FUSION BASED IMAGE QUALITY ENHANCEMENT

THESIS

SUBMITTED TO

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

LUCKNOW

BABASAHEB
BHIMRAO
AMBEDKAR
UNIVERSITY

LUCKNOW

प्रज्ञा शील करुणा

ESTABLISHED 1996

FOR THE DEGREE OF

Doctor of Philosophy

IN

COMPUTER SCIENCE

Submitted by

Sonam

(Enrolment No: 229/09)

Supervisor

Dr. Manoj Kumar

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE

SCHOOL FOR INFORMATION SCIENCE & TECHNOLOGY

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(A Central University; NAAC-'A' GRADE)

VIDYA VIHAR, RAE BARELI ROAD, LUCKNOW-226 025

2017

Candidate's Declaration

I hereby declare that I have completed research work for the full time prescribed and that the thesis embodies the results of my investigation conducted during the period I worked as Ph.D. research scholar. I further declare that to the best of my knowledge the thesis does not contain part of any work submitted for the award of any degree either in this University or any other Institute/University.

(Sonam)

Research Scholar

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY
LUCKNOW
प्रज्ञा शील करुणा
ESTABLISHED 1996

Babasaheb Bhimrao Ambedkar University
(A Central University)
Vidya Vihar, Rae Bareli Road, Lucknow - 226 025.
बाबासाहेब भीमराव अम्बेडकर विश्वविद्यालय
विद्या विहार, रायबरेली रोड, लखनऊ - 226 025

Certificate

This is to certify that the thesis titled **"Fusion Based Image Quality Enhancement"** submitted by **Ms. Sonam** is an original research work and has not been previously submitted in part or full for the award of any other degree or diploma to this or any other university.

The thesis submitted to Babasaheb Bhimrao Ambedkar University Lucknow satisfies all the requirements as stipulated in the *Doctor of Philosophy (Ph.D.) regulations-1999 as amended in 2010* and it is fit for submission and evaluation for the award of the degree of Doctor of Philosophy of the University.

Date: Supervisor

Head of the Department

Phone : +91 (0) 522 2440826, 2440827, 2441515, 2441615, 2441888, 2442190 (EPABX) **Fax :** +91 (0) 0522 2440821
Grams : BABVERSITY **E-mail :** root@dbraulko.ren.nic.in

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY
LUCKNOW
प्रज्ञा शील करुणा
ESTABLISHED 1996

Babasaheb Bhimrao Ambedkar University
(A Central University)
Vidya Vihar, Rae Bareli Road, Lucknow - 226 025.
बाबासाहेब भीमराव अम्बेडकर विश्वविद्यालय
विद्या विहार, रायबरेली रोड, लखनऊ - 226 025

Viva-Voce Certificate

This is to certify that thesis entitled **"Fusion Based Image Quality Enhancement"** submitted by **Ms. Sonam** under the supervision of **Dr. Manoj Kumar,** to the Department of Computer Science, Babasaheb Bhimrao Ambedkar University (A Central University) Lucknow, India, in the fulfillment of the requirement for the degree of Doctor of Philosophy in Computer Science has been approved after an oral examination of the same in collaboration with an external examiner.

Signature of External Examiner

Date:

Place: Lucknow

Head

Department of Computer Science

Babasaheb Bhimrao Ambedkar University

Phone : +91 (0) 522 2440826, 2440827, 2441515, 2441615, 2441888, 2442190 (EPABX) Fax : +91 (0) 0522 2440821
Grams : BABVERSITY E-mail : root@dbraulko.ren.nic.in

Abstract

Nowadays, image fusion plays a crucial role in computer vision and various areas of image processing. Image fusion synthesizes all the substantial information from more than one source images into a single image in such a way that the obtained fused image contains better description of the scene than any of the single source images. The obtained fused image yields more pertinent details and hence more suitable for machine or human perception. Image fusion has been used extensively in various disparate fields of image processing tasks such as, digital imaging, medical diagnosis, military, remote sensing, robotics, surveillance and so on. This thesis investigates the methods for image fusion as well as maintaining the original structure of the images. The aim of this work is to enhance the visual quality of images without any distortion and loss of information. Hence, image fusion is an essential tool to obtain the complete and appropriate information of multiple source images into one image.

In this thesis, our aim is to develop novel and improved image fusion techniques using pixel level image fusion approach which provides fused image having better visual quality without destroying any relevant information. The selection of pixel level image fusion is based on the comparison between all image fusion levels (pixel, feature and decision). From the comparison, it has been found that pixel level image fusion methods have an advantage of retaining more original information in the fused image. Furthermore, the methods are rather easy in implementation and computation, therefore this work has been performed using pixel level image fusion. Fusion process has been performed over the two different sets of multi-focus and multisensor images.

In the first approach, a discrete cosine transform (DCT) and sum-modified-laplacian (SML) based image fusion in stationary wavelet transform (SWT) domain is presented to obtain an image in which all objects of the scene are in focus. In this work, DCT is used because of its energy compaction property i.e., collection of most of the information in its few coefficients and SML focus measurement is used to select the sharp focused region.

The next approach is based on gradient method to preserve the sharp and directional information. In this work, we have proposed two image fusion techniques and in both the techniques gradient based fusion is performed to obtain the fused images having more sharp/directional information. The first technique is G-DWT in which the gradient rule is used to fuse multifocus images in DWT domain. The second technique is G-DTCWT, which is based on the gradient method in dual tree complex wavelet transform (DTCWT) domain due to its property of better shift invariance and good directionality over the DWT.

The low contrast image can not provide the complete interpretation of the scene. Therefore, contrast is necessary for an image which gives better visual information of the scene. Therefore in our next work an image fusion approach has been performed to enhance the contrast of the image. In this approach, fusion is performed based on directive contrast using discrete wavelet packet transform (DWPT).

The next approach is designed to develop an image fusion technique based multiresolution singular value decomposition (MSVD). The singular value decomposition (SVD) based fusion techniques are always computationally efficient, as only the few singular values can contain most of the information of the images. Therefore, this approach is based on singular value decomposition which uses the multiresolution property of SVD to obtain the fused images having better visual quality.

Due to the presence of noise in source images or inefficiency of fusion techniques, the fused images may always be not smooth. Hence, finally a cross bilateral filtering based fusion technique has been introduced which produces the smooth while

preserving the sharp changes in the fused images.

Finally, the thesis has been concluded with the future scope of work related to this field.

The performance of the fused images is evaluated qualitatively and quantitatively. Since, the visual criteria is not sufficient to measure the quality of images therefore we have also performed quantitative measurement criteria. Some standard metrics used for quantitative analysis are: peak-signal-to-noise-ratio, standard deviation, correlation coefficient, mutual information and so on. The proposed methods are compared with existing methods and state of the art methods. The results obtained from the proposed methods are better in terms of qualitative and quantitative assessments.

Acknowledgements

First and foremost, I would like to express my thank to God for providing me the great opportunity to study and work towards my Ph.D. degree under the guidance of Dr. Manoj Kumar, Assistant Professor, Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow. I appreciate very much for his precious guidance, support and attention throughout the successful completion of my research work. Without his consistent motivation, unconditional support and encouragement, this task would not have been possible for me.

I would like to thank to all research scholars who working for the development and advancement of the Science and Technology society. I feel proud and pleasure to become a part of this community.

The award of degree *Doctor of Philosophy* is one of the hardest deserving achievements. People struggle for it and achievement not easily found. During the entire research works some valuable people conceived their enormous positions in my heart. In this regard, I am grateful to the University and express my deep sense of gratitude to its Hon'ble Vice Chancellor for delivering this great opportunity to me.

I would also like to express my appreciation and gratitude to Prof. R.A. Khan, Dean SIST and Head, Department of Computer Science, Departmental Research Committee and all other faculty members of department namely Prof. Vipin Saxena, Prof. Sanjay K. Dwivedi, Dr. Deepa Raj, Dr. Narendra Kumar and Dr. Shalini Chandra for their motivation and support during the research. I convey my sincere gratitude to the Department of Computer Science for providing me the best

equipment, all essential facilities and healthy environment during my work. I also thanks to all administrative and supporting staffs of the University for providing the healthy environment and support.

I would also like to convey my special thanks to Dr. Manoj Diwakar, Department of Computer Science, Uttaranchal University, Dehradun and research scholars Ankita Vaish and Smita Agrawal who were always ready to help and support me. They have given me a healthy research environment, support and enjoyable university life during my research work.

I am deeply thankful to my parents, brother Amit Kumar, in-laws and all other family members for all their love, sacrifice and encouragement during my research work. They always supported me in all my pursuits.

Last but not least, my heartfelt thank to most of all for my loving, supportive and encouraging husband Kaushal Kishor whose faithful support and enduring patience during the final stages of this Ph.D. is so appreciated. Thanks for accompanying me in finishing up the degree and fulfilling my life with happiness. Without him, this thesis would not have been completed.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I thank University Grant Commission (UGC), New Delhi, India for the financial support to carry out this research work.

Lucknow (Sonam)

October, 2017

List of Figures

List of Tables

Table of Contents

CHAPTER 1

Introduction

Chapter 1

Introduction

This introductory chapter describes the motivation of the thesis, brief literature review and ends with a short outline of the thesis.

1.1 General Introduction

Due to an advancement in image processing systems, at the same time, multiple images of same scene can be obtained. However, to merge all the significant information by viewing and analyzing multiple images separately is not always convenient. This problem can be resolved by image fusion which is a process to combine the relevant information from two or more than two source images of a specific scene or view in order to construct the unified image. The produced image is perceptually enhanced and describes the whole information about specific scene and also contains more information by which we can easily interpret the scene than either of the source (input/original) images [48, 54]. The goal of image fusion is to obtain the fused image without any significant loss or distortion of information. The images to be integrated

are referred as input or source images, relevant information of the images refers to salient information and the resultant image as fused, composite or synthesized image. The image fusion has many advantages in scene analysis such as, reduced workload of human observer, increased accuracy in object detection and recognition, and increased system reliability, etc. Image fusion is an effective way for optimum utilization of large volumes of images from multiple sources. The quality of an image is mainly influenced by contrast, illumination, image sharpness and noise.

Genderen et al. [41] defined image fusion as: *"Image fusion is the combination of two or more different images to form a new image by using a certain algorithm"*. Image fusion is an effective way to integrate the significant information of two or more source images of same view or scene into an informative single image that is the most suitable for human visual perception and image processing tasks. The main aim of image fusion is to synthesize all the significant visual information from multiple source images in such a manner that the obtained image can provide more complete, stable and accurate description of the scene. The obtained image should be more informative and convenient for human visual interpretation or further image processing operations.

According to Smith et al. [121] an ideal image fusion algorithm extracts the information from source images and also satisfy three generic requirements:

1. The obtained fused image should not suppress the significant details contained in the source images.

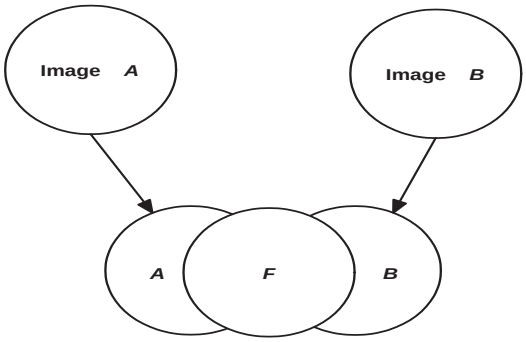

Figure 1.1: Graphical representation of image fusion process.

2. The fused image should not contain any artifacts or inconsistencies which can distract the decision of human observer or any subsequent image processing task.

3. The fused image should be capable to avoid the undesirable effects in the fused image.

If all the requirements listed above are fully satisfied then the complete enhanced image can be obtained. The image fusion process transfers the information from more than one information sets and integrate into a new one that retains all the relevant information from the original information sets.

The graphical representation of fusion process by Carroll et al. [23] is illustrated in Fig. 1.1 which uses two source images A and B to obtain a new fused image F by transferring the relevant information from the source images. The fused image F contains all the relevant information of both the source images [2]. The fusion of two source images of Lena is shown in Fig. 1.2. Fig. 1.2 (a)-(b) are the upper and lower

Figure 1.2: Fusion of Lena image. (a) Concentrated on upper part; (b) concentrated on lower part; (c) fused image.

part blurred images, respectively and Fig. 1.2 (c) is the fused image.

Applications of image fusion

In the last decades, image fusion has attracted an increasing amount of attention in a wide variety of application fields such as, surveillance, medical diagnosis, industrial applications and remote sensing.

- Surveillance: It covers applications identification, concealed weapon detection, mine detection, detection and tracking of targets and tactical situation assessment [12, 82, 88]. Fig. 1.3 illustrates the fusion of an infrared (IR) and visible image pairs, i.e., fusion of night and day time images. Fig. 1.3 (a) denotes a visible image which exhibits a high degree of textural information and Fig. 1.3 (b) represents infrared image which is sensitive to objects which have higher temperature than background and suitable to see in night without illumination and has the disadvantage of poor spatial resolutions. Due to low illumination,

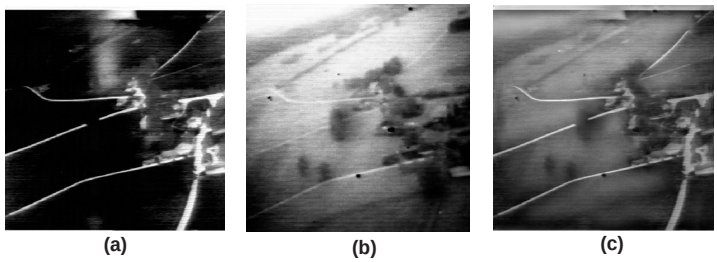

Figure 1.3: Example of surveillance image fusion. (a) Visible image; (b) infrared image; (c) fused image.

nighttime images have lower contrast and higher noise than their corresponding daytime images of the same scene. By the fusion of nighttime images with daytime images, the quality of nighttime images can be enhanced. The fusion of these image is able to give the most salient detail from both source images. Fig. 1.3 (c) shows the fused image obtained from the source images shown in Fig. 1.3 (a)-(b).

- Medical diagnosis: During the last few decades, medical image fusion has gained much attention to solve the medical issues. The main application of this field is to extract the clinical information (medical imaging, treatment planning and to assist in surgical procedures) from different sensors which are generally may not be visible in the images. The few biomedical sensors such as X-ray, ultrasound, magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET) which give the medical information by reflecting through the images of human body organs and cells are used in biomedical imaging [25, 56, 154].

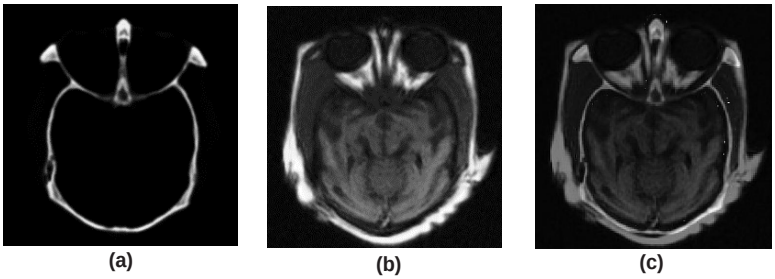

Figure 1.4: Example of medical image fusion. (a) CT; (b) MRI; (c) fused image.

Fig. 1.4 shows the example of image fusion of two medical images CT and MRI. CT images are the best suited for viewing the information of bone structures in the human body with high spatial resolutions. Whereas, MRI images are the best suited for capturing the soft issue structures in organs such as brain, heart and eyes. The fusion of CT and MRI images has achieved the advantage of improving clinical accuracy and visual quality [68]. The information provided by the CT and MRI images are shown in Fig. 1.4 (a) and 1.4 (b) respectively. These two different images can not provide all details of hard and soft tissues in a single image, therefore by combining these images we can obtain the whole information of both of the hard and soft tissues of a particular organ. The fused image of CT and MRI images is shown in Fig. 1.4 (c).

- Industrial applications: Image fusion is broadly used in industrial and civilian applications. Image fusion in industrial context is the combination of multifocus images. Images captured by the same camera but with diverse focal length have the problem of limited depth of focus. Due to the limitation of limited depth of

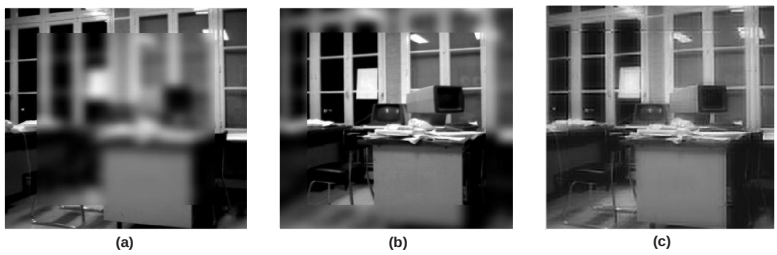

Figure 1.5: Example of industrial image fusion. (a) Middle side blurred image; (b) corner side blurred image; (c) fused image.

focus of cameras, it is often impossible to have all objects in focus within one shot. In order to address this problem, multifocus image fusion can be used to fuse the details of multiple source images of same scene but having different focus points to create an all in focus image which preserves the more significant information of the source images [9, 24]. Fig. 1.5 shows an example of fusion of two images having limited depths of focus. The fused image preserves more significant information than the original image therefore it is widely used in industrial fields.

- Remote sensing: Remote sensing images are captured from aircrafts and satellites to extract the information of the earth's surface. These systems are particularly deployed on satellites, provide a repetitive and consistent view of the earth providing valuable information about short and long term changes and the impact of human activities. Fig. 1.6 shows the fusion of panchromatic and multispectral images. The fusion application of panchromatic and multispectral images is known as pansharpening [124, 137]. The multispectral images used

Figure 1.6: Example of remote sensing image fusion. (a) Panchromatic image; (b) multispectral image; (c) fused image.

to distinguish some features spectrally but not spatially, whereas panchromatic images needed to distinguish features spatially but not spectrally. For image fusion, images captured from different sensors are merged into a composite image which yields high spatial and spectral resolution in the single fused image. Fig. 1.6 (a) shows panchromatic image and Fig. 1.6 (b) shows multispectral image and the fused image is shown in Fig. 1.6 (c).

According to the type of image capturing conditions, fusion can be performed over the various types of images [65, 107] such as, multifocus, multisensor, multiview and multiexposure, etc.

- **Multifocus image fusion**

  The multifocus images, acquired from the optical imaging cameras suffer from predicament of finite depth of focus of lens by which all objects may not be in focus. Consequently, if one object is in the focus other may be out of the focus in the scene, thus a portion of an image may get blurred. The images of complex scene may not provide the sufficient information of the scene therefore,

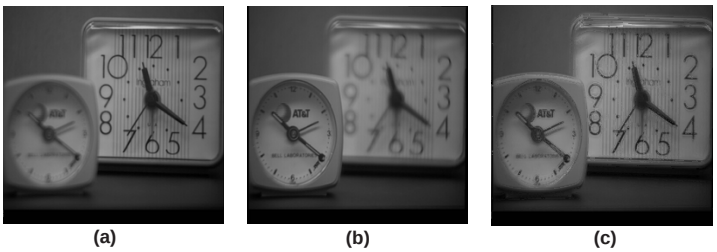

Figure 1.7: Example for multifocus image fusion. (a) Concentrated on right part; (b) concentrated on left part; (c) fused image.

it is difficult to obtain all objects in focus in a single image. One solution of this problem is multifocus image fusion, by which all objects can be brought into the focus in a single image. In this approach, multiple images of same scene are acquired from identical view points using different camera settings (focal lengths) and these images are fused with an appropriate fusion rule to produce an image with the hope that all objects will be focused in the fused image [31, 64, 76, 154, 155]. The example for multifocus image fusion is shown in Fig. 1.7.

- **Multisensor image fusion**

  Often, a single sensor image is unable to provide comprehensive details of an entire scene. The visible image provides spatial and spectral information and the target has same color and spatial details as its background then it becomes difficult to extract the details from background. If visible images are combined with the thermal images and target is to identify the colder or warmer than its background then it can be easily identified, even when its color and spatial

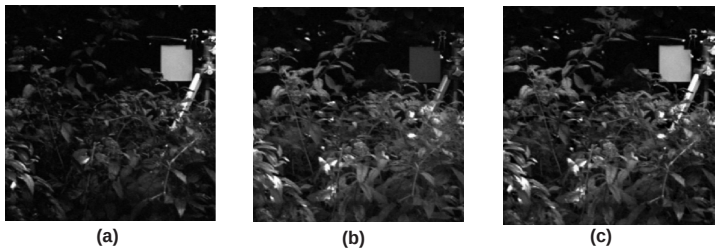

Figure 1.8: Example for multisensor image fusion. (a) Visible image; (b) infrared image; (c) fused image.

details are similar to those of its background. The single fused image can provide the whole information which can not be observed by any individual source image. The image fusion process reduces the amount of data to be processed without reducing the amount of significant details [41]. Here is an example of multisensor images is shown in Fig. 1.8. Fig. 1.8 (a) is visible image which capture the only bright objects and Fig. 1.8 (b) is an infrared image which can capture the thermal objects such as, plants and pedestrians. Fig. 1.8 (c) is the fused image obtained by fusing visible and infrared image which retains more information than any of the source images. The process of combining these two different sensor images into a single image is known as the multimodal/multisensor image fusion.

- **Multiview image fusion**

  Multiview images captured from same camera having same scene but with different view which points can not provide the clear image by which every object is generally not visible in the scene. The images of same scene having two depth

Figure 1.9: Example for multiview image fusion. (a) Foreground image; (b) background image; (c) fused image.

layers, background (far objects) and foreground (near objects). The images where the background images are focused on the background and foreground images are focused on the foreground, these two images can not provide both of the background and foreground information appropriately in a single image. If we combine these two images into a single image then this problem can be easily solved and this class of fusion is referred as multiview image fusion. The example for multiview image fusion is illustrated in Fig. 1.9, where Fig. 1.9 (a)-(b) are foreground and background images and Fig. 1.9 (c) is the fused image.

- **Multiexposure image fusion**

  Multiexposure images usually taken by ordinary digital cameras suffer from a lack of details in the under-exposed and over-exposed areas, if the camera has a different (low or high) exposure settings. The digital cameras are unable to capture a very high dynamic range (HDR) scenes due to their limited dynamic range. The several images of same scene are captured with varying exposure

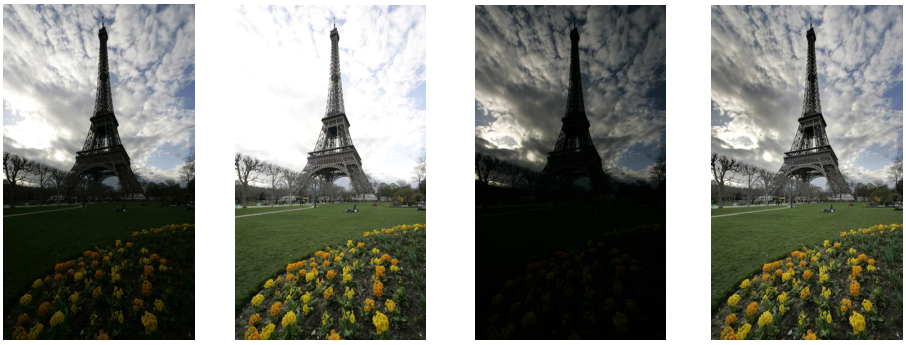

Figure 1.10: Example for multiexposure image fusion. (a) Normal exposure; (b) over exposure; (c) under exposure; (d) fused image.

settings are low dynamic range (LDR) images as input images. The set of these LDR images are used to generate a single HDR fused image. This type of fusion is known as multiexposure image fusion [158]. Fig. 1.10 (a)-(c) represent the LDR image sequences and Fig. 1.10 (d) is the fused image of the sequences.

Types of image fusion

Pohl and Genderen [108] categorized the image fusion process in three different levels of information representation: pixel level, feature level and decision level. These levels are introduced in the following subsections:

- **Pixel level image fusion**

  Pixel level fusion produces the fused image in which each pixel is determined from the corresponding set of pixel values in each source image to improve the performance of image processing operations. It is also called as signal level image fusion and known as the lowest level fusion. It is performed in pixel

by pixel basis [25, 90, 107]. In most of the image fusion applications, pixel level based fusion methods have been used due to the advantage of retaining original (pixel values) information. Furthermore, the methods are rather easy in implementation and computation. Pixel level image fusion has been used in a broad area of image processing but it is almost impossible to develop a universal fusion method which is applicable to all image fusion operations due to the diversity of images to be fused. The aim of pixel level image fusion is to represent the visual information contained in any number of source images into a single fused image without the distortion or loss of information. The representation of all the visual information from a set of source images into a single image is almost impossible. As a result, a more practical definition that is based on the faithful representation in the fused image of only the most important input information is generally accepted. A pixel level image fusion algorithm is required to identify the most significant information in the set of source images and to transfer the information without loss into the fused image. At the same time, the algorithm must ensure that there is no distortions or false information present in the fused image as a result of this process [104].

- **Feature level image fusion**

  Feature level image fusion is also known as object level image fusion. All features correspond to pixel intensities, texture, edges, colors are extracted from all the source images and fused in the resultant image [107]. Consequently, these

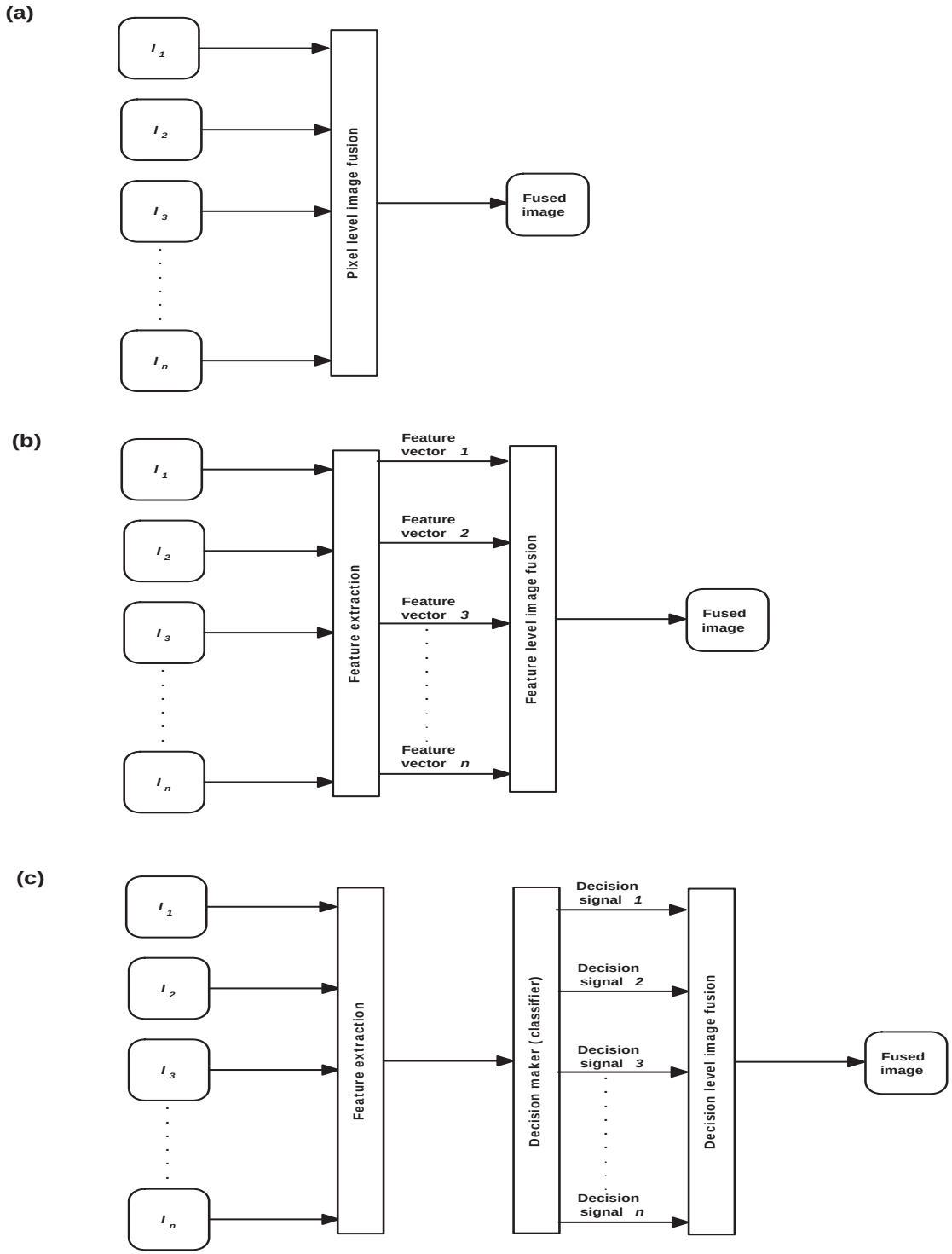

Figure 1.11: Relationship between different fusion levels. (a) Pixel level image fusion; (b) feature level image fusion; (c) decision level image fusion.

similar features from individual source images are combined using certain fusion methods in order to obtain a single output image.

- **Decision level image fusion**

  Decision level image fusion is also known as symbol level image fusion and known as the highest level image fusion. Fusion at decision level, fuses the results from multiple algorithms to provide the final fused image. Source images are processed individually for information extraction. The obtained information is then combined by applying decision rules to reinforce the common interpretation [127].

The selection of appropriate level depends on many different criteria such as, the characteristics of the physical sources, specific application and other factors such as execution time. These image fusion levels are related to each other. The fusion processes at these levels are illustrated in Fig. 1.11.

1.2 Motivation of the Study

Use of various imaging systems has increased the workload of human observers. Due to numerous imaging systems, images like multifocus or multisensor of same scene may have different information. The images may get blurred or degraded due to varying image capturing conditions or camera calibration thus, they may not provide the whole information of the scene in the single image. The images having less information may not provide an appropriate information of the scene. For instance,

medical images must be sharper for the diagnosis purpose. In some cases, we can not easily identify the objects in images due to their low contrast. Due to aforementioned conditions, it is quite difficult for a human operator to merge the reliable visual information by viewing and analyzing multiple images separately. Hence, a fusion system is required which can provide a single fused image by merging the relevant information present in the images of same scene obtained from same or different capturing conditions. The obtained fused image is more informative with better visual appearance than either of the single source images. Image fusion provides better interpretation of the scene, accurate information, enhanced resolution and robustness at low cost. Therefore, there is a need to develop novel and improved image fusion techniques to produce more informative fused image for multifocus and multisensor images, which are useful for better interpretation of the scene. Generally, the quality of the fused image depends on the application requirements. In many applications, the perception and interpretation of a fused image play an important role for human observer. Consequently, another way to assess the quality of fused image is its quantitative analysis. The qualitative analysis of the fused images are performed correctly but it is not sufficient due to the inconvenience and time consumption. Therefore, quantitative analysis can accurately measure the performance of the images.

The objective of this thesis is to develop the new and improved image fusion algorithms to enhance the visual quality of the images. We have concentrated only on pixel level image fusion because of the advantages of containing original information

and easy implementation. To increase the applicability of image fusion, the concentration is on developing the image fusion techniques which can give more information along with better visual quality. To evaluate the performance of fused images in term of quantitative analysis, various types of quality measurement metrics have also been used.

1.3 Brief Literature Review

The task of synthesizing the significant information from multiple images to create a single enhanced image is referred as image fusion. Pixel level image fusion had been classified by Stathaki [124] into two groups: spatial domain and transform domain based techniques. In spatial domain based techniques, fusion is performed by integrating all the pixel values of source images either in linear or non-linear manner. The fusion method that combines features from the source images in the spatial domain can be summarized as follows:

$$I_f(x,y) = f(I_1(x,y), I_2(x,y), ...I_n(x,y)) \tag{1.3.1}$$

where, n represents the number of source images and $f(\cdot)$ denotes the fusion rule. In spatial domain based fusion methods, intensity values of source images are directly fused. The resultant image can be obtained by the linear combination of source images. Fig. 1.12 illustrates the schematic diagram of the spatial domain based image fusion. All pixel values in the spatial domain may change by a small change in the single coefficient of the fused image in the transform domain. Several image fusion

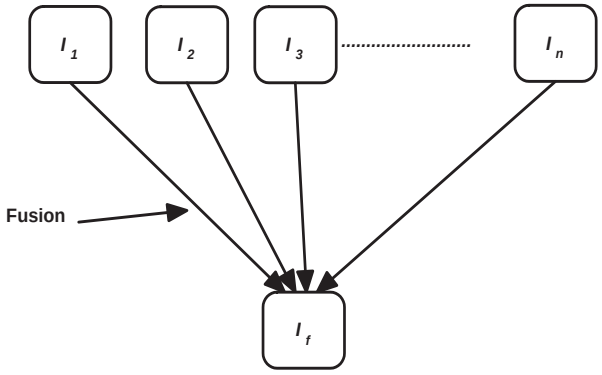

Figure 1.12: Spatial domain based fusion methods.

methods based on spatial domain have been proposed which can be classified as pixel [77, 143], window [32, 53] and region based methods [1, 166]. These methods are used to create a fused image by selecting the relevant image pixels or regions from input images. However, spatial domain based image fusion methods may lead to undesirable side effects such as, contrast reduction, blocking effects in the resultant composite image. To address these problems, transform domain based fusion techniques have been developed. The main motive behind the work in a transform domain is that the features of the source images are more clear in the transform domain. For the fusion in transform domain, the image structure is used rather than images pixels directly (independently). The principle behind the transform domain based image fusion methods is as follows: the images are decomposed into different multiresolution coefficients. These coefficients are combined together at each level to achieve the new coefficients of the resultant image. An inverse transform is performed appropriately over the new transformed coefficients to recover the fused image. Fig. 1.13 illustrates

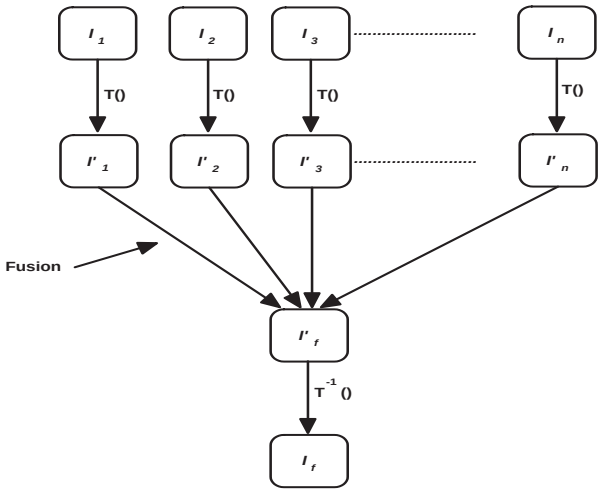

Figure 1.13: Transform domain based fusion methods.

the schematic diagram of the transform domain based image fusion. If we assume that $T(\cdot)$ and $T^{-1}(\cdot)$ denotes forward and inverse transform operator, respectively and $f(\cdot)$ fusion operator represents the merging of information from the multiple source images then the general transform domain based method is summarized as:

$$I_f(x,y) = T^{-1}(f(T(I_1(x,y), I_2(x,y), ...I_n(x,y)))) \tag{1.3.2}$$

The initial and the simplest image fusion techniques are the summation and averaging of the constituent source images. The pixel averaging method which fuses the source images by computing pixel by pixel average was presented in [162]. The method given in [21] produces an erroneous or unsatisfactory results. When the pixel averaging method is performed, it render the certain features with contrast reduction which are prominent in only one of the source images. Xu et al. [146] have presented a weighted averaging based fusion approach. In which, source images are directly fused and decomposition or transformation was not being performed over the images. Tao

et al. [128] have employed an image fusion method using PCA for multifocus images which is computationally efficient but unable to provide appropriate results for most of the fusion datasets. The Brovey transform (BT) [159], intensity hue saturation (IHS) [22, 144] and principal component analysis (PCA) [38, 59, 117], etc. also have been used to develop image fusion techniques. In recent years, various image fusion methods in spatial domain have been developed which produce better results. Few spatial domain based methods are image matting [75], gradient filtering [74], sparse representation [148], point spread function (PSF) [9] and dense SIFT feature based fusion methods [81]. Various focus measures such as, variance [86], spatial frequency [39], sum-modified-laplacian (SML) [97], Tenenbaum's algorithm (Tenengrad), energy of laplacian (EOL) [53] and energy of image gradient (EOG) [110] have also been used for multifocus image fusion purposes. The image fusion techniques based on pyramid concept was introduced in [4]. The first multiscale image fusion method based on Laplacian pyramid had been proposed by Burt [19] in 1984 for binocular fusion in human stereo vision. In this method, transformed images were combined using the maximum selection method and the composite image was obtained. Later, another image fusion method based on hierarchical image decomposition approach was presented by Burt et al. [20]. In which each of the source images was decomposed into its primitive sets. This technique was performed over a pyramidal transform over each of the source image. The representation of pyramidal decomposition was denoted as multiresolution representation, in which the successive images in the pyramid were

the sub-images of the source images and therefore this representation of images was also known as multi-resolution representation. Toet [131] have developed the ratio-of-low-pass pyramid based fusion algorithm instead of the laplacian pyramid to merge the visible and infrared images. Burt et al. [21] have introduced a gradient pyramid based fusion method of multisensor, multiexposure and multifocus images. The pyramid transform was applied over the source images which decomposes the images into multiscale decompositions. Generally, the averaging rule is being performed where the source images contain similar information and maximum saliency based fusion rules are used to select the feature pattern. The DWT based image fusion was first introduced by Ranchin et al. [111]. Li et al. [70] combined the obtained DWT transformed coefficients using an area based maximum selection method and consistency verification. DWT based image fusion techniques are found better than laplacian pyramid based techniques because the fused image does not contain artifacts. At the same time, Koren et al. [62] have presented a steerable dyadic wavelet transform based image fusion algorithm. The fusion process uses various decomposition bands of a laplacian pyramid instead of the source images. In order to resolve the problem of shift dependency in the DWT, Rockinger [114] has proposed an image fusion approach based on the undecimated wavelet transform (UWT) rather than DWT. Zhang et al. [162] had been presented a categorization of multiscale decomposition based image fusion to obtain a high quality fused image. To fuse the multiscale decomposed

coefficients, laplacian pyramid, DWT and discrete wavelet frame (DWF) were considered. Although, DWT is computationally efficient but it suffers from shift invariance, limited directionality and aliasing which are necessary for image fusion. To resolve these problems, Ray et al. [113] have introduced dual tree complex wavelet transform (DTCWT) for image fusion approach which yields better results than DWT based fusion methods. Li et al. [79] have proposed a multifocus image fusion approach using DWT and curvelet transform which gives the curves and edge information of images more accurately. Do et al. [36] have introduced contourlet transform which captures the edge information of the images and included laplacian pyramid and directional filter bank. The contourlet transform performs the multiresolution, directional and local image decomposition. In this transform, the extraction of point discontinuities is performed by laplacian pyramid and to link the point discontinuities a directional filter bank is used. The contourlet transform has been successfully used in medical imaging, remote sensing and surveillance due to its effectiveness in representing the spatial structures but it does not has the property of shift invariance. Therefore, to resolve this problem, another transform nonsubsampled contourlet transform [30, 160] a version of the contourlet transform has been introduced. In [136], low and high frequency coefficients of both of the transformed coefficients obtained from nonsubsampled contourlet transform were merged using averaging and standard deviation based methods which consume rather more time in the fusion process. After that, a shearlet transform [141] has been used for image fusion due to its advantage

of computational efficiency.

