International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Jan-2022)
http://dx.doi.org/10.12785/ijcds/110102

A Compendious Analysis of Feature-Extraction Algorithms to Frame Fusion Rules

Teku Sandhya Kumari[1], S. Koteswararao[2] and I. Santi Prabha[3]

[1]*Department of Electronics and Communication Engineering, Vignanś Institute of Engineering for Women, Visakhapatnam, India*
[2]*Department of Electronics and Communication Engineering, K. L.E. F (Deemed to be University), Guntur, India*
[3]*Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University, Kakinada, India*

Received 11 Dec. 2020, Revised 26 Jul. 2021, Accepted 13 Nov. 2021, Published 9 Jan. 2022

Abstract: Advancement in sensor technology provides the complete information captured by multiple sensors. To reduce the eye strain and workload from analyzing the scene with multiple images, the information is combined into a single image by the process called image fusion. In this paper, a compendious analysis of feature-extraction based fusion algorithms that define an appropriate fusion rule is reviewed. A state-of-art classification of feature-based fusion schemes is carried out and the extracted feature maps are presented. The qualitative analysis for different fusion methods are illustrated and compared. The quantitative fusion metrics are grouped as contrast, information, edge and visual based metrics and are evaluated. Finally, the conclusion and future directions are briefed out.

Keywords: Image fusion, feature-extraction, statistical features, structural features, saliency features, fusion rule, fusion metric

1. INTRODUCTION

Tremendous development in multi-sensor technology provides information of a captured scene from multiple sensors useful for industry automation, medical, military, night-vision and surveillance applications. The complementary information captured by multiple sensors increases the workload, degrades the performance and causes eyestrain of the observer if the captured scene is to be analyzed from multiple images. Image fusion is a digital processing process of implementing the mathematical techniques to combine complementary information from multiple sensors into a single composite image that provides good human perception ability for target detection and scene interpretation. Amassing accurate information in less time and less cost are the potential advantages of image fusion. Image features that cannot be perceived by a single sensor are easily distinguishable by image fusion. Hence, image fusion has a vast amount of research attention since last two decades.

Single-sensor image fusion focuses on combining the information captured by a single sensor at different focus points or different views that defines multi-focus or multi-view image fusion. Multiple-sensor image fusion deals with combining complementary information obtained from different sensors or modalities that define multi-sensor, multi-modality or multi-spectral fusion. Several algorithms for multi-focus, multi-sensor and multi-modality were emerged to solve the problems of expensive sensor requirement, information overload and to help clinicians. However, research on multi-sensor image fusion has gained a large attention in the last few decades. The objective of multi-sensor image fusion is to acquire the complete information from the source images without any artifacts introduced due to fusion process and to reduce redundancy.

Weighted average, select maximum and select minimum are the basic arithmetic operators that are employed as fusion rules. These arithmetic operators when used as fusion rules either directly in spatial domain or in transform domain to define the final fused image makes the algorithm simple, but results in loss of contrast, color distortion and blurring effects. Feature extraction based fusion rules for '*weighted average*' or '*select maximum*' scheme improves the ability of the algorithm compared to traditional methods. In the feature extraction based image fusion, the texture feature that provides a measure of smoothness, coarseness and regularity of the images are extracted and these are used to frame the fusion rule for combining. Extracting the important features to exactly interpret the image content and defining a suitable fusion rule plays a vital role to achieve the fusion objectives. Therefore, it is a critical challenge to propose a suitable parameter for feature extraction and define an appropriate fusion rule in order to attain the fusion requirements.

In this paper, the survey focuses on the features extracted from the source images to define the fusion rules.

E-mail address: teku.sandhya@gmail.com, rao.sk9@gmail.com, santiprabha@yahoo.com

The current work reviews the work of authors to identify appropriate parameters required to efficiently extract the features and then define arithmetic fusion rule for ensuring the complete information, edge symmetry and no artifacts in the final fused image.

2. FEATURE-EXTRACTION FUSION METHODS

A feature of an image is any attribute or distinguishing property of an image. Feature extraction in image processing is the process that extracts the desired characteristics of the acquired images. The features extracted by the process need to be more informative, reduce redundancy and are suitable for better human interpretation. Boundary and regional features are the two types of features that describe the external and internal characteristics suitable for further processing. Boundary features or shape features that can be extracted from the Fourier descriptors may not discriminate the objects in an image. Intensity, color and texture are the regional features that define the characteristics of a region. The texture of a region provides a measure of smoothness, coarseness and regularity. Texture analysis is widely used in image processing applications from segmentation to pattern recognition. In the image fusion scenario, the texture features of the input images are extracted and these are used to frame the fusion rule for combining. Recently, feature-based fusion using deep learning techniques gained more attention for accurate fusion and classification[1]

Various parameters have been defined by the researchers to extract the salient features. The work categorizes the fusion algorithms in literature based on domain specific and based on texture-feature extraction as illustrated in Fig. 1.

A. Domain based classification

The **spatial domain** method directly manipulates the pixels by defining a fusion rule. Feature-extraction based spatial domain techniques define the fusion rule for combining depending on the extracted features from the images. Spatial domain schemes are simple to implement but could not succeed in retaining the complete information, introduced artifacts and reduced the contrast of the final fused image.

Multi-Scale Transform (MST) such as Discrete Wavelet transform [2], Non- Sub sampled Contorlet Transform [3], Spectral Graph Wavelet Transform [4] and Curvelet Transform [5] etc. have been widely used in image fusion for representing the image at different scales. The scaling function of the transform decomposes the image to low-level and high-frequency level coefficients. In feature extraction based fusion, the features are extracted at the corresponding scales individually and are then used to construct a fusion rule. The transform domain techniques gained a lot of attention in achieving a high quality fused image. However, transform domain techniques are computationally expensive and selection of transform and level of transformation are the monotonous tasks. The transform domain method fails to preserve the edge and texture information.

Multi-Scale Decomposition (MSD) techniques employ edge-preserving smoothing filters to represent the image at various scales. Guided Filter [6], Neighbor distance filter [7], weighted least square [8] and iterative re-weighted filters [9], [10], [11] etc. have been used to efficiently decompose the image to various scales with edge and texture preservation. The salient features of these scaled images are extracted to improve the decision making fusion rule in feature extraction based fusion. However, these decomposition techniques have their deficiencies in accurately decomposing the image to structure and texture components that affect the final fused image. Hence, achieving promising results in spatial, transform and decomposition domain is still a continuous research for fusion.

B. Classification based on texture feature extraction

The traditional weighted average and select maximum fusion rules employed directly to the source images do not completely transfer the complementary information and also introduce noise. The characteristics of the source images captured by multiple sensors are different. Hence, the characteristics features of these images are used to define the rule instead of defining directly. Texture of an image is characterized by the spatial distribution of the pixel intensities within a neighborhood. The texture features or the characteristics of an image in our work are classified as statistical, saliency and structural approaches and the algorithms under these categories are reviewed.

C. Statistical Approach for Feature-Extraction

The characteristics of an image such as smooth, texture and edge regions are extracted by these statistical approaches. The statistical parameters considered to compute the texture features are listed as:

- Local Mean
- Local Standard Deviation or Local Variability
- Local variance
- Local Visibility
- Entropy
- Local Binary Pattern (LBP)
- Local Fractal Dimension (LFD)
- Principal Component Analysis
- Dispersion
- Kurtosis

Local mean, Local standard deviation and Local Variance are the fundamental local statistics computed to determine the texture features of an image region. These parameters consider the information of the neighboring pixels. Local mean computes the mean gray-scale pixel

