash Bhasin Award for his contribution to the field of electronics and information technology. His primary research contributes to the fields of computer vision, image processing, and pattern recognition.

He advised and guided the research work in the current study.