Recently, some comprehensive image fusion surveys have been given in [45, 73, 137]. Pajares et al. [100] introduced a tutorial of the wavelet based image fusion methods. They had presented a brief comparison of different pyramid merging methods, different levels and families of image fusion. Naidu et al. [94] have presented a multifocus image fusion technique based on PCA. This method is computationally efficient but it may not provide desirable results for most of the fusion datasets. Huang et al. [53] have given some focus measure based approaches for multifocus image fusion. In this method, source images are divided into blocks and various focus measures were used to select the best among the image blocks and it is found that the SML provides the superior performance than other focus measures. Hill et al. [50] adopted an image fusion technique based on the shift invariant and DTCWT. Parmar et al. [101] have introduced an image fusion scheme in transform domain in which two different methods are employed to fuse the transformed coefficients. In first approach, transformed coefficients are combined to obtain high quality fused image. In the second approach, a number of sparse coefficients are fused by applying the constrained threshold to enhance the rate of reconstruction. Yang et al. [149] have introduced a sparse representation based fusion method using orthogonal matching pursuit algorithm. Again, Yin et al. [156] presented a joint sparse representation (JSR) based fusion method. To extract the local salient features from images, an image fusion method based on sparse representation has been presented in [148]. Yao et al. [155] have developed

a hierarchical joint sparse representation based image fusion technique to integrate the sparse coefficients of source images. Recently, a guided filter based image fusion method [74] has been developed which preserves edges information and provides a deblurred fused image.

Tang [126] has suggested an image fusion method in DCT domain which produces blurring artifacts. Haghighat et al. [47] have given a fusion method based on the contrast. In this method, a multi-focus image fusion technique for visual sensor networks in DCT domain has been proposed. For activity level measurement, variance is calculated over 8×8 block DCT coefficients which is considered as a contrast criterion and consistency verification (CV) used to obtain a high quality output image. A contrast based image fusion approach has been employed in [130] using wavelet domain. The decomposed coefficients are combined using directive contrast method. Zhou et al. [164] have presented another image fusion method in gradient domain which detects definite focus regions and then identifies gradient weight near focused boundaries. This method has decreased the contrast between the boundaries of the source images. To preserve more texture details of the input images Han et al. [49] have given a fusion technique based on DCT sampling model and compressive sensing. Lu et al. [84] have proposed an image fusion technique using multiscale decomposition. The transformed coefficients obtained through multiscale decomposition were appropriately fused using maximum local energy rule to select the low frequency coefficients and sum modified laplacian to fuse the high frequency coefficients. Stanciu et al. [123]

have given two sum modified laplacian based image fusion techniques. In first and second method respectively, a sum modified laplacian maximum focus measurement and a sum modified laplacian weighted mean are used for creating the fused image. Another multiscale transform based image fusion method has been introduced in [83] using sharp frequency localized contourlet transform. The low frequency coefficients are calculated by using local energy method while sum modified laplacian is used to select high frequency coefficients. In [69], a lifting stationary wavelet transform and orientation information measure were used for image fusion purpose. To select low frequency coefficients, a new sum modified laplacian (NSML) is proposed which uses the focus measurement for orientation information. The high frequency coefficients are fused using feature contrast. Yang [152] has introduced another approach for image fusion which is helpful for multifocus images to evaluate the blurring of images and select more sharp detail based on DWT. Tseng et al. [135] have proposed a region based image fusion approach to improve the contrast and retain important features of source images. Tian et al. [129] have presented a multifocus image fusion method to evaluate the sharp information from source images by using bilateral sharpness criterion and phase coherence. A huge amount of work has been performed in the field of image fusion but still there is a scope and need to develop different image fusion algorithms for the better visual perception of the scene.

1.4 Some metrics for quantitative analysis

To evaluate the quality of the fused images only visual inspection is not sufficient for this reason quantitative analysis is also required [103, 158]. In this section, for quantitative analysis of some standard metrics has been discussed in detail which are further used to measure the quality of the fused images.

1.4.1 Mean and Standard Deviation

The mean or average pixel intensity *(API)* and standard deviation is given as:

$$\hat{\mu} = \frac{1}{mn}\sum_{p=1}^{m}\sum_{q=1}^{n} I_f(p,q) \tag{1.4.3}$$

$$S.D. = \sqrt{\frac{1}{mn-1}\sum_{p=1}^{m}\sum_{q=1}^{n}(I_f(p,q)-\hat{\mu})^2} \tag{1.4.4}$$

where, $\hat{\mu}$, *S.D.* and I_f represent mean, standard deviation and fused image, respectively. The *S.D.* criterion evaluates the contrast in an image and denotes the discrete image gray scale relative to the mean gray value. If the value of *S.D.* is high then the image gray level distribution of dispersion, image contrast is high which provides more information. Hence, higher value of the *S.D.* denotes the high contrast image.

1.4.2 Peak signal to noise ratio (PSNR)

PSNR is used to evaluate the error between the reference (original) and fused image.

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \tag{1.4.5}$$

where MSE is defined as

$$MSE = \frac{1}{mn}\sum_{p=0}^{m-1}\sum_{q=0}^{n-1}[I_r(p,q) - I_f(p,q)]^2$$

where, I_r and I_f represent the original and fused images, respectively. The high PSNR value denotes better fused image.

1.4.3 Correlation Coefficients

The correlation coefficient *(CC)* is often used to measure the degree of two linearly related variables. This measure has been used to evaluate the relation between reference and fused image. The correlation coefficient metric can be defined as [134, 145, 165]:

$$CC = \frac{\sum_{p=1}^{m}\sum_{q=1}^{n}(I_r - \mu(I_r)) \cdot (I_f - \mu(I_f))}{\sqrt{\sum_{p=1}^{m}\sum_{q=1}^{n}(I_r - \mu(I_r))^2 \cdot \sum_{p=1}^{m}\sum_{q=1}^{n}(I_f - \mu(I_f))^2}} \tag{1.4.6}$$

The value of correlation coefficient lies in $[0, 1]$. The value of correlation coefficient close to 1 denotes the more similar information in reference and fused image.

1.4.4 Mutual Information *(MI)*

Mutual information [39] is a metric measure reflects that the total amount of information that the fused image contains of the source images. It is defined as the sum of mutual information between each source image and the fused image. The mutual information I_{A,I_f} between source A and fused image I_f is given as follows:

$$I_{A,I_f} = \sum_{p,q} h_{A,I_f}(p,q) \log \frac{h_{A,I_f}(p,q)}{h_A(p)h_{I_f}(q)} \tag{1.4.7}$$

where, h_{A,I_f} is the jointly normalized histogram; h_A and h_{I_f} are normalized histogram of A and I_f, respectively. Similarly, I_{B,I_f} represents the mutual information between

other source image B and fused image I_f. Hence, the total *(MI)* between the source images A, B and fused image I_f is given as:

$$MI = I_{A,I_f} + I_{B,I_f} \tag{1.4.8}$$

The larger MI value represents better fusion result.

1.4.5 Fusion symmetry (FS)

It indicates how much symmetrical information of the fused image with respect to source image and is given as:

$$FS = 2 - \left| \frac{MI_{AI_f}}{MI_{total}} - 0.5 \right| \tag{1.4.9}$$

1.4.6 Spatial frequency (SF)

It estimates the overall information level in the regions of an image and is calculated as:

$$SF = \sqrt{RF^2 + CF^2} \tag{1.4.10}$$

where, RF and CF are row and column frequencies.

$$RF = \sqrt{\frac{\sum_{p=1}^{m}\sum_{q=1}^{n}(I_f(p,q) - I_f(p,q-1))^2}{mn}}$$

$$CF = \sqrt{\frac{\sum_{p=1}^{m}\sum_{q=1}^{n}(I_f(p,q) - I_f(p-1,q))^2}{mn}}$$

The large value of spatial frequency represents the large information in the image.

1.4.7 Average peak signal to noise ratio (APSNR)

This metric evaluates the error between source images and fused image and then these errors are used for calculating the average. The higher value represents better visual quality.

$$PSNR_{A,I_f} = 10\log_{10}\left(\frac{255^2}{\frac{1}{mn}\sum_{p=0}^{m-1}\sum_{q=0}^{n-1}[A(p,q) - I_f(p,q)]^2}\right) \tag{1.4.11}$$

Similarly, we can find $PSNR_{B,I_f}$ then average is computed as:

$$APSNR = (PSNR_{A,I_f} + PSNR_{B,I_f})/2 \tag{1.4.12}$$

1.4.8 Average gradient (AG)

It measures a degree of clarity and sharpness and is defined as [102]:

$$AG = \frac{\sum_{p=1}^{m}\sum_{q=1}^{n}((I_f(p,q) - I_f(p+1,q))^2 + (I_f(p,q) - I_f(p,q+1))^2))^{1/2}}{mn} \tag{1.4.13}$$

1.4.9 Entropy (EN)

It measures the amount of information presented in the fused image and is defined as:

$$EN = -\sum_{k=0}^{255} p_k \lg_2(p_k) \tag{1.4.14}$$

where, p_k is the probability of intensity value k in image. The larger values denote better result.

1.4.10 Q_0

The metric Q_0 [39] measures distortion of the fused image, it is a combination of three components as correlation, luminance and contrast. The metric Q_0 between the source image A and the fused image I_f is defined as follows:

$$Q_0(A, I_f) = \frac{4\sigma_{ij}\bar{i}\bar{j}}{(\sigma_i^2 + \sigma_j^2)(\bar{i}^2 + \bar{j}^2)} \tag{1.4.15}$$

where, σ_{ij} denotes the covariance, σ_i^2, σ_j^2 represent the variance and $\bar{i}, \bar{j}$ denote mean value of source image A and fused image I_f, respectively. $Q_0(A, B, I_f)$ represents the average value between $Q_0(A, I_f)$ and $Q_0(B, I_f)$, as given below:

$$Q_0(A, B, I_f) = \frac{Q_0(A, I_f) + Q_0(B, I_f)}{2} \tag{1.4.16}$$

The value of Q_0 lies between $[-1, 1]$ and it should be almost near to 1 for better quality.

1.4.11 Q_W

The metric Q_W [155, 110] between source images *A,B* and fused image I_f is defined as:

$$Q_W(A, B, I_f) = \sum_{w \epsilon W} c(w)(\lambda(w)Q_0(A, I_f|w) + (1 - \lambda(w))Q_0(B, I_f|w)) \tag{1.4.17}$$

where, $\lambda(w)$ is defined as:

$$\lambda(w) = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_B^2}$$

$\lambda(w)$ denotes the relative salience of A compared to B in the same window w and $c(w)$ indicates the normalized salience of the window *C(w)* which is defined as:

$$C(w) = max(\sigma_A^2, \sigma_B^2)$$

The Q_W contains the salience of information into account and the range lies between 0 to 1. 1 indicates the fused image retain all information from the source images.

1.4.12 Q_E

The metric Q_E [39] is defined as follows:

$$Q_E(A, B, I_f) = Q_W(A, B, I_f).Q_W(A', B', I_f')^{\alpha} \tag{1.4.18}$$

where, A', B' and I_f are the corresponding edge images of A, B, I_f, respectively. Parameter α reflects the contribution of the edge images compared to the original images which is set to 1. Q_E retains visual and edge information and larger value denotes the better fusion result.

1.4.13 Edge based similarity metric (Q^{AB/I_f})

Q^{AB/I_f} [147] metric evaluates the amount of edge information transferred from source images (A, B) images to fused image (I_f). It is calculated as:

$$Q^{AB/I_f} = \frac{\sum_{p=1}^{m}\sum_{q=1}^{n}(Q^{AI_f}(p,q).w^A(p,q) + Q^{BI_f}(p,q).w^B(p,q))}{\sum_{p=1}^{m}\sum_{q=1}^{n}(w^A(p,q) + w^B(p,q))} \tag{1.4.19}$$

where, $w^A(p,q)$ and $w^B(p,q)$ denote the weight for edge preservation values Q^{AI_f} and Q^{BI_f}. The definitions of Q^{AI_f} and Q^{BI_f} are same and given as:

$$Q^{AI_f}(p,q) = Q_g^{AI_f}(p,q).Q_\alpha^{AI_f}(p,q) \tag{1.4.20}$$

where $Q_g^{AI_f}(p,q)$ and $Q_\alpha^{AI_f}(p,q)$ represent the edge strength and orientation values at location p,q, respectively. Q^{AB/I_f} value close to zero denotes loss of input information and close to 1 denotes an ideal fused image.

1.5 Outline of the thesis

Chapter 2 proposes a novel image fusion technique based on DCT and SML. In this work, transformed coefficients using SWT are combined using DCT and SML focus measurement based fusion rule to obtain a focused image but it may not provide the appropriate sharp information of the scene. Therefore, in our next work, edges are preserved in fused images by using Gradient based image fusion rule.

Chapter 3 introduces a gradient based image fusion approach to obtain the sharp information of the resultant image. In this chapter, two image fusion techniques have been developed. First is Gradient based image fusion in DWT domain (G-DWT), while the second is Gradient based image fusion in dual tree complex wavelet transform (DT-CWT) domain. These techniques provide good image quality with better interpretation but these techniques may not be applicable on low contrast images. Therefore, to provide a better contrast in the fused images, a new image fusion technique has been developed using directive contrast.

Chapter 4 presents a contrast based fusion method using discrete wavelet packet transform (DWPT) domain. The correlation and directive contrast based fusion methods have been used to combine the details of transformed coefficients to produce

an image with better contrast. This work provides fused images having better sharp information while maintaining the contrast. However, in our next work, a novel fusion technique by utilizing the virtues of Multiresolution singular value decomposition (MSVD) has been developed.

Chapter 5 explores a MSVD based image fusion technique. The source images are transformed using DWT and transformed coefficients are fused using MSVD. In this work, the sharp and multidirectional details are obtained using MSVD but it may not provide a smooth fused image. Hence, to achieve smoothness cross bilateral filtering based method has been introduced.

Chapter 6 devotes a image fusion method based on cross bilateral filter (CBF) using discrete wavelet transform. Here, we have attempted to obtain a smooth, denoised and sharp fused image by performing the cross bilateral filtering. Therefore, a fused image having better edge information with good interpretation can be obtained.

Chapter 7 concludes this thesis with the analysis of the work presented in the previous chapters along with the brief discussion on the future scope.

CHAPTER 2

Discrete cosine transform and sum-modified-laplacian based image fusion

Chapter 2

Discrete cosine transform and sum-modified-laplacian based image fusion

This chapter proposes a new image fusion approach based on discrete cosine transform (DCT) and sum-modified-laplacian (SML). Here, DCT and SML are used for fusion of source images using stationary wavelet transform (SWT) domain. The aim of this work is to provide a focused image having better visual quality and better interpretation of the scene.

2.1 Introduction

In present scenario, image fusion has shown wide spectrum of applications in medical imaging, surveillance, robotics, machine vision, remote sensing and biometrics, etc. The aim of image fusion is to reconstruct a single image by synthesizing all significant details from multiple source images [18, 73]. The obtained new image is more

informative for computer processing operations and visual perception. Usually, the images captured from a camera have problem of a limited depth-of-field of lenses by which it is very difficult to obtain an image in which all objects are in focus. This type of problem can be resolved by using multifocus image fusion approach. The core idea behind the multifocus image fusion is to generate a fused image having every object in focus [68]. Generally, images can be fused better in transform domain. During the last few years, several transform domain techniques have been developed dedicated to the fusion field such as discrete wavelet transform (DWT) [28, 119], discrete cosine transform (DCT) [47, 93, 106], dual-tree complex wavelet transform (DTCWT) [35, 50, 116] and stationary wavelet transform (SWT) [7, 71]. As it is well known, pixel averaging is the simplest technique of spatial domain for image fusion and produces erroneous results such as reduced contrast and distorted images [98]. Further, another image fusion technique based on weighted average rule have been given by Xu et al. [146]. This technique is a spatial domain based fusion technique, hence, there is no any decomposition or transformation has been performed. The source images are directly averaged the defocused pixels to obtain the fused image. The wavelet transform based fusion technique [60] transforms the source image into a number of subband images having different resolutions and directional information. Various image fusion schemes have been introduced using the benefits of wavelets in image fusion and denoising. Lu et al. [84] have proposed fusion techniques to enhance the capability and reliability of fusion system using multiscale decomposition based

fusion method. For image fusion purpose, the transformed coefficients (low and high frequency coefficients) are combined using maximum local energy and SML fusion rule. At the end of fusion process, an inverse multiscale wavelet transform has been performed to obtain the final fused image. To achieve a fused image with better focus uniformity, image fusion method using sum modified laplacian rule has been presented by Stanciu et al. [123]. In [83], the cycle spinning sharp frequency localization contourlet transform has been used for multifocus image fusion purpose. In this method, the transformed coefficients were combined using local energy (LE) and sum modified laplacian rules. Another multimodal image fusion framework has been introduced using non subsampled contourlet transform (NSCT) [17], in which source images were enabled into low and high frequency transformed coefficients. Fused coefficients were achieved by performing the normalized Shannon entropy based on an activity measure and directive contrast rules which retains more significant texture information with better visual quality of the source images. An optimal image fusion method for multifocus images has been presented using differential evolution algorithms [8]. In this, firstly, input images are decomposed into blocks and sharp blocks have been selected by using sharpness criterion. Finally, all selected blocks were combined to obtain the fused image. An image fusion method using DCT domain has been given in [126] but it produces blurring. Therefore another image fusion method [47] using DCT domain has been introduced to enhance the contrast in the fused image. In this approach, variance for 8×8 block DCT coefficients was computed for activity level measurement

which was further considered as a contrast criterion and consistency verification to produce a better visual quality image. Furthermore, image fusion method using DCT sampling model and compressive sensing has been proposed to retain more textural information of the source images [49]. Again, various focused region detection based image fusion techniques [33, 85] have also been developed. These methods have an advantage of producing the fused image in which pixels of only focused regions were selected. To tackle the problem of multifocus images, a novel fusion scheme [69] based on lifting stationary wavelet transform (LSWT) and orientation information measure has been introduced. The low and high frequency coefficients have been fused by using the different fusion criteria. To combine the low frequency coefficients, a new sum-modified-laplacian (NSML) of the orientation of information measure was used while the high frequency coefficients were fused using a novel feature contrast of the orientation information measure and a sharp fused image was obtained. The above methods are used to merge multifocus images but it may fuse defocused parts into the resultant image. Therefore, we proposed a multifocus image fusion approach using focus measurement to obtain a focused image. To the best of our knowledge, it is found that high frequencies play a significant role to enhance the resolution. Therefore, for developing any fusion algorithm the separation of high frequencies from low frequencies is always beneficial. The wavelet transform is the most widely used tool for this purpose. For the 1-level decomposition, DWT produces one low and three high frequency subbands at orientation (0°, 45° and 90°). As compared with SWT,

it is found that SWT is alike DWT but does not downsample the input signal, thus it generates the output signal as the same size of input signal.

The aim of this work is to provide a better focused image by using the benefits of SWT, SML and DCT. In the proposed work, a multifocus image fusion approach using SML and DCT in SWT domain is presented. The key idea behind using SWT is that it provides better approximation part because of its shift invariant feature and it overcomes the problem of repeatability and provides robustness. The proposed image fusion algorithm has been performed using few necessary steps. First, the low and high frequency coefficients are obtained by transforming both of the source images using SWT. The low frequency coefficients corresponding to approximation part retains most of the information of an image whereas, high frequency coefficients denote detail parts which contain the directional information such as, horizontal, vertical and diagonal. After the decomposition, approximation and detail parts are fused using two different fusion rules. The approximation part is combined using blockwise DCT based variance rule whereas, detail parts are fused using blockwise SML based approach. The inverse SWT is performed over the obtained fused coefficients to get the final fused image.

The rest of content of this chapter is organized as follows: stationary wavelet transform, discrete cosine transform and sum-modified-laplacian are discussed in Section 2.2. In Section 2.3, proposed method is discussed. Experimental results and discussion are presented in Section 2.4. At last, we concluded the work in Section 2.5.

2.2 Preliminaries

In this section, the brief concepts of SWT, DCT and SML are described which is further used in the proposed fusion process.

2.2.1 Stationary wavelet transform

In 1995, Nason and Silverman [96] have given the concept of stationary wavelet transform (SWT). This transform has the properties of shift invariance, non-redundancy and provides better approximation in compare to discrete wavelet transform. It also reduces the visual artifacts and provides lower reconstruction error results in lesser time. As it is well known, the DWT can not yields these features therefore SWT has been broadly used in image fusion, denoising and edge detection, etc. SWT is also known as the undecimated wavelet transform and works like DWT but it does not downsample the source image and produces output image as that of the same size of source image and therefore a better detection can be executed [34, 87]. It is used to analyze an image into different resolution and provides diagonal, vertical and horizontal details of an image. SWT transforms the image into one low and three high frequency coefficients labeled as *(LL, HL, LH* and *HH)*. Over an image *I*, *l-1* level decomposition of SWT is applied and four coefficients LL_l, HL_l, LH_l and HH_l are obtained as depicted in Fig. 2.1. The low frequency coefficient LL_l is called as approximation part whereas, the remaining coefficients HL_l, LH_l and HH_l are called

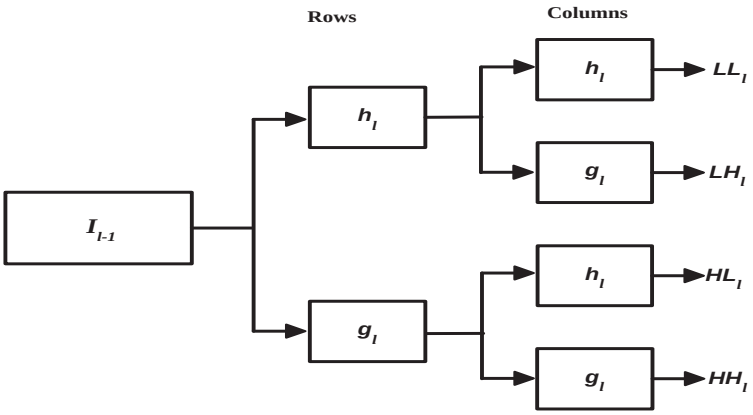

Figure 2.1: SWT decomposition.

as detail parts. The SWT decomposition [34] can be defined as:

$$A_{l,k_1,k_2} = \sum_{n_1}\sum_{n_2} h_0^{\uparrow 2^l}(n_1 - 2k_1)h_0^{\uparrow 2^l}(n_2 - 2k_2)A_{l-1,n_1,n_2}$$

$$H_{l,k_1,k_2} = \sum_{n_1}\sum_{n_2} h_0^{\uparrow 2^l}(n_1 - 2k_1)g_0^{\uparrow 2^l}(n_2 - 2k_2)A_{l-1,n_1,n_2}$$

$$V_{l,k_1,k_2} = \sum_{n_1}\sum_{n_2} g_0^{\uparrow 2^l}(n_1 - 2k_1)h_0^{\uparrow 2^l}(n_2 - 2k_2)A_{l-1,n_1,n_2}$$

$$D_{l,k_1,k_2} = \sum_{n_1}\sum_{n_2} g_0^{\uparrow 2^l}(n_1 - 2k_1)g_0^{\uparrow 2^l}(n_2 - 2k_2)A_{l-1,n_1,n_2} \tag{2.2.1}$$

where, A_{l,k_1,k_2} are the low frequency coefficients also known as coarse or approximation part and contain most of the information and $H_{l,k_1,k_2}, V_{l,k_1,k_2}$ and D_{l,k_1,k_2} are the high frequency coefficients and contain respectively, horizontal, vertical and diagonal information and these are called detail parts of the image. $h_0^{\uparrow 2^l}$ and $g_0^{\uparrow 2^l}$ indicate that $2^l - 1$ zeros are inserted between h_0 and g_0. To reconstruct the image, inverse SWT is used given by the following formula:

$$A_{l-1,n_1,n_2} = \frac{1}{4}\sum_{i=0}^{3}\Bigg\{\sum_{k_1}\sum_{k_2} h_1(n_1-2k_1-i)h_1(n_2-2k_2-i)A_{l,k_1,k_2} + \sum_{k_1}\sum_{k_2} h_1(n_1-2k_1-i)g_1(n_2-2k_2-i)H_{l,k_1,k_2} + \sum_{k_1}\sum_{k_2} g_1(n_1-2k_1-i)h_1(n_2-2k_2-i)V_{l,k_1,k_2} + \sum_{k_1}\sum_{k_2} g_1(n_1-2k_1-i)g_1(n_2-2k_2-i)D_{l,k_1,k_2}\Bigg\} \tag{2.2.2}$$

2.2.2 Discrete cosine transform

DCT is a non-redundant and invertible transformation, which de-correlates an image in frequency domain [47, 93]. Basically, DCT is used to define a sequence of real data points into its real spectrum. It has a property of good energy compaction, i.e. it contains most of the important detail of the image in only few coefficients, therefore, it is very popular in image processing. 2D-DCT [106] transform of $N \times N$ image *I(p,q)* is given as:

$$F(u,v) = \frac{2}{N}\alpha(u)\alpha(v)\sum_{q=0}^{N-1}\sum_{p=0}^{N-1} I(p,q) \times \cos\left[\frac{(2p+1)u\pi}{2N}\right] \times \cos\left[\frac{(2q+1)v\pi}{2N}\right] \tag{2.2.3}$$

where, $u, v = 0, 1, ..., N-1$ and

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{2}}, & if\ u = 0 \\ 1, & otherwise \end{cases}$$

The inverse DCT is given as:

$$I(p,q) = \sum_{v=0}^{N-1}\sum_{u=0}^{N-1}\frac{2}{N}\alpha(u)\alpha(v)F(u,v)\times\cos\left[\frac{(2p+1)u\pi}{2N}\right]\times\cos\left[\frac{(2q+1)v\pi}{2N}\right] \quad (2.2.4)$$

where, $p,q = 0,1...,N-1$.

2.2.3 Sum-modified-laplacian

In transform domain, high frequency sub-bands contain sharp information of the images. To fuse the high frequency sub-bands, the maximum fusion rule which selects high frequency sub-bands with maximum absolute value can be used but it suffers from the artifacts and does not consider the surrounding pixels. Thus, the maximum method is sensitive to edges and directional characteristics. To overcome the problem of artifacts in fusion, instead of selecting the pixel with a larger value SML focus measure is used. Recently, there are many clarity measurements, such as sum-modified-laplacian (SML), energy of laplacian (EOL), energy of gradient (EOG), Tenengrad and spatial frequency (SF) have been introduced [53, 110]. These are used to measure the degree of blurring images, the high frequency information such as, edges and boundaries are usually used as the basis. The focus measure defines the maximum focused area. Thus, it is popular in the area of multifocus image fusion, the focused source image produces maximum focus measure, whereas defocused image produces minimum focus measure. In [110], it is found that for fusion purpose, the

performance of SML is better than Tenengrad, SF, EOG and EOL. In [53], modified-laplacian (ML) can provide the salient features and sharp information of the image. The sum of the surrounding ML is denoted by SML operator which yields the local measures of the quality of image focus [97]. As in [109], it is found that SML is more efficient than ML in frequency domain, particularly for high frequences. Therefore, SML is used to fuse the high frequency coefficients. The definition of SML is given as [10]:

$$SML(p,q) = \sum_{i'=p-N}^{i'=p+N} \sum_{j'=q-N}^{j'=q+N} ML(i',j') \tag{2.2.5}$$

where, N determines the window of size $(2 \times N + 1) \times (2 \times N + 1)$ used to compute the focus measure. *ML(p, q)*, denotes the value of modified laplacian [83] located at *(p, q)* can be defined as:

$$\begin{aligned} ML(p,q) =& \mid 2I(p,q) - I(p-1,q) - I(p+1,q) \mid \\ &+ \mid 2I(p,q) - I(p,q-1) - I(p,q+1) \mid \end{aligned} \tag{2.2.6}$$

where, *I(p,q)* denotes the pixel value located at the position *(p,q)*.

2.3 Proposed method

The detail of our proposed fusion method for multi-focus images is discussed in this section. Fig. 2.2 illustrates the block diagram for proposed fusion method. In which, two images A and B are used, which are less informative and these two images are considered as source images. In the proposed work, *1-level* SWT is performed over the

given source images to obtain low and high frequency coefficients. These coefficients are fused with different fusion rules, which are described in the following subsections:

2.3.1 Fusion of low frequency coefficients:

The low frequency coefficients of both of the transformed images are fused using blockwise DCT based variance rule. In fusion process, all the low frequency coefficients are divided into 8×8 non-overlapping blocks. Over the each block, DCT is performed and DCT coefficients are obtained, as discussed in Section 2.2. Using these coefficients, variances are computed and compared for each image, as given below:

$$LL_{F,l}(p,q) = \begin{cases} x_1(p,q), \ if : \sigma^2_{A,l}(x_1(p,q)) > \sigma^2_{B,l}(y_1(p,q)) \\ y_1(p,q), \ otherwise \end{cases} \tag{2.3.7}$$

where, $x_1(p,q)$ and $y_1(p,q)$ are the DCT coefficients and σ^2 denotes variance. The variance of an image I at pixel *(p,q)* can be computed by the following formula:

$$\sigma^2 = \frac{1}{N^2} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} (I^2(p,q) - \hat{\mu}^2) \tag{2.3.8}$$

where, $\hat{\mu}$ represents mean of block $N \times N$ as given in chapter 1. Finally, inverse DCT is performed to obtain the fused low frequency subbands $LL_{new,l}(p,q)$.

2.3.2 Fusion of high frequency coefficients:

Over the high frequency coefficients, blockwise SML based fusion rule is performed, as discussed in Section 2.2. The obtained high frequency coefficients of the transformed images are divided into 8×8 non-overlapping blocks. The SML rule is

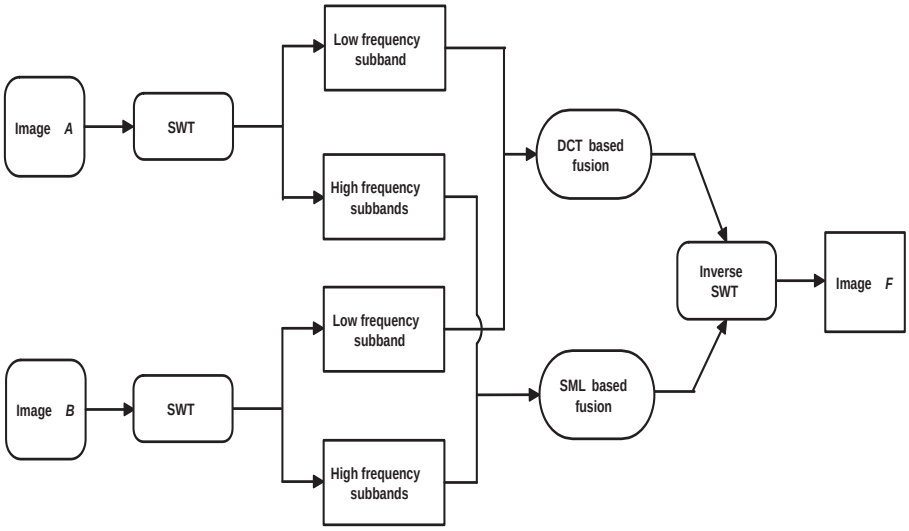

Figure 2.2: Block diagram for proposed fusion method.

followed by the maximum method which is performed over the each blocks. The SML based fusion rule for high frequency coefficients is described below:

$$HL_{new,l}(p,q) = \begin{cases} x_2(p,q), & if : SML_{A,l}(x_2(p,q)) \geq SML_{B,l}(y_2(p,q)) \\ y_2(p,q), & otherwise \end{cases} \tag{2.3.9}$$

$$LH_{new,l}(p,q) = \begin{cases} x_3(p,q), & if : SML_{A,l}(x_3(p,q)) \geq SML_{B,l}(y_3(p,q)) \\ y_3(p,q), & otherwise \end{cases} \tag{2.3.10}$$

$$HH_{new,l}(p,q) = \begin{cases} x_4(p,q), & if : SML_{A,l}(x_4(p,q)) \geq SML_{B,l}(y_4(p,q)) \\ y_4(p,q), & otherwise \end{cases} \tag{2.3.11}$$

where, $x_2(p,q)$, $x_3(p,q)$, $x_4(p,q)$ and $y_2(p,q)$, $y_3(p,q)$, $y_4(p,q)$ are the coefficients computed by performing SML, as discussed in Section 2.2.

The final synthesized image is achieved by performing inverse SWT over the fused coefficients ($LL_{new,l}, HL_{new,l}, LH_{new,l}$ and $HH_{new,l}$). The above discussion of the proposed method can be summarized in following Algorithm:

2.3.3 Proposed Image Fusion Algorithm

1. Take two source images A and B.

2. Perform SWT over each of the source images to obtain low and high frequency subbands.

3. Apply DCT based variance fusion rule to fuse low frequency subbands.

4. Apply SML fusion rule followed by maximum method to combine high frequency subbands.

5. Perform inverse SWT over the new coefficients and obtain the final fused image F.

2.4 Experimental results and discussion

To test the performance of our proposed fusion method various multifocus images of size 512×512 have been used. The images shown in Fig. 2.3 (a)- 2.6 (a) are original images considered as the reference images. In the proposed work, all reference images I_r are made blurred by performing gaussian blurring with 13×13 window size and standard deviation $\sigma = 5$ and the source images are obtained. Fig. 2.3 (b) and 2.4 (b) are upper blurred images and Fig. 2.3 (c) and 2.4 (c) are lower blurred images. The left part blurred images are depicted in Fig. 2.5 (b) and 2.6 (b). Fig. 2.5 (c) and 2.6 (c) are blurred images on the right parts.

Few existing methods such as, DWT [62], PCA [94] and MSVD [92] are used

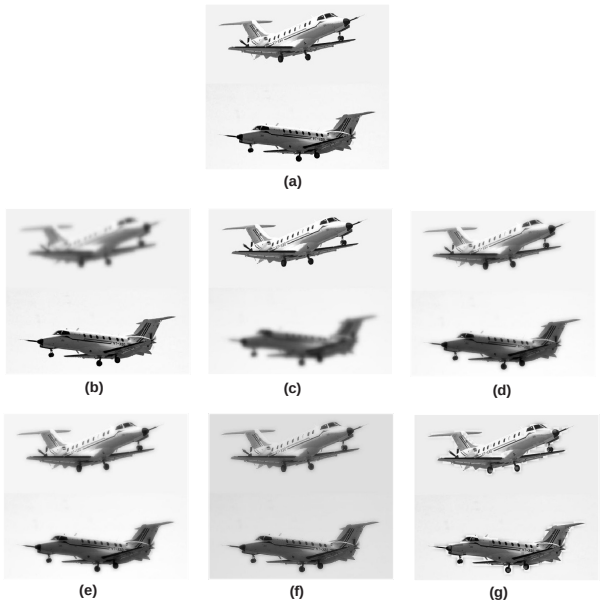

Figure 2.3: (a) Original dataset1 image; (b) upper part blurred image; (c) lower part blurred image; (d) DWT; (e) PCA; (f) MSVD; (g) proposed fusion method.

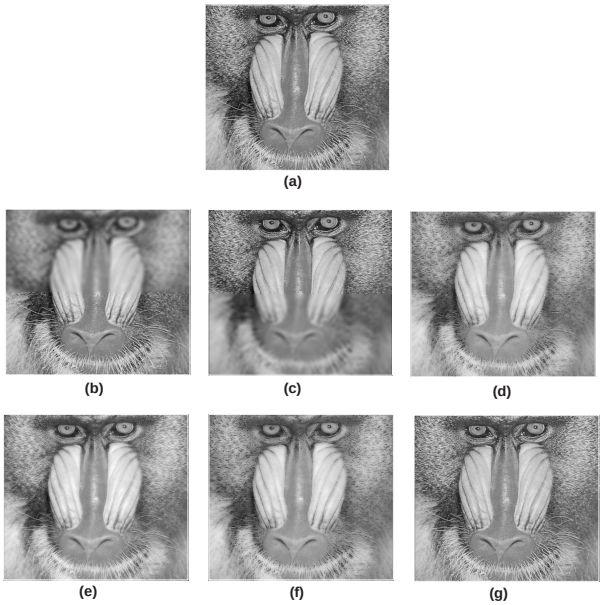

Figure 2.4: (a) Original mandrill image; (b) upper part blurred image; (c) lower part blurred image; (d) DWT; (e) PCA; (f) MSVD; (g) proposed fusion method.

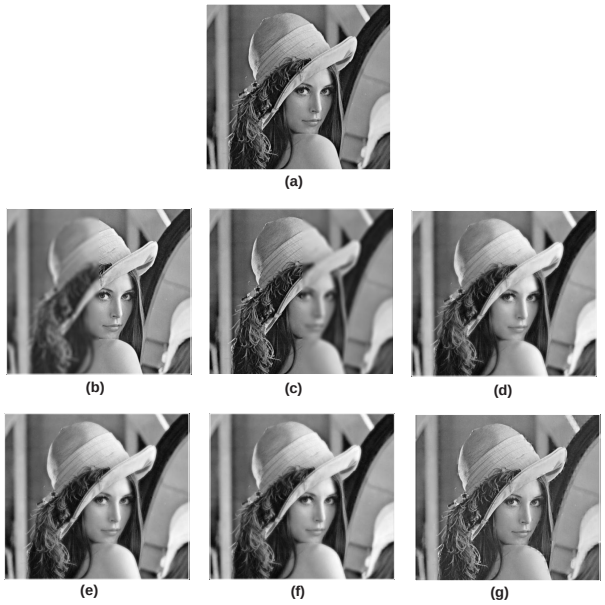

Figure 2.5: (a) Original lena image; (b) left part blurred image; (c) right part blurred image; (d) DWT; (e) PCA; (f) MSVD; (g) proposed fusion method.