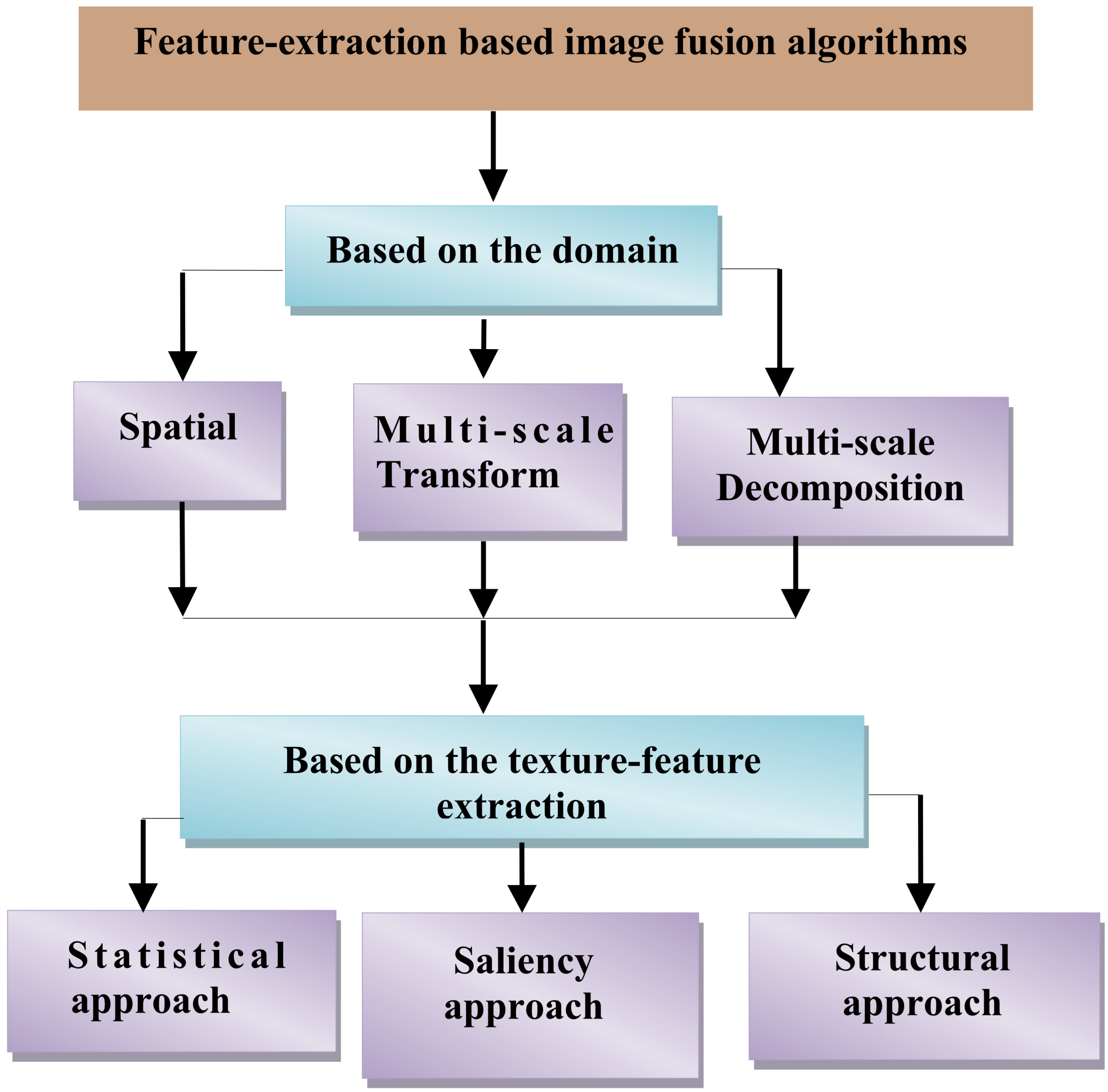

Figure 1. Classification of feature-extracted image fusion algorithms

values in an image region or patch. It describes the average brightness of the region and is useful to give the information regarding the pixel intensity distribution. Local standard deviation or local variability on the other hand measures the variability or smoothness in the gray level that is more informative than the local mean value. The smaller value indicates less variability of the gray level and larger value indicates more variability in the image region. Local variance is a statistical measure that quantifies the local activity of an image. The value of the variance computed within a window describes the smoothness or roughness of a region. Smaller values indicate smooth regions and larger values indicate rough regions. The background variations of the image are efficiently determined by this local variance measure.

A combination of region and pixel based image fusion algorithm using multi-resolution analysis is proposed in [12] where the activity level is measured by considering the absolute value of the pixel as the energy calculation of a surrounding region. In [13] the feature maps of the input images are extracted by considering the local mean and standard deviation. The feature maps of both local mean and standard deviation are used to define the *select maximum* rule for combining the approximate coefficients and only the standard deviation map is used to define the *select maximum* rule for detail coefficients in DWT domain. The feature map generated by using local standard deviation is however used in [14] to define *weighted average* rule for combining images in compressive sensing domain. Local variance is computed to the low-frequency coefficients decomposed by using a hybrid DWT and Curvelet transform and the fusion rule is defined to combine the MRI and CT medical images in [15]. Recently, the local activity or local variance is used as a fusion rule to construct an initial fused image in feature-oriented level fusion of multi-sensor images in [16]. The algorithm is computationally efficient and is a simple spatial domain method.

Local Visibility is a statistical parameter that defines the high visible areas of an image. The texture characteristics of an image are highly visible without any noise by considering this local visibility parameter. The high visibility content of the image is extracted by the local visibility parameter in the multi-sensor fusion of visible and infrared images in [17].

Entropy is an additional texture measure that defines the variability in the pixel intensities in a local window. The value of entropy is less for a smooth region and is large for texture region [18]. However, the texture region extracted from the entropy measure depends on the size of the window selected. Region based entropy priority maps for multi-focus images were used in [19] to extract the region properties and define intelligent rules for fusion compared to intensity-level fusion. The number of compressive measurements is adaptively adjusted for visible and infrared image integration using entropy in [20]. Entropy based *weighted average* is used to construct low frequency coefficient fused image which is decomposed by MSD technique proposed in [21], [22] constructed an initial fused image by the use of entropy as *weighted average* scheme after pre-processing the images to improve the fusion algorithm performance.

The mathematical formulation for computing the local mean m_i, j, standard deviation $s_{(i,j)}$, variance $v_{(i,j)}$, visibility '$v_{s_{(i,j)}}$' and entropy '$e_{(i,j)}$'of an image pixel intensity $f(i, j)$ located at $f(i, j)$ within a region $m \times n$ is

$$m_i, j = \frac{1}{mn}\sum_{k=1}^{m}\sum_{l=1}^{n} f(k,l) \tag{1}$$

$$m_i, j = \sqrt{\frac{1}{mn}\sum_{k=1}^{m}\sum_{l=1}^{n} (f(k,l) - m_i, j)^2} \tag{2}$$

$$v_i, j = \frac{1}{mn}\sum_{k=1}^{m}\sum_{l=1}^{n} (f(k,l) - m_i, j)^2 \tag{3}$$

$$v_{s_{(i,j)}} = \frac{1}{m*n}\sum_{k=1}^{m}\sum_{l=1}^{n} \alpha * \left(\frac{\left|f(k,l) - m_{i,j}\right|}{m_{i,j}}\right) where \alpha = \left(\frac{1}{m_{i,j}}\right)^{0.6} \tag{4}$$

'$m_{i,j}$' denotes local mean, $f_{k,l}$ is the local pixel intensity, 'm' and 'n' are the indices indicating window size.

$$e_{i,j} = -\sum_{k=1}^{m}\sum_{l=1}^{n} p(f(k,l))\, log_2 (p(f(k,l))) \tag{5}$$

where $p(f)$ represents the probability of occurrence of the pixel $f(i, j)$.

Local Binary Pattern (LBP) is an efficient and simple texture extractor that labels the pixels in binary by thresholding the neighborhood pixels $https : //scholarpedia.org/article/Local_Binary_Patterns$. It is a powerful tool used recently for fusion application. Considering a center pixel 'f_c' located at (X_C, Y_C) and pixels 'f_i' for $j = 0, 1, ...M - 1$ within a window consisting of 'M' number of pixels, the LBP of the pixel is computed as:

$$LBP(x_c, y_c) = \sum_{j=0}^{M-1} S\left(f_j - f_c\right) * 2^j \tag{6}$$

where

$$S\left(f_j - f_c\right) = \begin{cases} 0 & \left(f_j - f_c\right) < threshold \\ 1 & \left(f_j - f_c\right) \geq threshold \end{cases}$$

In order to preserve the edge information and the information of focused region, recently a new multi focus image fusion scheme is proposed by defining the weight maps using optimized LBP maps in [23].

Local Fractal Dimension (LFD) describes the local grey level changes in an image. It differentiates the local edges from noise and from segment interiors compared to Fractal Brownian noise model [24] that describes only the gray level change. The LFD measures the textural features in a local window. Several approaches have been defined to estimate LFD, out of which, the Blanket method for LFD estimation is popular. In this Blanket method, the area of the gray level surface is computed from the upper and lower surfaces at scale 'ϵ' and the linear relationship between this area and scale in logarithmic scale estimates the LFD. Fusion for night vision applications that make use of FD maps to construct the fused image is proposed in [25]. However, the local statistics accuracy in determining the texture features depends on the size of the window and the pixel intensity distribution within the window. Generally, smaller window sizes are more effective compared to large size image patches. Hence, the other statistical global approaches to extract texture features are Principal Component Analysis, Kurtosis and Dispersion used in literature for defining and optimizing the weights.

Principal Component Analysis (PCA) is a classical tool for analyzing the statistical nature of large data sets [26]. PCA considers the variances and co-variances of the pixel intensities. It transforms large number of inter-correlated variables to a small set of variables called principal components [27]. The features that contain variance of the input data to a maximum possible extent are defined by the principal components. The first principal component derived gives the maximum variance. The second principal component is obtained by considering the feature with maximum variance and is orthogonal to the first principal component. This procedure is repeated to obtain the number of principal components required. In image fusion scenario, the principal components of the input images are computed and are considered as weights.

PCA has been widely used by several researchers for the past two decades in image fusion process. PCA along with total variation approach is proposed earlier in [28] to iteratively modify the weights for fusion. Recently PCA has been widely used as a feature extraction metric that defines the weighted average fusion rule in DWT and hybrid domain [29], [30]. Recently, multi scale decomposition techniques such as anistropic diffusion [31], [32], fourth order partial differential equation [33], were frequently used to decompose the image to detail and base layers. The PCA is then used to extract the feature maps to define the fusion rule for detail layer fusion in multi-sensor and multi-modality fields. The suggested methods using PCA as a feature map improved the quality of the fused image.

Dispersion is a statistical parameter described as the ratio of 4th order central moment to the variance of an image. The variability or spread of the histogram of an image is measured. A normalized 4th order central moment called **Kurtosis** determines the peakedness of the probability distribution. Dispersion as a saliency measure to minimize the cost function that optimally updates the weights required to define the weighted average fusion rule is proposed in [34]. Instability and bias due to lack of prior information is a major drawback of this approach. Kurtosis Maximization method proposed by the same author avoids the limitations. Optimized weights for fusion are selected based on the maximization of Kurtosis.

The statistical approach for feature extraction has a profound area for defining the fusion rules using the basic arithmetic operators. New statistical parameters such as range of the histogram, average and centre symmetric LBP [35] considered as essential features for disease classification in health care applications can be recommended to fusion techniques. Feature extraction by Gray Level Co-occurrence Matrix that involves various statistical features for accurate classification is employed in the recent research for decision fusion [36], [37].

Selection of appropriate statistical parameters to characterize the properties of an image region has a significant effect on the fusion result. The state-of-art methods proposed earlier using statistical properties showed improved quantitative and qualitative performance. However, the scope for the use of these statistical parameters for feature-level, feature-oriented level fusion is still in progress. The simulation results of feature maps extracted in MATLAB for a standard data set image downloaded from $https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029$ by using the statistical approaches described above are shown in Fig. 2.