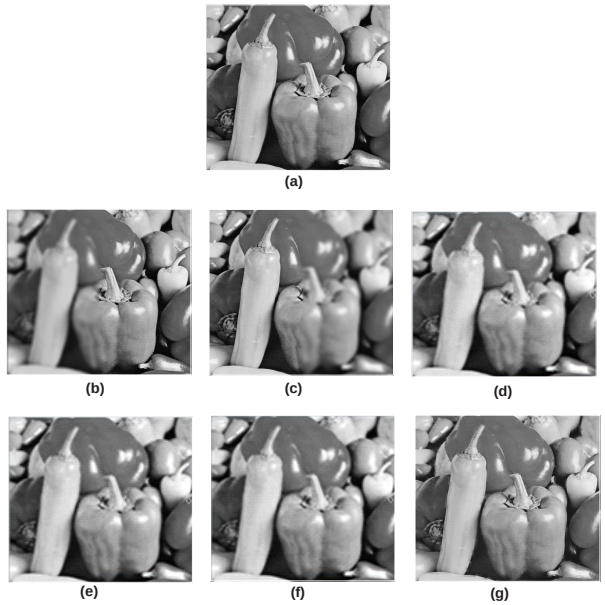

Figure 2.6: (a) Original pepper image; (b) left part blurred image; (c) right part blurred image; (d) DWT; (e) PCA; (f) MSVD; (g) proposed fusion method.

Table 2.1: Quantitative evaluation for multifocus images

Source Images	**Evaluation indices**	**[62]**	**[94]**	**[92]**	**Proposed method**
Dataset1	Mean	221.43	221.43	221.43	220.71
	S.D.	52.457	51.873	52.454	56.154
	PSNR	34.081	33.395	34.070	38.582
	CC	0.9864	0.9842	0.9863	0.9945
	$Q^{AB/F}$	0.5862	0.6283	0.6265	0.7415
Mandrill	Mean	128.97	128.97	128.97	129.10
	S.D.	38.647	35.618	38.646	42.337
	PSNR	34.120	31.826	34.118	40.088
	CC	0.9735	0.9585	0.9735	0.9929
	$Q^{AB/F}$	0.5827	0.4764	0.6156	0.8224
Lena	Mean	123.40	123.40	123.40	122.77
	S.D.	45.529	45.173	45.529	47.296
	PSNR	36.689	36.147	36.691	36.338
	CC	0.9884	0.9870	0.9884	0.9870
	$Q^{AB/F}$	0.6260	0.5905	0.6253	0.7193
Peppers	Mean	119.15	119.15	119.15	118.87
	S.D.	51.231	50.832	50.979	52.706
	PSNR	36.677	36.079	35.342	36.691
	CC	0.9910	0.9898	0.9875	0.9904
	$Q^{AB/F}$	0.6097	0.6002	0.5630	0.6602

to compare the results with proposed method. The results obtained by existing methods are shown in Figs. 2.3 (d-f)-2.6 (d-f). The results of the proposed method are given in Figs. 2.3 (g)-2.6 (g). In terms of visual perception, it can be easily seen that the results obtained from the proposed method are better than others. But to quantitatively evaluate the performance of these methods, some metrics have been used such as, PSNR, standard deviation, correlation coefficients, mean and $Q^{AB/F}$ which are already discussed in Chapter 1. The obtained results corresponding to

these metrices are illustrated in Table. 2.1. From the table, it can be analyzed that the results obtained from the proposed methods are better than others in most of the cases.

2.5 Conclusions

In this chapter, a novel multifocus image fusion technique based on SML and DCT in SWT domain has been discussed. The main objective of this work is to produce a better focused image having better visual perception. Two different rules are used to fuse low and high frequency coefficients. DCT based variance fusion rule has been employed to fuse low frequency coefficients, whereas SML based fusion approach has used to fuse high frequency coefficients. The proposed methods and existing methods have been tested on several multifocus images. To evaluate the performance of proposed quantitatively several performance metrics such as mean, PSNR, *S.D.*, CC and $Q_{AB/F}$ are computed and compared.

From the experimental results, it can be seen that the proposed method retains more visual information and better focus images for the better interpretation of the scene. The proposed method preserves more significant information and improves the visual quality than other existing methods in both quantitative and qualitative analysis. But to preserve the more adequate sharp information of multi-focus images a gradient based approach has been proposed in the next chapter.

CHAPTER 3

Gradient based image fusion

Chapter 3

Gradient based image fusion

This work attempts to provide better fused images by preserving sharp details of the source images. The work includes two image fusion techniques based on gradient fusion rule. First is gradient based fusion in DWT domain (G-DWT), while the second is gradient based fusion in dual tree complex wavelet transform (G-DTCWT) domain.

3.1 Introduction

In the last few decades, several spatial domain [9, 32, 75, 143, 166] and transform domain [101, 136, 149, 155] based image fusion methods have been proposed. Few spatial domain techniques such as weighted averaging, linear fusion, principal component analysis [117], independent component analysis [42, 55], sharpness criteria are used to merge images. However, spatial domain based fusion methods often produce erroneous results because they usually produce spectral distortions in the fused image. Recently, many transform based image fusion methods have been developed due

to their multiresolution decomposition property. The spatial domain based methods select the minimum or maximum intensity pixels from the source images. These methods are suitable for the interest of dark and bright features. The maximum method [143] is widely used technique for fusion. This method is sensitive in terms of noise and artifacts. A wavelet transform and weighted averaging based medical image fusion technique has been developed by Shangli et al. [118]. The multiresolution fusion techniques based on wavelet transform, provides a better spatial and spectral localized information of image than other multiresolution techniques [6, 27]. A sharp image is suitable for the human perception and also desirable in various machine vision tasks. Thus, to obtain the sharp fused images a fusion tool is required. By increasing the sharpness of images the spatial resolution also increases [26]. A new fusion method based on DWT [152] had been introduced for multi-focus images to produce a sharp fused image. The low frequency sub-bands were fused using the maximum sharpness focus measure method, while high frequency sub-bands were combined by a maximum neighboring energy method. Hill et al. [50] have presented a dual tree complex wavelet transform to fuse the source images. Due to advantages of good shift invariance and directionality than DWT, it is more useful. To extract the edge details, DTCWT has been developed which contains approximate shift invariance and high directionality and these advantages of DTCWT make it suitable to fuse the images. A novel multifocus image fusion method based on the concept of hybrid wavelet and classifier has been given in [157]. In this method, source images are

decomposed into six directional frequency subbands using DTCWT and to classify the subbands, support vector machine (SVM) is used. Using DTCWT [67], in region based image fusion technique a feature map was generated through segmentation of the features from source images to merge region by region images. To measure the clarity of information in image, image sharpness can be represented in different ways. In the image, high frequencies are corresponding to the high details. The sharpness is associated with the quality of edges and boundaries which provides relevant high details. The key challenge of multi-focus image fusion problem is to evaluate the blurring of each image and then select more relevant information from the processed image to increase the visual quality of the image.

Here, we have proposed two fusion methods in which sharp details are obtained from the transformed coefficients with the aim to increase the sharpness in the fused image. This chapter is based on gradient based image fusion. In which, the sharpness is evaluated by selecting the important information from the sharp details. The objective of this chapter is to improve the quality and visual appearance of image by combining all sharp details of individual source images.

The remainder of this chapter is organized as follows. The theory of DWT and DTCWT is discussed in Section 3.2. Proposed scheme is explained in Section 3.3. Experimental results are discussed in Section 3.4. Finally, concluding remarks are given in Section 3.5.

3.2 Basic concepts of DWT and DTCWT

In this chapter, discrete wavelet transform and dual tree complex wavelet transform are discussed in the following subsections:

3.2.1 Discrete Wavelet Transform

Due to the multiresolution property, discrete wavelet transform (DWT) is widely used in image processing. DWT is a technique, which converts an image from spatial domain to frequency domain [70, 77]. It is used to analyze an image at different resolutions. We can obtain horizontal, vertical and diagonal information of the images using DWT. At first level decomposition, DWT decomposes the image into two parts: approximation and detailed parts. Approximation part *(LL)* contains one low frequency subband and detailed parts contain three high frequency subbands are labeled as *(LH, HL* and *HH)*, as shown in Fig. 3.1 (a). Most of the information of image is contained in approximation part. For second level decomposition, approximation part is further decomposed into four frequency subbands, as shown in Fig. 3.1 (b). The decomposition levels can be increased as per the requirement.

2-D DWT [44] for image *f(x, y)* of size $m \times n$ is defined as

$$W_{\phi}(j_0, u, v) = \frac{1}{\sqrt{mn}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) \phi_{j_0,u,v}(x, y) \tag{3.2.1}$$

$$W_{\psi}(j, u, v) = \frac{1}{\sqrt{mn}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) \psi_{j,u,v}(x, y) \tag{3.2.2}$$

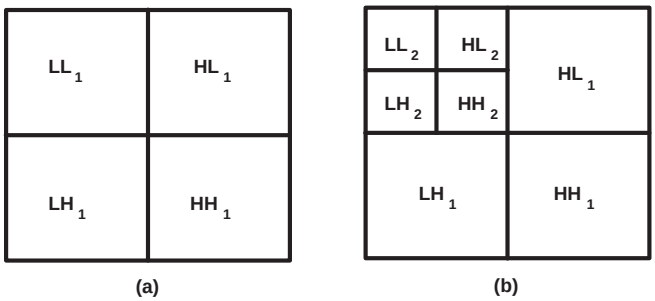

Figure 3.1: DWT decomposition. (a) First level 2-D DWT decomposition; (b) second level 2-D DWT decomposition.

where, Eqs. (3.2.1) and (3.2.2) are corresponding to approximation and detailed coefficients of image *f(x, y)*.

Conversely, an inverse DWT is used to reconstruct the image. For the above given Eqs. (3.2.1) and (3.2.2), the inverse DWT is given as

$$f(x,\ y) = \frac{1}{\sqrt{mn}} \sum_{u} \sum_{v} W_{\phi}(j_0, u, v)\phi_{j_0,u,v}(x, y) + \frac{1}{\sqrt{mn}} \sum_{j=j_0}^{\infty} \sum_{u} \sum_{v} W_{\psi}(j, u, v)\psi_{j,u,v}(x, y) \tag{3.2.3}$$

The first and second level DWT decomposition of Lena image are shown in Fig. 3.2. The original Lena image is given in Fig. 3.2 (a) whereas, Fig. 3.2 (b) and 3.2 (c) show the first and second level DWT decomposition of the original image.

3.2.2 Dual tree complex wavelet transform

The motivation for developing DTCWT is to overcome the disadvantages of discrete wavelet transform. DWT has several disadvantages such as, lack of shift invariance, aliasing and poor directionality due to large changes in wavelet coefficients and down-sampling. Because of shift variance, small changes in the input signals can cause large

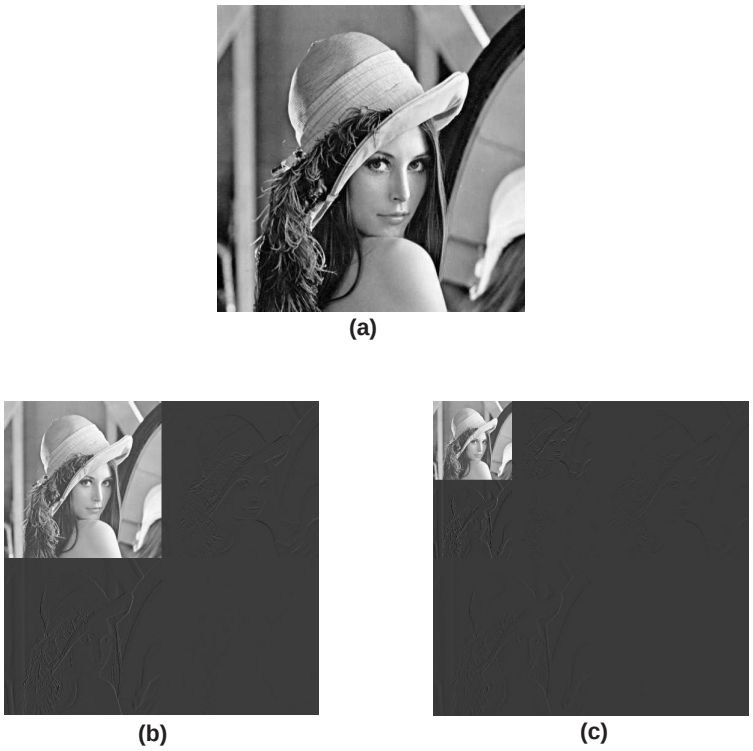

Figure 3.2: DWT decomposition of Lena image. (a) Lena image; (b) first level 2-D DWT decomposition; (c) second level 2-D DWT decomposition.

changes in the DWT coefficients. DWT has inability to distinguish between positive and negative frequencies because of poor directionality [67, 139]. These disadvantages of DWT can be overcome by using complex wavelet transform (CWT). Selesnick et al. [116] and Kingsbury [61] proposed a DTCWT, which yields approximate shift invariance and directional selectivity. These advantages of DTCWT provide better image fusion results than DWT [157]. Fig. 3.3 shows the decomposition and reconstruction of DTCWT. In which, DTCWT decomposes the input signal into real and imaginary parts. The obtained real and imaginary coefficients are used to compute amplitude and phase information. The DTCWT uses two DWTs over the same data

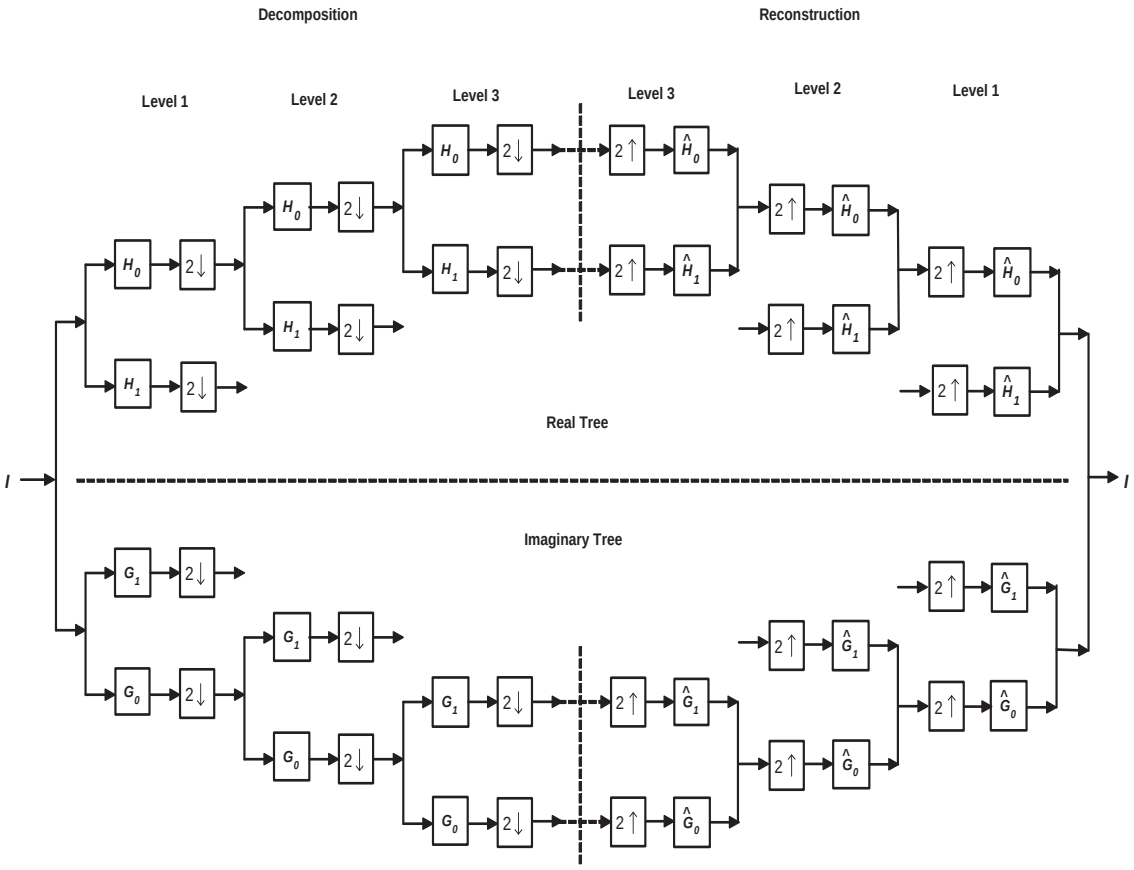

Figure 3.3: 3-level DTCWT decomposition and reconstruction of DTCWT coefficients with filters H_0 and H_1 for real part decomposition, G_0 and G_1 for imaginary part decomposition, $\hat{H}_0$ and $\hat{H}_1$ for real part reconstruction, $\hat{G}_0$ and $\hat{G}_1$ for imaginary part reconstruction.

and filters are designed as the upper DWT represents real part and the lower DWT refers as imaginary part [10, 11].

At each level of decomposition, DTCWT decomposes an image into two low and six distinct high frequency sub-bands at $\pm 15°$, $\pm 45°$ and $\pm 75°$ orientation. 2-D DTCWT [157] splits an image I to different scales y^j as:

$$y = \{x^j, y^1, y^2, y^3....., y^J\} \tag{3.2.4}$$

$$y^j = \begin{cases} y^j_{real,1}(l,k), y^j_{real,2}(l,k), ..., y^j_{real,6}(l,k) \\ y^j_{imag,1}(l,k), y^j_{imag,2}(l,k), ..., y^j_{imag,6}(l,k) \end{cases}$$

where x^j and y^j represent low and high frequency sub-bands, where y^j can be obtained from the combination of six real and imaginary directional sub-bands and (l,k) denotes spatial position of the coefficients and $d = 1, 2, ...6$ represents the orientation.

3.3 The proposed image fusion scheme

In this Section, the brief discussion of two proposed image fusion methods using gradient method is presented. In both of the fusion methods gradient method is employed to obtain the fused image. In the first method, gradient based approach for image fusion in discrete wavelet transform domain has been given. In the second method, an image fusion method using gradient rule in dual tree complex wavelet transform has been presented.

3.3.1 Gradient based image fusion in DWT domain (G-DWT)

In this subsection, an approach for image fusion in wavelet domain has been presented where the source images are of same view but with different focuses. The complete overview of proposed method for G-DWT is depicted in Fig. 3.4 and which can be described in the following Algorithm:

3.3.2 Algorithm for G-DWT method

1. Take two multi-focus source images A and B.

2. Perform DWT to decompose the source images into low and high frequency sub-bands.

3. Apply pixel averaging fusion method on low frequency coefficients of both transformed images in approximation parts.

4. Apply gradient based fusion rule on high frequency coefficients of both transformed images in detail parts.

5. Perform an inverse DWT on fused wavelet coefficients to achieve the final fused image (I_f).

In G-DWT method, two multi-focus images are taken. Over the each source image DWT is performed by which images are decomposed into low-pass *(L)* and high-pass *(H)* frequency sub-bands. The main idea of our proposed scheme is to improve the sharpness of image and reconstruct the enhanced image by selecting the sharp details from high frequency sub-bands. For that, high frequency coefficients of both transformed images are fused using gradient based method. The gradient coefficients of both of the transformed images are calculated using high frequency sub-bands and compared. If gradient coefficients of one image is greater than other image then pixel wise averaging method is performed otherwise maximum method is opted. Similarly, on low frequency sub-bands, simple pixel averaging method is performed. Finally, the resultant fused image is reconstructed by performing inverse DWT on new fused

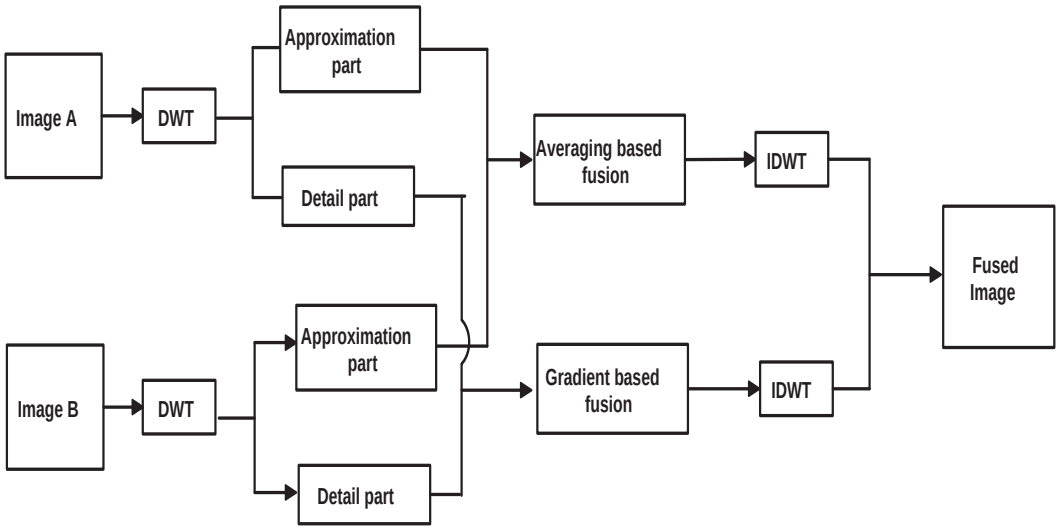

Figure 3.4: Block diagram for G-DWT fusion method.

wavelet coefficients. The sharpness focus measure [152] is computed as:

$$\nabla G(p) = [\nabla G_x(p)^2 + \nabla G_y(p)^2]^{1/2} \tag{3.3.5}$$

where $\nabla G_x(p), \nabla G_y(p)$ can be defined as:

$$\nabla G_x(p) = \Big\{ -D(s-1,t-1,u,v) - 2D(s-1,t,u,v) - D(s-1,t+1,u,v) \\ +D(s+1,t-1,u,v) + 2D(s+1,t,u,v) + D(s+1,t+1,u,v) \Big\}$$

$$\nabla G_y(p) = \Big\{ D(s-1,t-1,u,v) + 2D(s,t-1,u,v) + D(s+1,t-1,u,v) \\ -D(s-1,t+1,u,v) - 2D(s,t+1,u,v) - D(s+1,t+1,u,v) \Big\}$$

In this method, let *p=(s, t, u ,v)* be represent the index of a particular multiscale decomposition coefficient, where s and t represent spatial position, u and v denote

decomposition level and frequency band of multi-scale decomposition representation. However, G-DWT fusion method provides better fusion results but edges are not much clear for multisensor/multimodal images. Therefore, our next approach is based on gradient based method using DTCWT which is described in the following section:

3.3.3 Gradient based image fusion in DTCWT domain (G-DTCWT)

The complete overview of G-DTCWT method is depicted in Fig. 3.5 and the proposed fusion framework can be summarized in Algorithm as follows:

3.3.4 Algorithm for G-DTCWT method

1. Take two source images A and B.

2. Apply j-level DTCWT decomposition on the source images to achieve low x^j and six high frequency $y^j(j = 1, 2, ..., J)$ sub-bands at each level.

$$A : (x_A^j, y_{A,d}^j) \quad , \quad B : (x_B^j, y_{B,d}^j) \tag{3.3.6}$$

   where, x_*^j and $y_{*,d}^j$ are low and high frequency sub-bands at level j in the orientation d and $*$ denotes image A or B.

3. Fuse the high frequency sub-bands of source images by selecting largest absolute value.

$$y_{F,d}^j(l,k) = \begin{cases} y_{A,d}^j(l,k), & y_{A,d}^j(l,k) \geq y_{B,d}^j(l,k) \\ y_{B,d}^j(l,k), & otherwise \end{cases} \tag{3.3.7}$$

4. Perform gradient rule on low frequency sub-bands (x_A^j, x_B^j) of both of the transformed images and obtain gradient coefficients $(x_{A'}^j, x_{B'}^j)$. The gradient coefficients [152] are computed from the Eq. 3.3.5.

5. Compare the $x_{A'}^j$ and $x_{B'}^j$ by a step function and obtain the decision map

$$D(l,k) = \begin{cases} 1, & if(x_{A'}^j(l,k) > x_{B'}^j(l,k)) \\ 0, & otherwise \end{cases} \tag{3.3.8}$$

6. Select the pixels from the (x_A^j, x_B^j) using decision map $D(l,k)$ to obtain new fused low coefficients x_F^j.

$$x_F^j(l,k) = D(l,k)x_A^j + (1 - D(l,k))x_B^j \tag{3.3.9}$$

7. Reconstruct the final fused image F by performing j-level inverse DTCWT on the new high $(y_{F,d}^j(l,k))$ and low $(x_F^j(l,k))$ frequency sub-bands.

Here, in G-DTCWT fusion method, two source images A and B of same size are taken to produce a single fused image F. In the proposed framework, to obtain the low and high frequency sub-bands DTCWT is performed on the source images. The high frequency sub-bands usually include sharp details such as edges, boundaries and texture of the image. The most popular way of selection for high frequency sub-bands is the selection of the largest absolute values, therefore to fuse high frequency sub-bands maximum method is used. The low frequency sub-bands represent the approximation part which contains most of the information of image. The simplest rule is pixel averaging method to produce composite bands but it may not provide

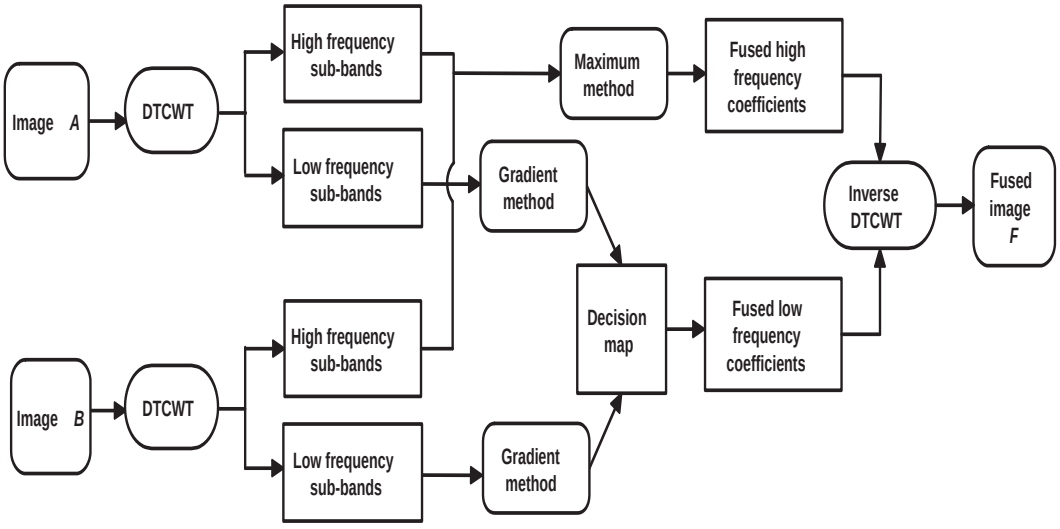

Figure 3.5: Schematic diagram for G-DTCWT fusion framework.

the high quality images because of the contrast reduction. Although, the averaging method is widely used for the fusion of approximation parts but the gradient based method may fuse the directional and smooth changes also. Therefore, gradient based fusion rule is employed to fuse the approximation parts. Finally, to achieve the final fused image an inverse DTCWT is performed on the fused coefficients.

3.4 Experimental results

In this section, the proposed scheme is tested on some standard images such as mandrill, lena and barbara each of size 512×512. Fig. 3.6 (a)-3.8 (a) are considered as the ideal reference images denoted as (I_r). Fig. 3.6 (b)- 3.8 (b) and 3.11 (a) are upper side blurred images, while other Fig. 3.6.(c)- 3.8 (c) and 3.11 (b) are lower side blurred images. Fig. 3.12 (a)-(b) are left and right part blurred images. Fig. 3.11

and Fig. 3.9 are dataset1 and dataset2 (FLIR and LLTV) images. Fig. 3.10 (a)-(b) are medical (CT and MRI) images. Fig. 3.9 (a)-(b) and Fig. 3.11 (a)-(b) images are obtained from the support of Dr. V.P.S Naidu [92]. Fig. 3.10 (a)-(b) and Fig. 3.12 (a)-(b) images are obtained from the link http://www.metapix.de/download.htm. In G-DWT based method, Haar as the wavelet basis is used upto 2-level decomposition. But before applying DWT, reference images are blurred with Gaussian filter of 13×13 window with standard deviation (σ) =5. The results obtained from G-DWT method are compared against some existing method such as minimum method [143], maximum method [125] and directive contrast method [14] and the results from the existing methods are shown in Fig. 3.6 (d-f)-3.8 (d-f). The results obtained from the proposed method G-DWT are shown in Fig. 3.6 (g)-3.8 (g). Further, the G-DTCWT proposed method is compared with some existing methods such as DWT (avg-max) [70, 100], PCA [94] and DTCWT [120] and the results obtained from existing methods are shown in Fig. 3.9 (c-e)-3.12 (c-e). The results obtained from G-DWT and G-DTCWT proposed methods are illustrated in Fig. 3.9 (f-g)-3.12 (f-g) and it can be easily analyzed from the experimental results that obtained results from G-DTCWT method provide better results than existing methods.

From the visual inspection, it can be easily seen that the results obtained from proposed methods are better than existing methods. The proposed fusion methods are compared in two aspects of quality measurements i.e., subjective and objective.

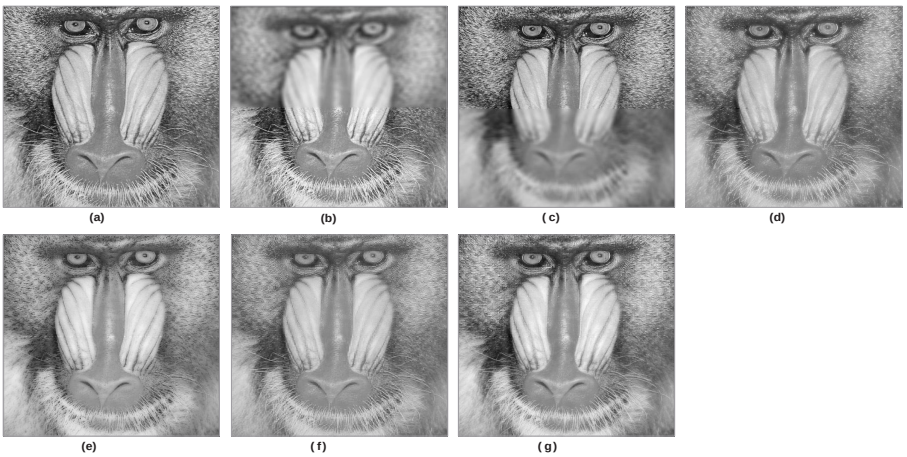

Figure 3.6: Fusion results for mandrill image. (a) Reference image; (b) upper side blurred image; (c) lower side blurred image; (d) maximum method; (e) minimum method; (f) directive contrast method; (g) G-DWT method.

Figure 3.7: Fusion results for lena image. (a) Reference image; (b) upper side blurred image; (c) lower side blurred image; (d) maximum method; (e) minimum method; (f) directive contrast method; (g) G-DWT method.

Figure 3.8: Fusion results for barbara image. (a) Reference image; (b) upper side blurred image; (c) lower side blurred image; (d) maximum method; (e) minimum method; (f) directive contrast method; (g) G-DWT method.

Table 3.1: Objective evaluation of the fused images.

Proposed method	**Source images**	**Fusion rules**	**PSNR w.r.t reference image**
G-DWT	Mandrill	[143]	28.3171
		[125]	29.3383
		[14]	28.6002
		G-DWT	29.8569
	Lena	[143]	32.1539
		[125]	34.09539
		[14]	34.2217
		G-DWT	35.7386
	Barbara	[143]	29.9535
		[125]	31.3625
		[14]	31.9447
		G-DWT	33.1251

Table 3.2: Performance measurement of the fused images.

Source images	**Evaluation indices**	[70, 100]	[94]	[120]	G-DWT	G-DTCWT
FLIR and LLTV	Mean	84.3786	151.312	84.378	99.451	99.663
	S.D.	48.1265	96.4685	63.2763	74.082	74.1559
	APSNR	17.7146	16.3124	17.2712	17.564	17.7941
	$Q^{AB/F}$	0.4365	0.3987	0.5511	0.5142	0.5238
CT and MRI	Mean	32.0820	51.8274	31.8523	46.259	46.5972
	S.D.	35.9304	54.1734	58.1705	59.236	60.1984
	APSNR	16.0520	15.0043	15.7646	15.265	14.6722
	$Q^{AB/F}$	0.5035	0.6518	0.7255	0.6345	0.7883
Saras	Mean	227.666	227.666	227.666	227.666	227.666
	S.D.	46.3984	45.9007	49.9141	49.924	50.0429
	APSNR	31.6325	31.7550	31.9686	31.996	31.9658
	$Q^{AB/F}$	0.5941	0.6135	0.7368	0.7125	0.7371
Clock	Mean	97.0389	97.037	97.038	97.264	97.556
	S.D.	49.4888	49.3160	51.7481	51.9264	51.9456
	APSNR	36.2591	33.6275	34.1183	34.154	33.9563
	$Q^{AB/F}$	0.6131	0.3020	0.6688	0.6688	0.6691

For objective evaluation of proposed methods some quality measure such as PSNR, mean, *S.D.*, *APSNR* and $Q^{AB/F}$ are used. From the results it can be analyzed that the proposed fusion method produce better informative synthesized image than other methods. The results obtained from the above existing methods and both proposed methods are illustrated in Table. 3.1 and 3.2.

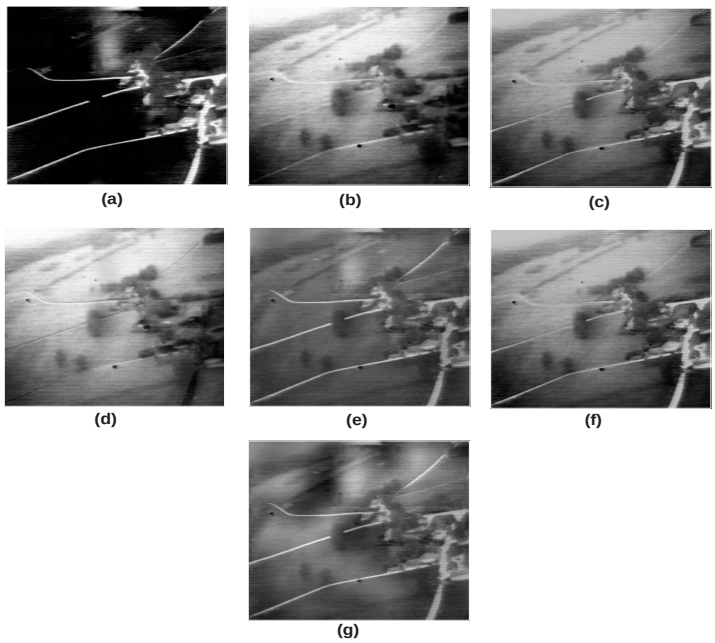

Figure 3.9: (a) FLIR; (b) LLTV; (c) DWT; (d) PCA; (e) DT-CWT; (f) G-DWT; (g) G-DTCWT method.

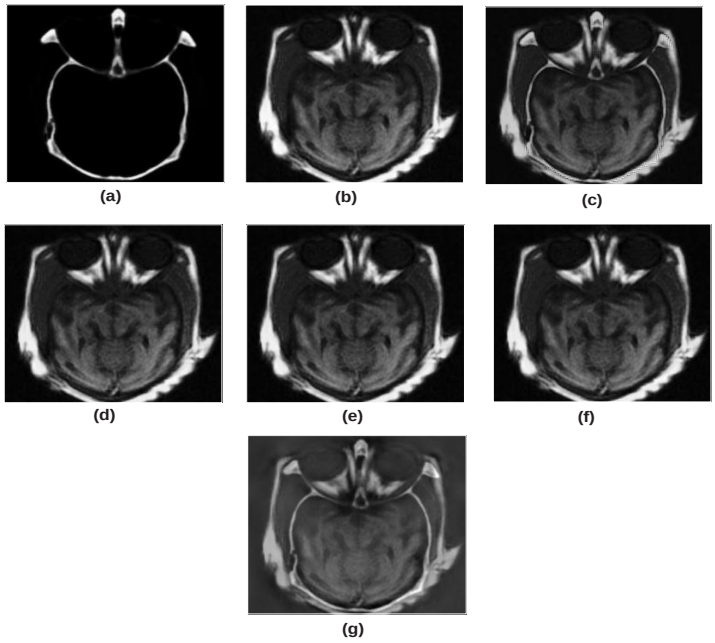

Figure 3.10: (a) CT; (b) MRI; (c) DWT; (d) PCA; (e) DTCWT; (f) G-DWT; (g) G-DTCWT method.

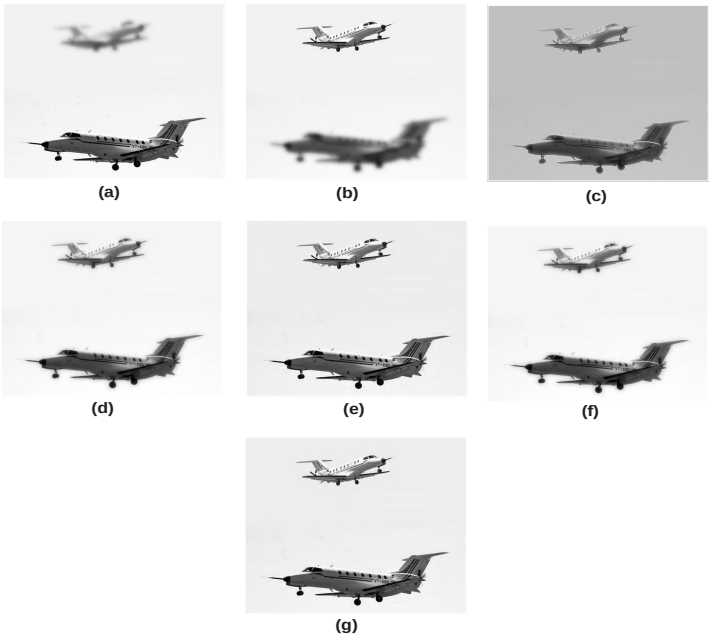

Figure 3.11: (a) Image blurred on upper part; (b) image blurred on lower part; (c) DWT; (d) PCA; (e) DTCWT; (f) G-DWT; (g) G-DTCWT method.