D. Saliency approach for feature extraction

Saliency of an image highlights the significant regions that attract the human vision compared to other regions. Saliency term is introduced for extracting the visual features in 1980's. The limitations of saliency detection methods such as low resolution, undefined boundaries and reduction in spatial frequency range etc. lead to the development of various saliency measures for the improvement in the detection of salient objects for various image processing applications. Recently, saliency based feature extraction for image fusion has been widely preferred for defining the set of weights for combining. The weights defined by the saliency metric measures the contribution of every pixel to the final fused image. The saliency of an image is defined in various ways by several researchers. Some of the saliency measures used for fusion process is listed as:

- Spatial Frequency
- Frequency tuned saliency
- Maximum symmetric surround saliency
- Visual perception saliency
- Pixel value based Saliency
- Mean square error based saliency
- Gradient based saliency
- Spectral Residual based saliency

Spatial Frequency (SF) is a saliency metric that measures the entire activity of an image. It is an effective feature extraction technique that describes the row and column frequency distribution. The low and high frequency visual information of an image with high visual perception can be extracted by computing the regional characteristics using SF. For an image region of size M x N with pixel 'x_i, j' as the centre pixel, the frequency distribution of an image along row (RF) and column (CF) are determined and then the region SF is computed as:

$$RF = \sqrt{\frac{1}{mn}\sum_{i=1}^{M}\sum_{j=2}^{N}\left(x_{i,j} - x_{i-1,j}\right)^2} \tag{7}$$

$$CF = \sqrt{\frac{1}{mn}\sum_{j=1}^{N}\sum_{i=2}^{M}\left(x_{i,j} - x_{i-1,j}\right)^2} \tag{8}$$

$$SF = \sqrt{RF^2 + CF^2} \tag{9}$$

SF with high value describes the features such as fine details and edges whereas the low value of SF describe the smoothness or global shape. The periodic distributions of the pixels in a region are estimated by this parameter.

A computationally simple algorithm in spatial domain for multi-focus fusion is proposed in [38].The method divides

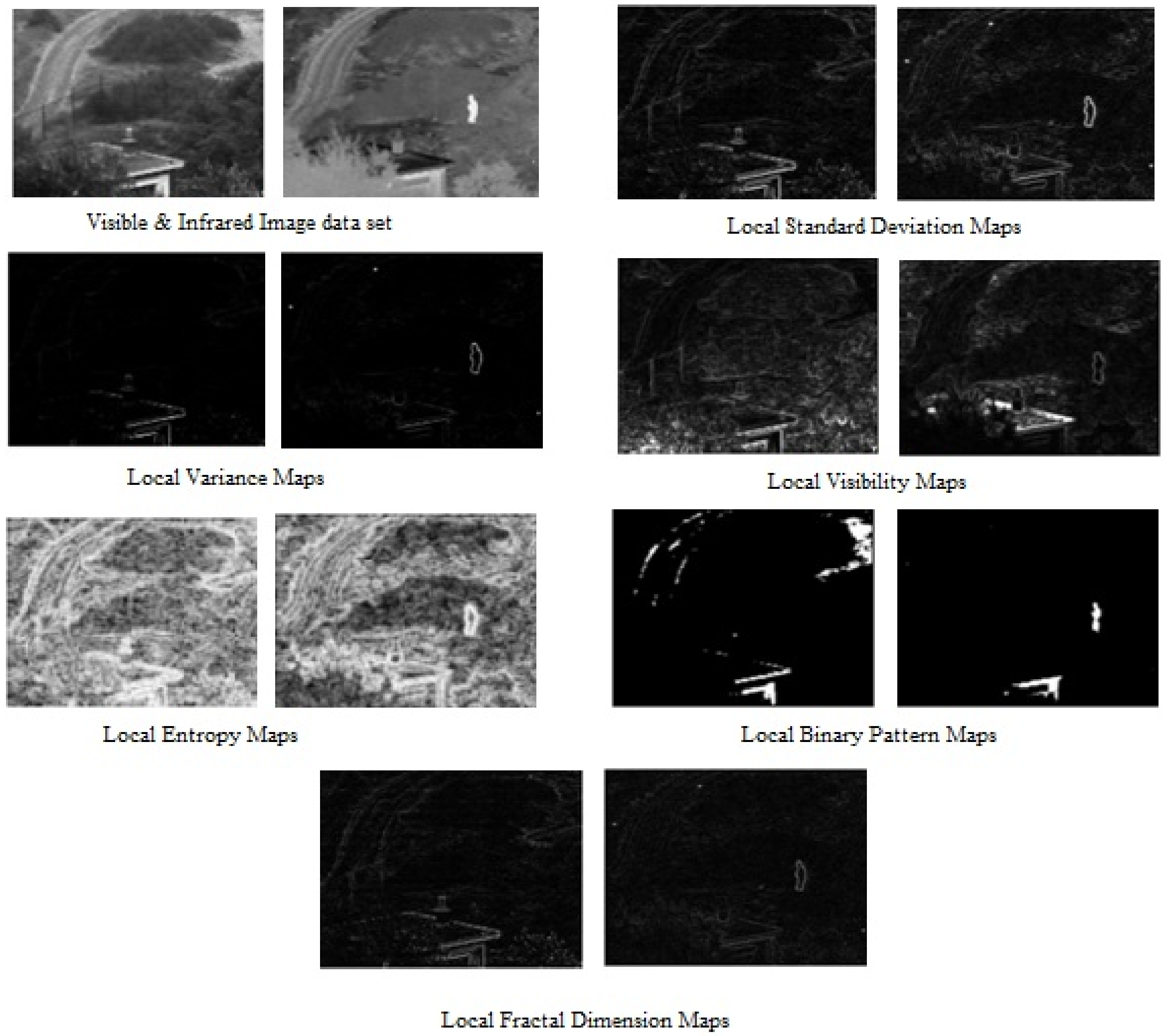

Figure 2. Feature Maps using Statistical Approach

the images into blocks and the SF of each block is computed. The computed SF is used to define the *select maximum* fusion rule to get a composite image. The algorithm is tested for the fusion of images captured at different focuses used for real-time application.

Later, the same metric SF has been used by the same author for region- based image fusion where the SF defines the decision maps required for fusion. In [39], a multi-focus fusion algorithm in Discrete Cosine Transform (DCT) domain is proposed in which the SF of the DCT blocks is computed and *select maximum* rule is defined to get the fused DCT block. However, recently the metric is used in [40] for extracting the local features. In this method, SWT decomposes the source images to low and high frequency components which are further divided into sub-images using fuzzy sets. The SF of these sub-images is computed to frame the fusion rule for sub-image fusion. Further, the inverse SWT reconstructs the fused image. Global and local layers of the two images are obtained by neighbor distance filter proposed in [6] and *weighted average* rule is framed based on the SF calculated for the local layers that reduces the blocking effect in the final image.

Frequency tuned saliency is proposed in [41] to detect the salient objects in an image with clear boundaries and with full resolution. In order to retain a wide band of frequencies, the Difference of Gaussian (DoG) band pass filter is used to obtain the Gaussian blurred image to compute saliency. DoG filter is appropriate for detecting the changes in the intensity levels of an image defined as:

$$DoG\,(i, j) = \frac{1}{2\pi}\left[\frac{1}{\sigma_1^2}e^{-\left(\frac{i^2+j^2}{\sigma_1^2}\right)} - \frac{1}{\sigma_2^2}e - \left(\frac{i^2+j^2}{\sigma_2^2}\right)\right] \qquad (10)$$

where $\sigma_2 < \sigma_1$ represents the standard deviations of the Gaussian function that determines the high and low spatial frequencies. Thus the frequency-tuned saliency of an image

'I' is computed by taking the absolute difference between the mean image 'I_μ' and the DoG Gaussian blurred image given as:

$$S(i,j) = \left|I_\mu - I_{DoG}(i,j)\right| \tag{11}$$

In order to achieve the saliency detection objective of making the computation efficient, the above equation is modified by the same author as:

$$S(i,j) = \left\|I_\mu - I_{DoG}(i,j)\right\| \tag{12}$$

where '$\|\ \|$' represents the Euclidean distance. However, the Eq. 12 results in asymmetric surround of pixels because of considering the complete image as global surround. Hence, the local salient detection is performed by considering a **Maximum Symmetric Surround Saliency (MSSS)** proposed in [42], where the mean image is obtained as:

$$I_\mu(i,j) = \frac{1}{w}\sum_{x=i-i_0}^{i+i_0}\sum_{y=j-j_0}^{j+j_0} I(x,y) \tag{13}$$

with size as M and N respectively. However, the area of the sub-image and the offsets are computed as:

$$w = (2i_0+1)(2j_0+1) \tag{14}$$

$$i_0 = min\{i, M-i\}$$

$$j_0 = min\{j, N-J\}$$

MSSS represented by S(i, j), thus extracts the salient objects even for complex backgrounds highlighting the salient regions. These saliency measures have been utilized to extract the salient features of differently focused regions to define the fusion rule in image fusion. An efficient method for multi-focus, multi-sensor, multi-spectral fusion is proposed in [43], where the saliency maps of decomposed detail layers are computed by taking the absolute value of the Laplacian filtered image convolved with Gaussian filtered image to construct weight maps. The saliency maps constructed by this method present a good description of detail layer information [44] proposed an edge-preserved multi-scale decomposition approach to decompose the image to detail and base layers. The frequency tuned saliency maps of the detail layers are extracted and are used as weights for detail layers fusion of visible and infrared images. Further, the binary weight maps are constructed from the feature maps extracted by the frequency-tuned saliency maps in [45] for the fusion of multi-sensor images in spatial domain. In [46], the saliency maps constructed by frequency tuned saliency calculated for multiple local windows are used and the fused images are obtained by weighted average rule framed by these maps at each level. The final fused image is obtained from the synthesis of all the fused images. In [47], MSSS maps for detail layers that are decomposed by simple average filters is used for fusion. The fusion result obtained by this method is computationally simple and combines both the focused and unfocused regions of two sources images into a single image with more clarity. Later, the same author introduced a new saliency measure that measures the absolute difference between the mean image 'I_μ' and the median image 'I_{med}' called visual perception saliency described mathematically as:

$$S(i,j) = \left|I_\mu - I_{med}\right| \tag{15}$$

The saliency maps obtained by visual saliency are used to frame the weight maps and are further used for the fusion of infrared and visible image detail layers obtained by simple average decomposition in the method proposed in [48].

Pixel value based Saliency is calculated for every pixel in an image surrounded by the neighboring pixels in a local window. The value is obtained by taking the absolute difference between the pixel intensity that is considered for computing and the neighbor pixel intensities in a window. The sum of all these differences gives the saliency measure expressed mathematically as:

$$S_p = \sum_{\forall,q,\varepsilon,w} F(p,q) \tag{16}$$

Where,

$$F(p,q) = \left|f_p - f_q\right|$$

'f_p' , is the pixel intensity at which the saliency is to be determined, 'f_q' is the neighbor pixel intensity within a window 'w'. The pixel based saliency computed for constructing the weight maps in image fusion defines different weights to each pixel and hence improvement in the visual quality is observed. This is proved in the method proposed in [49] for the fusion of visible and infrared images using multi-scale decomposition. The saliency maps for the base layers decomposed by using edge-preserved decomposition filter were computed by using this pixel value based saliency in [50].