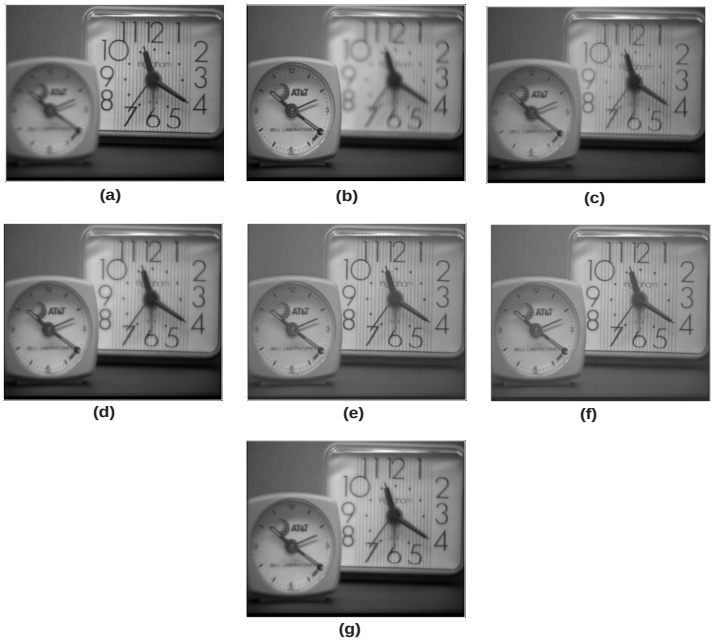

Figure 3.12: Clock image. (a) Image blurred on left part; (b) image blurred on right part; (c) DWT; (d) PCA; (e) DTCWT; (f) G-DWT; (g) G-DTCWT method.

3.5 Conclusions

In this chapter, two novel, simple, effective and different fusion approaches (G-DWT and G-DTCWT) based on gradient rule have been discussed. The first proposed image fusion method (G-DWT) is based on gradient method using discrete wavelet transform. The next proposed image fusion method (G-DTCWT) is based on gradient method in dual tree complex wavelet transform. Both of the proposed methods provide better outcomes than other existing methods. The main contribution of these works is to increase the sharpness of fused image. The G-DWT method gives better results for multifocus images and erroneous results for multisensor/multimodal images. Whereas, G-DTCWT method provides better results for multifocus as well as multisensor images. Apart from subjective evaluation, objective evaluation has also been performed using some standard metrics. The proposed techniques provide good quality image with better interpretation but these techniques may not be applicable on low contrast images. Therefore, next work is developed for better contrast in the fused images.

CHAPTER 4

Directive contrast based image fusion

Chapter 4

Directive contrast based image fusion

The work discussed in the previous chapter provides a good quality image with better visual interpretation but it may not be acceptable for low contrast images. Hence, a new image fusion technique has been developed here using directive contrast to provide a better contrast fused image.

4.1 Introduction

Various methods based on transform domain have been explored which produce good results with less computation and in less time and proved suitable for fusion purposes. At different resolutions, transform domain methods contain unique information and also provide directional information in all decompositions. The basic aim of transform domain based methods is to perform multiresolution decomposition on

each source image and generate a composite image by combining all these decompositions using certain fusion rules and finally, inverse multiresolution transform has been performed to reconstruct the fused image [164]. The simplest fusion method is average method [78], in which pixel by pixel average values are calculated. But this method, reduces the contrast. For the problem of low contrast, Tian et al. [130] have introduced a contrast based image fusion algorithm in wavelet transform domain. In this algorithm, a novel concept of directive contrast has been given for the fusion of visible and infrared images. The multiresolution architectures are obtained from the source images using DWT. After that corresponding subbands of each source image are selected using directive contrast method to obtain the multiresolution architecture of the resultant image. At the end, an inverse wavelet transform is performed to reconstruct the resultant fused image. Furthermore, an another wavelet packet transform (WPT) domain based fusion technique is introduced to increase the contrast of the fused image [14]. The low frequency coefficients are combined using median rule and high frequency coefficients are combined through directive contrast method in wavelet packet transform domain. For medical image fusion, a directive contrast based image fusion approach using NSCT domain has been proposed to extract the most significant detail from source images into the fused image [16]. The medical images are considered as source images and can be transformed using NSCT which decomposes an image into low and high frequency subbands. These frequency subbands are fused using phase congruency and directive contrast based fusion methods. After combining

all the new fused frequency subbands an inverse NSCT is performed to obtain a fused image. In [70], DWT based fusion has been performed using maximum selection rule. This scheme selects the largest absolute wavelet coefficients at each location from the source images and integrates these into a single fused image. In PCA based fusion technique [94], source images are transformed into uncorrelated images and principal components are evaluated from eigenvalues. These principal components represent variance of the pixels and contribute to the weights used for fusion process. Li et al. [69] have proposed a fusion technique using lifting stationary wavelet transform domain. In this method, to fuse the low and high frequency subbands new sum modified laplacian (NSML) focus measure and local feature contrast of the orientation information has been used. This method provides the sharp boundaries and salient features of an image. Using the fuzzy transform (FTR) domain, an another new image fusion technique has been introduced based on the concept of directional contrast [95]. In which, source images were divided into several non overlapping blocks. The obtained sub-blocks were combined using directional contrast based fusion rule in FTR domain. By using inverse FTR, fused sub-blocks were transformed into original block size. Further, to reconstruct the final fused image, these transformed blocks were merged using maximum selection based fusion rule.

In this chapter, an effective multi-focus image fusion based on directive contrast scheme is proposed. Due to the advantages of easy implementation and time efficient, the work is using pixel-level image fusion. The objective of this chapter is to fuse two

multi-focus (defocused) images of same scene into a single image with entire scene in focus for better visual information. In this chapter, a local correlation and directive contrast based image fusion method in discrete wavelet packet transform domain is proposed.

The rest of the chapter is organized as follows. Section 4.2 describes the basic concepts of wavelet packet transform and directive contrast. In Section 4.3, the proposed fusion method is explained. Experimental results and discussions are given in Section 4.4. Finally, Section 4.5 concludes the chapter.

4.2 Basic theory of WPT and directive contrast

In this chapter, wavelet packet transform and directive contrast are used to fuse the multifocus images. The basic theories of wavelet packet transform and directive contrast are discussed in the following subsections.

4.2.1 Wavelet Packet Transform

The discrete wavelet packet transform (DWPT) is the generalization of the DWT and provides a more flexible tool for time-scale analysis of the data [138, 140]. It retains all the properties of wavelet transform because the wavelet basis is in the repertoire of bases available with the wavelet packet transform [14]. For the first level decomposition, when DWT is applied over an image, it decomposes the image into low and high frequency subbands. Where, the low frequency subband is usually referred as the approximation part and the high frequency subbands as the detail

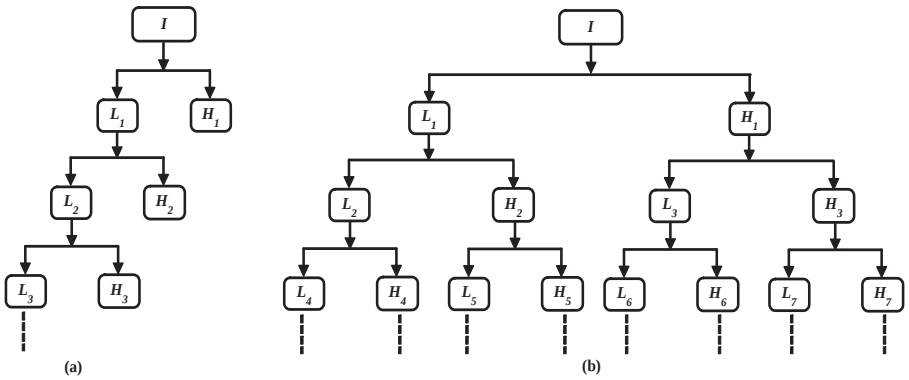

Figure 4.1: (a) Wavelet decomposition; (b) wavelet packet decomposition.

parts containing horizontal, vertical and diagonal information. For the second level decomposition, low frequencies are further decomposed into another set of low and high frequency subbands and the process can be repeated upto the desired level [5], as given in Fig. 4.1 (a). During the WPT decomposition, process is iterated on both low and high frequency subbands as shown in Fig. 4.1 (b). Hence, wavelet packet transform decomposes the frequency space into more number of frequency subbands and provide better frequency localization of the signals. Finally, the image can be reconstructed by performing inverse WPT.

4.2.2 Directive Contrast

The concept of local luminance contrast of an image has been developed by Toet et al. [132]. The contrast can be defined as [46, 130]:

$$R = (L - L_B)/L_B = L_H/L_B \tag{4.2.1}$$

where, R and L represent the contrast of the image and local grey level (luminance),

L_B local brightness of the background which corresponds to local low frequency component and $L_H = L - L_B$ corresponds to local high frequency components.

The directive contrast for *l-level* can be defined as the ratio of high frequency subbands $(H_{l,p}, V_{l,p}, D_{l,p})$ and low frequency subbands $A_{l,p}$ [14].

$$Horizontal\ contrast : R^H_{l,p} = \frac{H_{l,p}}{A_{l,p}} \tag{4.2.2}$$

$$Vertical\ contrast : R^V_{l,p} = \frac{V_{l,p}}{A_{l,p}} \tag{4.2.3}$$

$$Diagonal\ contrast : R^D_{l,p} = \frac{D_{l,p}}{A_{l,p}} \tag{4.2.4}$$

where, $1 \leq p \leq 2^{l+1}$.

In this work, directive contrast is used for high frequency subbands of discrete wavelet packet transform (DWPT). Where, DWPT decomposes the image into four frequency subbands, one low frequency component and three high frequency components. The level wise different high frequencies are selected for finding directive contrast.

4.3 Proposed Method

In this section, the proposed fusion method is described in detail. For multi-focus image fusion, two multi-focus images are required which are considered as source images in our method. Let us consider, each source images of size $m \times n$. The proposed fusion method can be summarized by using the following Algorithm:

4.3.1 Algorithm for proposed method

1. Take two source images X and Y.

2. Perform DWPT over the source images which decompose the images into low and high frequency subbands.

3. Apply correlation based fusion method on low frequency subbands ($A^X_{l,p}$, $A^Y_{l,p}$) of both the transformed images.

4. Apply directive contrast based fusion rule on high frequency coefficients (($H^X_{l,p}, V^X_{l,p}, D^X_{l,p}$) and ($H^Y_{l,p}, V^Y_{l,p}, D^Y_{l,p}$)) of both transformed images.

5. Perform an inverse DWPT on fused wavelet coefficients to achieve the final fused image (I_f).

In the proposed work, DWPT is applied upto 2-level decomposition over the source images X and Y. It decomposes X and Y images into one low ($A^X_{l,p}$, $A^Y_{l,p}$) and three high frequency coefficients (($H^X_{l,p}, V^X_{l,p}, D^X_{l,p}$), ($H^Y_{l,p}, V^Y_{l,p}, D^Y_{l,p}$)) respectively, as discussed in Section 4.2. The block diagram of proposed method is depicted in Fig. 4.2. The low and high frequencies obtained by DWPT are fused as given in following sections.

4.3.2 Fusion of low frequency coefficients

The obtained low frequency coefficients ($A^X_{l,p}$, $A^Y_{l,p}$) are fused using local correlation based approach. In this method, block wise correlation coefficients are computed

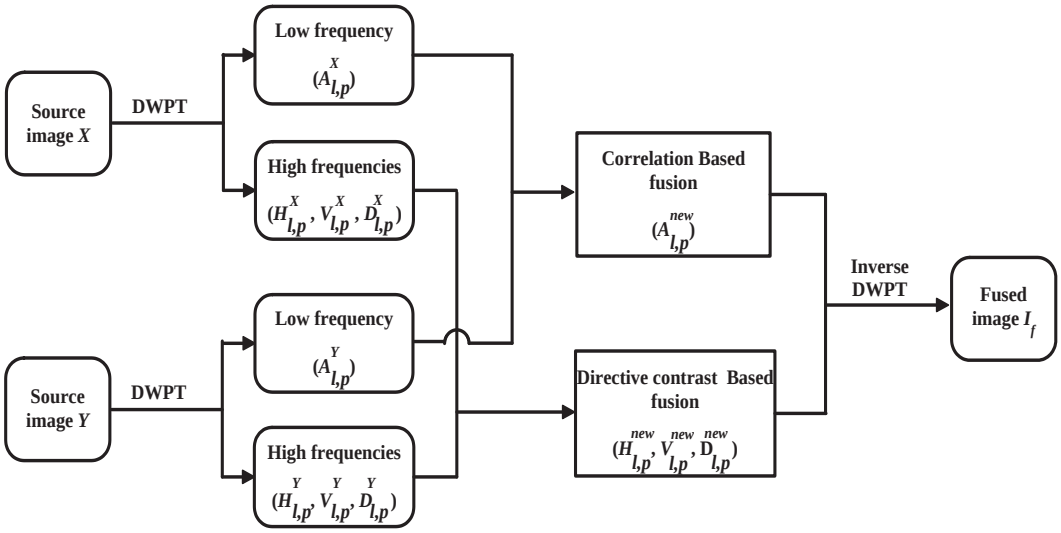

Figure 4.2: Schematic diagram for proposed fusion method.

from these low frequencies using 8×8 block size. These correlation coefficients are computed using the formula which have already discussed in Chapter 1.

In correlation based fusion strategy, obtained correlation value (CC) is compared with the threshold value (T). Here, in this work the threshold value T is considered as 0.6. If the value of correlation coefficients is less than or equal to the threshold value, then maximum method is performed. In maximum method, fusion is performed by selecting the largest values from both of the transformed images. Otherwise averaging method is employed, which computes the average value using both the transformed images to perform fusion. Correlation based fusion strategy is given as follows:

$$A^{new}_{l,p} = \begin{cases} max(A^X_{l,p}, A^Y_{l,p}), & if \quad (CC \leq 0.6) \\ avg(A^X_{l,p}, A^Y_{l,p}), & otherwise \end{cases} \tag{4.3.5}$$

where, max and avg stand for maximum and average values, respectively and $A^{new}_{l,p}$ for the fused coefficients.

4.3.3 Fusion of high frequency coefficients

The human visual system is highly sensitive to local luminance contrast. Therefore, the directive contrast discussed in Section 4.2 provides a suitable criterion. The high frequency coefficients $(H^X_{l,p}, V^X_{l,p},\ D^X_{l,p})$ and $(H^Y_{l,p}, V^Y_{l,p},\ D^Y_{l,p})$ from both of the transformed images are fused using directive contrast method. The horizontal, vertical and diagonal contrasts can be obtained as [14]:

$$H^{new}_{l,p} = \begin{cases} H^X_{l,p}, & if \ \ |R^{H,X}_{l,p}| \geq |R^{H,Y}_{l,p}| \\ H^Y_{l,p}, & otherwise \end{cases} \tag{4.3.6}$$

$$V^{new}_{l,\ p} = \begin{cases} V^X_{l,p}, & if \ \ |R^{V,X}_{l,p}| \geq |R^{V,Y}_{l,p}| \\ V^Y_{l,p}, & otherwise \end{cases} \tag{4.3.7}$$

$$D^{new}_{l,p} = \begin{cases} D^X_{l,p}, & if \ \ |R^{D,X}_{l,p}| \geq |R^{D,Y}_{l,p}| \\ D^Y_{l,p}, & otherwise \end{cases} \tag{4.3.8}$$

Obtained fused wavelet coefficients from low $(A^X_{l,p},\ A^Y_{l,p})$ and high frequency coefficients $((H^X_{l,p}, V^X_{l,p}, D^X_{l,p}),\ (H^Y_{l,p}, V^Y_{l,p}, D^Y_{l,p}))$ are represented as $A^{new}_{l,p}$ and $H^{new}_{l,p}, V^{new}_{l,p}, D^{new}_{l,p}$. These four new fused coefficients are used for inverse DWPT. After performing inverse DWPT, a fused image I_f can be obtained.

4.4 Experimental results and discussion

The proposed method has been applied on pepper, mandrill, pepsi images of size 512×512 and office image of size 256×256. Among these, pepper and mandrill images are considered as reference images I_r. But before applying the proposed

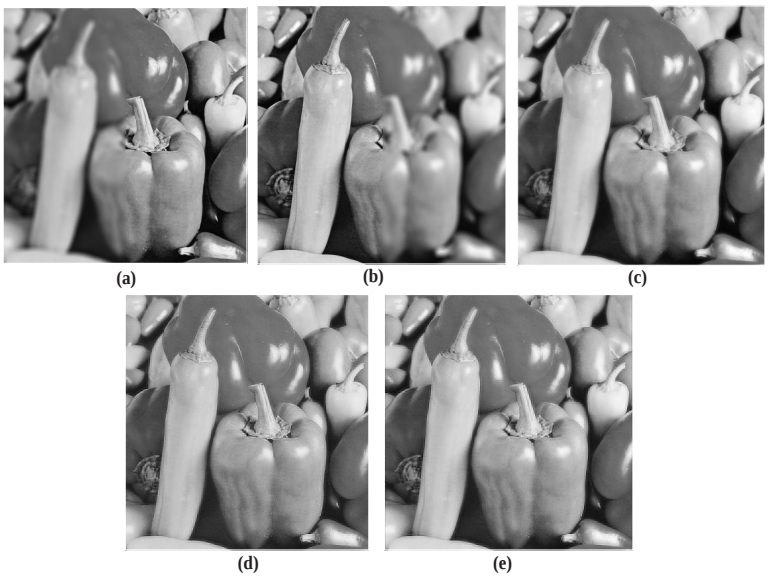

Figure 4.3: Fusion of pepper images. (a) Image blurred on the left part; (b) image blurred on the right part; (c) PCA; (d) directive contrast; (e) proposed method.

method, these two reference images are blurred by convolving a gaussian filter using 13×13 window and standard deviation $\sigma = 5$ and obtained images are referred as the source images. The source images are shown in Fig. 4.3 (a-b)-4.6 (a-b).

The pepper images, which are highly concentrated on right and left part are given in Fig. 4.3 (a) and 4.3 (b). Mandrill images are shown in Fig. 4.4 (a) and 4.4 (b) are blurred on upper and lower parts. In pepsi images, barcode and container are focused as given in Fig. 4.5 (a) and 4.5 (b). The office images are blurred on middle and corner as shown in Fig. 4.6 (a) and 4.6 (b), respectively. Over the source images, proposed fusion method is applied. The experimental results obtained by proposed method are compared with PCA [94] and directive contrast [14] based image fusion techniques. The obtained results by existing methods are shown in Fig. 4.3 (c-d)-4.6 (c-d). The results obtained by proposed method are shown in Fig. 4.3 (e)-4.6 (e).

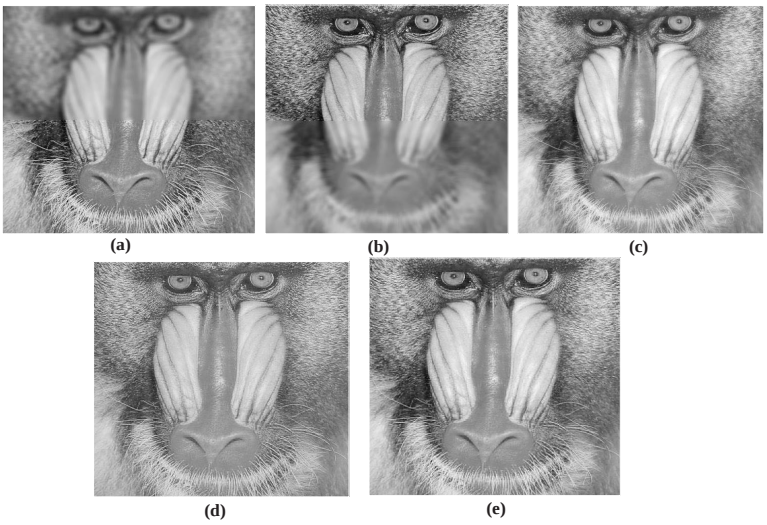

Figure 4.4: Fusion of mandrill images. (a) Image blurred on the upper part; (b) image blurred on the lower part; (c) PCA; (d) directive contrast; (e) proposed method.

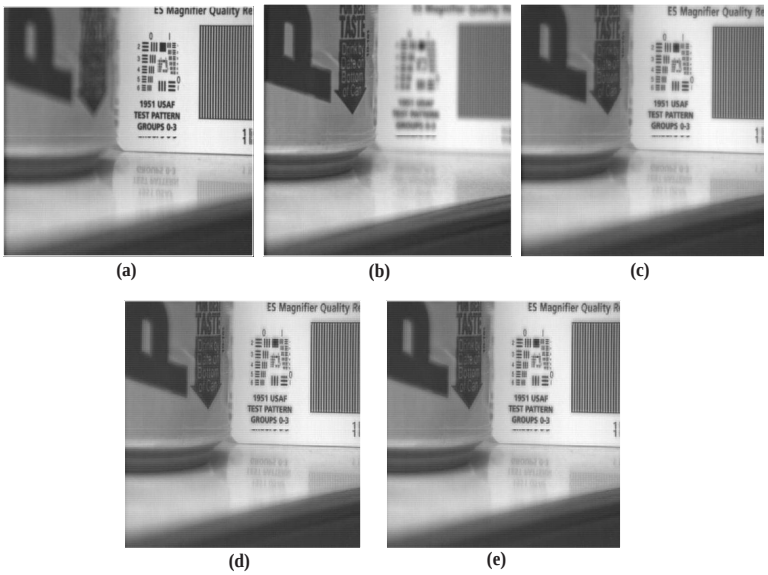

Figure 4.5: Fusion of pepsi images. (a) Image focus on barcode; (b) image focus on container; (c) fused image by PCA; (d) fused image by directive contrast; (e) fused image by proposed method.

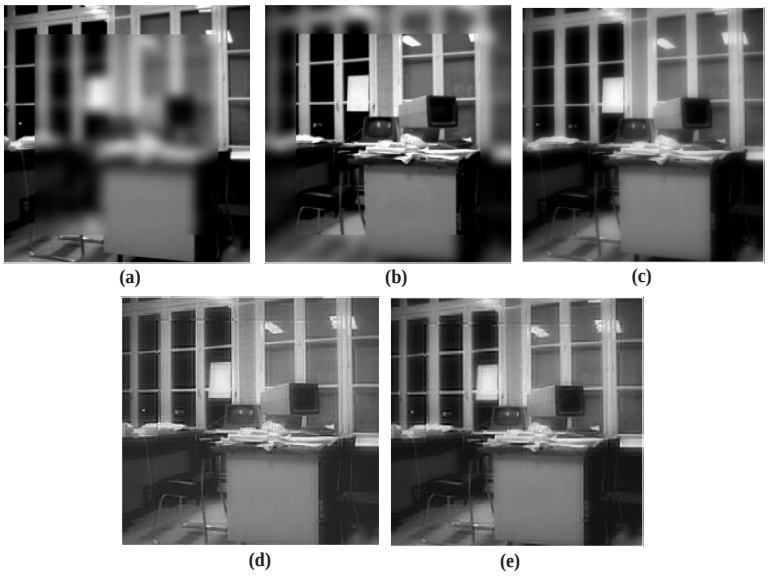

Figure 4.6: Fusion of office images. (a) Middle side blurred; (b) corner side blurred image; (c) PCA; (d) directive contrast; (e) proposed method.

From the visual results it can be observed that the fused images obtained from the proposed method are better quality images in term of contrast than other existing methods in most of the cases. But only visual inspection is not sufficient to measure the quality of fused images. In addition to human vision analysis, the fused images have also been compared through quantitative analysis. To evaluate the performance of existing methods and proposed method quantitatively, several parameters such as mean, standard deviation, correlation coefficients and PSNR are also computed. The obtained results are shown in Table. 4.1 and it can observed that the results of proposed method are better than other existing methods in most of the cases.

Table 4.1: Evaluation metrics using fused images

Source Images	Evaluation indices	PCA	Directive contrast	Proposed method
Peppers	Mean	119.1268	119.1555	119.5538
	S.D.	50.8325	51.1246	51.2989
	CC	0.9643	0.9847	0.9889
	PSNR	34.2197	34.5824	34.9147
Mandrill	Mean	128.4620	128.9738	129.5436
	S.D.	35.6187	38.0478	38.1671
	CC	0.9103	0.9189	0.9399
	PSNR	30.2266	30.2827	30.7864
Pepsi	Mean	97.5710	97.5725	97.8974
	S.D.	43.9852	44.3332	44.4092
Office	Mean	80.3662	80.3571	80.5871
	S.D.	59.4153	60.7419	60.9577

4.5 Conclusions

In this chapter, a multi-focus image fusion technique based on local correlation and directive contrast has been discussed using discrete wavelet packet transform domain. In the discussed work, DWPT has been used which provides better localization of low as well as high frequency subbands. After performing the DWPT decomposition, obtained low frequency coefficients are combined using local correlation and high frequency coefficients are fused using directive contrast method. Using these fusion techniques, we obtained an improved fused image in which the contrast and details from each original image are enhanced. Several pairs of multi-focus images are used to test the performance of the image fusion method. The experimental results demonstrate that the proposed method is better than existing methods qualitatively

as well as quantitatively.

This work described here for fused images having better contrast characteristic but may not provide appropriate sharp details and overall information of the source images. Therefore, the next approach is based on multiresolution singular value decomposition based image fusion.

CHAPTER 5

Multiresolution singular value decomposition based image fusion

Chapter 5

Multiresolution singular value decomposition based image fusion

The previous chapter was based on the concept of contrast enhancement to produce a fused image having better contrast but that method is usually does not provide the sufficient sharp details over the entire image. Therefore, in this chapter, a novel fusion technique by utilizing the virtues of multiresolution singular value decomposition (MSVD) has been developed to obtain the more sharp details in the fused image.

5.1 Introduction

In the recent development of technologies, image fusion has become an important and useful application for image analysis, image enhancement and computer vision [48, 89]. Image fusion is the process of integrating all significant information from two or more images of same scene into a single composite image. The main objective of

image fusion is to make the fused image more informative with better visual appearance. In image fusion, it is tried to obtain fused image without any distortion and loss of information. In [80], the performances of image fusion algorithms using multiresolution transforms were compared with CVT and contourlet to obtain better fusion results. Another, multifocus image fusion technique is given using Dual-Channel Pulse Coupled Neural Networks (Dual-Channel PCNN) [143] which incorporates focus measure of source images to compute weighted coefficients. In [148], multifocus image fusion and restoration is presented using sparse-representation. In this method, sparse coefficients were calculated from the source images and these coefficients were combined using maximum fusion rule. Similarly, another sparse representation image fusion technique [149] has been proposed based on overlapping patches instead of the whole image and simultaneously the orthogonal matching pursuit technique was used which decomposes the source images into the same subset of dictionary bases. Image fusion based on contourlet packet was introduced followed by a Nonsubsampled Directional Filter Bank (NSDFB) [151] which provided a more accurate fused image than wavelet packet method. The new multifocus image fusion method based on sharpness criterion [129] is given to enhance the sharp information and remove blur details of the images. In [108], multisensor image fusion in remote sensing was presented, it described mainly pixel based image fusion of earth observation satellite data. A multi-resolution image fusion [99] based approach was proposed to combine

the high and low resolution images data by adding some wavelet planes to the low resolution intensity component and obtained a better fused image. Again, DWT based fusion [70] was proposed, in which the maximum selection rule was used. This simple scheme just selects the largest absolute wavelet coefficients at each location from the input images. Similarly, another fusion technique was introduced based on wavelets and principal component analysis [94]. In which, source images were decomposed into low and high frequency subbands using DWT and obtained coefficients were transformed into uncorrelated coefficient and the eigenvalues are evaluated from principal components (PCs). The scheme selects the two highest principal components and represent these as weights for fusion rule. Finally, inverse DWT was performed to achieve the final fused image.

In [92], a multiresolution singular value decomposition based scheme was proposed which gave better results than DWT based scheme [70] both in terms of quality and execution time. In this scheme, maximum selection rule was performed over detailed coefficients and averaging was performed over approximation coefficients to get the final fused image. In DWT, multiresolution decomposition was based on low and high pass filters. Low pass filters outputs to approximation part which contains most of the information of image whereas, high pass filter outputs to detailed subbands which contain directional (horizontal, vertical and diagonal) information. However, on the other hand, the decomposition in MSVD was based on singular/eigenvalues. Here, the approximation part was corresponding to larger eigenvalues and contains

most of the information whereas, the detailed coefficient correspond to edges, texture, boundaries and other sharp changes in image.

With this motivation, we have proposed a new image fusion scheme using DWT and MSVD. In this work, the DWT is applied on both the source images and approximation and detailed components are obtained. Since the detailed components contain directional information therefore it is quite logical to apply the gradient based fusion rule directly on the detailed coefficients of both the images. However, the obtained approximation part is further decomposed into low and high frequency coefficients using MSVD. Again, the obtained low frequency coefficients contain the most of the average information, therefore the averaging method is the best suited for fusion, whereas the detailed coefficients are responsible for sharper changes and the maximum selection method is quite appropriate to fuse the detailed coefficients. Finally, the inverse MSVD and inverse DWT are performed appropriately on fused coefficients to get the fused image.

The remainder of the chapter is organized as follows: basic theory of MSVD is discussed in Section 5.2. Proposed scheme of image fusion is introduced in Section 5.3. Experimental results followed by discussion are given in Section 5.4. Finally, concluding remarks are given in Section 5.5.

5.2 Multiresolution Singular Value Decomposition (MSVD)

A singular value decomposition [15, 43, 57] of a matrix A of size $m \times n$ can be written as: $A = USV^T$, where U and V are left and right singular vectors of size $m \times m$ and $n \times n$ respectively. U and V are orthogonal matrices and S is a diagonal matrix of size $m \times n$ containing eigenvalues in non-increasing order.

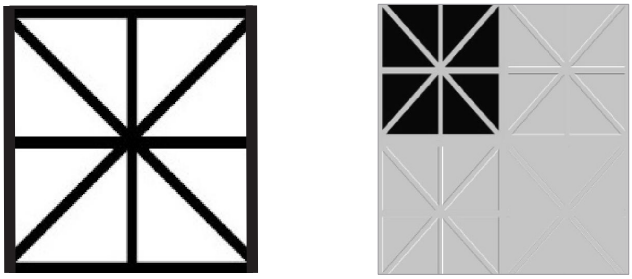

Figure 5.1: (a) Original image; (b) first level MSVD decomposition.

Fig. 5.1 (a) shows the original synthetic image. At first level decomposition, MSVD decomposes the image into two parts: approximation and detailed parts. Approximation part contains one low frequency subband *(LL)* and detailed parts contain three high frequency subbands *(LH, HL* and *HH)*. Fig. 5.1 (b) shows its first level decomposition using MSVD.

For a matrix A of size $m \times n$ the MSVD [163] is computed as follows:

1. Divide the matrix A into non-overlapping blocks of size $k \times l$ and write each block in the form of a vector.

2. Stack the vectors in columns and form the matrix A_1 of size $kl \times mn/kl$.

3. Compute the scatter matrix, $T_1 = \bar{A_1}\bar{A_1}^T$ of size $kl \times kl$.

4. Find the orthogonal matrix U_1 for scatter matrix T_1.

5. Compute the diagonal matrix S_1 containing the squares of eigenvalues in decreasing order as, $S_1^2 = U_1^T T_1 U_1$.

6. Construct matrix $\hat{A} = {U_1}^T \bar{A_1}$.

$\hat{A}$ contains approximation and detailed parts. First row of $\hat{A}$ is considered as approximation part containing larger singular values. Similarly, the remaining rows of $\hat{A}$ are regarded as the detailed parts containing remaining singular values, i.e. $\Phi_1 = \hat{A}(1,:)$ and $\Psi_1^i = \hat{A}(i,:)$, where, Ψ_1^i is the i^{th} detailed subbands at level 1 [92]. The first level MSVD decomposition of a synthetic image is shown in Fig. 5.1 (b) and the decomposition is based on eigenvalues.

5.3 Proposed Scheme

In this section, the proposed scheme is described in brief. In the proposed scheme, two less informative (multifocus and multisensor) images are used to obtain the fused image having more information. For multifocus and multisensor image fusion, two multifocus and multisensor images are considered as source images. In the proposed work, DWT is applied over the source images X and Y which decomposes each of the source images into low (LL_l^X and LL_l^Y) and high frequency subbands ((HL_l^X, LH_l^X

and HH_l^X) and (HL_l^Y, LH_l^Y and HH_l^Y)) respectively, which is already discussed in Chapter 3.

Over the low frequency subbands (LL_l^X and LL_l^Y) of both of the transformed images, MSVD is performed. When MSVD is applied, LL_l^X and LL_l^Y frequency subbands of both of the transformed images are further decomposed into low frequency ($\Phi_l^{LL^X}$ and $\Phi_l^{LL^Y}$) and high frequency subbands (($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$)) respectively, as discussed in Section 5.2. Low frequency subbands refers to approximation part and high frequency subbands treated as detailed parts. Obtained low frequency coefficients ($\Phi_l^{LL^X}$, $\Phi_l^{LL^Y}$) of both of the transformed images are fused using pixel averaging method. In pixel averaging fusion method, average pixel values from both of the transformed images are selected to fuse. Whereas, high frequency coefficients ($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$) of both of the transformed images are combined using the maximum method. In the maximum method, corresponding maximum values from both of the transformed images are selected to fuse. Obtained orthogonal matrices U_l^X and U_l^Y, as discussed in Section 5.2 are fused using averaging fusion method. Further, over the obtained fused coefficients from low and high frequency subbands, an inverse MSVD is performed and set of new fused coefficient is obtained and represented as ω_a.

Remaining high frequency coefficients (HL_l^X, LH_l^X, HH_l^X) and (HL_l^Y, LH_l^Y, HH_l^Y), as given in Fig. 5.2 of both of the transformed images are fused using gradient based sharpness focus method [152, 153]. The gradient coefficients are computed as

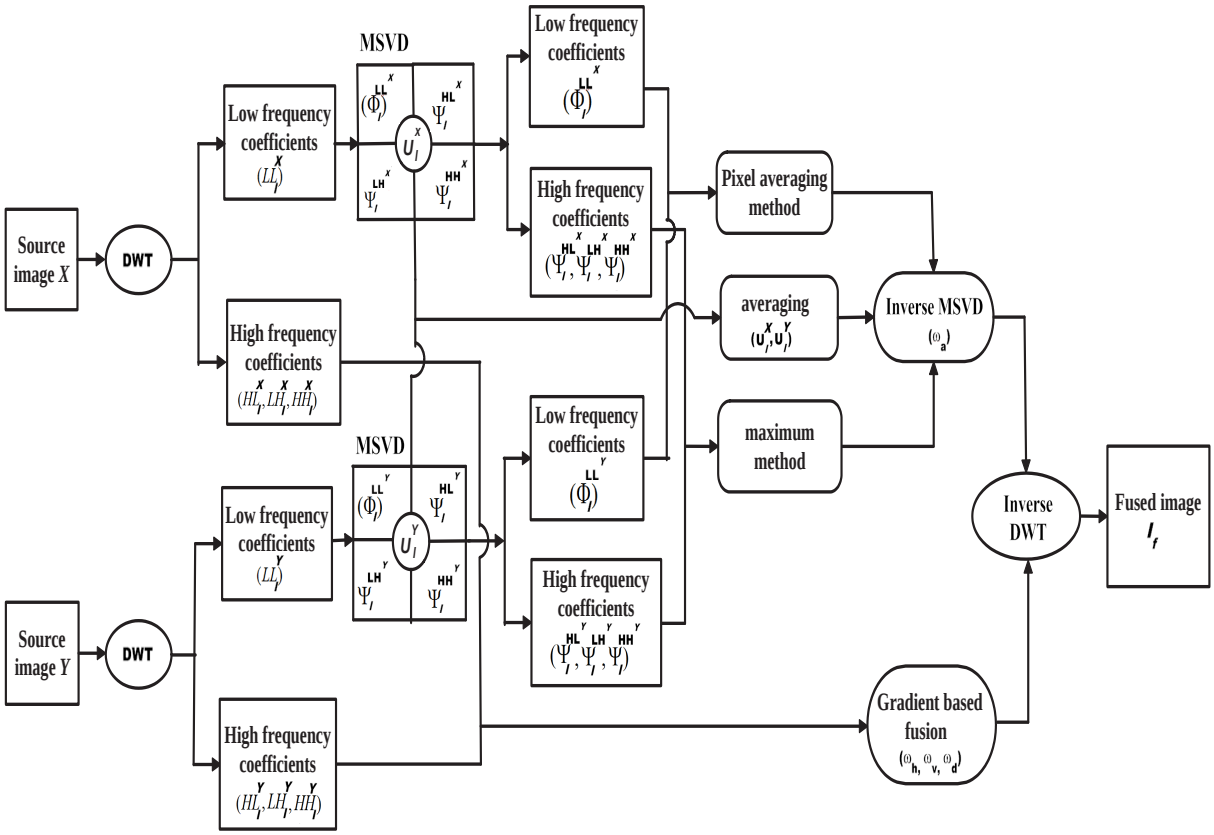

Figure 5.2: Block diagram of proposed scheme.

discussed in Chapter 3. Obtained gradient coefficients of both of the transformed images are fused using pixel averaging method and new fused coefficients (ω_h, ω_v and ω_d) are obtained. Finally, to reconstruct the final fused image (I_f), inverse DWT is performed over the obtained fused coefficients ($\omega_a, \omega_h, \omega_v$ and ω_d). The block diagram of the proposed scheme is given in Fig. 5.2. The proposed method is summarized in the following Algorithm:

5.3.1 Proposed fusion scheme algorithm

1. Take two source images X and Y.

2. Apply DWT over the source images, which decomposes each of the images into four subbands: (LL_l^X, HL_l^X, LH_l^X and HH_l^X) and (LL_l^Y, HL_l^Y, LH_l^Y and HH_l^Y).

3. Perform MSVD on approximation part (LL_l^X, LL_l^Y) of both of the transformed images and obtain approximation ($\Phi_l^{LL^X}$ and $\Phi_l^{LL^Y}$) and detailed parts (($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$)) (as given in Section 5.2).

4. Apply pixel averaging method on $\Phi_l^{LL^X}$ and $\Phi_l^{LL^Y}$ of both of the transformed images and maximum method on high frequency coefficients ($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$).

5. Apply averaging method on U_l^X and U_l^Y (as discussed in Section 5.2).

6. Apply inverse MSVD on fused coefficients and obtain fused coefficient ω_a.

7. Apply gradient based fusion on detailed parts (HL_l^X, LH_l^X and HH_l^X) and (HL_l^Y, LH_l^Y and HH_l^Y) using averaging method and obtain fused coefficients $(\omega_h, \omega_v, \omega_d)$.

8. Perform inverse DWT, over all the fused coefficients $(\omega_a, \omega_h, \omega_v, \omega_d)$ to reconstruct the final fused image (I_f).