Mean Square Error (MSE) based Saliency and **Gradient based saliency** measures were introduced in [51] with an intention that the MSE describes the average deviation of the data from the mean intensity that detects the salient objects in an image. However, Gradient reflects the frequency variations and predicts the defined objects. The saliency calculation for these approaches is given as:

$$S_p = \sum_{\forall,q,\varepsilon,w}\sqrt{\frac{1}{|w|}\sum_{q,\varepsilon,w}\left(I_q - I_{pw}\right)} \tag{17}$$

$$S_p = \frac{1}{|w|}\sum_{\forall,q,\varepsilon,w}\sqrt{\frac{\left(\Delta_x I_p\right)^2 + \left(\Delta_y I_p\right)^2}{2}} \tag{18}$$

where $\Delta_x I_p$ and $\Delta_y I_p$ are the differences between the pixel intensity at that position and the pixel at its right along the row and column direction in a window 'w',I_{pw} is the mean value of the pixels located within the window.

However, **Spectral Residual Saliency (SRS)** is another saliency measure in literature [52] that extracts the image residual in spectral domain and then the saliency is measured. The mathematical steps required to compute SRS includes the computation of amplitude spectrum '$A(f)$', phase spectrum '$P(f)$', Logarithmic spectrum '$L(f)$' and the residual spectrum '$R(f)$' of an image 'I' expressed as:

$$A(f) = real(FFT(I)) \tag{19}$$

$$P(f) = Img(FFT(I)) \tag{20}$$

$$L(f) = log(A(f)) \tag{21}$$

$$R(f) = L(f) - h_n(f) * L(f) \tag{22}$$

Where

$h_f(f) = averagefilter(I)$

Hence, the saliency is computed from the residual spectrum as:

$$S(p) = g(p) * FFT^{-1}\left[(exp(R(f) + P(f)))^2\right] \tag{23}$$

where 'g_p' is the Gaussian filtered image.

The saliency maps of MSE, gradient and SRS approaches have been utilized for framing the select maximum rule in the fusion of differently focused source images. However, the combined SF and saliency feature maps have been used in [53] where the SF maps define weighted average rule for low-frequency components and select maximum rule for high-frequency components that are decomposed by NSCT. However, the saliency maps are used to detect the salient objects that are used as priority to combine and construct final fused image.

The selection of an appropriate saliency measure for feature extraction effects the fusion process as the parameter selected should appropriately extract the target regions as salient objects for combining. However, accurate salient object detection discriminating the background objects, framing an intelligent fusion rule were few challenges that are still in consideration by the researchers for further improvement in the performance. The simulation results of saliency based feature maps extracted is shown in Fig.3.

E. Structural approach for feature extraction

The human eye is highly sensitive to noise in edge and smooth regions rather than texture regions according to human visual system characteristics. Hence, it is necessary that the edge and texture information is preserved while processing an image. The process of edge detection reduces the amount of data to be processed and preserves the structural information of the source images. Hence, feature extraction by edge detection is a tool used by several researchers to construct feature maps and frame the fusion rules in order to highlight the edge characteristics and rich texture details. Thus, the structural approach to extract the features is to detect the edges. The basic edge detectors include **gradient** operators and **Canny's edge filters** etc. that have been used in image fusion to construct feature maps. In general, the gradient of an image is defined mathematically as:

$$\nabla f = |g_x| + \left|g_y\right| = |f(x,y) - f(x+1,y)| + |f(x,y) - f(x,y+1)| \tag{24}$$

where $f(x,y)$ is the pixel located at (x,y). Based on the general definition, in [54], a local gradient within a local window of size 'w' is defined as:

$$G(p) = \sum_{q,\epsilon,w} |f(q) - f(p)|^2 \tag{25}$$

This local gradient is used as an activity measure to define the weight maps and the combined weighted average and select maximum rules were utilized for combining the visible and infrared images. The method not only highlights the features but also enriches the edge and strength details of the fused image. Considering that the larger value of gradient results in blurring of the image details, In [55], a new cross gradient feature map extraction by computing the gradient between the pixels in the low-frequency component '$f_{lf}(x,y)$' and the adjacent pixel of the band-pass coefficient '$f_{bp}(x,y)$' decomposed by NSCT. The cross gradient of a region is calculated as:

$$CG(x,y) = \frac{1}{M*N} \sum_{i=1}^{M}\sum_{j=1}^{N} \sqrt{\frac{\left[g_x(x+i,y+j)\right]^2 + \left[g_{xy}(x+i,y+j)\right]^2}{2}} \tag{26}$$

where

$$g_x(x,y) = \left|f_{lf(x+1,y)-f_{bp}(x,y)}\right| \tag{27}$$

$$g_y(x,y) = \left|f_{lf(x,y+1)-f_{bp}(x,y)}\right| \tag{28}$$

The smaller values of cross gradient represent the sharper visual details clearly and hence the minimum cross gradient obtained is considered for select maximum rule for the fusion of sub band images. In [56], the source images are decomposed to base and detail layers by an edge preserved multi-scale decomposition filters and the edge characteristics of the detail layers are extracted by **Sum-Modified Laplacian (SML)** edge detector to combine the details layers using select maximum rule. The SML in a local window of size M X N centered by pixel f(x, y) is given by:

$$SML(X,Y) = \sum_{i=1}^{M}\sum_{j=1}^{N} M(i,j) \tag{29}$$

where

$$M(i,j) = |2*f(x,y) - f(x-1,y) - f(x+1,y)| + |2*f(x,y) - f(x,y-1) - f(x,y+1)|$$

In order to maintain the structural information of the object boundaries, in [57] a new multi-focus fusion technique in which the edge features of the wavelet coefficients are extracted using Canny filters at each level and a feature-oriented level fusion is performed.

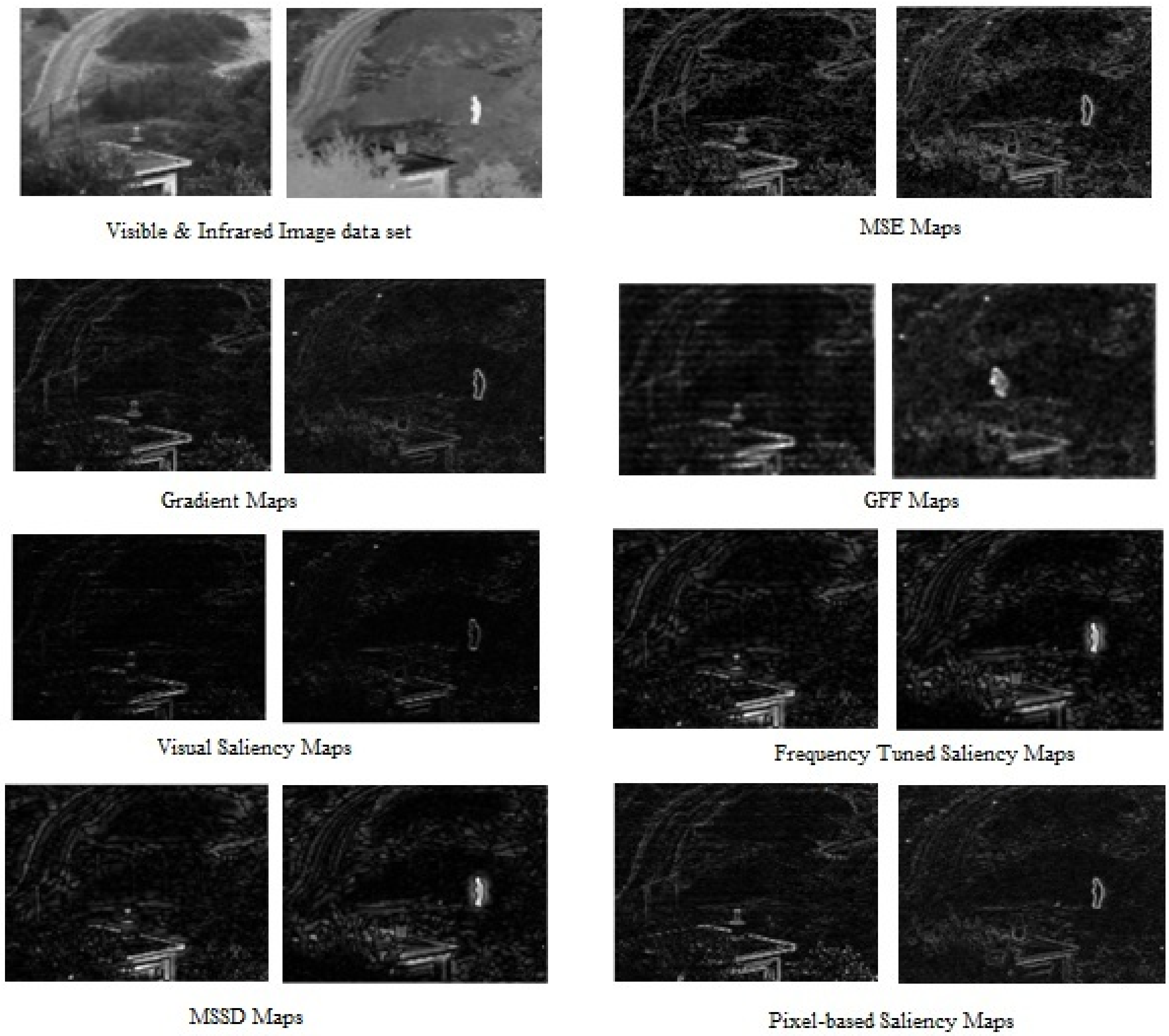

Figure 3. Feature Maps using Saliency Approach

However, the basic edge detectors has a shortfall of detecting the false positives i.e. the points detected as edges may not be perceived as edges by the human vision and these are sensitive to illumination variations. Hence, the structural features are defined by the energy functions because any local maximum energy can be defined as a feature [58]. The **energy functions** are used as feature extractors in image fusion to form feature maps and further to construct weight maps. In general, the energy of an image is computed as:

$$E = \sum_{x,y} f^2(x, y) \tag{30}$$

Fusion algorithms based on energy based feature extraction for combining the high frequency coefficients of DWT using *select maximum* rule is proposed in [59], low-frequency components of DWT using *select maximum* rule in [60], low-frequency coefficients of Uniform Discrete Curvelet Transform using weighted average rule in [61] and the edge maps obtained for high frequency coefficients in NSCT domain in [62]. In [63] the author proposed night vision fusion that utilized the energy maps to frame a *select maximum* rule to high frequency coefficients obtained by Framelet transform.