5.4 Experimental results

The proposed scheme is tested on several test images of size 512×512 shown in Fig. 5.3 (a)-5.6 (a) and also of size 256×256 shown in Fig. 5.7 (a) and 5.8 (a). All these images are considered as reference images (I_r). In the proposed work, DWT method with Haar as the wavelet basis is used. Before applying DWT, each of the reference images in Fig. 5.3 (a), 5.4 (a), 5.7 (a) and 5.8 (a) are convolved by a gaussian blurring of 13×13 window with standard deviations $\sigma = 5$ and multi-focus (defocus) source images X and Y are obtained. Fig. 5.5 (a)-(c) are considered as dataset1 source images and Fig. 5.6 (a) and 5.6 (b) are FLIR and LLTV source images considered as dataset2 images [92].

Fig. 5.7 (a) and 5.8 (a) are the TNO's UN camp and head monument visual images, which are online available at `http://www.deakin.edu.au/~mhossny/fusion/`. Fig. 5.3 (b), 5.7 (b) and 5.8 (b) are blurred on left part, whereas Fig. 5.3 (c), 5.7 (c) and 5.8 (c) are blurred on right part. Fig. 5.4 (b), 5.5 (b) are upper side blurred images and Fig. 5.4 (c), 5.5 (c) are lower side blurred images, respectively. The

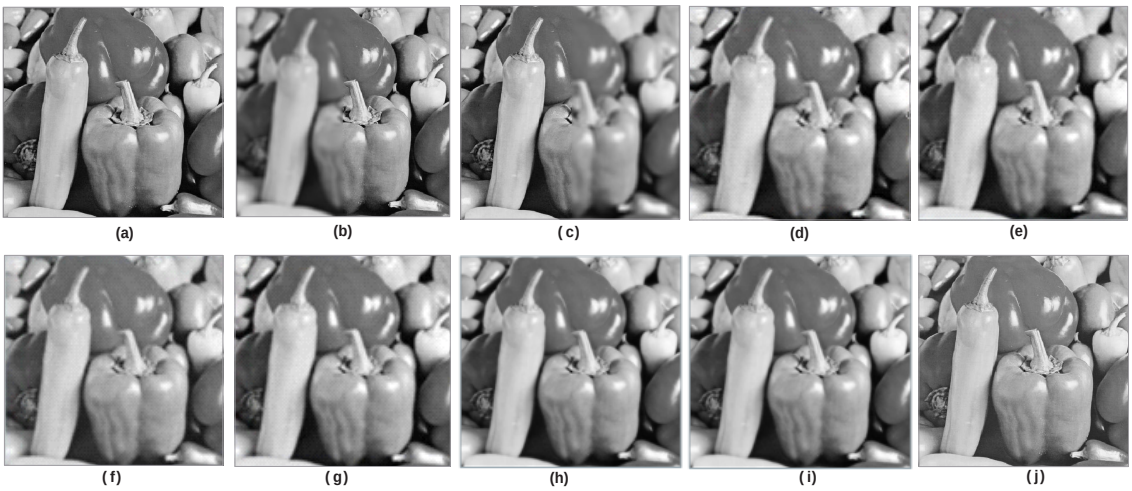

Figure 5.3: Pepper image. (a) Reference image (original image); (b) image blurred on the left side; (c) image blurred on the right side; (d) fused image by DWT; (e) fused image by DWT+PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

Figure 5.4: Lena image. (a) Reference image (original image); (b) blurred on upper part; (c) blurred on lower part; (d) fused image by DWT; (e) fused image by DWT+PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

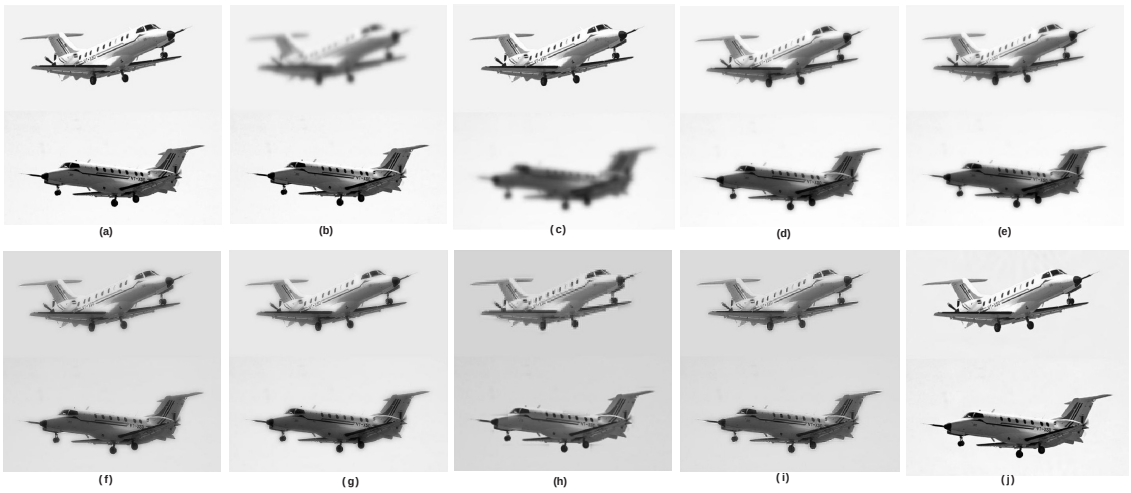

Figure 5.5: Dataset1 image. (a) Reference image (original image); (b) blurred on upper part; (c) blurred on lower part; (d) fused image by DWT; (e) fused image by DWT+PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

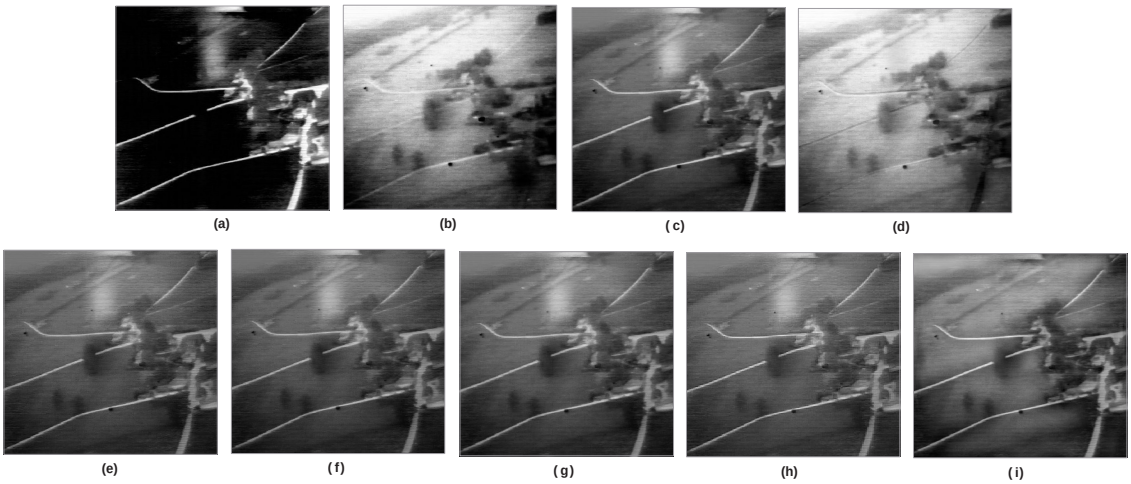

Figure 5.6: Dataset2 image. (a) FLIR image; (b) LLTV image; (c) fused image by DWT; (d) fused image by DWT+PCA; (e) fused image by MSVD; (f) fused image by SWT; (g) fused image by CVT; (h) fused image by NSCT; (i) fused image by proposed scheme.

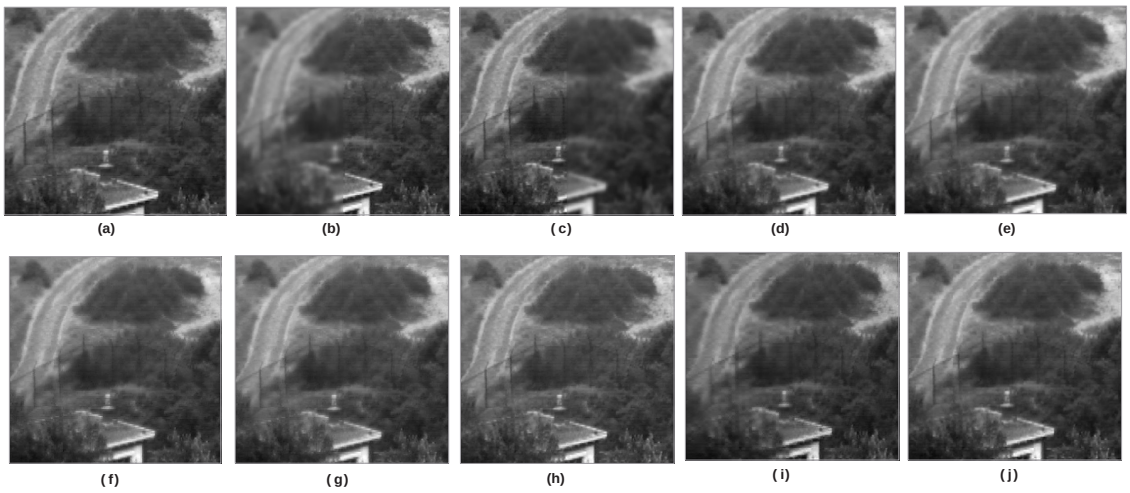

Figure 5.7: TNO's UN camp image. (a) Reference image (original image); (b) blurred on left part; (c) blurred on right part; (d) fused image by DWT; (e) fused image by DWT+PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

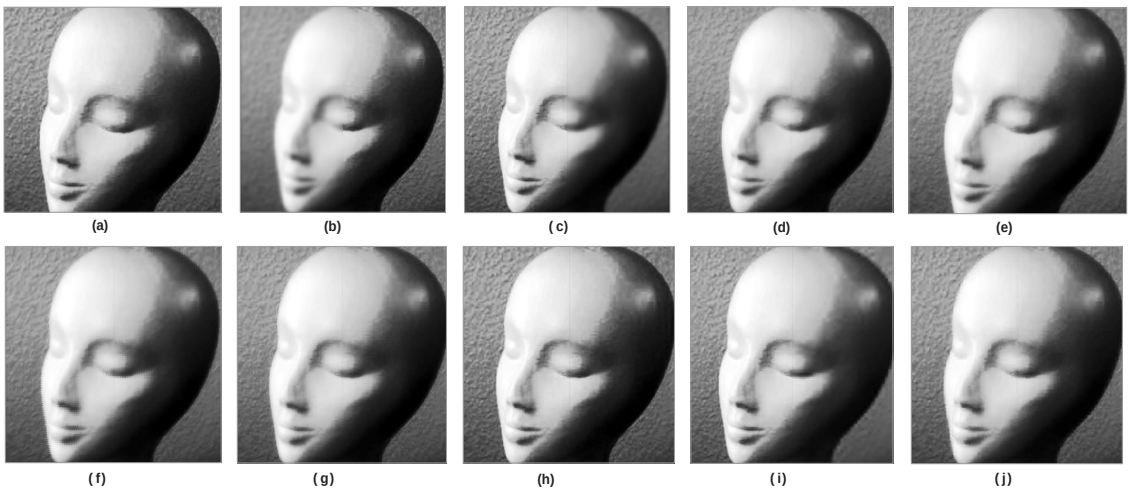

Figure 5.8: Head monument image. (a) Reference image (original image); (b) blurred on left part; (c) blurred on right part; (d) fused image by DWT; (e) fused image by DWT+PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

Table 5.1: Quantitative analysis

Source Images	**Evaluation indices**	**DWT** [70]	**DWT+PCA** [94]	**MSVD** [92]	**SWT** [52]	**CVT** [3]	**NSCT** [160]	**Proposed scheme**
Pepper	PSNR	36.0793	36.0791	35.3424	35.5148	35.5214	35.8210	36.5174
	CC	0.9898	0.9898	0.9875	0.9805	0.9762	0.9889	0.9905
	Mean	119.155	119.155	119.155	119.155	119.156	119.155	119.165
	S.D.	50.8323	50.8325	50.9791	50.9519	51.2063	51.3940	51.4211
	MI	1.2101	1.1998	1.2768	1.3006	1.2920	1.3010	1.3018
	Q_0	0.6832	0.6925	0.6953	0.6128	0.6872	0.6957	0.6872
	Q_W	0.8264	0.8098	0.8295	0.8716	0.8752	0.8836	0.8765
	Q_E	0.6164	0.6112	0.6268	0.6241	0.6105	0.6254	0.6321
	$Q^{AB/F}$	0.5347	0.5329	0.5630	0.5436	0.5421	0.5384	0.5437
Lena	PSNR	34.1621	34.1847	32.3791	34.8267	34.8694	34.8945	34.9193
	CC	0.9887	0.9888	0.9822	0.9801	0.9877	0.9827	0.9902
	Mean	121.609	121.608	121.609	121.609	121.508	121.697	121.609
	S.D.	60.3515	60.3597	60.1866	60.5962	60.9821	60.7543	61.1822
	MI	1.2458	1.2450	1.2416	1.2844	1.2759	1.2089	1.2271
	Q_0	0.8645	0.8598	0.8591	0.9331	0.9289	0.9367	0.9414
	Q_W	0.7562	0.7512	0.7591	0.8767	0.8541	0.8793	0.8811
	Q_E	0.6378	0.6351	0.6435	0.6421	0.6492	0.6487	0.6460
	$Q^{AB/F}$	0.4362	0.4355	0.4731	0.4531	0.4803	0.4989	0.5613
Dataset1	PSNR	36.7513	36.8220	38.8281	37.7920	38.7982	38.8724	38.9818
	CC	0.9847	0.9854	0.9871	0.9876	0.9758	0.9849	0.9882
	Mean	227.666	227.871	227.666	227.666	227.568	227.825	227.666
	S.D.	45.8628	45.9310	46.3967	46.1644	46.5231	46.7903	46.8027
	MI	1.4331	1.4068	1.4291	1.4319	1.4297	1.4367	1.4373
	Q_0	0.3543	0.3532	0.3734	0.3745	0.3758	0.3428	0.3797
	Q_W	0.4353	0.4216	0.4437	0.4528	0.4376	0.4392	0.4439
	Q_E	0.3201	0.3254	0.3633	0.3096	0.3257	0.3467	0.3581
	$Q^{AB/F}$	0.5634	0.5207	0.5553	0.5126	0.5167	0.5948	0.6277
Dataset2	Mean	84.3786	84.3786	84.3786	84.3786	84.3773	84.4671	84.8658
	S.D.	47.3890	47.3890	49.1255	48.8580	48.5708	49.7903	49.2007
	MI	1.3965	1.3018	1.3520	1.3991	1.4071	1.4023	1.4080
	Q_0	0.3967	0.3807	0.3862	0.3831	0.3810	0.3806	0.3826
	Q_W	0.4325	0.4336	0.4381	0.4310	0.4326	0.4374	0.4307
	Q_E	0.3029	0.3031	0.3047	0.3051	0.3068	0.3073	0.3082
	$Q^{AB/F}$	0.5528	0.5541	0.5520	0.5516	0.5529	0.5574	0.5519
TNO	PSNR	34.8460	34.8562	34.8844	35.7054	35.8325	35.9077	35.9116
	CC	0.9931	0.9931	0.9930	0.9943	0.9758	0.9849	0.9914
	Mean	87.3152	87.3154	87.3153	87.3152	87.5621	87.5639	87.6153
	S.D.	36.5409	36.5431	36.7410	36.7115	36.6742	36.7309	36.9588
	MI	1.3295	1.3295	1.3479	1.3596	1.3498	1.3671	1.3450
	Q_0	0.8497	0.8497	0.8137	0.8276	0.8420	0.8437	0.8542
	Q_W	0.9315	0.9315	0.9494	0.9559	0.9576	0.9792	0.9752
	Q_E	0.5721	0.5692	0.5771	0.5748	0.5622	0.5767	0.5782
	$Q^{AB/F}$	0.6856	0.6857	0.6803	0.7033	0.6841	0.6948	0.7313
Head	PSNR	33.3781	33.8152	34.6210	34.5709	35.5421	35.5714	35.6509
	CC	0.9973	0.9905	0.9898	0.9976	0.9922	0.9959	0.9982
	Mean	126.876	126.872	127.376	127.839	127.865	127.675	127.893
	S.D.	73.6487	73.5140	73.5871	73.4594	73.5291	73.5933	73.6107
	MI	1.4179	1.4140	1.4276	1.4273	1.4297	1.4266	1.4313
	Q_0	0.8437	0.8250	0.8374	0.8405	0.8354	0.8420	0.8531
	Q_W	0.7187	0.7189	0.7270	0.7123	0.7296	0.7232	0.7203
	Q_E	0.5520	0.5429	0.5593	0.5297	0.5487	0.5447	0.5481
	$Q^{AB/F}$	0.7187	0.7172	0.7205	0.7124	0.7287	0.7248	0.7236

results of our proposed scheme is compared with DWT [70], DWT with PCA [94], MSVD [92], SWT [52], CVT [3], NSCT [160] based fusion methods. The results of the respective existing schemes are shown in Fig. 5.3 (d-i)-5.8 (d-i). Results obtained by our proposed scheme are given in Fig. 5.3 (j)-5.8 (j). It can be visually seen that the resultant images from proposed scheme contain better information than other existing schemes. To measure the quality of fused images quantitatively, some parameters are used such as PSNR, CC, mean, S.D., *MI*, Q_0, Q_W, Q_E and $Q^{AB/F}$. The higher values of these parameters stand for better fusion result. The quantitative results are shown in Table 5.1. It is clear that in most of the cases, results obtained from the proposed scheme are better as compared to other existing schemes.

5.5 Conclusions

In this chapter, a fusion scheme based on multiresolution singular value decomposition has been discussed. The experimental results of proposed scheme are performed on several pairs of multifocus and multisensor images which demonstrate that the proposed scheme preserves more significant details, provides sharp details, rich texture and also improves the visual quality of the fused image than other fusion schemes. Apart from qualitative measurement, quantitative measurement is performed to evaluate the performance of proposed method by using some quantitative parameters. The results of the proposed methods are better in most of the cases. However, the scheme is slightly more time consuming.

This work provides the sharp and multidirectional details of the images using MSVD. But it is unable to obtain a smooth, de-noised having sharp details simultaneously in the fused image. Therefore, the next work is based on cross bilateral filter which removes noise and provides smooth information and also preserves edges and sharp information of the source images.

CHAPTER 6

Cross bilateral filter based image fusion

Chapter 6

Cross bilateral filter based image fusion

In the previous work, the sharp and multidirectional details were obtained using MSVD but it can not provide smooth details along with edge information in the fused image. Here, in this work, we have attempted to obtain a smooth, de-noised and sharp fused image.

6.1 Introduction

The concept of image filtering has become most popular in image processing to reduce the noise or extract useful information from the images. Few types of filters like Sobel filter, laplacian filter and gaussian filter are broadly used in image sharpening/blurring, feature extraction and edge detection [44]. The comparative study on several image fusion algorithms has been presented by Wang et al. [144]. The pixel by pixel averaging is the simple and intuitive image fusion method and due

to this it has attracted many researchers in the last decades [66, 70, 102]. But this method suffers from a number of undesirable effects, such as reduced contrast. To address this problem, multi-resolution analysis based methods have been proposed, which consist of the following three basic steps. First, the source images are decomposed into multi-resolution representations with low and high frequency information and corresponding transformed coefficients are obtained. Second, the obtained transformed coefficients are combined together according to some fusion rules. Finally, the inverse transform is performed over the fused coefficients to reconstruct the fused image [75]. Generally, multi-resolution based methods provide better results than the other transform methods. However, the other transformations also preserve the same salient features such as edges and lines and have been used in fusion. For an instance, a DCT based image fusion is introduced for fusion instead of pyramids or wavelet [91]. Again, a new multi-resolution DCT decomposition based image fusion approach has been given to reduce the computational complexity without any loss of image information [66]. Many image fusion techniques have been developed such as, maximum, minimum, average methods, principal component analysis (PCA) [38, 59, 94], intensity-hue-saturation [144]. In above discussed methods, average method provides the average information of the images but weighted average methods provide the information according to their weights and produce better fusion results.

Various filtering techniques have been introduced to produce the smooth images and preserve different levels of structures. They are most useful in image processing and also perform operations to remove the unwanted information. There are few examples of filters such as, bilateral filter (BF) [150], guided filter (GF) [74], contourlet filter [142], weighted median filters [161] and so on are average based filters and mainly used to aware edges. These filters have the similar motive to retain high contrast edges and removes low contrast details. To remove the high contrast information median [72, 115], mode [58] and weighted median filters [112] are used. These filters are used to compute weighted median in local patches than average which produce higher computational cost results and remove salt and pepper noise. A variant of bilateral filter like joint/cross bilateral filter uses a second image to shape the filter kernel and operate on the first image has been introduced in [37] which are used to tackle the issue of combining the information of images captured with flash or no flash under ambient illumination. Fattal et al. [40] have introduced a bilateral filter for multiscale decomposition of multilight image collections to enhance the shape detail. For the fusion of multispectral video a temporal joint bilateral filter and dual bilateral filter was proposed by Bennett et al. [13]. Further, the fusion of hyperspectral images using bilateral filtering was developed in [63] which is used to separate weak and fine edges details by subtracting the filtered image using bilateral filtering from the original image. The fusion of IR and visible images an another variant of bilateral filter which uses center and neighboring pixels from IR visible image to find the filter kernel and

operates on IR image had been developed for human detection [29]. Furthermore, a multiscale directional bilateral filter which is a combination of bilateral filter and directional filter bank was introduced for the fusion of multisensor images to use capability of bilateral filter to preserve edge information and directional information of directional filter bank capturing capability [51]. In [65], source images are fused by weighted average from the image details, where these details are extracted from the source images using CBF to improve the fusion performance.

In this chapter, we proposed a novel pixel level image fusion scheme based on discrete wavelet transform and cross bilateral filter. The low frequency subbands are fused using pixel averaging method because approximation part contain most of the average information. Meanwhile, the high frequency subbands give the sharp details, therefore weighted average fusion rule is performed over the detail parts, where the weights of two different images contribute the information for fusion according to their weight value. The main contribution of this chapter is to enhance the visual quality of the images by integrating all significant details of the source images using cross bilateral filter.

The rest of this chapter is organized as follows. In Section 6.2, the basic theory of CBF is described. Proposed methodology is explained in Section 6.3. Experimental results followed by discussion are presented in Section 6.4. Section 6.5 concludes the proposed work.

6.2 Cross bilateral filter

The Gaussian filter is the most popular local and linear filter which provides smoothness and removes noise present in the images. It is based on the weighted average of the intensity of the adjacent pixels. This filter removes the noise and provides smoothness but meantime also removes the sharp details. Therefore, to address this problem, Smith et al. [122] was the first person who introduced bilateral filter under the name SUSAN. Again, it was discovered by Tomasi et al.[133] who have named it *bilateral filter* which is the mostly used name in present days. Bilateral filter is a local but non-linear and non-iterative filter which smoothes images while preserving edges and other sharp details. Due to advantage of preserving edge information, bilateral filtering has been widely used in image processing applications. The bilateral filter (BF) is the combination of spatial and range filters [35, 51, 66]. The spatial filter kernel is treated as a classical low pass filter used to obtain geometric closeness between the neighboring pixels, whereas the range filter kernel is treated like an edge-stopping function used for gray-level similarity between the neighboring pixels, which attenuates the filter kernel when the intensity differences between pixels are large. Both filter kernels are based on Gaussian distribution and the weights obtained from these filters depend not only on Euclidian distance but also on the distance in gray or color spaces.

For an image X, the output of bilateral filter at a pixel location p is defined as:

$$X_f(p) = \frac{1}{W}\sum_{q\in s} G_{\sigma_s}(\|p-q\|) \times G_{\sigma_r}(\mid X(p) - X(q) \mid)X(q) \tag{6.2.1}$$

where, $W = \sum_{q\in s} G_{\sigma_s}(\|p-q\|)G_{\sigma_r}(\mid X(p) - X(q) \mid)$ is a normalization factor, $\|p-q\|$ is the Euclidian distance between p and q. $G_{\sigma_s}(\|p-q\|) = e^{-\frac{\|p-q\|^2}{2\sigma_s^2}}$ is a geometric closeness function and $G_{\sigma_r}(\mid X(p) - X(q) \mid) = e^{-\frac{|X(p)-X(q)|^2}{2\sigma_r^2}}$ is a gray-level similarity or edge-stopping function. G_{σ_s} is a spatial Gaussian function that decreases symmetrically as the distance from the center increases. G_{σ_r} is also a Gaussian function that decreases with the increase of the intensities difference between $X(p)$ and $X(q)$. s is a spatial neighborhood of the pixel p. Parameters σ_s and σ_r are standard deviations of Gaussian functions G_{σ_s} and G_{σ_r} respectively and determine the amount of filtering for the image X.

Cross/joint bilateral filter (CBF) considers both gray-level similarities and geometric closeness of neighboring pixels in image X to shape the filter kernel and filters the image Y. The output of CBF [65, 66, 105] for image Y at a pixel location p is calculated as:

$$Y_{CBF}(p) = \frac{1}{W}\sum_{q\in s} G_{\sigma_s}(\|p-q\|) \times G_{\sigma_r}(\mid X(p) - X(q) \mid)Y(q) \tag{6.2.2}$$

where, $W = \sum_{q\in s} G_{\sigma_s}(\|p-q\|)G_{\sigma_r}(\mid X(p) - X(q) \mid)$ is a normalization factor and $G_{\sigma_r}(\mid X(p) - X(q) \mid) = e^{-\frac{|X(p)-X(q)|^2}{2\sigma_r^2}}$ is a gray-level similarity or edge-stopping function.

For images X and Y, the detail image is obtained by subtracting CBF output from

their respective source image and is given as $X_D = X - X_{CBF}$ and $Y_D = Y - Y_{CBF}$, respectively. These detail images are used to find the weights by measuring the strength of details as reported in [66].

6.3 Proposed methodology

In this section, we briefly discussed the proposed method. In the proposed method, two images having less information are used to produce a single composite image with enhanced visual quality. Here, the two source images either multi-focus or multisensor are used for fusion. The DWT is performed over both the source images to decompose the images into low and high frequency subbands as discussed in Chapter 3. The low frequency subband labeled as *LL* which is also called approximation part, contains the average information of the images. On the other hand, the high frequency subbands corresponds to *(LH,HL* and *HH)* , i.e., detail parts which contain the information of edges and sharp changes. The low frequency subbands of both of the transformed images are fused using pixel averaging method to get the overall average information from both the source images in a single fused image. The high frequency subbands of both of the transformed images are fused according to their weights in respective images to get the sharp information in the fused image. Thus, the fused image has overall average information as well as dominant sharp information chosen from both the source images. The block diagram of the proposed method is illustrated in Fig. 6.1.

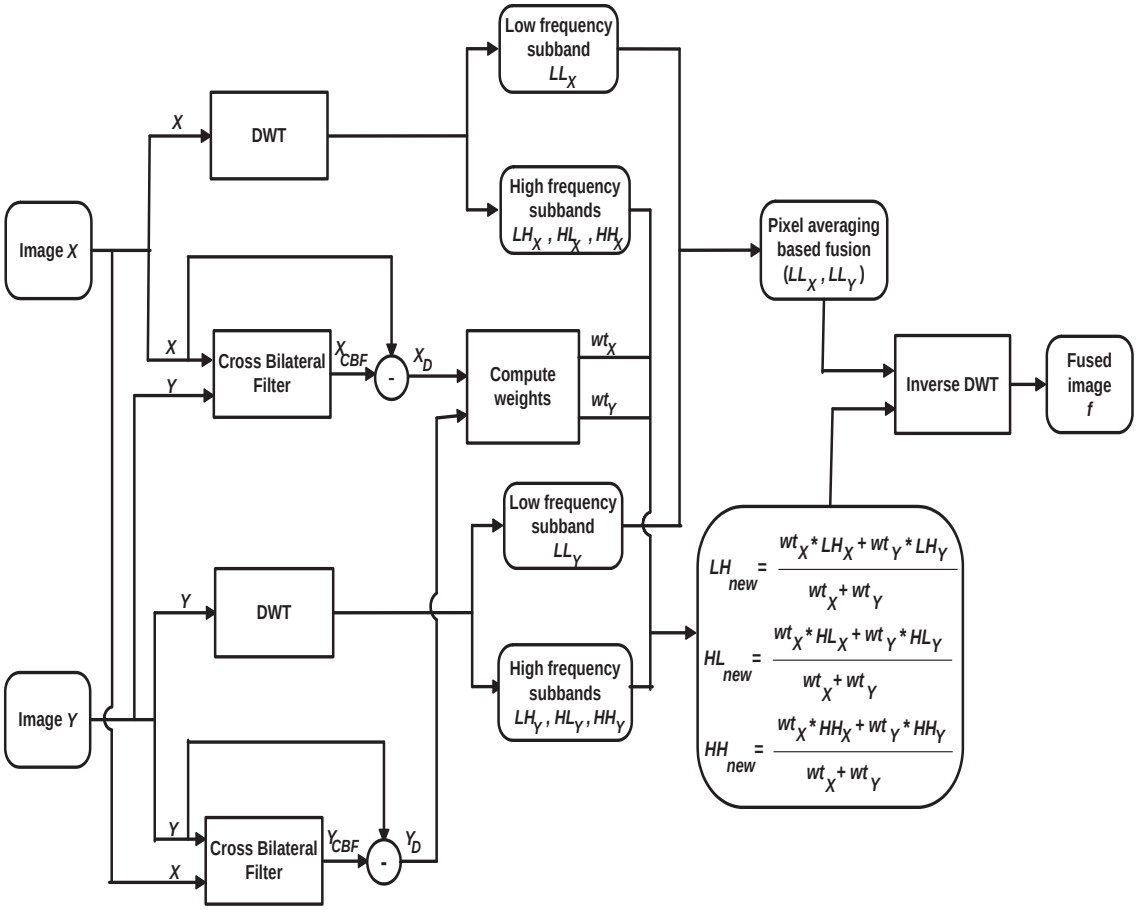

Figure 6.1: Block diagram for proposed image fusion framework.

The steps of the proposed method can be described in the following Algorithm:

6.3.1 Proposed fusion methodology algorithm

1. Take two source images X and Y.

2. Perform DWT over the images X and Y which decomposes the images into low and high frequency subbands.

3. Apply pixel averaging method to fuse the low frequency subbands and obtain the average fused coefficient (LL_{new}).

4. Combine high frequency subbands by their weighted average of both transformed images. The weights are computed as in [65]:

   a) Compute the covariance matrix as for both the images and then calculate horizontal and vertical detail strengths.

   b) Integrate both detail strengths and obtain weight wt_X and wt_Y from images X and Y.

   c) Fuse the high frequency subbands with their respective weights using:

$$HL_{new} = \frac{wt_X * HL_X + wt_Y * HL_Y}{wt_X + wt_Y} \tag{6.3.3}$$

   Similarly, we can obtain fused wavelet coefficients (LH_{new} and HH_{new}).

5. Perform inverse DWT over the obtained fused coefficients ($LL_{new}, LH_{new}, HL_{new}$ and HH_{new}) to reconstruct the fused image *f*.

6.4 Experimental results and analysis

The proposed method is carried out on several pairs of multi-focus and multisensor images as shown in Fig. 6.2-6.6. Each image is size of 512×512. Fig. 6.2 (a) and 6.2 (b) are blurred on left and right part images. Fig. 6.3 (a) and 6.3 (b) are upper side and lower side blurred images are called as dataset1. Fig. 6.4 (a) and 6.4 (b) are FLIR and LLTV images are called as dataset2. Dataset1 and dataset2 images are taken from [92]. Fig. 6.5 (a) and 6.5 (b) are hoed images which are blurred on middle and boundary part. CT and MRI images of size 256×256 are shown in Fig. 6.6 (a) and 6.6 (b) are acquired from the link `http://www.metapix.de/examples_r.htm`. The parameters $\sigma_s = 1.8$ and $\sigma_r = 25$ are used for the proposed method with window size 5×5.

The results of proposed methods are compared with some existing methods such as, DWT [70], PCA with DWT [94], SWT [52] and CBF [65] based fusion methods. The results obtained from above existing methods are shown in Fig. 6.2 (c-f)-6.6 (c-f). However, the results obtained by proposed method are shown in Fig. 6.2 (g)-6.6 (g). Visually it can be seen that in most of the cases, the fused images obtained from proposed method are better than other existing methods. All the results of existing and proposed methods are tabulated in Table 6.1. It can be found that in most of the cases the results of the proposed method have shown better performance over the other existing methods.

Table 6.1: Image fusion performance measure

Input images	**Fusion rules**	**S.D.**	**CC**	**API**	**AG**	**EN**	**MI**	**FS**	**SF**	$Q^{XY/f}$
Clock	[70]	49.315	0.9721	97.038	3.879	3.8869	5.728	1.8904	6.2841	0.8341
	[94]	49.315	0.9754	97.038	3.878	4.3445	5.875	1.8722	6.2817	0.8338
	[52]	49.408	0.9872	97.038	4.471	4.4882	6.492	1.9021	7.4981	0.8700
	[65]	49.892	0.9869	96.548	5.5265	7.2755	7.3415	1.9600	10.142	0.8982
	Proposed	49.417	0.9890	97.080	4.4044	7.2761	6.8667	1.9934	8.0545	0.8583
Dataset1	[70]	51.872	0.9746	221.43	3.1869	2.5333	5.0198	1.9587	10.222	0.7958
	[94]	51.873	0.9798	221.43	3.1871	2.8173	5.0368	1.9590	10.223	0.7959
	[52]	52.172	0.9876	221.43	4.2307	3.5672	5.3640	1.9526	13.578	0.8698
	[65]	52.561	0.9886	220.58	4.9215	4.2667	5.4154	1.9635	16.926	0.9284
	Proposed	55.242	0.9904	221.45	5.8073	4.7166	5.0230	1.9766	18.656	0.8451
Dataset2	[70]	40.394	0.5703	84.631	7.0495	1.9383	3.0799	1.7832	10.612	0.6249
	[94]	41.220	0.5770	84.631	9.7021	2.0432	3.0765	1.7704	21.908	0.4599
	[52]	41.087	0.5689	84.631	9.3246	2.0486	3.0832	1.7543	14.019	0.7226
	[65]	41.358	0.5317	84.665	9.7106	7.3465	3.0842	1.7735	16.804	0.7305
	Proposed	41.429	0.5815	86.642	7.7882	7.8537	3.4738	1.8188	17.431	0.8310
Hoed	[70]	56.559	0.9521	95.951	13.2480	1.9248	4.5982	1.8954	17.496	0.7168
	[94]	56.568	0.9460	95.951	13.2807	2.0431	4.7601	1.8643	17.593	0.7188
	[52]	57.947	0.9479	95.951	28.0801	7.6941	4.7803	1.8092	30.798	0.8957
	[65]	58.097	0.9628	95.958	22.8185	7.6988	4.6434	1.9987	30.905	0.8775
	Proposed	61.849	0.9566	96.173	25.2017	7.7399	6.9666	1.9952	33.964	0.9478
CT and MRI	[70]	34.883	0.6502	32.082	5.5005	1.6771	3.4601	1.6204	10.271	0.6440
	[94]	35.155	0.6521	32.082	6.6750	1.8607	3.4028	1.6082	13.160	0.6140
	[52]	35.107	0.6246	32.082	6.1676	2.0452	3.4122	1.6091	11.346	0.6914
	[65]	35.785	0.6905	32.166	7.2431	5.9698	3.4311	1.6554	14.717	0.7143
	Proposed	37.862	0.6545	41.414	11.5408	6.7736	5.5999	1.6122	21.095	0.9116

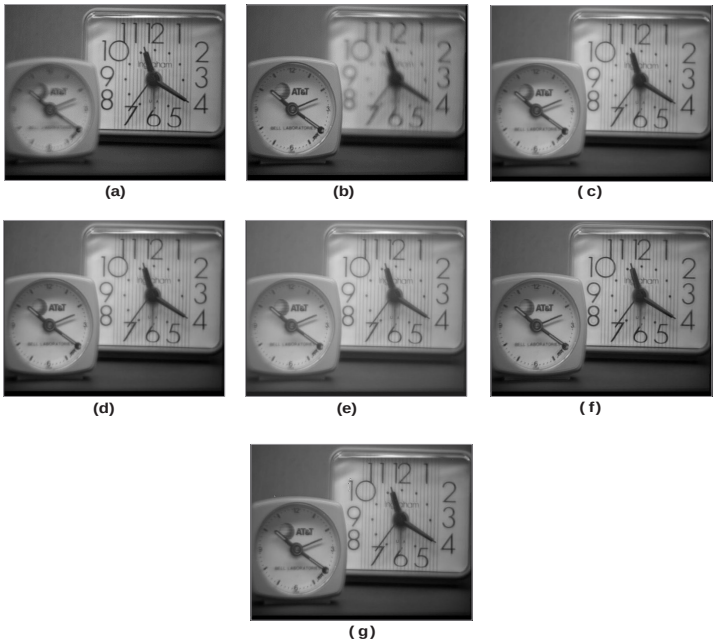

Figure 6.2: Fusion results for clock image. (a) left part blurred; (b) right part blurred; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image by CBF; (g) fused image by proposed method.

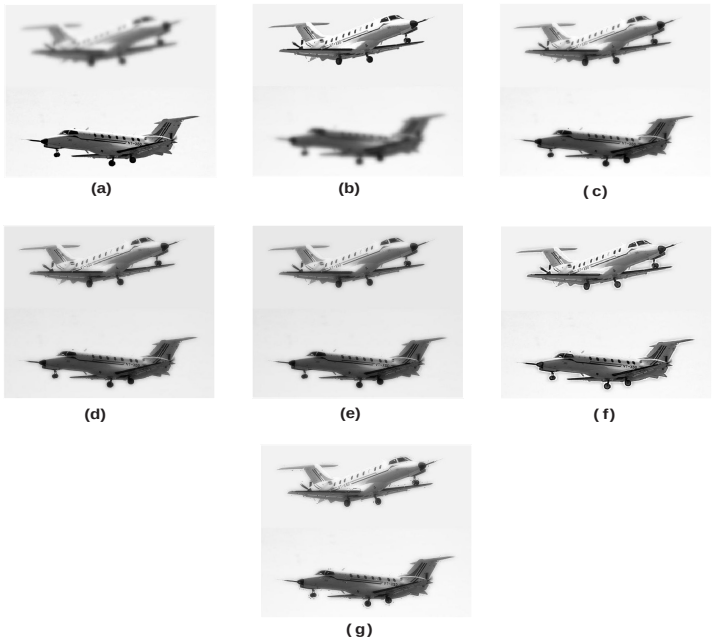

Figure 6.3: Fusion results for dataset 1. (a) upper part blurred; (b) lower part blurred; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image CBF; (g) fused image by proposed method.