Another structural feature extraction measure considered to construct weight maps is **Phase Congurency(PC)** . The detection and localization of edges is obtained at the points of maximum energy called Phase Congruency. An edge is said to be detected at a point at which the phase of all the components becomes zero. The significant edge has a PC value of ‘1’ and for insignificant edges its value is ‘0’ [64]. The phase congruency can distinguish different kinds of features when compared to the features generated by edge operators such as Laplacian operator. Hence, the PC maps are been used in image fusion process to frame the fusion rules. In [65] a visible and infrared image fusion technique is proposed that incorporates the PC maps to define the weighted average rule to multi-scale decomposed image coefficients. However, the PC cannot reflect the changes of the local luminance. It is contrast invariant. Hence, the textured regions of an image are analyzed by the local analysis of the luminance decrease measured by **local contrast**. Local contrast of an image is highly correlated to the image gradient. The local contrast of an image is defined as the ratio of local high frequency luminance to the local low frequency luminance. However, the definition of local contrast gives an inherent correlation between the low and high frequency. Hence a modified contrast called **directive contrast** was introduced in literature [66].

The structural features extracted by these edge, energy and contrast based functions convey different characteristics of image information. Hence, the collective use of these functions was employed in literature as a novel activity measure to get the visual characteristics. In [67], a multi-modal medical image fusion is proposed by extracting the characteristics of low frequency coefficients using the regional energy and the high frequency characteristics are extracted by using the gradient operator in DWT domain to construct weighted average fusion rule. In [68], multi focus image fusion using energy and gradient feature maps is proposed. Local energy, local contrast and gradient are the feature extractors used to acquire the feature maps of low frequency coefficients and cross contrast to band pass frequency components. The feature maps are used to define the select maximum rule and combine the information from the complementary images in NSCT domain proposed in [69]. PC and directive contrast measures are used to define two different fusion rules to combine MRI and CT images in NSCT domain in [70]. In [71] a novel activity by using PC is proposed where local contrast and energy feature extractors define the fusion rule for IR and visible fusion. The simulation results of generated structural maps using the mathematical formulae defined above are shown in Fig. 4.

However, the feature maps are also constructed by the combined statistical, saliency and structural approaches to improve the fusion performance. The combination of mean, standard deviation, entropy and gradient were used in [72], entropy, visibility and local contrast measures in [73], local mean and local energy for the fusion of low and high frequency coefficients in Discrete wavelet packet transform domain in [74], PC and FD combination is used to define the fusion rule for multi-sensor fusion for night vision applications by [75] and energy, variance combination in [76] in DCT domain for multi-sensor fusion. Geometric structural features such as attribute angle, length, distance etc. in combination to statistical features are the recent features used by the authors in machine learning algorithms [77].

Optimal fusion result by the feature extraction based fusion algorithms in literature rely on the domain selected, type of activity measure used and the fusion rule defined. Selection of the statistical, saliency or structural approach to define the fusion rule attains the fusion requirements. However, the selection of window size, activity measure, appropriate filter, fusion rule and computational complexities were few measures that limit the performance of the algorithm. Hence, an activity measure for feature extraction is still a continuous research task to attain the goal of image fusion.

3. EXPERIMENTAL ANALYSIS

Simulation results of few statistical, saliency and structural based fusion algorithms available in literature are verified in MATLAB. In this section, 10 different approaches PCA [76], DWT [78], DWT+PCA [29] , DWT+LS [73], AI+KL [47], FD [24], representing statistical approaches, WPT+DC [79] , GFF [41], NDF [49] representing saliency approaches and DWT+E [59] is the structural approach for image fusion that define the fusion rules considered for evaluation. The qualitative and quantitative analysis for three different data set images such as Multi-Sensor (MS), Multi-Focus (MF) and Multi-Modality (MM) fusion examples $http : //home.ustc.edu.cn/liuvu1/$ are evaluated and are shown in Fig. 5 and Fig. 6 respectively.

Qualitative and quantitative analysis of fusion algorithm describes the method efficacy and also the accomplishment of the fusion objectives The qualitative analysis in Fig. 5 shows that the statistical approach for feature extraction and fusion rule construction performs better fusion in relation to the contrast and complete information. However, the performance is better if the feature extraction is performed in multi-scale transform or decomposition domain. It is also observed that the feature-extraction based fusion process is best suited for multi-sensor fusion used for target detection and scene interpretation where the images obtained from visible and infrared sensors are combined to a single image. Qualitative analysis is subjective that depends on human perception on the fusion performance and hence it is not the only metric for fusion evaluation. A number of quantitative metrics were available in literature to evaluate the fusion scheme. The quantitative metrics used for the evaluation of the algorithms in this paper are categorized as:

- Contrast metrics
- Information metrics
- Edge metrics
- Visual quality metrics.

Contrast metrics that describe the overall contrast of the fused image are Mean Pixel Intensity (MPI) and Standard Deviation (SD) given by:

$$MPI = \frac{1}{mn} \sum_{i=1}^{M} \sum_{j=1}^{N} I(i, j) \tag{31}$$

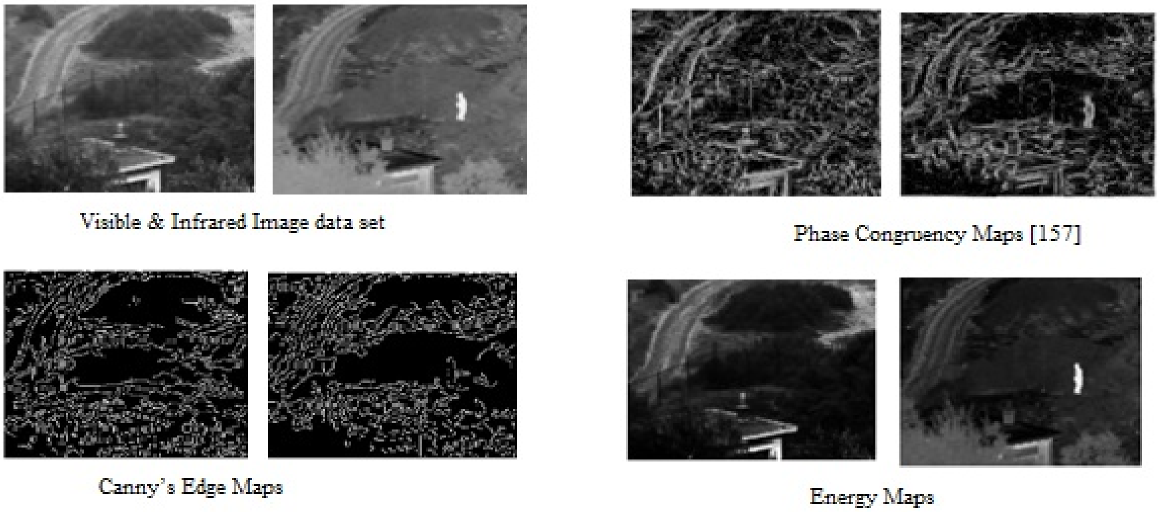

Figure 4. Feature Maps using Structural Approach

TABLE I. Table 1: Comparative analysis of contrast and Information metrics

Method	Domain	Features	Contrast Metrics						Information Metrics						Highlights
			MPI			SD			E			QMI			
			MS	MF	MM	MS	MF	MM	MS	MF	MM	MS	MF	MM	
PCA	Spatial	Statistical	78.87	123.00	49.98	**48.90**	45.45	51.65	**7.45**	6.97	5.95	0.45	0.47	0.79	Preserves mean contrast and details of MS fused image
WPT+DC	MST	Structural	90.15	120.84	51.01	32.21	48.39	53.92	6.92	7.11	**6.87**	0.42	0.42	0.65	Edge details are enhanced
GFF	MSD	Saliency	94.81	124.68	54.20	29.27	**51.40**	56.30	6.77	**7.23**	6.22	0.44	0.48	0.80	Preserves mean contrast and details of MF fused image
FD	Spatial	Statistical	89.29	124.00	12.99	40.83	46.82	43.00	7.18	6.95	1.89	**0.50**	**0.50**	**0.87**	Information amount is high for all the image sets
DWT+E	MST	Structural	**254.06**	**254.35**	25.26	15.49	12.88	**76.17**	0.04	0.03	0.47	0.11	0.15	0.84	High average brightness and contrast

TABLE II. Comparative analysis of visual and edge metrics

Method	Domain	Features	Visual Metrics						Edge Metrics									Highlights
			SF			VIF			AG			FS			Q0			
			MS	MF	MM	MS	MF	MM	MS	MF	MM	MS	MF	MM	MS	MF	MM	
PCA	Spatial	Statistical	12.782	7.742	12.695	0.211	0.730	0.287	6.066	2.728	5.181	0.022	0.002	0.128	0.390	**0.997**	**0.729**	Good edge similarity
DWT+PCA	MST	Statistical	7.321	8.122	10.072	0.321	0.737	0.228	3.898	2.784	3.409	0.020	0.006	0.120	**0.511**	0.996	0.643	Less structural Distortion
DWT+LS	MST	Statistical	12.785	13.958	13.002	0.233	0.491	0.245	5.316	4.251	5.404	0.014	0.005	**0.076**	0.473	0.926	0.678	Good edge symmetry
WPT+DC	MST	Structural	15.490	15.014	18.202	0.107	0.427	0.190	7.107	**5.328**	7.708	0.017	0.004	0.079	0.216	0.828	0.614	Edge clarity & sharpness
GFF	MSD	Saliency	12.797	13.252	18.749	0.235	**0.926**	0.444	6.531	4.736	7.539	0.002	0.014	0.124	0.394	0.961	0.661	Preserves edge symmetry
NDF	MSD	Saliency	**17.482**	**15.433**	23.463	**0.428**	0.804	**0.562**	**9.161**	5.539	9.040	0.020	0.009	0.122	0.473	0.948	0.604	Better visual perception, edge clarity and better activity level.
AI+KL	MSD	Statistical	8.035	7.902	9.763	0.317	0.730	0.331	4.398	2.814	3.926	0.020	0.001	0.110	0.509	0.996	0.728	Preserves edge symmetry
DWT+E	MST	Structural	15.519	14.502	**46.615**	0.000	0.003	0.070	0.666	0.574	5.792	0.203	0.066	0.158	-0.008	-0.013	0.185	Good activity level in medical images

$$SD = \sqrt{\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [I(i,j) - MPI]^2}{MN}} \tag{32}$$

where MN is the size of the image and $I(i,j)$ is the intensity value of image I at location(i,j). Higher the values indicate better is the contrast.

Information metrics that describes the maximum amount of information transfer to the output image from the two input images are defined by Entropy (E) and Normalised Mutual Information (QMI) metrics defined as:

$$E = \sum_{i=1}^{M} \sum_{j=1}^{N} p\left(I_{i,j}\right) lnp\left(I_{i,j}\right) \tag{33}$$

where $p\left(I_{i,j}\right)$ is the probability of occurrence of the pixel at (i,j). The fused image quality 'F' in accordance to the input images say 'A' and 'B' can be determined by QMI given as:

$$QMI = 2\left[\frac{MI_{AF}}{E_A + E_F} + \frac{MI_{BF}}{E_B + E_F}\right] \tag{34}$$

where M_{AF} and M_{AF} is the mutual information shared between the source images A, B and fused image F expressed as:

$$MI_{AF} = \sum_{a,b} p_{AF}(a,b) \, log \frac{p_{AF}(a,b)}{p_A(a,b) \, p_F(a,b)} \tag{35}$$

$$MI_{BF} = \sum_{a,b} p_{BF}(a,b) \, log \frac{p_{BF}(a,b)}{p_B(a,b) p_F(a,b)}$$

where p_{AF} and p_{BF} are the joint probabilities and p_A, p_B and p_F are the marginal probabilities, E_A, E_B and E_F are the entropies.

Edge metrics describe the edge preserved details, edge symmetry and structural similarity of the final image. The edge metrics considered for evaluation are Average Gradient

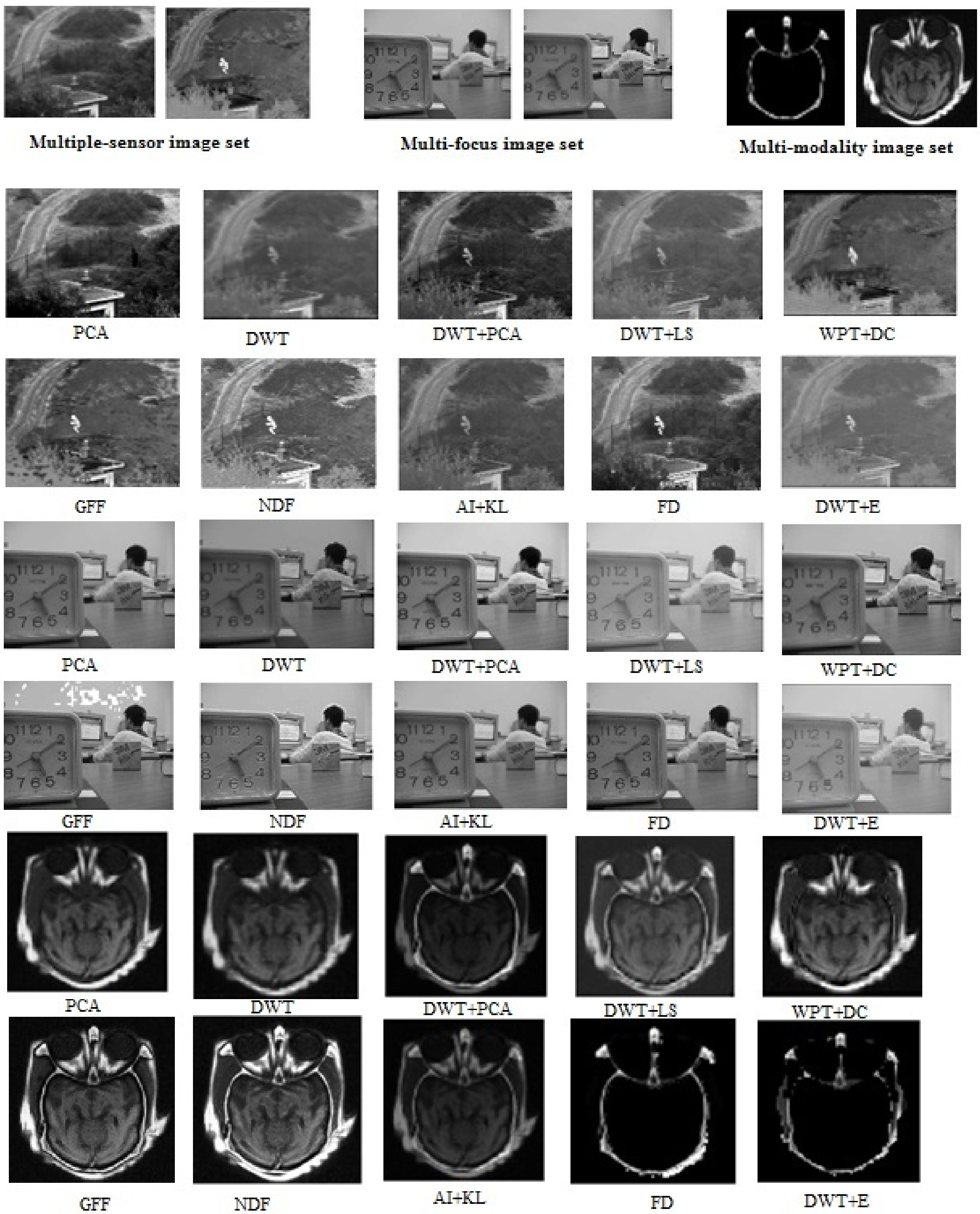

Figure 5. Illustration of statistical, saliency and structural feature extraction based fusion algorithms considered for evaluation

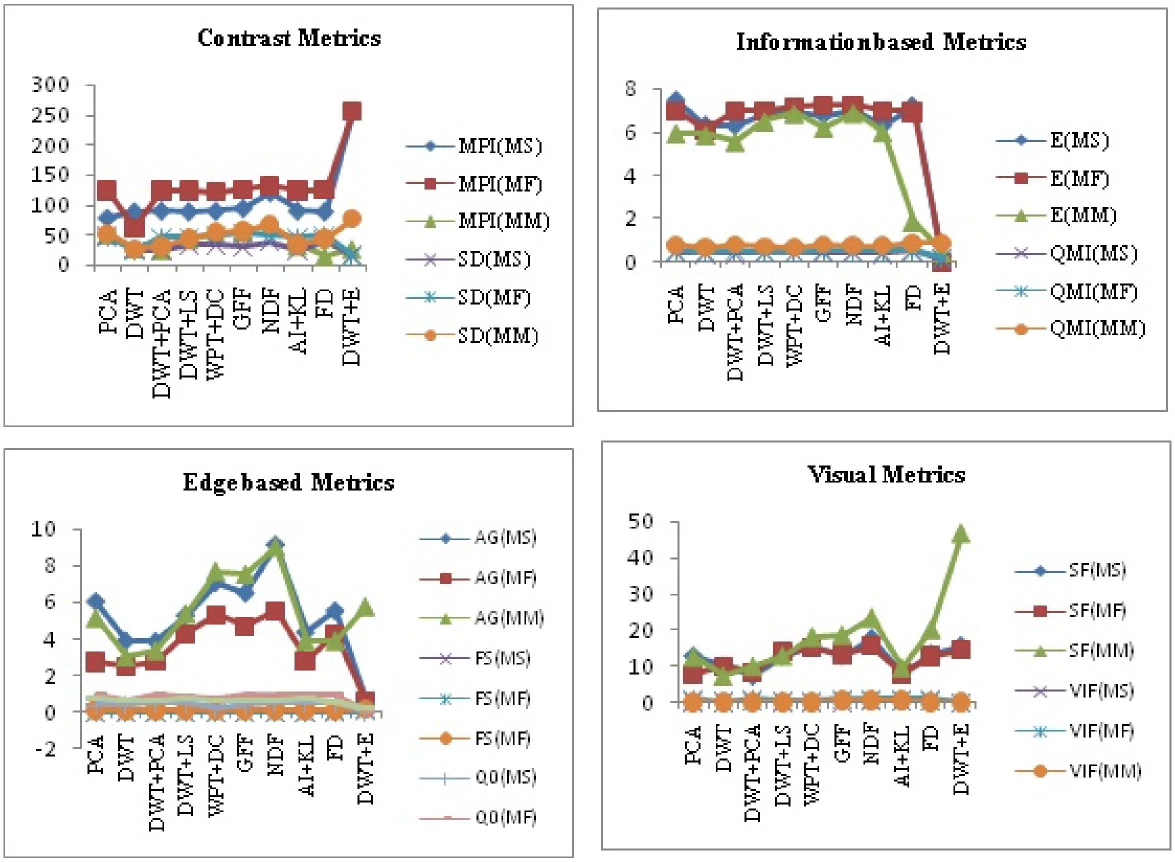

Figure 6. Quantitative fusion metrics for methods considered for evaluation

(AG) and Fusion Symmetry (FS). Edge metrics describe the clarity and sharpness. AG, Symmetry with respect to input images is measured by FS, structural distortion is quantified by Q_0 defined as:

$$AG = \sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{\frac{[I(i+1,j)-I(i,j)]^2+[I(i,j+1)-I(i,j)]^2}{2}} \tag{36}$$

$$FS = abs\left(\frac{MI_{AF}}{MI_{AF}+MI_{BF}}-0.5\right) \tag{37}$$

Visual quality metrics such as Spatial Frequency (SF) and Visual Information Fidelity (VIF) describe the visual quality that leads to perfect detection and interpretation of the scene captured $http://hansy.weebly.com/image-fusion-metric-ifm.html$. The global activity level of the final image is described by SF as given in Eq. (9). The higher the values the better is the visual quality.