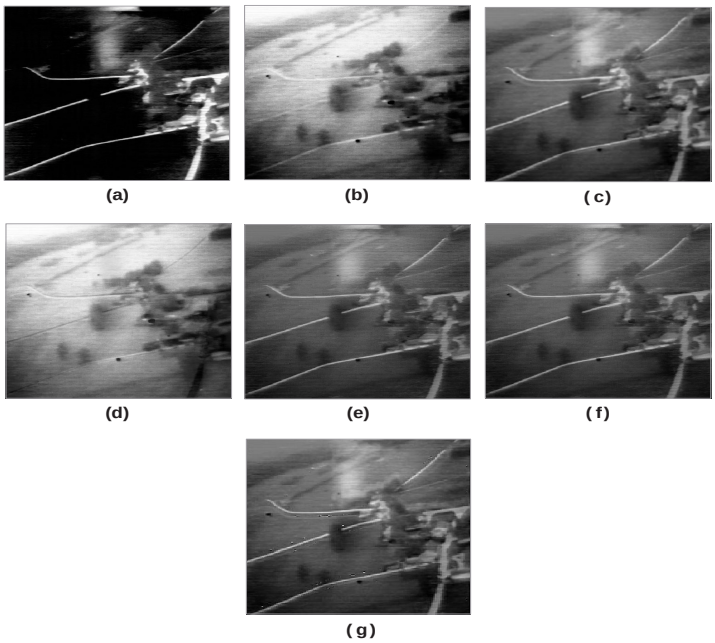

Figure 6.4: Fusion results for dataset 2. (a) FLIR image; (b) LLTV image; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image by CBF; (g) fused image by proposed method.

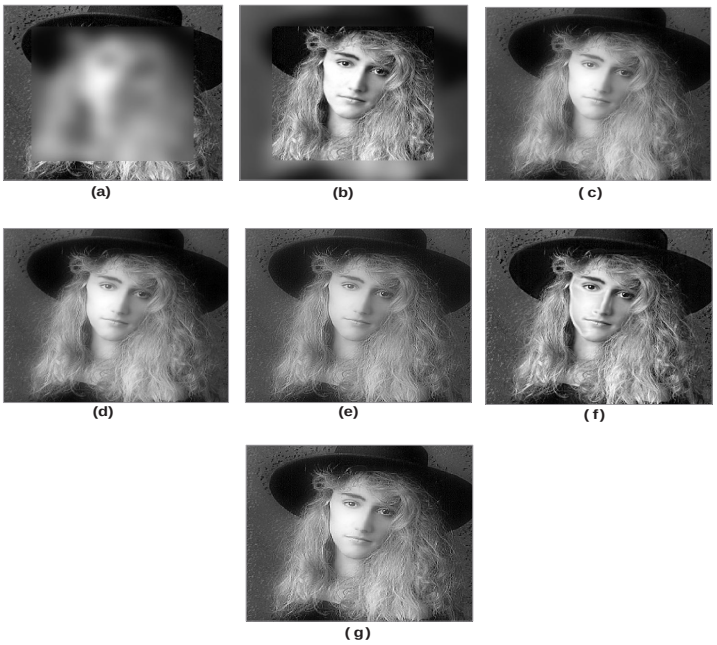

Figure 6.5: Fusion results for hoed image. (a) middle part blurred; (b) boundary part blurred; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image CBF; (g) fused image by proposed method.

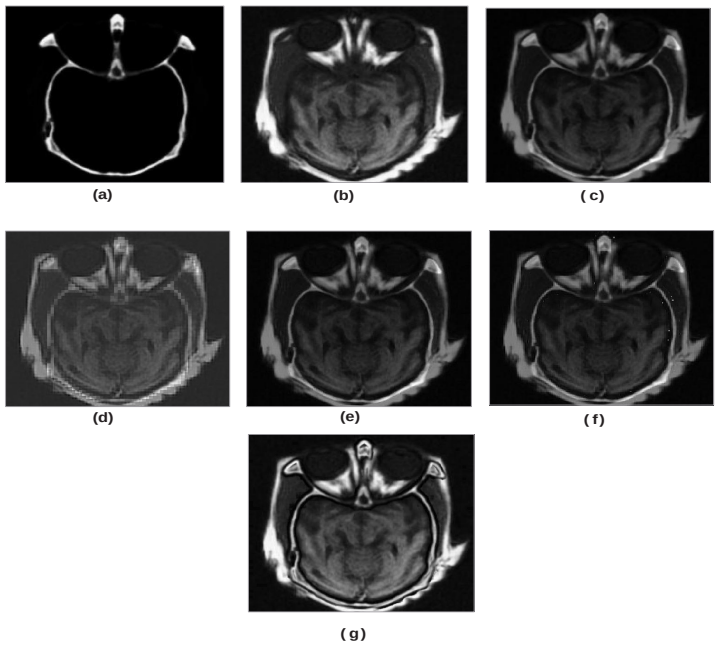

Figure 6.6: Fusion results for CT and MRI image. (a) CT image; (b) MRI image; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image CBF; (g) fused image by proposed method.

6.5 Conclusions

In this chapter, a fusion method of two multi-focus or multisensor images based on DWT and CBF algorithm has been discussed. In the proposed framework, the average and sharp details of the source images are integrated very cleverly into a single fused image. The performance of the proposed fusion method is tested against other fusion methods using several pairs of test images. Fusion performance has been evaluated using many quantitative measurement criterions. Through the experimental results, it is found that the proposed method preserves more sharp information while eliminating artifacts and has shown better performance than other existing fusion methods visually as well as quantitatively.

CHAPTER 7

Conclusions and Future scope

Chapter 7

Conclusions and Future scope

This chapter concludes the investigations of image quality enhancement through the proposed methods along with future scope in this field.

7.1 Contributions

The thesis comprises a set of studies on image fusion techniques developed for enhancing the quality of fused images. The thesis briefly summarized the main findings of the theoretical and practical work performed in the course of this work. In this study, several novel and improved image fusion techniques have been proposed. The main focus of the study is to obtain the images of good quality using fusion techniques. Both the spatial and transform domain based approaches have been used in order to propose improved fusion techniques.

The first work starts by using discrete cosine transform and sum-modified-laplacian. DCT has the property of energy compaction therefore it gives most of the information of image into a very few number of transformed coefficients and SML provides a

focused image which helps in efficient fusion purpose. This technique provides better fusion results for defocused images. But it can not provide the sharp information which is required for the visual analysis for human or machine.

Therefore, another image fusion techniques has been developed using the concept of gradient. In this work, two gradient based techniques for image fusion have been proposed. In the first technique, DWT domain has been used to transform the images and gradient rule is used for fusion purpose. This work can not provide the better shift invariance and good directionality due to the use of DWT. Therefore, next technique using dual tree complex wavelet transform has been developed due to the property of shift invariance and good directionality over the DWT. Although, these methods provide better edge information with sharp details but in case of low contrast image, it may not enhance the contrast in the fused images.

The next work has been developed by using the directive contrast and correlation based fusion concepts. In this work, directional contrast has been calculated by which an image with better contrast can be obtained. The experimental results provide good fusion results with better visual information but still there is a challenge to obtain overall better image along with sharp details.

Therefore, a more efficient and novel image fusion technique has been developed using MSVD in combination with DWT by utilizing the properties of MSVD. This scheme is proven to be good over the existing methods, but can not provide a fused image having smoothness and edges information at the same time. Therefore, next

work has been developed using cross bilateral filtering concept.

In this next work, a novel technique by integrating the complementary information of two source images has been developed using cross bilateral filter, which has the property of noise removal and also preserves the edge information.

The code for the entire work of this thesis have been developed using MATLAB. The methods given in this work are tested on several multifocus and multisensor images and the effectiveness of proposed methods is compared with existing fusion methods. In the work of the thesis, the experimental results of proposed methods are superior than existing methods in terms of qualitative as well as quantitative assessment. For quantitative analysis, several mathematical parameters have been used such as *S.D., PSNR, CC*, mean, Q^{AB/I_f}, *MI* and so on. Experimental results demonstrate that the thesis work provide fused image with better visual quality which may be useful for many image processing applications.

7.2 Future scope

The proposed methods are novel and improved than existing methods in terms of better visual quality. This work has its contribution in the area of image fusion using pixel level image fusion for multifocus and multisensor images in the spatial and transform domain but still there are open areas of feature and decision level based image fusion for heterogeneous set of images and also in terms of speed and extension of the image fusion techniques. Several aspects of the proposed algorithms can still

be improved. The extended versions of multiresolution transforms such as, framelet, shearlet, tetrolet, etc. can also be used to improve the quality of images. Many times only the quality of fused images is not of that importance but we are also concerned about more application perspective such as, face recognition, security, denoising and so on. Therefore, the developed proposed schemes can play a prominent role while considering new aspects of fusion.

References

References

[1] Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., and Cohen, M. Interactive digital photomontage. *ACM Transactions on Graphics (TOG) 23*, 3 (2004), 292–300.

[2] Al-Azzawi, N., and Abdullah, W. W. Medical image fusion schemes using contourlet transform and pca bases. *Proc. Image fusion and its applications* (2011), 93–110.

[3] Ali, F., El-Dokany, I., Saad, A., and El-Samie, F. A. A curvelet transform approach for the fusion of mr and ct images. *Journal of Modern Optics, Taylor and Francis 57*, 4 (2010), 273–286.

[4] Allerton, D., and Clare, A. Sensor fusion methods for synthetic vision systems. *Proc. Digital Avionics Systems* (2004), 1–13.

[5] Amiri, G., and Asadi, A. Comparison of different methods of wavelet and wavelet packet transform in processing ground motion records. *International Journal of Civil Engineering 7*, 4 (2009), 248–257.

[6] Amolins, K., Zhang, Y., and Dare, P. Wavelet based image fusion techniques - an introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing 62*, 4 (2007), 249–263.

[7] Asgari, S., Mehrnia, A., and Moussavi, M. Automatic detection of atrialfibrillation using stationary wavelet transform and support vector machine. *Computers in Biology and Medicine 60* (2015), 132–142.

[8] Aslantas, V., and Kurban, R. Fusion of multi-focus images using differential evolution algorithm. *Expert Systems with Applications 37*, 12 (2010), 8861–8870.

[9] Aslantas, V., and Toprak, A. A pixel based multi-focus image fusion method. *Optics Communications 332* (2014), 350–358.

[10] Bai, X., Zhang, Y., Zhou, F., and Xue, B. Quadtree-based multi-focus image fusion using a weighted focus-measure. *Information Fusion 22* (2015), 105–118.

[11] Bal, U. Dual tree complex wavelet transform based denoising of optical microscopy images. *Biomedical Optics Express 3*, 12 (2012), 3231–3239.

[12] Bastiere, A. Methods for multisensor classification of airborne targets integrating evidence theory. *Aerospace Science and Technology 2*, 6 (1998), 401–411.

[13] Bennett, E., Mason, J., and McMillan, L. Multispectral bilateral video fusion. *IEEE Transactions on Image Processing 16*, 5 (2007), 1185–1194.

[14] Bhatnagar, G., and Raman, B. A new image fusion technique based on directive contrast. *Electronics letters on computer vision and image analysis 8*, 2 (2009), 18–38.

[15] Bhatnagar, G., Saha, A., Wu, Q., and Atrey, P. Analysis and extension of multiresolution singular value decomposition. *Information Sciences 277* (2014), 247–262.

[16] Bhatnagar, G., Wu, Q. J., and Liu, Z. Directive contrast based multimodal medical image fusion in nsct domain. *IEEE Transactions on Multimedia 15*, 5 (2013), 1014–1024.

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

[18] Blum, R., and Liu, Z. *Multi-sensor Image Fusion and Its Applications*. 2005.

[19] Burt, P. The pyramid as a structure for efficient computation. *Proc. Multiresolution Image Processing and Analysis, Springer 12* (1984), 6–35.

[20] Burt, P., and Adelson, E. Merging images through pattern decomposition. *Proc. SPIE, Applications of Digital Image Processing VIII 0575* (1985), 173–181.

[21] Burt, P., and Kolczynski, R. Enhanced image capture through fusion. *Proc. 4th International Conference on Computer Vision, IEEE* (1993), 173–182.

[22] Carper, W., Lillesand, T., and Kiefer, R. The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing 56*, 4 (1990), 459–467.

[23] Carroll, J., Ruskey, F., and Weston, M. Which n-venn diagrams can be drawn with convex k-gons? *Discrete and Computing Geometry 37*, 4 (2007), 619–628.

[24] Chai, Y., Li, H., and Li, Z. Multifocus image fusion scheme using focused region detection and multiresolution. *Optics Communications 284*, 19 (2011), 4376–4389.

[25] Chaudhuri, S., and Kotwal, K. *Hyperspectral image fusion.* 2013.

[26] Chaudhury, K., Sage, D., and Unser, M. Fast o(1) bilateral filtering using trigonometric range kernels. *IEEE Transactions on Image Processing 20*, 12 (2011), 3376–3382.

[27] Chen, Y., Chen, L., Gu, H., and Wang, K. Technology for multi-focus image fusion based on wavelet transform. *Proc. Third International Workshop on Advanced Computational Intelligence, IEEE* (2010), 405–408.

[28] Chipman, L., Orr, T., and Graham, L. Wavelets and image fusion. *Proc. SPIE, Wavelet Applications in Signal and Image Processing III 2569* (1995), 248–251.

[29] Choi, E., and Park, D. Human detection using image fusion of thermal and visible image with new joint bilateral filter. *Proc. 5th International Conference on Computer Sciences and Convergence Information Technology (ICCIT), IEEE* (2010), 882–885.

[30] da Cunha, A., Jianping, Z., and Do, M. The nonsubsampled contourlet transform: theory, design and applications. *IEEE Transactions on Image Processing 15*, 10 (2006), 3089–3101.

[31] De, I., and Chanda, B. A simple and efficient algorithm for multifocus image fusion using morphological wavelets. *Signal Processing 86*, 5 (2006), 924–936.

[32] De, I., and Chanda, B. Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure. *Information Fusion 14*, 2 (2013), 136–146.

[33] De, I., Chanda, B., and Chattopadhyay, B. Enhancing effective depth-of-field by image fusio using mathematical morphology. *Image and Vision Computing 24*, 12 (2006), 1278–1287.

[34] Deng, J., Yue, H., Zhuo, Z., Yan, H., Liu, D., and Li, H. A stationary wavelet transform based approach to registration of planning ct and setup cone beam-ct images in radiotherapy. *Journal of Medical Systems* (2014), 38–40.

[35] Diwakar, M., Sonam, and Kumar, M. Ct image denoising based on complex wavelet transform using local adaptive thresholding and bilateral filtering. *Proc. Third International Symposium on Women in Computing and Informatics, ACM Digital Library* (2015), 297–302.

[36] Do, M., and Vetterli, M. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing 14*, 12 (2005), 2091–2106.

[37] Eisemann, E., and Durand, F. Flash photography enhancement via intrinsic relighting. *ACM Transactions on Graphics (TOG) 23*, 3 (2004), 673–678.

[38] Eleyan, A. Enhanced face recognition using data fusion. *International Journal of Intelligent systems and applications (IJISA) 5*, 1 (2013), 98–103.

[39] Eskicioglu, A., and Fisher, P. Image quality measures and their performance. *IEEE Transactions on Communications 43*, 12 (1995), 2959–2965.

[40] Fattal, R., Agrawala, M., and Rusinkiewicz, S. Multiscale shape and detail enhancement from multi-light image collections. *ACM Transactions on Graphics (TOG) 26*, 3 (2007), 1–9.

[41] Genderen, J. V., and Pohl, C. Image fusion: Issues, techniques and applications. *Proc. EARSeL Workshop on Intelligent Image Fusion* (1994), 18–24.

[42] Ghahremani, M., and Ghassemian, H. Remote-sensing image fusion based on curvelets and ica. *Internation Journal of Remote Sensing 36*, 16 (2015), 4131–4143.

[43] Golub, G., and Reinsch, C. Singular value decomposition and least squares solutions. *Numerische Mathematik 14*, 5 (1970), 403–420.

[44] Gonzalez, R., and Woods, R. *Digital Image Processing.* 2013.

[45] Goshtasby, A., and Nikolov, S. Image fusion: Advances in the state of the art. *Information Fusion 8*, 2 (2007), 114–118.

[46] Guixi, L., Wenjin, C., and Wenjie, L. An image fusion method based on directional contrast and area-based standard deviation. *Proc. of SPIE, Electronic Imaging and Multimedia Technology IV 5637* (2005), 50–56.

[47] Haghighat, M., Aghagolzadeh, A., and Seyedarabi, H. Multi-focus image fusion for visual sensor networks in dct domain. *Computers and Electrical Engineering 37*, 5 (2011), 789–797.

[48] Hall, L., and Llinas, J. An introduction to multisensor data fusion. *Proc. International Symposium on Circuits and Systems, IEEE 85* (1998), 6–23.

[49] Han, J., Loffeld, O., Hartmann, K., and Wang, R. Multi image fusion based on compressive sensing. *Proc. International Conference on Audio Language and Image Processing (ICALIP), IEEE* (2010), 1463–1469.

[50] Hill, P., Canagarajah, N., and Bull, D. Image fusion using complex wavelets. *Proc. British Machine Vision Conference* (2002), 487–496.

[51] Hu, J., and Li, S. The multiscale directional bilateral filter and its application to multisensor image fusion. *Information Fusion 13*, 3 (2012), 196–206.

[52] Huafeng, L., Shanbi, W., and Chai, Y. Multifocus image fusion scheme based on feature contrast in the lifting stationary wavelet domain. *EURASIP Journal on Advances in Signal Processing 39* (2012), 1–16.

[53] Huang, W., and Jing, Z. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters 28*, 4 (2007), 493–500.

[54] Hunger, S., Karrasch, P., and Wessollek, C. Evaluating the potential of image fusion of multispectral and radar remote sensing data for the assessment of water body structure. *Proc. SPIE, Remote Sensing for Agriculture, Ecosystems and Hydrology XVIII 9998* (2016), 1–11.

[55] Hyvrinen, A. Survey on independent component analysis. *Neural Computing Surveys 2* (1999), 94–128.

[56] James, A., and Dasarathy, B. Medical image fusion: A survey of the state of the art. *Information Fusion 19*, 1 (2014), 4–19.

[57] Kakarala, R., and Ogunbona, P. Signal analysis using a multiresolution form of the singular value decomposition. *IEEE Transactions on Image Processing 10*, 5 (2001), 724–735.

[58] Kass, M., and Solomon, J. Smoothed local histogram filters. *ACM Transactions on Graphics (TOG) 29*, 4 (2010), 1–10.

[59] Kaur, S., and Dadhwal, H. Biorthogonal wavelet transform using bilateral filter and adaptive histogram equalization. *International Journal of Intelligent systems and applications (IJISA) 3* (2015), 37–43.

[60] Kim, Y., Lee, C., Han, D., Kim, Y., and Kim, Y. Improved additive-wavelet image fusion. *IEEE Transactions on Geoscience and Remote Sensing 8*, 2 (2011), 263–267.

[61] Kingsbury, N. Image processing with complex wavelets. *Proc. Wavelets: The Key to Intermittent Information, Oxford University Press* (1999), 165–185.

[62] Koren, I., Laine, A., and Taylor, F. Image fusion using steerable dyadic wavelet transform. *Proc. International Conference on Image Processing, IEEE* (1995), 232–235.

[63] Kotwal, K., and Chaudhuri, S. Visualization of hyperspectral images using bilateral filtering. *IEEE Transactions on Geoscience and Remote Sensing 48*, 5 (2010), 2308–2316.

[64] Kubota, A., Aizawa, K., and Chen, T. Reconstructing dense light field from array of multifocus images for novel view synthesis. *IEEE Transactions on Image Processing 16*, 1 (2008), 269–279.

[65] Kumar, B. S. Image fusion based on pixel significance using cross bilateral filter. *Signal, Image and Video Processing 9*, 5 (2015), 1193–1204.

[66] Kumar, B. S., Swamy, M., and Ahmad, M. O. Multiresolution dct decomposition for multifocus image fusion. *Proc. Canadian Conference on Electrical and Computer Engineering (CCECE), IEEE* (2013), 1–4.

[67] Lewis, J., OCallaghan, R., Nikolov, S., Bull, D., and Canagarajah, N. Pixel and region based image fusion with complex wavelets. *Information Fusion 8*, 2 (2007), 119–130.

[68] Li, H., Chai, Y., and Li, Z. A new fusion scheme for multifocus images based on focused pixels detection. *Machine vision and applications 24*, 6 (2013), 1167–1181.

[69] Li, H., Chai, Y., Ling, R., and Yin, H. Multifocus image fusion scheme using feature contrast of orientation information measure in lifting stationary

wavelet domain. *Journal of Information Science and Engineering 29*, 2 (2013), 227–247.

[70] Li, H., Manjunath, B., and Mitra, S. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing 57*, 3 (1995), 235–245.

[71] Li, S. Multisensor remote sensing image fusion using stationary wavelet transform: effects of basis and decomposition level. *International Journal of Wavelets, Multiresolution and Information Processing 6*, 1 (2008), 37–50.

[72] Li, S., and Kang, X. Fast multi-exposure image fusion with median filter and recursive filter. *IEEE Transactions on Consumer Electronics 58*, 2 (2012), 626–632.

[73] Li, S., Kang, X., Fang, L., Hu, J., and Yin, H. Pixel-level image fusion: A survey of the state of the art. *Information Fusion 33* (2017), 100–112.

[74] Li, S., Kang, X., and Hu, J. Image fusion with guided filtering. *IEEE Transactions on Image Processing 22*, 7 (2013), 2864–2875.

[75] Li, S., Kang, X., Hu, J., and Yang, B. Image matting for fusion of multi-focus images in dynamic scenes. *Information Fusion 14*, 2 (2013), 147–162.

[76] Li, S., Kwok, J., Tsang, I., and Wang, Y. Fusing images with different focuses using support vector machines. *IEEE Transactions on Neural Networks 15*, 6 (2004), 1555–1561.

[77] Li, S., Kwok, J., and Wang, Y. Multifocus image fusion using artificial neural networks. *Pattern Recognition Letters 23*, 8 (2002), 985–997.

[78] Li, S., and Yang, B. Multifocus image fusion by combining curvelet and wavelet transform. *Pattern Recognition Letters 29*, 9 (2008), 1295–1301.

[79] Li, S., and Yang, B. Multifocus image fusion using region segmentation and spatial frequency. *Image and Vision Computing 26*, 7 (2008), 971–979.

[80] Li, S., Yang, B., and Hu, J. Performance comparison of different multi-resolution transforms for image fusion. *Information Fusion 12*, 2 (2011), 74–84.

[81] Liu, Y., Liu, S., and Wang, Z. Multi-focus image fusion with dense sift. *Information Fusion 23* (2015), 139–155.

[82] Liu, Z., Yin, H., and Fang, B. A novel fusion scheme for visible and infrared images based on compressive sensing. *Optics Communications 335* (2015), 168–177.

[83] Lu, H., Hu, X., Zhang, L., Yang, S., and Serikawa, S. Local energy based image fusion in sharp frequency localized contourlet transform. *Journal of Computational Information Systems 6*, 12 (2010), 3997–4005.

[84] Lu, H., Zhang, L., and Serikawa, S. Maximum local energy: An effective approach for multisensor image fusion in beyond wavelet transform domain. *Computers and Mathematics with Applications 64*, 5 (2012), 996–1003.

[85] Luo, X., Zhang, J., and Dai, Q. A regional image fusion based on similarity characteristics. *Signal Processing 92*, 5 (2012), 1268–1280.

[86] Maurer, U., Smailagic, A., Siewiorek, D., and Deisher, M. Activity recognition and monitoring using multiple sensors on different body positions. *International workshop on wearable and implantable, body sensor networks* (2006), 113–116.

[87] Merah, M., Abdelmalik, T., and Larbi, B. R-peaks detection based on stationary wavelet transform. *Computer Methods and Programs in Biomedicine 121*, 3 (2015), 149–160.

[88] Michael, D. M. Data fusion for vehicle-borne mine detection. *Proc. International Conference on Detection of Abandoned Land Mines* (1996), 167–171.

[89] Mitchell, H. *Image Fusion: Theories, techniques and applications*. 2010.

[90] Mitianoudis, N., and Stathaki, T. Pixel-based and region-based image fusion schemes using ica bases. *Information Fusion 8*, 2 (2007), 131–142.

[91] Naidu, V. Discrete cosine transform-based image fusion. *Defence Science Journal 60*, 1 (2010), 48–54.

[92] Naidu, V. Image fusion technique using multi-resolution singular value decomposition. *Defence Science Journal 61*, 5 (2011), 479–484.

[93] Naidu, V. Hybrid ddct-pca based multi sensor image fusion. *Journal of Optics 43*, 1 (2014), 48–61.

[94] Naidu, V., and Raol, J. Pixel-level image fusion using wavelets and principle component analysis. *Defence Science Journal 58*, 3 (2008), 338–352.

[95] Nandal, A., and Rosales, H. Enhanced image fusion using directional contrast rules in fuzzy transform domain. *Nandal and Rosales SpringerPlus 5*, 1 (2016), 1–17.

[96] Nason, G., and Silverman, B. The stationary wavelet transform and some statistical applications. *Wavelets and Statistics 103* (1995), 281–299.

[97] Nayar, S., and Nakagawa, Y. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence 16*, 8 (1994), 824–831.

[98] Nikolov, S., Hill, P., Bull, D., and Canagarajah, N. Wavelets for image fusion. *Proc. Wavelets in Signal and Image Analysis, Computational Imaging and Vision Series* (2001), 213–244.

[99] Nunez, J., Otazu, X., and Fors, O. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing 37*, 3 (1999), 1204–1211.

[100] Pajares, G., and de la Cruz, J. A wavelet-based image fusion tutorial. *Pattern Recognition 37*, 9 (2004), 1855–1872.

[101] Parmar, K., and Kher, R. A comparative analysis of multimodality medical image fusion methods. *Proc. Sixth Asia Modelling Symposium (AMS), IEEE* (2012), 93–97.

[102] Petrovic, V. Multisesor pixel-level image fusion, 2001.

[103] Petrovic, V. Subjective tests for image fusion evaluation and objective metric validation. *Information Fusion 8*, 2 (2007), 208–216.

[104] Petrovic, V., and Xydeas, C. Gradient-based multiresolution image fusion. *IEEE Transactions on Image Processing 13*, 2 (2004), 228–237.

[105] Petscnigg, G., Agrawala, M., Hoppe, H., Szeliski, R., Cohen, M., and Toyama, K. Digital photography with flash and no-flash image pairs. *ACM Transactions on Graphics (TOG) 23*, 3 (2004), 664–672.

[106] Phamila, A., and Amutha, R. Low complexity multifocus image fusion in discrete cosine transform domain. *Optica Applicata XLIII*, 4 (2013), 693–706.

[107] Piella, G. Adaptive wavelets and their applications to image fusion and compression, 2003.

[108] Pohl, C., and Genderen, J. Review article multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing 19*, 5 (1998), 823–854.

[109] QU, X., Yan, J., and Yang, G. Multifocus image fusion method of sharp frequency localized contourlet transform domain based on sum-modified-laplacian. *Optics and Precision Engineering 17*, 5 (2009), 1203–1212.

[110] QU, X., Yan, J., and Yang, G. Sum-modified-laplacian based multifocus image fusion method in cycle spinning sharp frequency localized contourlet transform domain. *Optics and Precision Engineering* (2009), 1–10.

[111] Ranchin, T., and Wald, L. The wavelet transform for the analysis of remotely sensed images. *International Journal of Remote Sensing 14*, 3 (1993), 615–619.

[112] Rashidha, R., and Simon, P. A modified directional weighted median filter using second order difference based detection for impulse noise removal. *Proc. Third International Conference on Advanced Computing (ICoAC), IEEE* (2012), 56–61.

[113] Ray, L., and Adhami, R. Dual tree discrete wavelet transform with application to image fusion. *Proc. 38th Southeastern Symposium on System Theory, IEEE* (2006), 430–433.

[114] Rockinger, O. Image sequence fusion using a shift-invariant wavelet transform. *Proc. International Conference on Image Processing, IEEE* (1997), 288–291.

[115] Sarkar, A., Changder, S., and Mandal, J. Directional multi-scaled fusion based median filter for removal of rvin. *Proc. International Conference on Computational Intelligence in Data Mining (CIDM), Springer 2* (2014), 637–646.

[116] Selesnick, I., Baraniuk, R., and Kingsbury, N. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine 22*, 6 (2005), 123–151.

[117] Shahdoosti, H., and Ghassemian, H. Combining the spectral pca and spatial pca fusion methods by an optimal filter. *Information Fusion 27* (2016), 150–160.

[118] Shangli, C., Junmin, H., and Zhongwei, L. Medical images of pet/ct weighted fusion based on wavelet transform. *Proc. Second International Conference on Bioinformatics and Biomedical Engineering (ICBBE), IEEE* (2008), 2523–2525.

[119] Singh, R., and Khare, A. Fusion of multimodal medical images using daubechies complex wavelet transform a multiresolution approach. *Information Fusion 19* (2014), 49–60.

[120] Singh, R., Srivastava, R., Prakash, O., and Khare, A. Dtcwt based multimodal medical image fusion. *Proc. International conference on Signal, Image and Video processing, Springer* (2012), 403–407.

[121] Smith, M., and Heather, J. A review of image fusion technology in 2005. *Proc. SPIE, Thermosense XXVII 5782* (2005), 29–45.

[122] Smith, S., and Brady, J. Susan a new approach to low level image processing. *International Journal of Computer Vision 23*, 1 (1997), 45–78.

[123] Stanciu, S., Dragulinescu, M., and Stanciu, G. Sum-modified-laplacian fusion methods experimented on image stacks of photonic quantum ring laser devices collected by confocal scanning laser microscopy. *U.P.B. Sci. Bull. 73*, 2 (2011), 139–146.

[124] Stathaki, T. *Image fusion: Algorithms and applications.* 2008.

[125] Tamilselvan, K., and Murugesan, G. Survey and analysis of various image fusion techniques for clinical ct and mri images. *Proc. International Journal of Imaging Systems and Technology 24*, 2 (2014), 193–202.

[126] Tang, J. A contrast based image fusion technique in the dct domain. *Digital Signal Processing 14*, 3 (2004), 218–226.

[127] Tao, Q., and Veldhuis, R. Threshold-optimized decision-level fusion and its application to biometrics. *Pattern Recognition 42*, 5 (2009), 823–836.

[128] Tao, W., Chenchen, Z., and Zengchang, Q. Multifocus image fusion based on robust principal component analysis. *Pattern Recognition Letters 34*, 9 (2013), 1001–1008.

[129] Tian, J., Chen, L., Ma, L., and Yu, W. Multi-focus image fusion using a bilateral gradient-based sharpness criterion. *Optics Communications 284*, 1 (2011), 80–87.

[130] Tian, P., and Ni, G. Contrast based image fusion using the discrete wavelet transform. *Optical Engineering 39*, 8 (2000), 2075–2082.

[131] Toet, A. Image fusion by a ratio of low-pass pyramid. *Pattern Recognition Letters 9*, 4 (1989), 245–253.

[132] Toet, A., Ruyven, L., and Valeton, J. Merging thermal and visual images by a contrast pyramid. *Optical Engineering 28*, 7 (1989), 789–792.

[133] Tomasi, C., and Manduchi, R. Bilateral filtering for gray and color images. *Proc. Sixth International Conference on Computer Vision, ACM* (1998), 839–846.

[134] Tsagaris, V., Anastassopoulos, V., and Lampropoulos, G. Fusion of hyperspectral data using segmented pct for color representation and classification. *IEEE Transactions on Geoscience and Remote Sensing 43*, 10 (2005), 2365–2375.

[135] Tseng, D., Liu, Y., and Chou, C. Image fusion with contrast improving and feature preserving. *Mathematical Problems in Engineering* (2014), 1–14.

[136] Upla, K., Joshi, M., and Gajjar, P. An edge preserving multiresolution fusion: use of contourlet transform and mrf prior. *IEEE Transactions on Geoscience and Remote Sensing 53*, 6 (2015), 3210–3220.

[137] Vivone, G., Alparone, L., Chanussot, J., Mura, M. D., Garzelli, A., Licciardi, G., Restaino, R., and Wald, L. A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing 53*, 5 (2015), 2565–2586.

[138] Walczak, B., Bogaert, B., and Massart, D. Application of wavelet packet transform in pattern recognition of near-ir data. *Analytical Chemistry 68*, 10 (1996), 1742–1747.

[139] Wan, T., Canagarajah, N., and Achim, A. Segmentation-driven image fusion based on alpha-stable modeling of wavelet coefficients. *IEEE Transactions on Multimedia 11*, 4 (2009), 624–633.

[140] Wang, H., Peng, J., and Wu, W. A fusion algorithm of remote sensing based on discrete wavelet packet. *Proc. International Conference on Machine learning and cybernetics, IEEE* (2003), 2557–2562.

[141] Wang, L., Li, B., and Tian, L. Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients. *Information Fusion 19*, 1 (2014), 20–28.

[142] Wang, P., Tian, H., and Zheng, W. A novel image fusion method based on frft-nsct. *Mathematical Problems in Engineering 2013* (2013), 1–9.

[143] Wang, Z., Ma, Y., and Gu, J. Multi-focus image fusion using pcnn. *Pattern Recognition 43*, 6 (2010), 2003–2016.

[144] Wang, Z., Ziou, D., Armenakis, C., and Li, D. A comparative analysis of image fusion methods. *IEEE transactions on Geoscience and Remote Sensing 43*, 6 (2005), 1391–1402.

[145] Wenzhong, S., ChangQing, Z., Yan, T., and Janet, N. Wavelet-based image fusion and quality assessment. *International Journal of Applied Earth Observation and Geoinformation 6* (2005), 241–251.

[146] Xu, M., Chen, H., and Varshney, P. An image fusion approach based on markov random fields. *IEEE Transactions on Geoscience and Remote Sensing 49*, 12 (2011), 5116–5127.

[147] Xydeas, C., and Petrovic, V. Objective image fusion performance measure. *Electronics Letters 36*, 4 (2000), 308–309.

[148] Yang, B., and Li, S. Multifocus image fusion and restoration with sparse representation. *IEEE Transactions on Instrumentation and Measurement 59*, 4 (2010), 884–892.

[149] Yang, B., and Li, S. Pixel-level image fusion with simultaneous orthogonal matching pursuit. *Information Fusion 13*, 1 (2012), 10–19.

[150] Yang, Q. Recursive bilateral filtering. *Proc. European Conference on Computer Vision (ECCV), Springer* (2012), 399–413.

[151] Yang, S., Wang, M., and Jiao, L. Image fusion based on a new contourlet packet. *Information Fusion 11*, 2 (2010), 78–84.

[152] Yang, Y. A novel dwt based multi-focus image fusion method. *Proc. International Conference on Advances in Engineering 24* (2011), 177–181.

[153] Yang, Y., Huang, S., Gao, J., and Qian, Z. Multi-focus image fusion using an effective discrete wavelet transform based algorithm. *Measurement science review 14*, 2 (2014), 102–108.

[154] Yang, Y., Park, D. S., Huang, S., and Rao, N. Medical image fusion via an effective wavelet-based approach. *EURASIP Journal on Advances in Signal Processing 10* (2010), 1–13.

[155] Yao, Y., Guo, P., Xin, X., and Jiang, Z. Image fusion by hierarchical joint sparse representation. *Cognitive Computation 6*, 3 (2014), 281–292.

[156] Yin, H., and Li, S. Multimodal image fusion with joint sparsity model. *Optical Engineering 50*, 6 (2011), 1–10.

[157] Yu, B., Jia, B., Ding, L., Cai, Z., Wu, Q., Law, R., Huang, J., Song, L., and Fu, S. Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion. *Neurocomputing 182* (2016), 1–9.

[158] Zeng, K., Ma, K., Hassen, R., and Wang, Z. Perceptual evaluation of multi-exposure image fusion algorithms. *Proc. 6th International Workshop on Quality Multimedia Experience (QoMEX), IEEE* (2014), 7–12.

[159] Zhang, N., and Wu, Q. Effects of brovey transform and wavelet transform on the information capacity of spot -5 imagery. *Proc. SPIE, International Symposium on Photoelectronic Detection and Imaging: Image Processing 6623* (2008), 1–5.

[160] Zhang, Q., and Guo, B. Multifocus image fusion using the nonsubsampled contourlet transform. *Signal Processing 89*, 7 (2009), 1334–1346.

[161] Zhang, Q., Xu, L., and Jia, J. 100+ times faster weighted median filter. *Proc. International Conference on Computer Vision and Pattern Recognition (CVPR), IEEE* (2014), 2830–2837.

[162] Zhang, Z., and Blum, R. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proc. of the IEEE 87*, 8 (1999), 1315–1326.

[163] Zhao, X., Ye, B., and Chen, T. Theory of multi-resolution singular value decomposition and its application to signal processing and fault diagnosis. *Journal of Mechanical Engineering 46*, 20 (2010), 64–75.

[164] Zhou, Z., Li, S., and Wang, B. Multi-scale weighted gradient-based fusion for multi-focus images. *Information Fusion 20* (2014), 60–72.

[165] Zhu, Y., Varshney, P., and Chen, H. Evaluation of ica based fusion of hyperspectral images for color display. *Proc. International Conference on Information Fusion, IEEE* (2007), 1–7.

[166] Zribi, M. Non-parametric and region-based image fusion with bootstrap sampling. *Information Fusion 11*, 2 (2010), 85–94.

List of Publications

List of Publications

1. Sonam and Kumar, M. Discrete Wavelet Transform and Cross Bilateral Filter based Image Fusion. International Journal of Intelligent Systems and Applications (IJISA), MECS, 2017; Vol. 9 (1), pp. 37-45.

2. Sonam and Kumar, M. An Effective Image Fusion Technique based on Multiresolution Singular Value Decomposition, INFOCOMP Journal of Computer Science, 2015; Vol. 14 (2), pp. 31-43.

3. Sonam and Kumar, M. Wavelet Transform based Noisy and Defocused Image Enhancement using Fusion. International Journal of Electronics, Electrical and Computational System (IJEECS), 2015; Vol. 4, pp. 208-217.

4. Sonam and Kumar, M. An Efficient Multi-focus Image Fusion Approach Based on DWT. In Proceedings of the Fifth International Conference on Signal, Networks, Computing, and Systems (ICSNCS), Lecture Notes in Electrical Engineering, 2016; Vol. 395 (1), pp. 45-52, Springer, India.