The quantitative metrics are evaluated and from the graphical representation of these metrics shown in Fig. 6, the comparison is performed. A detail comparative analysis for contrast, information, visual and edge metrics to the image sets considered for evaluation is presented in Table1 and Table 2 respectively. The highlights of each technique are elaborated. The contrast based qualitative metrics (MPI and SD) is moderate and are high using statistical and saliency approaches.

Though the values of these metrics are high for structural approach, it indicates over saturation that leads to misinterpretation. The information based metrics (E and QMI) are almost high for statistical and saliency based feature- extraction approach. These metrics indicate the complete information in the amassed image. The edge based metrics (AG and F) values are observed to be better for saliency based approach that defines the edge symmetry and structural preservation. The visual metrics (SF and VIF) are better with all the three approaches that show high visual quality. Hence, the analysis shows that selection of approach, selection of parameter, selection of fusion rule defines the fusion results.

4. CONCLUSIONS AND FUTURE DIRECTIONS

This paper discusses the significance of feature extraction for image fusion. A novel classification of feature extraction based fusion algorithms build on the review of feature maps constructed for defining the fusion rules is presented. The mathematical description of the defined statistical, saliency and structural feature measures

are briefly discussed. The feature- extraction based fusion rules framed avoids the blurring and contrast loss effects compared to traditional mathematical fusion rules. The concluding remarks of fusion methods in literature are outlined that describes the suitability of parameters for a specific fusion application. The quantitative metrics considered for evaluation in the paper define the success of fusion algorithm in terms of contrast, edge symmetry, no loss and no artefacts effects.
The sensor sensitivity to low lighting conditions and low dynamic range sensors reduces the image quality and has poor contrast, limited dynamic range and also several other reported problems. The image captured by such low-light sensors reduces the overall contrast of the image in the fusion process that leads to misclassification or misinterpretation. However, few techniques that perform the enhancement process after fusion are in literature.
But, the computational complexity of the entire process is one of the major limitations. Therefore, it is necessary to develop joint Image Enhancement and fusion algorithms that improve the robustness of the fusion process.
Robustness of the fusion process can be further improved by considering the effect of various noises that degrade the performance of the sensor during acquisition. Image de-noising and fusion can be explored in future. However, hybrid domain methods can be implemented in future to overcome the limitations of several existing fusion methods. Research on new statistical, saliency and structural parameters for feature extraction can be extended for formulating efficient fusion rules. Selection of feature parameters and fusion rules in future should be developed such that an informative fused image is produced without redundancy especially for specific applications. The increase in multi-sensor technology for industry automation provides a large data set images which requires a large feature set. Deep learning fusion methods can be pursued in future for better feature extraction, classification and fusion.

References

[1] S. Almabdy and L. Elrefaei, “Feature extraction and fusion for face recognition systems using pre-trained convolutional neural networks,” *International Journal of Computing and Digital Systems*, vol. 9, pp. 1–7, 2021.

[2] P. R. Hill, C. N. Canagarajah, and D. R. Bull, “Image fusion using complex wavelets.” in *BMVC*, 2002, pp. 1–10.

[3] K. He, D. Zhou, X. Zhang, R. Nie, Q. Wang, and X. Jin, “Infrared and visible image fusion based on target extraction in the nonsubsampled contourlet transform domain,” *Journal of Applied Remote Sensing*, vol. 11, no. 1, p. 015011, 2017.

[4] X. Yan, H. Qin, J. Li, H. Zhou, and J.-g. Zong, “Infrared and visible image fusion with spectral graph wavelet transform,” *JOSA A*, vol. 32, no. 9, pp. 1643–1652, 2015.

[5] P. Shah, S. Merchant, and U. B. Desai, “Fusion of surveillance images in infrared and visible band using curvelet, wavelet and wavelet packet transform,” *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 8, no. 02, pp. 271–292, 2010.

[6] S. Li, X. Kang, and J. Hu, “Image fusion with guided filtering,” *IEEE Transactions on Image processing*, vol. 22, no. 7, pp. 2864–2875, 2013.

[7] V. Chaudhary and V. Kumar, “Block-based image fusion using multi-scale analysis to enhance depth of field and dynamic range,” *Signal, Image and Video Processing*, vol. 12, no. 2, pp. 271–279, 2018.

[8] Y. Jiang and M. Wang, “Image fusion using multiscale edge-preserving decomposition based on weighted least squares filter,” *IET image Processing*, vol. 8, no. 3, pp. 183–190, 2014.

[9] H. Li, H. Qiu, Z. Yu, and Y. Zhang, “Infrared and visible image fusion scheme based on nsct and low-level visual features,” *Infrared Physics & Technology*, vol. 76, pp. 174–184, 2016.

[10] C. Liu, Y. Qi, and W. Ding, “Infrared and visible image fusion method based on saliency detection in sparse domain,” *Infrared Physics & Technology*, vol. 83, pp. 94–102, 2017.

[11] G. Piella, “A region-based multiresolution image fusion algorithm,” in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002.(IEEE Cat. No. 02EX5997)*, vol. 2. IEEE, 2002, pp. 1557–1564.

[12] S. Arivazhagan, L. Ganesan, and T. S. Kumar, “A modified statistical approach for image fusion using wavelet transform,” *Signal, image and video processing*, vol. 3, no. 2, pp. 137–144, 2009.

[13] X. Li and S.-Y. Qin, “Efficient fusion for infrared and visible images based on compressive sensing principle,” *IET Image Processing*, vol. 5, no. 2, pp. 141–147, 2011.

[14] J. Joseph and A. Barhatte, “Medical image fusion based on wavelet transform and fast curvelet transform,” *International Journal of Engineering Development and Research*, vol. 2, no. 1, pp. 284–8, 2014.

[15] G. Bhatnagar and Z. Liu, “Multi-sensor fusion based on local activity measure,” *IEEE Sensors Journal*, vol. 17, no. 22, pp. 7487–7496, 2017.

[16] T. Shibata, M. Tanaka, and M. Okutomi, “Versatile visible and near-infrared image fusion based on high visibility area selection,” *Journal of Electronic Imaging*, vol. 25, no. 1, p. 013016, 2016.

[17] J. Zhao, X. Gao, Y. Chen, H. Feng, and D. Wang, “Multi-window visual saliency extraction for fusion of visible and infrared images,” *Infrared Physics & Technology*, vol. 76, pp. 295–302, 2016.

[18] J. Lewis, R. O’callaghan, S. Nikolov, D. Bull, and C. Canagarajah, “Region-based image fusion using complex wavelets,” in *Proc. 7th International Conference on Information Fusion*, vol. 1. Citeseer, 2004, pp. 555–562.

[19] A. Jameel, A. Ghafoor, and M. M. Riaz, “Adaptive compressive fusion for visible/ir sensors,” *IEEE Sensors Journal*, vol. 14, no. 7, pp. 2230–2231, 2014.

[20] W. Zhiming and B. Hong, “Entropy based multi-resolution visible-infrared image fusion,” in *2012 IEEE Symposium on Electrical & Electronics Engineering (EEESYM)*. IEEE, 2012, pp. 639–643.

[21] H. Xu, Y. Wang, Y. Wu, and Y. Qian, "Infrared and multi-type images fusion algorithm based on contrast pyramid transform," *Infrared Physics & Technology*, vol. 78, pp. 133–146, 2016.

[22] W. Yin, W. Zhao, D. You, and D. Wang, "Local binary pattern metric-based multi-focus image fusion," *Optics & Laser Technology*, vol. 110, pp. 62–68, 2019.

[23] K. D. Toennies and J. A. Schnabel, "Edge detection using the local fractal dimension," in *Proceedings of IEEE Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 1994, pp. 34–39.

[24] G. Bhatnagar and Q. J. Wu, "A fractal dimension based framework for night vision fusion," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 1, pp. 220–227, 2018.

[25] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[26] A. Hyvärinen, J. Hurri, and P. O. Hoyer, *Natural image statistics: A probabilistic approach to early computational vision.* Springer Science & Business Media, 2009, vol. 39.

[27] M. Kumar and S. Dass, "A total variation-based algorithm for pixel-level image fusion," *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 2137–2143, 2009.

[28] R. Sardana, M. Dutta, and S. S. Saini, "Design of hybrid technique based on rgb contents and entropy of the image and its comparative analysis with existing technique," *International Journal of Computer Applications*, vol. 975, p. 8887, 2015.

[29] R. Vijayarajan and S. Muttan, "Discrete wavelet transform based principal component averaging fusion for medical images," *AEU-International Journal of Electronics and Communications*, vol. 69, no. 6, pp. 896–902, 2015.

[30] D. P. Bavirisetti and R. Dhuli, "Fusion of infrared and visible sensor images based on anisotropic diffusion and karhunen-loeve transform," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 203–209, 2015.

[31] S. Yin and Y. Zhang, "Singular value decomposition-based anisotropic diffusion for fusion of infrared and visible images," *International Journal of Image and data fusion*, vol. 10, no. 2, pp. 146–163, 2019.

[32] D. P. Bavirisetti, G. Xiao, and G. Liu, "Multi-sensor image fusion based on fourth order partial differential equations," in *2017 20th International conference on information fusion (Fusion)*. IEEE, 2017, pp. 1–9.

[33] L. Oudre, "Study of two new spatial domain image fusion methods," Ph.D. dissertation, Department of Electrical and Electronic Engineering, Imperial College London, 2007.

[34] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[35] N. Joseph, D. Murugan, B. J. Thomas *et al.*, "Patient specific brain tumor segmentation using context sensitive feature extraction in mr images," *International Journal of Computing and Digital Systems*, vol. 9, no. 6, pp. 1091–1097, 2020.

[36] R. Narayanaswamy and S. G. Sadasivam, "Multimodal prediction of breast cancer using radiogenomics and clinical trails with decision fusion," *International Journal of Computing and Digital System*, 2021.

[37] J. J. Shanthamalar and R. G. Ramani, "A novel approach for glaucoma disease identification through optic nerve head feature extraction and random tree classification," *International Journal of Computing and Digital Systems*, vol. 10, 2021.

[38] S. Li, J. T. Kwok, and Y. Wang, "Combination of images with diverse focuses using the spatial frequency," *Information fusion*, vol. 2, no. 3, pp. 169–176, 2001.

[39] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang, and Y. Zhang, "Multi-focus image fusion based on spatial frequency in discrete cosine transform domain," *IEEE signal processing letters*, vol. 22, no. 2, pp. 220–224, 2014.

[40] Q. Jiang, X. Jin, S.-J. Lee, and S. Yao, "A novel multi-focus image fusion method based on stationary wavelet transform and local features of fuzzy sets," *IEEE Access*, vol. 5, pp. 20 286–20 302, 2017.

[41] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image processing*, vol. 22, no. 7, pp. 2864–2875, 2013.

[42] R. Achanta and S. Süsstrunk, "Saliency detection using maximum symmetric surround," in *2010 IEEE International Conference on Image Processing*. IEEE, 2010, pp. 2653–2656.

[43] G. Cui, H. Feng, Z. Xu, Q. Li, and Y. Chen, "Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition," *Optics Communications*, vol. 341, pp. 199–209, 2015.

[44] J. Zhu, W. Jin, L. Li, Z. Han, and X. Wang, "Fusion of the low-light-level visible and infrared images for night-vision context enhancement," *Chinese Optics Letters*, vol. 16, no. 1, p. 013501, 2018.

[45] J. Zhao, G. Cui, X. Gong, Y. Zang, S. Tao, and D. Wang, "Fusion of visible and infrared images using global entropy and gradient constrained regularization," *Infrared Physics & Technology*, vol. 81, pp. 201–209, 2017.

[46] D. P. Bavirisetti and R. Dhuli, "Multi-focus image fusion using maximum symmetric surround saliency detection," *ELCVIA: electronic letters on computer vision and image analysis*, vol. 14, no. 2, pp. 58–73, 2015.

[47] D. Bavirisetti and R. Dhuli, "Two-scale image fusion of visible and infrared images using saliency detection," *Infrared Physics & Technology*, vol. 76, pp. 52–64, 2016.

[48] J. Zhao, Q. Zhou, Y. Chen, H. Feng, Z. Xu, and Q. Li, "Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition," *Infrared physics & technology*, vol. 56, pp. 93–99, 2013.

[49] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8–17, 2017.

[50] Z. Liu, X. fei Li, and X. Zhang, "A real time multi-focus image fu-

sion framework ⋆," *The Journal of Information and Computational Science*, vol. 11, pp. 1295–1302, 2014.

[51] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.

[52] F. Meng, M. Song, B. Guo, R. Shi, and D. Shan, "Image fusion based on object region detection and non-subsampled contourlet transform," *Computers & Electrical Engineering*, vol. 62, pp. 375–383, 2017.

[53] J. Mou, W. Gao, and Z. Song, "Image fusion based on non-negative matrix factorization and infrared feature extraction," in *2013 6th International Congress on Image and Signal Processing (CISP)*, vol. 2. IEEE, 2013, pp. 1046–1050.

[54] J. Adu, S. Xie, and J. Gan, "Image fusion based on visual salient features and the cross-contrast," *Journal of Visual Communication and Image Representation*, vol. 40, pp. 218–224, 2016.

[55] Y. Jiang and M. Wang, "Image fusion using multiscale edge-preserving decomposition based on weighted least squares filter," *IET image Processing*, vol. 8, no. 3, pp. 183–190, 2014.

[56] R. Nava, B. Escalante-Ramírez, and G. Cristóbal, "A novel multi-focus image fusion algorithm based on feature extraction and wavelets," in *Optical and Digital Image Processing*, vol. 7000. International Society for Optics and Photonics, 2008, p. 700028.

[57] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern recognition letters*, vol. 6, no. 5, pp. 303–313, 1987.

[58] Y. Yang, "A novel dwt based multi-focus image fusion method," *Procedia engineering*, vol. 24, pp. 177–181, 2011.

[59] L. Zhan, Y. Zhuang, and L. Huang, "Infrared and visible images fusion method based on discrete wavelet transform," *J. Comput*, vol. 28, no. 2, pp. 57–71, 2017.

[60] Q. Li, J. Du, and L. Xu, "Visible and infrared video fusion using uniform discrete curvelet transform and spatial-temporal information," *Chinese Journal of Electronics*, vol. 24, no. 4, pp. 761–766, 2015.

[61] L. Kun, G. Lei, L. Huihui, and C. Jingsong, "Fusion of infrared and visible light images based on region segmentation," *Chinese Journal of Aeronautics*, vol. 22, no. 1, pp. 75–80, 2009.

[62] G. Bhatnagar, Q. J. Wu, and B. Raman, "Navigation and surveillance using night vision and image fusion," in *2011 IEEE Symposium on Industrial Electronics and Applications*. IEEE, 2011, pp. 342–347.

[63] P. Kovesi *et al.*, "Image features from phase congruency," *Videre: Journal of computer vision research*, vol. 1, no. 3, pp. 1–26, 1999.

[64] W. Gan, X. Wu, W. Wu, X. Yang, C. Ren, X. He, and K. Liu, "Infrared and visible image fusion with the use of multi-scale edge-preserving decomposition and guided image filter," *Infrared Physics & Technology*, vol. 72, pp. 37–51, 2015.

[65] A. Beghdadi and A. Le Negrate, "Contrast enhancement technique based on local detection of edges," *Computer vision, graphics, and image processing*, vol. 46, no. 2, pp. 162–174, 1989.

[66] G. Bhatnagar, Q. J. Wu, and Z. Liu, "A new contrast based multimodal medical image fusion framework," *Neurocomputing*, vol. 157, pp. 143–152, 2015.

[67] X. Zhang, Y. Zheng, Y. Peng, W. Liu, and C. Yang, "Research on multi-mode medical image fusion algorithm based on wavelet transform and the edge characteristics of images," in *2009 2nd International Congress on Image and Signal Processing*. IEEE, 2009, pp. 1–4.

[68] Y. Zhang, L. Chen, Z. Zhao, and J. Jia, "Multi-focus image fusion based on image decomposition and quad tree decomposition," *Journal of Computers*, vol. 25, no. 3, 2014.

[69] J. Adu, J. Gan, Y. Wang, and J. Huang, "Image fusion based on nonsubsampled contourlet transform for infrared and visible light image," *Infrared Physics & Technology*, vol. 61, pp. 94–100, 2013.

[70] G. Bhatnagar, Q. J. Wu, and Z. Liu, "Directive contrast based multimodal medical image fusion in nsct domain," *IEEE transactions on multimedia*, vol. 15, no. 5, pp. 1014–1024, 2013.

[71] H. Li, H. Qiu, Z. Yu, and Y. Zhang, "Infrared and visible image fusion scheme based on nsct and low-level visual features," *Infrared Physics & Technology*, vol. 76, pp. 174–184, 2016.

[72] L.-j. Wang, J. Han, Y. Zhang, and L.-f. Bai, "Image fusion via feature residual and statistical matching," *IET Computer Vision*, vol. 10, no. 6, pp. 551–558, 2016.

[73] A. V. Vanmali and V. M. Gadre, "Visible and nir image fusion using weight-map-guided laplacian–gaussian pyramid for improving scene visibility," *Sādhanā*, vol. 42, no. 7, pp. 1063–1082, 2017.

[74] P. Shah, S. N. Merchant, and U. B. Desai, "An efficient adaptive fusion scheme for multifocus images in wavelet domain using statistical properties of neighborhood," in *14th International Conference on Information Fusion*. IEEE, 2011, pp. 1–7.

[75] G. Bhatnagar and Z. Liu, "A novel image fusion framework for night-vision navigation and surveillance," *Signal, Image and Video Processing*, vol. 9, no. 1, pp. 165–175, 2015.

[76] V. Naidu, "Discrete cosine transform based image fusion techniques," *Journal of Communication, Navigation and Signal Processing*, vol. 1, no. 1, pp. 35–45, 2012.

[77] I. Ariawan, Y. Herdiyeni, and I. Z. Siregar, "Geometry feature extraction of shorea leaf venation based on digital image and classification using random forest," *International Journal of Computing and Digital System*, 2021.

[78] H. Li, B. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical models and image processing*, vol. 57, no. 3, pp. 235–245, 1995.

[79] R. Balasubramanian and G. Bhatnagar, "A new contrast based image fusion using wavelet packets," *arXiv preprint arXiv:0812.0759*, 2008.

T. Sandhya Kumari Mrs. T. Sandhya Kumari is an Associate Professor in the Department of ECE, Vignan's Institute of Engineering for Women, Visakhapatnam. She did her B.E. degree from S.R.K.R. Engineering College, Bhimavaram, affiliated to Andhra University and M. Tech degree from JNT University, Kakinada. She is presently pursuing Ph.D from JNT ,University, Kakinada. She has published two papers in International Journals and in National Conferences. Her areas of interest are Multi Modal Image Fusion, Image Enhancement and Image Restoration.

Dr. S. Koteswara Rao Dr. S. Koteswara Rao was awarded Ph. D in Digital Statistical Signal Processing from Andhra University in the year 2010. He worked at NSTL holding several positions from Junior Scientific Officer to Scientist-G throughout his service. He has carried out several research projects related to DRDO and developed signal processing algorithms for defense projects. He has 140 research publications to his credit and has authored one International Text Book. He is a Fellow member of IETE and was a reviewer to reputed International Journals. Currently, he is a Professor of ECE Department at K.L. E. F. Deemed to be University, Guntur District. He is guiding several Ph. D scholars in the fields of Statistical Signal Processing, Adaptive Signal Processing, Image Processing and Bio-Medical Signal Processing.

Dr I. Santi Prabha Dr I. Santi Prabha is a Professor of ECE Department. She was awarded with Ph.D. in Speech signal processing by JNT University, Kakinada in 2005. She has 30 years experience of teaching, has produced several Ph.D. and is guiding 5 Ph. D scholars. She is a member of Indian Society for Technical Education, Fellow member of IETE and Fellow in Institution of Engineers (India). She published more than 70 technical papers in National and International journals and conferences. Her areas of interest include Speech Signal Processing and Image Processing.