5. Sonam and Kumar, M. A Local correlation and Directive contrast based Image Fusion. In Proceedings of the International Conference on Computer Vision and Image Processing (CVIP), Advances in Intelligent Systems and Computing, 2016; Vol. 459, pp. 419-430, Springer, India.

6. Sonam and Kumar, M. A Novel Hybrid Approach for Image Fusion Using DCT and SML. In Proceedings of the Fourth International Conference on Parallel,

Distributed and Grid Computing (PDGC), 2016; pp. 337-342, IEEE, India.

7. Sonam and Kumar, M. An Efficient Image Fusion Technique based on DTCWT. Communicated in International Conference on Advances in Computing and Data Sciences (ICACDS), 2018.

An Effective Image Fusion Technique based on Multiresolution Singular Value Decomposition

SONAM GAUTAM[1]
MANOJ KUMAR[2]

[1,2]Department of Computer Science,
[1,2]Babasaheb Bhimrao Ambedkar University,
Vidya Vihar, Raebareli Road, Lucknow (U.P.), India.
[1]sonam870115@gmail.com
[2]mkjnuiitr@gmail.com

Abstract. The main objective of image fusion is to combine the multiple images of same scene into a single image having more information and better visual appearance, without introducing any artifacts. In this paper, a new multifocus and multisensor image fusion scheme based on Multiresolution Singular Value Decomposition (MSVD) and gradient based sharpness approach using Discrete Wavelet Transform (DWT) is proposed. The proposed scheme is conducted into two parts. In first part, each of the source images is decomposed using DWT and high frequency DWT coefficients are directly fused by gradient based sharpness method. In second part, low frequency DWT coefficients are further decomposed using MSVD and fused by averaging and maximum methods. To reconstruct the final fused image, inverse DWT is performed. The resultant fused image is enhanced and complete as compared to any of the source images. The experimental results obtained by proposed scheme are compared with the existing schemes. The experimental results and comparison demonstrate the effectiveness of proposed scheme.

Keywords: Discrete wavelet transform, multiresolution singular value decomposition, peak-signal-to-noise-ratio, correlation coefficients, standard deviation.

(Received October 8th, 2015 / Accepted February 4th, 2016)

1 Introduction

In the recent development of technologies, image fusion has become an important and useful application for image analysis, image enhancement and computer vision [1]. Image fusion is the process of integrating all significant information from two or more images of same scene into a single composite image. The main objective of image fusion is to make the fused image more informative with better visual appearance. In image fusion, it is tried to obtain fused image without any distortion and loss of information. Image fusion has been successfully used in various fields such as remote sensing, military, medical diagnosis, robotics, surveillance, etc. Generally, a single image of complex scene does not contain enough information for appropriate analysis of the scene. If one image is focused on object and other is out-of-focus, then to get entire scene in focus in a single image is a difficult task. Although, combining the details of the source images appropriately, a single image with more information can be obtained. The process of combining the details of different focusing target images into a single image with entire scene in focus is known as the multi-focus image fusion.

If the images of same scene are taken by different sensors then, each of the source images gives some relevant information according to its sensor ability. To improve the quality and information of images, the information of different sensor images is combined using fusion process. For example, CT (Computed Tomography) and MRI (Magnetic Resonance Imaging)

images are captured by different sensors to diagnose the diseases. CT images are best suited for viewing the information of bone, blood vessels and soft tissues, whereas MRI suited for viewing much more information of soft tissues. Individually, these source images may not give sufficient information for a single body part having bones and tissues. Therefore, the details of CT and MRI are combined to achieve a single fused image with all relevant information of the part. This kind of fusion process is known as multi-sensor image fusion.

Image fusion algorithms are broadly classified into three categories: pixel level, feature level and decision level. Pixel level fusion defines the process to combine relevant information from each source images pixel by pixel into a single composite image containing more source details. Pixel level fusion is the lowest level of image fusion. It is easy to implement and time efficient, therefore most of the image fusion techniques are based on pixel level method [2]. In feature level fusion, by extracting the features such as color, edges and texture from all source images, fusion is performed based on features with some certain selection criteria [3]. Decision level fusion is the highest level of image fusion. It combines the results from multiple algorithms to provide a final decision for fusion [4]. Fusion techniques are generally performed in two domains: spatial and transform [5]. In spatial domain, pixel by pixel fusion is performed over the all source images and a single fused image is achieved. Some spatial domain techniques of fusion are: average method [6] [7], Principal Component Analysis (PCA) method [8] [9], Intensity-Hue-Saturation (IHS) method [10] [11], High Pass Filter (HPF) [12], etc. Generally, spatial domain methods produce several undesired effects, such as distortion and reduced contrast. To overcome these problems, transform domain approaches were proposed, which provide directional information in all decomposition and contain unique information at different resolutions. Various kind of multi-resolution transform approaches have been proposed for image fusion, including DWT [13], Stationary Wavelet Transform (SWT) [33], Curvelet Transform (CVT) [30], Nonsubsampled Contourlet Transform (NSCT) [31], etc. In [22], the performances of image fusion algorithms using multi-resolution transforms are compared with CVT and contourlet to obtain better fusion results. Another, multifocus image fusion technique is given using Dual-Channel Pulse Coupled Neural Networks (Dual-Channel PCNN) [23] which incorporates focus measure of source images to compute weighted coefficients. In [24], multifocus image fusion and restoration is presented using sparse-representation. In this method, sparse coefficients are calculated from the source images and these coefficients are combined using maximum fusion rule. Similarly, another sparse representation image fusion technique [25] has been proposed based on overlapping patches instead of the whole image and simultaneously the orthogonal matching pursuit technique is used which decomposes the source images into the same subset of dictionary bases. Image fusion based on contourlet packet is introduced followed by a Nonsubsampled Directional Filter Bank (NSDFB) [26] which provides a more accurate fused image than wavelet packet method. The new multifocus image fusion based on sharpness criterion [27] is given to enhance the sharp information and remove blur details of the images. In [28], multisensor image fusion in remote sensing is presented, it describes mainly pixel based image fusion of Earth observation satellite data. A multi-resolution image fusion [29] based approach is proposed to combine the high and low resolution images data by adding some wavelet planes to the low resolution intensity component and obtained a better fused image. DWT based fusion [13] is proposed, in which the maximum selection rule is used. This simple scheme just selects the largest absolute wavelet coefficients at each location from the input images. Similarly, another fusion technique is introduced based on wavelets and principal component analysis [9]. In which source images are decomposed into low and high frequency subbands using DWT and obtained coefficients are transformed into uncorrelated coefficient and the eigenvalues are evaluated from principal components (PCs). The scheme selects two highest principal components and represent these as weights for fusion rule. Finally, inverse DWT is performed to achieve a final fused image.

In [14], a multiresolution singular value decomposition based scheme is proposed which gives better results than DWT based scheme [13] both in terms of quality and execution time. In this scheme, maximum selection rule is performed over detailed coefficients and averaging is performed over approximation coefficients to get the final fused image. In DWT, multiresolution decomposition is based on low and high pass filters. Low pass filters outputs to approximation part which contains most of the information of image whereas, high pass filter outputs to detailed subbands which contain directional (horizontal, vertical and diagonal) information. However, on the other hand, the decomposition in MSVD is based on singular/eigenvalues. Here, the approximation part is corresponding to larger eigenvalues and contains most of

the information whereas, the detailed coefficient correspond to edges, texture, boundaries and other sharp changes in image [14]. With this motivation, we propose a new image fusion scheme using DWT and MSVD. In this scheme, the DWT is applied on both the source images and approximation and detailed components are obtained. Since the detailed components contain directional information therefore it is quite logical to apply the gradient based fusion rule directly on the detailed coefficients of both the images. However, the obtained approximation part is further decomposed into low and high frequency coefficients using MSVD. Again, the obtained low frequency coefficients contain the most of the average information, therefore the averaging method is the best suited for fusion, whereas the detailed coefficients are responsible for sharper changes and the maximum selection method is quite appropriate to fuse the detailed coefficients. Finally, the inverse MSVD and inverse DWT is performed on fused coefficients to get the fused image.

The remainder of the paper is organized as follows: basic theory of DWT and MSVD is discussed in Section 2. Proposed scheme of image fusion is introduced in Section 3. Experimental results followed by discussion are given in Section 4. Finally, concluding remarks are presented in Section 5.

2 Basic theory of DWT and MSVD

In this paper, discrete wavelet transform and multiresolution singular value decomposition are used for multifocus and multisensor image fusion. In this section, basic theories of DWT and MSVD are discussed.

2.1 Discrete Wavelet Transform (DWT)

Due to the multiresolution property, discrete wavelet transform is widely used in image processing [13]. DWT is a technique, which converts an image from spatial domain to frequency domain. It is used to analyze an image at different resolutions. We can obtain horizontal, vertical and diagonal information of the images using DWT. At first level decomposition, DWT decomposes the image into two parts: approximation and detailed parts. Approximation part contains one low frequency subband (LL) and detailed parts contain three high frequency subbands (LH, HL and HH), as shown in Fig. 1(a). Most of the information of image is contained in approximation part. For second level decomposition, approximation part is further decomposed into four frequency subbands, as shown in Fig. 1(b). The decomposition levels can be increased as per the requirement.

2-D DWT [19] for image *f(x, y)* of size $m \times n$ is defined as

$$W_\phi(j_0,u,v) = \frac{1}{\sqrt{mn}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y)\phi_{j_0,u,v}(x,y) \quad (1)$$

$$W_\psi(j,u,v) = \frac{1}{\sqrt{mn}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y)\psi_{j,u,v}(x,y) \quad (2)$$

where, Eqs. (1) and (2) are respectively approximation and detailed coefficients of image *f(x, y)*. Finally, inverse DWT is used to reconstruct the image. For the above given Eqs. (1) and (2), the inverse DWT is given as

$$f(x,y) = \frac{1}{\sqrt{mn}} \sum_u \sum_v W_\phi(j_0,u,v)\phi_{j_0,u,v}(x,y) + \frac{1}{\sqrt{mn}} \sum_{j=j_0}^{\infty} \sum_u \sum_v W_\psi(j,u,v)\psi_{j,u,v}(x,y) \quad (3)$$

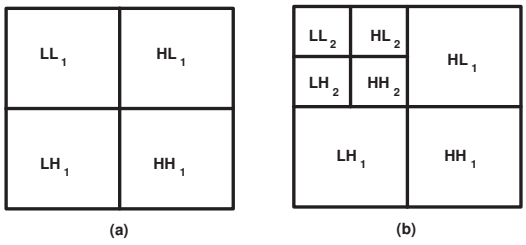

Figure 1: (a) First level 2-D DWT decomposition; (b) second level 2-D DWT decomposition.

Fig. 2(a) shows the original synthetic image and Fig. 2(b) shows its first level decomposition using DWT. The directional information (HL, LH and HH) is more clear in detailed subbands in DWT based decomposition.

2.2 Multiresolution Singular Value Decomposition (MSVD)

A singular value decomposition [15, 16, 17] of a matrix *A* of size $m \times n$ can be written as: $A = USV^T$, where *U* and *V* are *left* and *right singular vectors* of size $m \times m$ and $n \times n$ respectively. *U* and *V* are *orthogonal* matrices and *S* is a diagonal matrix of size $m \times n$ containing eigenvalues in non-increasing order.

For a matrix *A* of size $m \times n$ the MSVD [18] is computed as follows:

1. Divide the matrix *A* into non-overlapping blocks of size $k \times l$ and write each block in the form of a vector.

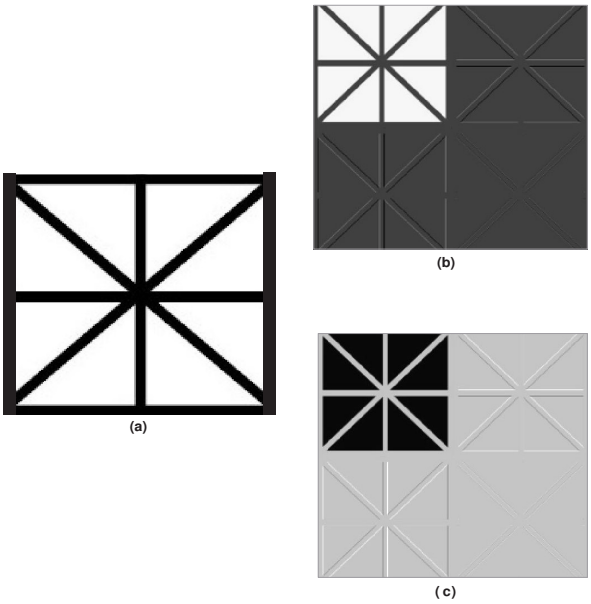

Figure 2: (a) Original image; (b) first level DWT decomposition; (c) first level MSVD decomposition.

2. Stack the vectors in columns and form the matrix A_1 of size $kl \times mn/kl$.

3. Compute the scatter matrix, $T_1 = \bar{A_1}\bar{A_1}^T$ of size $kl \times kl$.

4. Find the orthogonal matrix U_1 for scatter matrix T_1.

5. Compute the diagonal matrix S_1 containing the squares of eigenvalues in decreasing order as, $S_1^2 = U_1^T T U_1$.

6. Construct matrix $\hat{A} = {U_1}^T \bar{A_1}$.

$\hat{A}$ contains approximation and detailed parts. First row of $\hat{A}$ is considered as approximation part containing larger singular values. Similarly, the remaining rows of $\hat{A}$ are regarded as the detailed parts containing remaining singular values, i.e. $\Phi_1 = \hat{A}(1,:)$ and $\Psi_1^i = \hat{A}(i,:)$, where, Ψ_1^i is the i^{th} detailed subbands at level 1. The first level MSVD decomposition of a synthetic image is shown in Fig. 2(c). Here, the decomposition is based on eigenvalues.

3 Proposed Scheme

In this section, the proposed scheme is described in brief. In the proposed scheme, two less informative (multifocus and multisensor) images are used to obtain a fused image having more information. For multifocus and multisensor image fusion, two multifocus and multisensor images are considered as source images. In the proposed work, DWT is applied upto *l-level* over the source images *X* and *Y*, as given in Fig. 3, which decomposes each of the source images into low (LL_l^X and LL_l^Y) and high frequency subbands ((HL_l^X, LH_l^X and HH_l^X) and (HL_l^Y, LH_l^Y and HH_l^Y)) respectively, as discussed in Section 2.1.

Over the low frequency subbands (LL_l^X and LL_l^Y) of both of the transformed images, MSVD is performed. When MSVD is applied, LL_l^X and LL_l^Y of both of the transformed images are further decomposed into low frequency ($\Phi_l^{LL^X}$ and $\Phi_l^{LL^Y}$) and high frequency subbands (($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$)) respectively, as discussed in Section 2.2. Low frequency subbands refers to approximation part and high frequency subbands treated as detailed parts. Obtained low frequency coefficients ($\Phi_l^{LL^X}$, $\Phi_l^{LL^Y}$) of both of the transformed images are fused using pixel averaging method. In pixel averaging fusion method, average pixel values from both of the transformed images are selected to fuse. Whereas, high frequency coefficients ($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$) of both of the transformed images are combined using maximum method. In maximum method, corresponding maximum values from both of the transformed images are selected to fuse. Obtained orthogonal matrices U_l^X and U_l^Y, as discussed in Section 2.2 are fused using averaging fusion method. Further, over the obtained fused coefficients from low and high frequency subbands, an inverse MSVD is performed and set of new fused coefficient is obtained and represented as ω_a.

Remaining high frequency coefficients (HL_l^X, LH_l^X, HH_l^X) and (HL_l^Y, LH_l^Y, HH_l^Y), as given in Fig. 3 of both of the transformed images are fused using gradient based sharpness focus method [20] [21]. The gradient coefficients are computed as follows:

$$\nabla G(z) = [\nabla G_p(z)^2 + \nabla G_q(z)^2]^{1/2} \tag{4}$$

where, $\nabla G_p(z), \nabla G_q(z)$ can be defined as:

$$\nabla G_p(z) = \Big\{ -D(p-1,q-1,r,s) - 2D(p-1,q,r,s) - D(p-1,q+1,r,s) + D(p+1,q-1,r,s) + 2D(p+1,q,r,s) + D(p+1,q+1,r,s) \Big\}$$

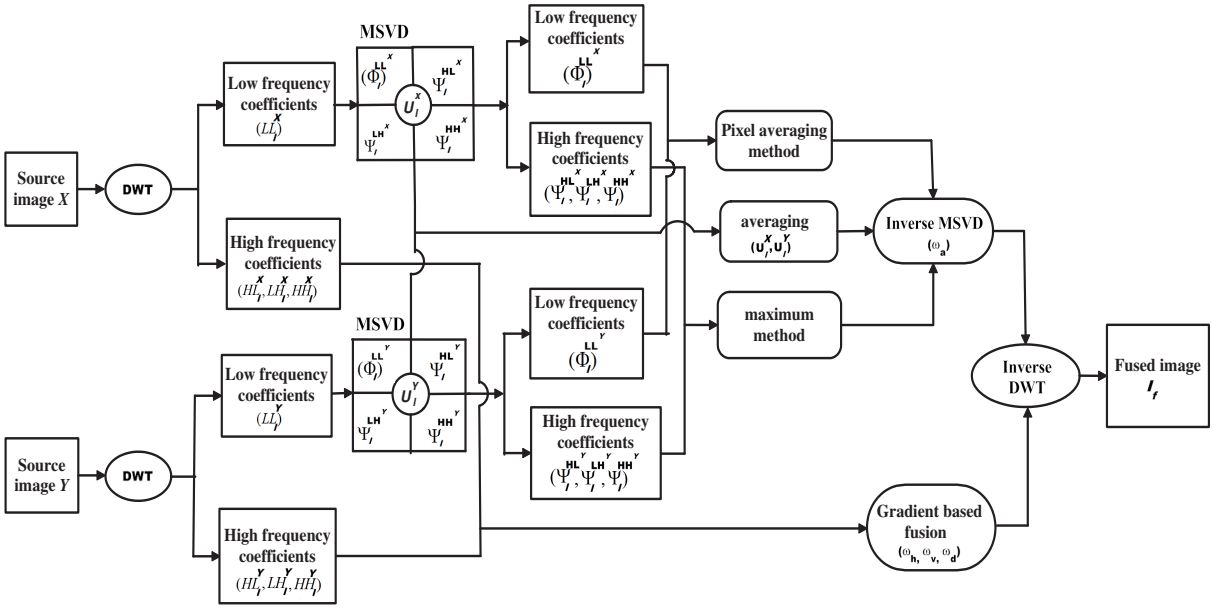

Figure 3: Block diagram of proposed scheme

$$\nabla G_q(z) = \Big\{ D(p-1,q-1,r,s) + 2D(p,q-1,r,s) + D(p+1,q-1,r,s) - D(p-1,q+1,r,s) - 2D(p,q+1,r,s) - D(p+1,q+1,r,s) \Big\}$$

In this method, let $z = (p, q, r, s)$ be represent the index of a particular multiscale decomposition coefficient, where (p, q) represents the spatial position, r decomposition level and s frequency band of multiscale decomposition.

Obtained gradient coefficients of both of the transformed images are fused using pixel averaging method and new fused coefficients (ω_h, ω_v and ω_d) are obtained. Finally, to reconstruct the final fused image (I_f), inverse DWT is performed over the obtained all fused coefficients ($\omega_a, \omega_h, \omega_v$ and ω_d). The block diagram of the proposed scheme is given in Fig. 3. The algorithm is summarized in the following steps:

Algorithm:

1. Take two source images *X* and *Y*.
2. Apply DWT over the source images, which decomposes each of the images into four subbands: (LL_l^X, HL_l^X, LH_l^X and HH_l^X) and (LL_l^Y, HL_l^Y, LH_l^Y and HH_l^Y).
3. Perform MSVD on approximation part (LL_l^X, LL_l^Y) of both of the transformed images and obtain approximation ($\Phi_l^{LL^X}$ and $\Phi_l^{LL^Y}$) and detailed parts (($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$)) (given in Section 2.2).
4. Apply pixel averaging method on $\Phi_l^{LL^X}$ and $\Phi_l^{LL^Y}$ of both of the transformed images and maximum method on high frequency coefficients ($\Psi_l^{HL^X}, \Psi_l^{LH^X}, \Psi_l^{HH^X}$) and ($\Psi_l^{HL^Y}, \Psi_l^{LH^Y}, \Psi_l^{HH^Y}$).
5. Apply averaging method on U_l^X and U_l^Y (discussed in Section 2.2).
6. Apply inverse MSVD on fused coefficients and obtain fused coefficient ω_a.
7. Apply gradient based fusion on detailed parts (HL_l^X, LH_l^X and HH_l^X) and (HL_l^Y, LH_l^Y and HH_l^Y) using averaging method and obtain fused coefficients ($\omega_h, \omega_v, \omega_d$).
8. Perform inverse DWT, over the all fused coefficients ($\omega_a, \omega_h, \omega_v, \omega_d$) to reconstruct the final fused image (I_f).

4 Experimental results

The proposed scheme is tested on several test images of size 512×512 shown in Fig. 4(a), 5(a), 6(a), 7(a) and also of size 256×256 shown in Fig. 8(a) and 9(a). All these images are considered as reference images

(I_r). In the proposed work, DWT method with Haar as the wavelet basis is used. Before applying DWT, each of the reference images in Fig. 4(a), 5(a), 8(a) and 9(a) are convolved by a gaussian blurring of 13×13 window with standard deviations $\sigma = 5$ and multi-focus(defocus) source images *X* and *Y* are obtained. Fig. 6(a), 6(b), 6(c) are considered as dataset1 source images and 7(a) and 7(b) are FLIR and LLTV source images as dataset2 [14]. Fig. 8(a) and 9(a) are the TNO's UN camp and head monument visual images, which are online available at `http://www.deakin.edu.au/~mhossny/fusion/`.

Fig. 4(b), 8(b) and 9(b) are blurred on left part, whereas Fig. 4(c), 8(c) and 9(c) are blurred on right part. Fig. 5(b), 6(b) are upper side blurred images and Fig. 5(c), 6(c) are lower side blurred images, respectively. The results of our proposed scheme is compared with DWT [13], DWT with PCA [9], MSVD [14], SWT [33], CVT [30], NSCT [31] based fusion methods. The results of the respective existing schemes are shown in Fig. 4(d-i), 5(d-i), 6(d-i), 7(d-i), 8(d-i), and 9(d-i).

Results obtained by our proposed scheme are given in Fig. 4(j), 5(j), 6(j), 7(j), 8(j), and 9(j). It can be visually seen that the resultant images from proposed scheme contain better information than other existing schemes. But only visual inspection is not sufficient to measure the quality of images. To measure the quality of fused images quantitatively some parameters are used such as PSNR, correlation coefficients (C.C.), mean, standard deviation (S.D.), mutual information *(MI)*, Q_0, Q_W, Q_E and $Q^{AB/F}$. The higher values of these parameters stand for better fusion result. The superiority of the results is highlighted by bold letters and shown in Table 1. It is clear that in most of the cases, results by the proposed scheme are better as compared to other existing schemes.

The above metrics are defined as follows:

1. PSNR [34] quantitatively evaluates the error between one of the reference and fused images. MSE indicates how much error fused image conveyed in the reference image.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \tag{5}$$

where, MSE is defined as:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I_r(i,j) - I_f(i,j)]^2$$

where, I_r and I_f denote the ideal reference image and fused image. PSNR values of proposed scheme and existing schemes for various test images are calculated. Higher PSNR values indicate better visual quality. From Table 1, it is evident that the proposed scheme is achieved better PSNR results with good visual information as comparison to other existing schemes.

2. The correlation coefficient [34] measures the degree in which two variables are linearly related. The value of correlation coefficient lying between $[0, 1]$. Correlation coefficient between $m \times n$ is defined as:

$$\rho = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} [I_r(i,j) - I_f(i,j)]^2}{mn}} \tag{6}$$

Correlation coefficient values for proposed scheme and existing schemes are calculated between the reference and fused images. The correlation coefficient values near to 1 indicates the reference image and fused image contain almost similar information. It is observed from Table 1 that the proposed scheme has scored better correlation result.

3. The mean and standard deviation [34] are defined as:

$$\hat{\mu} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} I_f(i,j) \tag{7}$$

$$\sigma = \sqrt{\frac{1}{mn-1} \sum_{i=1}^{m} \sum_{j=1}^{n} (I_f(i,j) - \hat{\mu})^2} \tag{8}$$

For various test images, standard deviation is calculated for proposed and existing schemes. Higher standard deviation value indicate high contrast image. Hence, it is observed from Table 1 that the proposed scheme provides better results in comparison to others.

4. Mutual information *(MI)* [22] is a metric measure reflects that the total amount of information that the fused image contains of source images. It is defined as the sum of mutual information between each source image and the fused image. The mutual information I_{X,I_f} between source *X* and fused image I_f is given as follows:

$$I_{X,I_f} = \sum_{i,j} h_{X,I_f}(i,j) \log \frac{h_{X,I_f}(i,j)}{h_X(i) h_{I_f}(j)} \tag{9}$$

where, h_{X,I_f} is the jointly normalized histogram; h_X and h_{I_f} are normalized histogram of *X* and I_f,

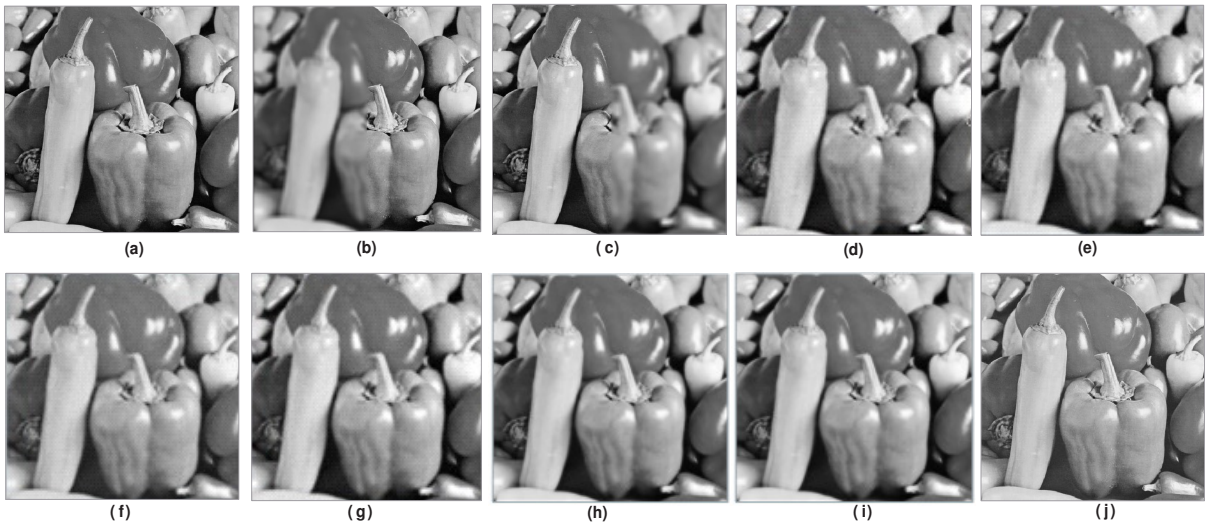

Figure 4: Fusion results for pepper image: (a) Reference image (original image); (b) image blurred on the left; (c) image blurred on the right; (d) fused image by DWT; (e) fused image by DWT + PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

Figure 5: Fusion results for lena image: (a) Reference image (original image); (b) blurred on upper part; (c) blurred on lower part; (d) fused image by DWT; (e) fused image by DWT + PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

respectively. Similarly, I_{Y,I_f} represents the mutual information between other source image *Y* and fused image I_f. Hence, the total mutual information *(MI)* between the source images X, Y and fused image I_f is given as:

$$MI = I_{X,I_f} + I_{Y,I_f} \tag{10}$$

The larger *MI* value represents better fusion result. From Table 1, it is observed that the results obtained from proposed scheme are better than others.

5. The metric Q_0 [22] measures distortion of the fused image, it is a combination of three components as correlation, luminance and contrast [32]. The metric Q_0 between the source image *X* and the fused image I_f is defined as follows:

$$Q_0(X, I_f) = \frac{4\sigma_{ij}\bar{i}\bar{j}}{(\sigma_i^2 + \sigma_j^2)(\bar{i}^2 + \bar{j}^2)} \tag{11}$$

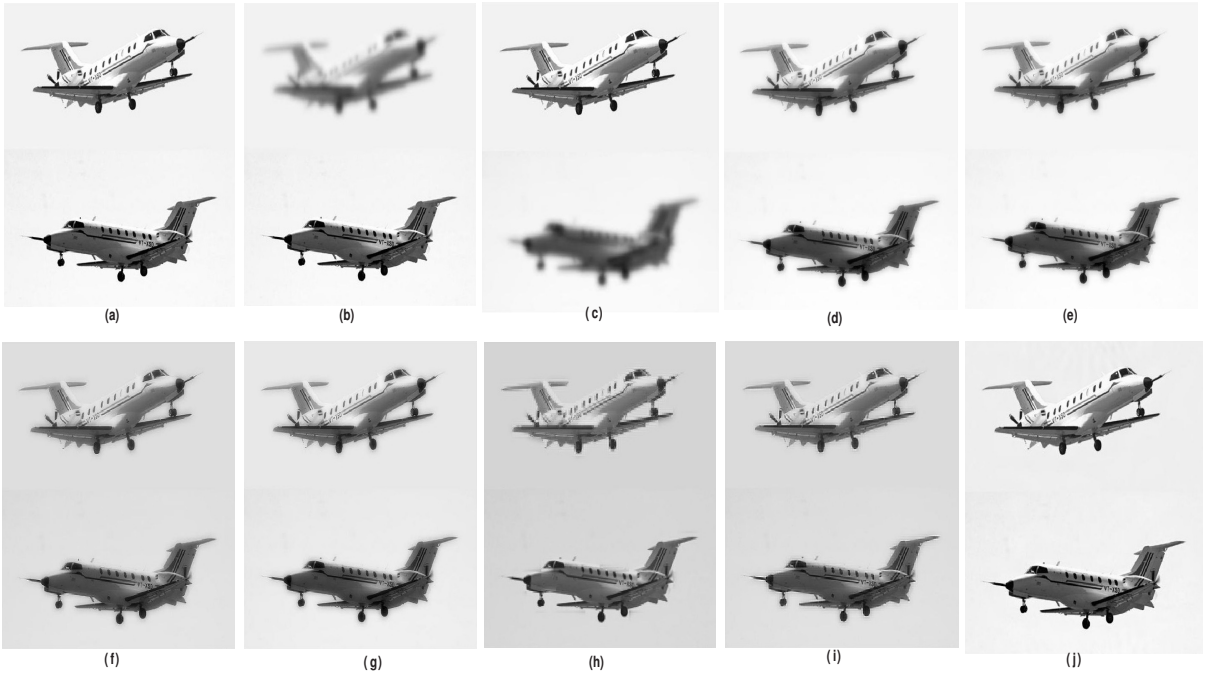

Figure 6: (a) Reference image (original image); (b) blurred on upper part; (c) blurred on lower part; (d) fused image by DWT; (e) fused image by DWT + PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

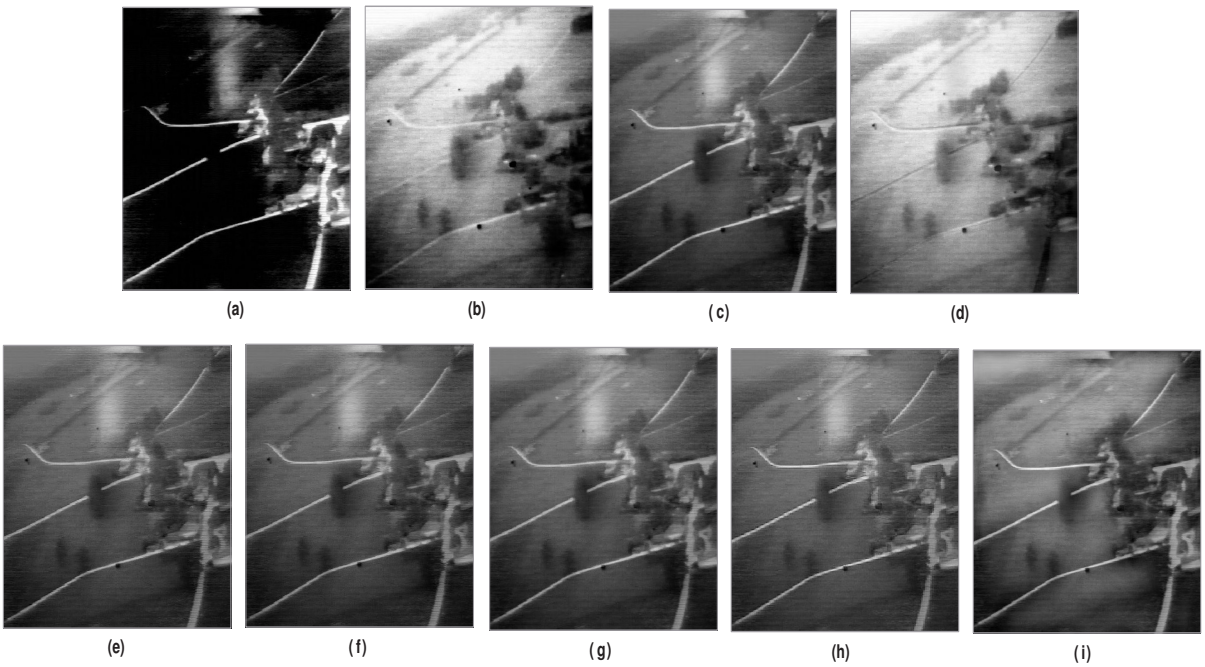

Figure 7: (a) FLIR image; (b) LLTV image; (c) fused image by DWT; (d) fused image by DWT + PCA; (e) fused image by MSVD; (f) fused image by SWT; (g) fused image by CVT; (h) fused image by NSCT; (i) fused image by proposed scheme.

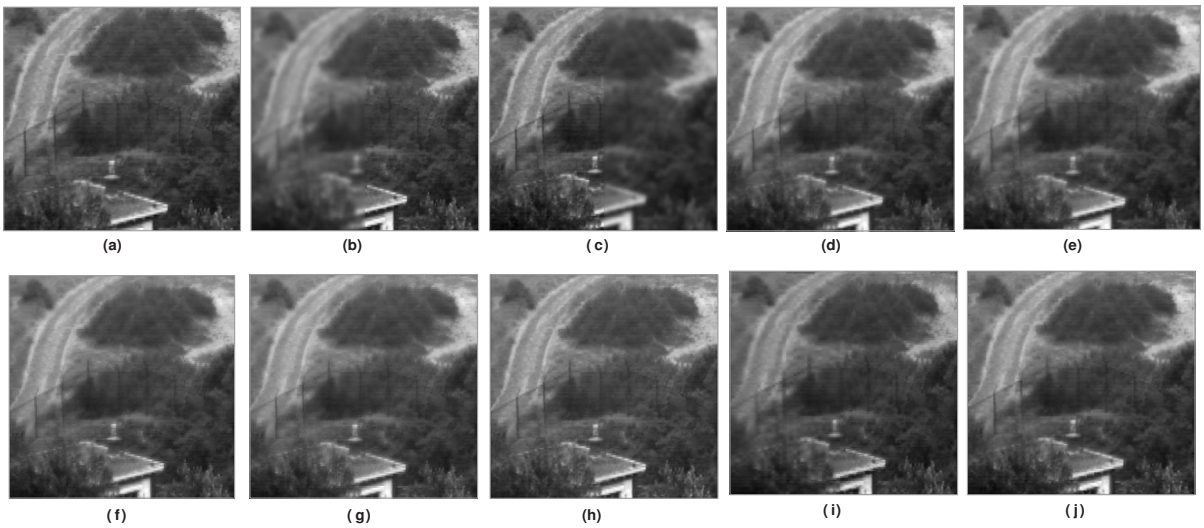

Figure 8: Fusion results for TNO's UN camp image: (a) Reference image (original image); (b) blurred on left part; (c) blurred on right part; (d) fused image by DWT; (e) fused image by DWT + PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

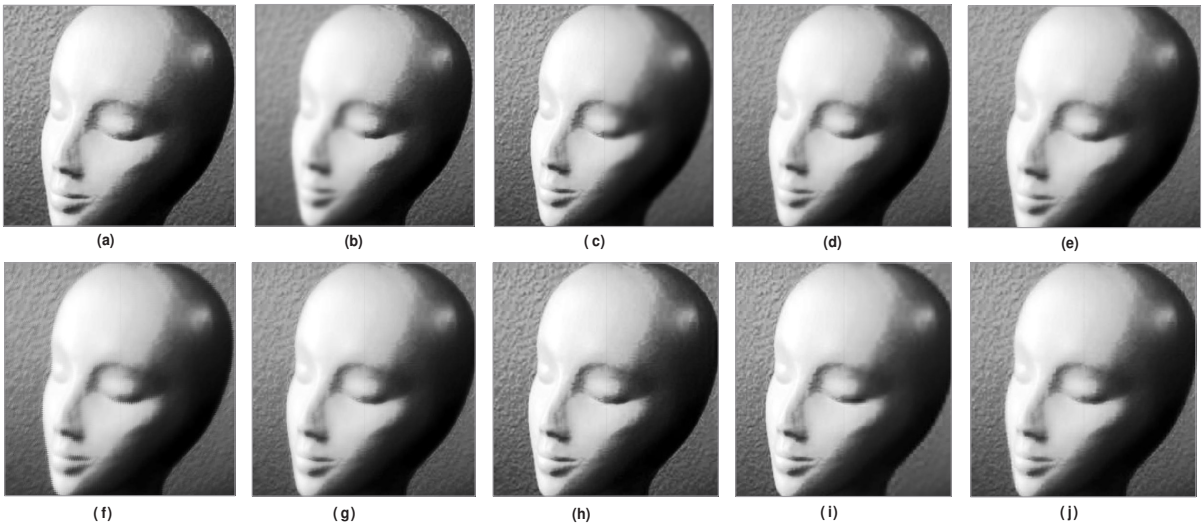

Figure 9: Fusion results for head monument image: (a) Reference image (original image); (b) blurred on left part; (c) blurred on right part; (d) fused image by DWT; (e) fused image by DWT + PCA; (f) fused image by MSVD; (g) fused image by SWT; (h) fused image by CVT; (i) fused image by NSCT; (j) fused image by proposed scheme.

where, σ_{ij} denotes the covariance, σ_i^2, σ_j^2 represent the variance and $\bar{i}, \bar{j}$ denote mean value of source image *X* and fused image I_f, respectively. $Q_0(X, Y, I_f)$ represents the average value between $Q_0(X, I_f)$ and $Q_0(Y, I_f)$, as given below:

$$Q_0(X, Y, I_f) = \frac{Q_0(X, I_f) + Q_0(Y, I_f)}{2} \quad (12)$$

The value of Q_0 lies between $[-1, 1]$ and it should be almost near to 1 for better quality. The resultant value of Q_0 between source and fused images for proposed scheme are shown in Table 1. It is observed from proposed schemes that the values of Q_0 are almost close to 1, which represent the effectiveness of the proposed scheme.

6. The metric Q_W [9, 25] between source images *X,Y*

Table 1: Comparison of image fusion performance of proposed scheme against 6 existing schemes

Source Images	**Evaluation indices**	**DWT** [13]	**DWT + PCA** [9]	**MSVD** [14]	**SWT** [33]	**CVT** [30]	**NSCT** [31]	**Proposed scheme**
Pepper	PSNR	36.0793	36.0791	35.3424	35.5148	35.5214	35.8210	**36.5174**
	C.C.	0.9898	0.9898	0.9875	0.9805	0.9762	0.9889	**0.9905**
	Mean	119.1555	119.1555	119.1555	119.1555	119.1562	119.1555	**119.1655**
	S.D.	50.8323	50.8325	50.9791	50.9519	51.2063	51.3940	**51.4211**
	MI	1.2101	1.1998	1.2768	1.3006	1.2920	1.3010	**1.3018**
	Q_0	0.6832	0.6925	0.6953	0.6128	0.6872	**0.6957**	0.6872
	Q_W	0.8264	0.8098	0.8295	0.8716	0.8752	**0.8836**	0.8765
	Q_E	0.6164	0.6112	0.6268	0.6241	0.6105	0.6254	**0.6321**
	$Q^{AB/F}$	0.5347	0.5329	**0.5630**	0.5436	0.5421	0.5384	0.5437
Lena	PSNR	34.1621	34.1847	32.3791	34.8267	34.8694	34.8945	**34.9193**
	C.C.	0.9887	0.9888	0.9822	0.9801	0.9877	0.9827	**0.9902**
	Mean	121.6091	121.6083	121.6090	121.6093	121.5089	**121.6978**	121.6094
	S.D.	60.3515	60.3597	60.1866	60.5962	60.9821	60.7543	**61.1822**
	MI	1.2458	1.2450	1.2416	**1.2844**	1.2759	1.2089	1.2271
	Q_0	0.8645	0.8598	0.8591	0.9331	0.9289	0.9367	**0.9414**
	Q_W	0.7562	0.7512	0.7591	0.8767	0.8541	0.8793	**0.8811**
	Q_E	0.6378	0.6351	0.6435	0.6421	**0.6492**	0.6487	0.6460
	$Q^{AB/F}$	0.4362	0.4355	0.4731	0.4531	0.4803	0.4989	**0.5613**
Dataset1	PSNR	36.7513	36.8220	38.8281	37.7920	38.7982	38.8724	**38.9818**
	C.C.	0.9847	0.9854	0.9871	0.9876	0.9758	0.9849	**0.9882**
	Mean	227.6663	**227.8712**	227.6663	227.6663	227.5684	227.8259	227.6663
	S.D.	45.8628	45.9310	46.3967	46.1644	46.5231	46.7903	**46.8027**
	MI	1.4331	1.4068	1.4291	1.4319	1.4297	1.4367	**1.4373**
	Q_0	0.3543	0.3532	0.3734	0.3745	0.3758	0.3428	**0.3797**
	Q_W	0.4353	0.4216	0.4437	**0.4528**	0.4376	0.4392	0.4439
	Q_E	0.3201	0.3254	**0.3633**	0.3096	0.3257	0.3467	0.3581
	$Q^{AB/F}$	0.5634	0.5207	0.5553	0.5126	0.5167	0.5948	**0.6277** *
Dataset2	PSNR	–	–	–	–	–	–	–
	C.C.	–	–	–	–	–	–	–
	Mean	84.3786	84.3786	84.3786	84.3786	84.3773	84.4671	**84.8658**
	S.D.	47.3890	47.3890	49.1255	48.8580	48.5708	**49.7903**	49.2007
	MI	1.3965	1.3018	1.3520	1.3991	1.4071	1.4023	**1.4080**
	Q_0	**0.3967**	0.3807	0.3862	0.3831	0.3810	0.3806	0.3826
	Q_W	0.4325	0.4336	**0.4381**	0.4310	0.4326	0.4374	0.4307
	Q_E	0.3029	0.3031	0.3047	0.3051	0.3068	0.3073	**0.3082**
	$Q^{AB/F}$	0.5528	0.5541	0.5520	0.5516	0.5529	**0.5574**	0.5519
TNO	PSNR	34.8460	34.8562	34.8844	35.7054	35.8325	35.9077	**35.9116**
	C.C.	0.9931	0.9931	0.9930	**0.9943**	0.9758	0.9849	0.9914
	Mean	87.3152	87.3154	87.3153	87.3152	87.5621	87.5639	**87.6153**
	S.D.	36.5409	36.5431	36.7410	36.7115	36.6742	36.7309	**36.9588**
	MI	1.3295	1.3295	1.3479	1.3596	1.3498	**1.3671**	1.3450
	Q_0	0.8497	0.8497	0.8137	0.8276	0.8420	0.8437	**0.8542**
	Q_W	0.9315	0.9315	0.9494	0.9559	0.9576	**0.9792**	0.9752
	Q_E	0.5721	0.5692	0.5771	0.5748	0.5622	0.5767	**0.5782**
	$Q^{AB/F}$	0.6856	0.6857	0.6803	0.7033	0.6841	0.6948	**0.7313**
Head	PSNR	33.3781	33.8152	34.6210	34.5709	35.5421	35.5714	**35.6509**
	C.C.	0.9973	0.9905	0.9898	0.9976	0.9922	0.9959	**0.9982**
	Mean	126.8768	126.8727	127.3763	127.8391	127.8653	127.6759	**127.8937**
	S.D.	**73.6487**	73.5140	73.5871	73.4594	73.5291	73.5933	73.6107
	MI	1.4179	1.4140	1.4276	1.4273	1.4297	1.4266	**1.4313**
	Q_0	0.8437	0.8250	0.8374	0.8405	0.8354	0.8420	**0.8531**
	Q_W	0.7187	0.7189	0.7270	0.7123	**0.7296**	0.7232	0.7203
	Q_E	0.5520	0.5429	**0.5593**	0.5297	0.5487	0.5447	0.5481
	$Q^{AB/F}$	0.7187	0.7172	0.7205	0.7124	**0.7287**	0.7248	0.7236

For the dataset2 original images are not available to compare our results, it is indicated by $'-'$ in Table 1.

and fused image I_f is defined as:

$$Q_W(X,Y,I_f) = \sum_{w \epsilon W} c(w)(\lambda(w)Q_0(X,I_f|w) + (1-\lambda(w))Q_0(Y,I_f|w)) \quad (13)$$

where, $\lambda(w)$ is defined as:

$$\lambda(w) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}$$

$\lambda(w)$ denotes the relative salience of *X* compared to *Y* in the same window *w* and $c(w)$ indicates the normalized salience of the window *C(w)* which is defined as:

$$C(w) = max(\sigma_X^2, \sigma_Y^2)$$

The Q_W contains the salience of information into account. The range of Q_W is 0 to 1. One indicates the fused image retain all information from the source images. It is observed from Table 1 that the value of Q_W between source and fused images are achieve almost close to 1.

7. The metric Q_E [22] is defined as follows:

$$Q_E(X,Y,I_f) = Q_W(X,Y,I_f).Q_W(X',Y',I_f')^{\alpha} \quad (14)$$

where, X', Y' and I_f are the corresponding edge images of X, Y, I_f, respectively. Parameter α reflects the contribution of the edge images compared to the original images which is set to 1. Q_E retains visual and edge information. The larger value of Q_E denotes the better fusion result. Most of the result of proposed scheme provides better fusion result as shown in Table 1.

8. The metric $Q^{AB/F}$ [25] measures the amount of edge information transferred from source images to the fused images. $Q^{AB/F}$ is defined as:

$$Q^{AB/F} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M}(xx.w_x + yy.w_y)}{\sum_{n=1}^{N}\sum_{m=1}^{M}(w_x + w_y)} \quad (15)$$

where $xx = Q^{XI_f}(n,m)$, $yy = Q^{YI_f}(n,m)$, $w_x = w^X(n,m)$ and $w_y = w^Y(n,m)$.

xx denotes the edge strength and orientation values and w^x is the influence parameter of xx. Similarly, yy is taken. *n, m* denote the image location and *N, M* are the size of images, respectively. The value of $Q^{AB/F}$ near to 1 denotes better quality. From Table 1, the proposed scheme gives value almost close to 1.

5 Conclusions

In this paper, a fusion scheme based on multiresolution singular value decomposition and gradient based sharpness approach is proposed. The experimental results of proposed scheme are performed on several pairs of multifocus and multisensor images which demonstrate that the proposed scheme preserves more significant details, provides sharp details, rich texture and also improves the visual quality of the fused image than other fusion schemes. Apart from qualitative measurement the values of other quantitative measurement metrics such as PSNR, correlation coefficients, mean, standard deviation, mutual information, Q_0, Q_W, Q_E and $Q^{AB/F}$ are also better in most of the cases. The proposed scheme can be used to fuse the multifocus and multisensor images. However, the proposed scheme is slightly more time consuming.

Acknowledgement

We are gratefully acknowledge the support by Dr. V.P.S. Naidu for providing us images of the dataset1 and dataset2 which are used in this paper.

References

[1] Hall, David L., Llinas, J. An introduction to multisensor data fusion, Proc. IEEE, 85 (1), p.6-23, 1997.

[2] Mitianoudis, N., Stathaki, T. Pixel-based and region-based image fusion schemes using ICA bases, Inf. Fusion, 8 (2), p.131-42, 2007.

[3] M, Sasikala., N, Kumaravel. A comparative analysis of feature-based image fusion method, Inf. Tech. J., 6 (8), p.1224-1230, 2007.

[4] Tao, Q., Veldhuis, R. Threshold-optimized decision-level fusion and its application to biometrics, Pattern Recogn, 42 (5), p.823-836, 2009.

[5] Goshtasby, A.A., S, Niklov. Guest editorial: Image fusion: advances in the state of the art, Inf. Fusion, 8 (2), p.114-118, 2007.

[6] T, Pu., G, Ni. Contrast based image fusion using the discrete wavelet transform, Opt. Eng., 39 (8), p.2075-2082, 2000.

[7] Pajaes, G., Cruz, J. A wavelet-based image fusion tutorial, Pattern Recognition, 37 (9), p.1855-72, 2004.

[8] Senthil Kumar, S., Mahesh Bharath, K., Muttan, S. Implementation of Max principle with PCA in image fusion for surveillance and navigation application, Electron Lett Comput Vis Image Anal, 10 (1), p.1-10, 2011.

[9] Naidu, V.P.S., Raol, J.R. Pixel-level image fusion using wavelets and principle component analysis, Defence Science Journal, 58 (3), p.338-352, 2008.

[10] Koutsias, N., Karteris, M., Chuvieco, E. The use of intensity-Hue-Saturation transform of Landsat-5 mapper data for burned land mapping, Photogramm. Eng. Remote Sens., 66 (7), p.829-839, 2000.

[11] Z, Wang., D, Ziou., C, Armenakis., D, Li., Q, Li. A Comparative Analysis of Image Fusion Methods, IEEE transactions on Geoscience and Remote Sensing, 43 (6), p.1391-1402, 2005.

[12] Chavez, P.S., Sides, S.C., Anderson, J.A. Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, Photogrammetric Engineering and Remote Sensing, 57 (3), p.295-303, 1991.

[13] H, Li., B.S., Manjunath., S.K., Mitra. Multisensor image fusion using the wavelet transform, Graphical Models and Image Processing, 57 (3), p.235-245, 1995.

[14] Naidu, V.P.S. Image fusion technique using multiresolution singular value decomposition, Defence Sci. J., 61 (5), p.479-484, 2011.

[15] Kakarala, R., P.O., Ogunbona. Signal analysis using a multiresolution form of the singular value decomposition, IEEE Trans. Image Process., 10 (5), p.724-735, 2001.

[16] Bhatnagar, G., Saha, A., Wu, Q.M.J., Atrey, P.K. Analysis and extension of multiresolution singular value decomposition, Inf. Sciences, 277, p.247-262, 2014.

[17] Golub, G.H., Reinsch, C. Singular value decomposition and least squares solutions, Numer. Math, 14 (5), p.403-420, 1970 .

[18] Zhao, X., Ye, B., Chen, T. Theory of multiresolution singular value decomposition and its application to signal processing and fault diagnosis, J. Mech.Eng. 46 (20), p.64-75, 2010.

[19] Gonzalez, R.C., Woods, R.E. Digital Image Processing, third edition, Pearson Education, 2013.

[20] Yang, Y., Huang, S., Gao, J., Qian, Z. Multi-focus image fusion using an effective discrete wavelet transform based algorithm, Measurement science review, 14 (2), p.102-108, 2014.

[21] Yang, Y. A novel DWT based multi-focus image fusion method, Procedia Eng., 24, p.177-181, 2011.

[22] Li, Shutao., Yang, Bin., Hu, Jianwen. Performance comparison of different multi-resolution transforms for image fusion, Inf. Fusion, 12 (2), p.74-84, 2011.

[23] Wang, Zhaobin., Ma, Yide., Gu, Jason. Multifocus image fusion using PCNN, Pattern Recogn, 43 (6), p.2003-2016, 2010.

[24] Yang, Bin., Li, Shutao. Multifocus Image Fusion and Restoration With Sparse Representation, IEEE Trans. on Instrumentation and Measurement, 59 (4), p.884-892, 2010.

[25] Yang, Bin., Li, Shutao. Pixel-level image fusion with simultaneous orthogonal matching pursuit, Inf. Fusion, 13 (1), p.10-19, 2012.

[26] Yang, Shuyuan., Wang, Min., Jiao, Licheng. Image fusion based on a new contourlet packet, Inf. Fusion, 11 (2), p.78-84, 2010.

[27] Tian, Jing., Chen, Li., Ma, Lihong. Multi-focus image fusion using a bilateral gradient-based sharpness criterion, Opt. Comm., 284 (1), p.80-87, 2011.

[28] Pohl, C., Van Genderen, JL. Multisensor image fusion in remote sensing: concepts, methods and applications, International Journal of Remote Sensing, 19 (5), p.823-854, 1998.

[29] Nunez, J., Otazu, X., Fors, O. Multiresolution-based image fusion with additive wavelet decomposition, IEEE Trans. on Geoscience and Remote Sensing, 37 (3), p.1204-1211, 1999.

[30] Ali, F.E., El-Dokany, I.M., Saad, A.A., Abd El-Samie, F.E. A curvelet transform approach for the fusion of MR and CT images, Journal of Modern Optics, Taylor and Francis. 57 (4), p.273-286, 2010.

[31] Zang, Q., Guo, B.L. Multifocus image fusion using the nonsubsampled contourlet transform, Signal Processing, 89, p.1334-1346, 2009.

[32] Wang, Z., Bovik, A.C. A universal Image Quality Index, IEEE Signal Process. Lett. 9 (3), p.81-84, 2002.

[33] Huafeng, Li., Shanbi, W., Chai, Yi. Multifocus image fusion scheme based on feature contrast in the lifting stationary wavelet domain, EURASIP Journal on Advances in Signal Processing. 39, p.1-16, 2012.

[34] Bhatnagar, G., Raman B. A new image fusion technique based on directive contrast, Electronic letters on computer vision and image analysis, 8 (2), p.18-38, 2009.

I.J. Intelligent Systems and Applications, **2017, 1, 37-45**
Published Online January 2017 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2017.01.04

Discrete Wavelet Transform and Cross Bilateral Filter based Image Fusion

Sonam
Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow-226025, India
E-mail: sonam870115@gmail.com

Manoj Kumar
Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow-226025, India
E-mail: mkjnuiitr@gmail.com

***Abstract*—**The main objective of image fusion is to obtain an enhanced image with more relevant information by integrating complimentary information from two source images. In this paper, a novel image fusion algorithm based on discrete wavelet transform (DWT) and cross bilateral filter (CBF) is proposed. In the proposed framework, source images are decomposed into low and high frequency subbands using DWT. The low frequency subbands of the transformed images are combined using pixel averaging method. Meanwhile, the high frequency subbands of the transformed images are fused with weighted average fusion rule where, the weights are computed using CBF on both the images. Finally, to reconstruct the fused image inverse DWT is performed over the fused coefficients. The proposed method has been extensively tested on several pairs of multi-focus and multisensor images. To compare the results of proposed method with different existing methods, a variety of image fusion quality metrics are employed for the qualitative measurement. The analysis of comparison results demonstrates that the proposed method exhibits better results than many other fusion methods, qualitatively as well as quantitatively.

***Index Terms*—**Image Fusion, Discrete Wavelet Transform, Cross Bilateral Filter, Standard Deviation, Correlation Coefficients.

I. INTRODUCTION

Nowadays, image fusion has gained much attention in image processing and computer vision. Image fusion is the process of combining the significant visual information from various source images of the same scene or view to produce a single enhanced image without any distortion or less information in images. The obtained fused image is more useful for human visual perception and further image processing operations [1]. Image fusion has been used in many disparate fields such as, medical imaging [2,3], military [4], surveillance [5], robotics and remote sensing [6], etc.

Generally, an image having less information does not provide an appropriate analysis of the scene. Consider, two images where one image is focused on some part of the scene and rest of part is in out-of-focus and another image is focused on the portion which is defocused in the first image however, defocused region which is focused in the first image. To get the entire scene in focus in a single image is a difficult task. Therefore, by integrating the relevant details from both the source images, a fused image with complete information can be obtained. This whole process of image fusion is called as multi-focus image fusion and these source images are known as multi-focus images.

The images captured from different sensors are multisensor images such as, CT (Computed Tomography) and MRI (Magnetic Resonance Imaging). These images may provide a different kind of information of the same organ to diagnose the diseases. CT images give the information of bone, blood vessels and hard tissues, whereas, MRI images give the information of the soft tissues. These two different images individually provide the different information according to their sensor ability but if we combined these images into a single image then all the relevant information (bone and soft tissues) of the organ can be presented in a single image which may provide more beneficiaries for diagnosis purpose. The fusion of multisensor images is known as the multisensor image fusion.

Image fusion process can be conducted at different levels, depending on represented information and applications. These levels are categorized into signal or pixel level, object or feature level and symbol or decision level [7]. In pixel level image fusion, the visual information of the source images is fused with their respective pixels to generate a single fused image. This level represents the lowest level of fusion. Feature level fusion defines the process of combining the features such as, edges, texture or color that have already been extracted from the source images. Finally, the decision level fusion represents the highest level of fusion, which combines the results from multiple algorithms to obtain a final fused image. Among these, pixel level fusion is widely used in most image fusion applications due to the advantage of containing original information, easy implementation and low time consumption.

Further, fusion methods can be categorized into two domains. One is the spatial domain based methods and

the other is the transform domain based methods [11]. The spatial domain methods perform the fusion process in spatial domain directly. The spatial domain methods are time efficient and easy to implement. On the other hand, the transform domain based methods use transformations like, discrete wavelet transform [8], stationary wavelet transform [12], dual tree complex wavelet transform [13], and so on. Recently, some other transform domain based fusion methods are also introduced, such as curvelet transform [14], contourlet transform [15], nonsubsampled contourlet transform (NSCT) [16], multi-resolution singular value decomposition (MSVD) [17] and so on.

The pixel by pixel averaging is the simplest image fusion method and due to this it has attracted many researchers in the last decades [8-10]. But this method suffers from a number of undesirable effects, such as reduced contrast. To address this problem, multi-resolution analysis based methods have been proposed, which consist of the following three basic steps. First, the source images are decomposed into multi-resolution representations with low and high frequency information and corresponding transformed coefficients are obtained. Second, the obtained transformed coefficients are combined together according to some fusion rules. Finally, the inverse transform is performed over the fused coefficients to reconstruct the fused image [18]. Generally, multi-resolution based methods provide better results than the other transform methods. However, the other transformations also preserve the same salient features such as edges and lines and are used in fusion. For an instance, a discrete cosine transform (DCT) based image fusion is introduced for fusion instead of pyramids or wavelet [19]. Again, a new multi-resolution DCT decomposition based image fusion approach has been given to reduce the computational complexity without any loss of image information [10].

Many image fusion techniques have been developed such as, maximum, minimum, average methods, principal component analysis (PCA) [26,30,31], intensity-hue-saturation [28]. In above discussed methods, average method provides the average information of the images but weighted average methods provide the information according to their weights and produce better fusion results. In [24], source images are fused by weighted average from the image details, where these details are extracted from the source images using CBF to improve the fusion performance. Hence, we proposed a novel pixel level image fusion scheme based on discrete wavelet transform and cross bilateral filter. The low frequency subbands are fused using pixel averaging method because approximation part contain most of the average information. Meanwhile, the high frequency subbands give the sharp details, therefore weighted average fusion rule is performed over the detail parts, where the weights of two different images contribute the information for fusion according to their weight value. The main contribution of this paper is to enhance the visual quality of the images by integrating all significant details of the source images using the pixel averaging and weighted average fusion rule.

The rest of this paper is organized as follows. In Section II, the basic theory of DWT and CBF are described. Proposed methodology is explained in Section III. Experimental results followed by discussion are presented in Section IV. Section V concludes the proposed work.

II. Basic Theories of DWT and CBF

In this section, a brief review of the basic theories of discrete wavelet transform and cross bilateral filter is presented.

A. Discrete Wavelet Transform

Discrete wavelet transform contains the multi-resolution analysis property, therefore it is widely used in image processing. It converts an image from spatial domain to frequency domain [8,27,20]. DWT decomposes the image into low and high frequency subbands. The low frequency subbands corresponds to approximation part, which contains average information of the entire image and is represented as (LL) subband. Whereas, the high frequency subbands are considered as detail parts containing the sharp information of images. The detail parts consist of three high frequency subbands (LH, HL and HH) as shown in Fig. 1 (a). For second level decomposition, only LL subband is further decomposed into four frequency subbands, whereas LH, HL and HH subbands remain as such, as given in Fig. 1 (b). The decomposition levels can be increased as per the requirement.

A 2-D DWT [20] for image $f(x,y)$ of size $m \times n$ is defined as:

$$W_{\phi}(j_0,u,v) = \frac{1}{\sqrt{mn}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \phi_{j_0,u,v}(x,y) \quad (1)$$

$$W_{\psi}(j,u,v) = \frac{1}{\sqrt{mn}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \phi_{j,u,v}(x,y) \quad (2)$$

where, Eqs. (1) and (2) are approximation and detail coefficients of image *f(x, y)*. Conversely, the image is reconstructed by performing inverse DWT. For the above given Eqs. (1) and (2), the inverse DWT is given as:

$$f(x,y) = \frac{1}{\sqrt{mn}} \sum_{u} \sum_{v} W_{\varphi}(j_0,u,v) \varphi_{j_0,u,v}(x,y) + \frac{1}{\sqrt{mn}} \sum_{j=j_0}^{\infty} \sum_{u} \sum_{v} W_{\psi}(j,u,v) \psi_{j,u,v}(x,y) \quad (3)$$

The DWT is applied over the image and the obtained results are shown in Fig. 2. Fig. 2 (a) shows the original lena image whereas, Fig. 2 (b) and 2 (c) show the first level decomposition and second level decomposition of the image after applying DWT.

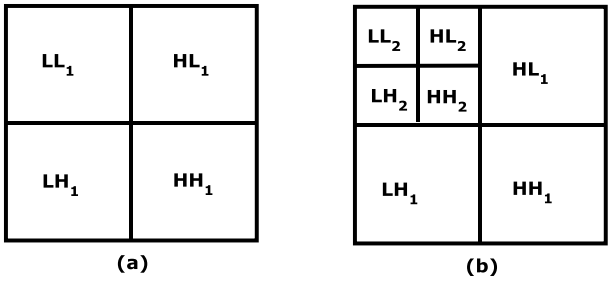

Fig.1. (a) First level DWT decomposition; (b) second level DWT decomposition.

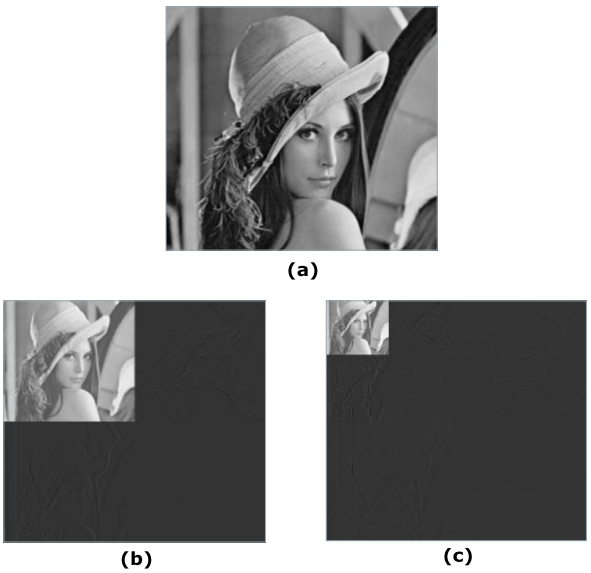

Fig.2. (a) Original lena image; (b) first level DWT decomposition; (c) second level DWT decomposition.

B. Cross Bilateral Filter

The Gaussian filter is the most popular filter which provides smoothness and removes noise present in the images. It is based on the weighted average of the intensity of the adjacent pixels. This filter removes the noise and provides smoothness but meantime also removes the sharp details. Therefore, to address this problem, the bilateral filter was introduced by Tomasi and Manduchi [21]. It is a local, nonlinear and non-iterative filter which smoothes images while preserving edges and other sharp details. The bilateral filtering has been widely used in image processing applications. The spatial filter kernel is treated as a classical low pass filter used to obtain geometric closeness between the neighboring pixels, whereas the range filter kernel is treated like an edge-stopping function used for gray-level similarity between the neighboring pixels, which attenuates the filter kernel when the intensity differences between pixels are large. Both filter kernels are based on Gaussian distribution and the weights obtained from these filters depend not only on Euclidian distance but also on the distance in gray or color spaces. The bilateral filter (BF) is the combination of spatial and range filters [22-24].

For an image *X*, the output of bilateral filter at a pixel location *p* is defined as:

$$X_f(p) = \frac{1}{W}\sum_{\mathbf{q}\in S} G_{\sigma_s}(\|p-q\|)\times G_{\sigma_r}(|X(p)-X(q)|)X(q) \quad (4)$$

where ,

$$W = \sum_{q\in s} G_{\sigma_s}(\|p-q\|)G_{\sigma_r}(|X(p)-X(q)|)$$

is a normalization factor, $\|p-q\|$ is the Euclidian distance between *p* and *q*.

$$G_{\sigma_s}(\|p-q\|) = e^{-\frac{\|p-q\|^2}{2\sigma_s^2}}$$

is a geometric closeness function and

$$G_{\sigma_r}(|X(p)-X(q)|) = e^{-\frac{|X(p)-X(q)|^2}{2\sigma_r^2}}$$

is a gray-level similarity or edge-stopping function. G_{σ_s} is a spatial Gaussian function that decreases symmetrically as the distance from the center increases. G_{σ_r} is also a Gaussian function that decreases with the increase of the intensities difference between *X(p)* and *X(q)*. *s* is a spatial neighborhood of the pixel *p*. Parameters σ_s and σ_r are standard deviations of Gaussian functions G_{σ_s} and G_{σ_r} respectively and determine the amount of filtering for the image *X*.

Cross bilateral filter (CBF) considers both gray-level similarities and geometric closeness of neighboring pixels in image *X* to shape the filter kernel and filters the image *Y*. The output of CBF [24,25] for image *Y* at a pixel location *p* is calculated as:

$$Y_{CBF}(p) = \frac{1}{W}\sum_{\mathbf{q}\in S} G_{\sigma_s}(\|p-q\|)\times G_{\sigma_r}(|X(p)-X(q)|)Y(q) \quad (5)$$

where,

$$W = \sum_{q\in s} G_{\sigma_s}(\|p-q\|)G_{\sigma_r}(|X(p)-X(q)|)$$

is a normalization factor and

$$G_{\sigma_r}(|X(p)-X(q)|) = e^{-\frac{|X(p)-X(q)|^2}{2\sigma_r^2}}$$

is a gray-level similarity or edge-stopping function.

For images *X* and *Y*, the detail image is obtained by subtracting CBF output from their respective source

image and is given as $X_D = X - X_{CBF}$ and $Y_D = Y - Y_{CBF}$, respectively. These detail images are used to find the weights by measuring the strength of details as reported in [24].

III. Proposed Methedology

In this section, we briefly discuss the proposed method. In the proposed method, two images having less information are used to produce a single composite image with enhanced visual quality. Here, the two source images either multi-focus or multisensor are used for fusion. The DWT is performed over both the source images to decompose the images into low and high frequency subbands. The low frequency subbands, which is also called approximation part, contains the average information of the images. The low frequency subbands of both of the transformed images are fused using pixel averaging method to get the overall average information from both the source image in a single fused image. On the other hand, the high frequency subbands are called detail parts which contain the information of edges and sharp changes. The high frequency subbands of both of the transformed images are fused according to their weights in respective images to get the sharp information in the fused image. Thus, the fused image has overall average information as well as dominant sharp information chosen from both the source images. The block diagram of the proposed method is shown in Fig. 3.

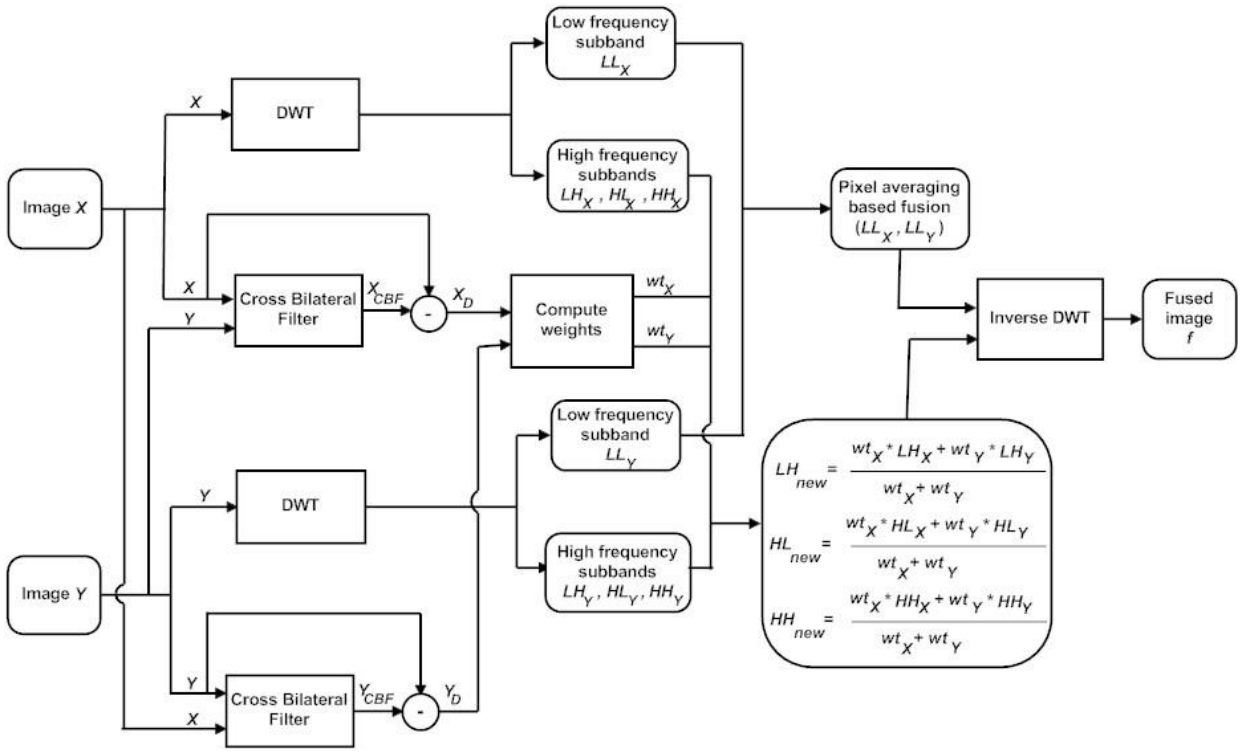

Fig.3. Block diagram of the proposed method

The steps of the proposed method are summarized in Table 1.

Table 1

Algorithm: Proposed fusion method

1. Take two source images *X* and *Y*.
2. Perform DWT over the images *X* and *Y* which decomposes the images into low and high frequency subbands.
3. *Apply pixel averaging method to fuse the low frequency subbands and obtain the average fused coefficient (LL_{new}).*
4. Combine high frequency subbands by their weighted average of both transformed images. The weights are computed as in [24]:

   a) Compute the covariance matrix as for both the images and then calculate horizontal and vertical detail strengths.

   b) Integrate both detail strengths and obtain weight wt_X and wt_Y from images *X* and *Y*.

   c) Fuse the high frequency subbands with their respective weights using:

$$HL_{new} = \frac{wt_X * HL_X + wt_Y * HL_Y}{wt_X + wt_Y} \quad (6)$$

   Similarly, we can obtain fused wavelet coefficients (LH_{new} and HH_{new}).

5. *Perfom inverse DWT over the obtained fused coefficients (*LL_{new}, LH_{new}, HL_{new} *and* HH_{new}*) to reconstruct the fused image* f.

IV. Experimental Results and Discussion

The proposed method is carried out on several pairs of multi-focus and multisensor images as shown in Fig. 4, 5, 6, 7 and 8. Each image is size of 512×512. Fig. 4 (a) and 4 (b) are blurred on left and right part images. Fig. 5 (a) and 5 (b) are upper side and lower side blurred images are called as dataset1. Fig. 6 (a) and 6 (b) are FLIR and LLTV images are called as dataset2. Dataset1 and dataset2 images are taken from [17]. Fig. 7 (a) and 7 (b) are hoed images which are blurred on middle and boundary part. CT and MRI images of size 256×256 are shown in Fig. 8 (a) and 8 (b) are acquired from the link http://www.metapix.de/examples_r.htm. The parameters σ_s=1.8 and σ_r=25 are used for the proposed method with window size 5×5. The results of proposed methods are compared with some existing methods such as, DWT [8], PCA with DWT [26], SWT [12] and CBF [24] based fusion methods. The results obtained from above existing methods are shown in Fig. 4 (c-f), 5 (c-f), 6 (c-f), 7 (c-f) and 8 (c-f). However, the results obtained by proposed method are shown in Fig. 4 (g), 5 (g), 6 (g), 7 (g) and 8 (g). Visually it can be seen that in most of the cases, the fused images obtained from proposed method are better than other existing methods.

But to evaluate the quality of fused image only visual inspection is not sufficient. Therefore, for quantitative measurement of the fused images some metrics are used. All the results of existing and proposed methods are tabulated in Table 2. It is found that in most of the cases the results of the proposed method have shown better performance than others and the best results are bolded. The metrics used in this paper are defined as:

A. Average pixel intensity (API) or mean (μ)

It measures an index of contrast, which is defined as:

$$API = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j)}{mn} \tag{7}$$

where, *f(i,j)* represents fused image and the size of image is $m\times n$

B. Standard deviation (SD)

It is defined as:

$$SD = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j)-\mu)^2}{mn}} \tag{8}$$

It reflects the spread in data and higher value represents better fusion result.

C. Average gradient (AG)

It measures a degree of clarity and sharpness and is defined as:

$$AG = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}((f(i,j)-f(i+1,j))^2+(f(i,j)-f(i,j+1))^2)^{1/2}}{mn} \tag{9}$$

D. Entropy (EN)

It measures the amount of information presented in the fused image and is defined as:

$$EN = -\sum_{k=0}^{255} p_k \log_2(p_k) \tag{10}$$

where, p_k is the probability of intensity value k in image. The larger values denote better result.

E. Mutual information (MI)

It is defined as the sum of mutual information between source images and fused image and measures the degree of dependence of two images. The mutual information MI_{Xf} between source image X and fused image f is given as follows:

$$MI_{Xf} = \sum_{i,j} h_{Xf}(i,j)\log_2\frac{h_{Xf}(i,j)}{h_X(i)h_f(i,j)} \tag{11}$$

where, h_{Xf} is the joint histogram between X and f, respectively. Similarly, MI_{Yf} represents the mutual information between source images Y and fused image f. The total mutual information between source images X, Y and fused image f is defined as:

$$MI_{total} = MI_{Xf} + MI_{Yf} \tag{12}$$

The higher MI_{total} value implies better fusion results.

F. Fusion symmetry (FS)

It indicates how much symmetrical information of the fused image with respect to source image and is given as:

$$FS = 2 - \left|\frac{MI_{Xf}}{MI_{total}} - 0.5\right| \tag{13}$$

G. Correlation coefficient (CC)

It estimates a relevance of fused image to source image, larger value represents better fusion results which is defined as:

$$CC = (r_{Xf} + r_{Yf})/2 \tag{14}$$

where,

$$r_{Xf} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(X(i,j)-\overline{X})(f(i,j)-\mu)}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(X(i,j)-\overline{X})^2\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j)-\mu)^2\right)}}$$

and

$$r_{Yf} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(Y(i,j)-\bar{Y})(f(i,j)-\mu)}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(Y(i,j)-\bar{Y})^2\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j)-\mu)^2\right)}}$$

H. Spatial frequency (SF)

It estimates the overall information level in the regions of an image and is calculated as:

$$SF = \sqrt{RF^2 + CF^2} \tag{15}$$

where, RF and CF are row and column frequency

$$RF = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j)-f(i,j-1))^2}{mn}}$$

$$CF = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j)-f(i-1,j))^2}{mn}}$$

The large value of spatial frequency represents the large information in the image.

I. $Q^{XY/f}$

It evaluates the total information transferred from the source images to fused image [29]. Mathematically, it is defined as:

$$Q^{XY/f} = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{[Q^{Xf}(i,j)w^X(i,j)+Q^{Yf}(i,j)w^Y(i,j)]}{\sum_{i=1}^{m}\sum_{j=1}^{n}[w^X(i,j)+w^Y(i,j)]} \tag{16}$$

where, X, Y are source images and f is fused image. The definitions of Q^{Xf} and Q^{Yf} are same and given as:

$$Q^{Xf}(i,j) = Q_g^{Xf}(i,j)\cdot Q_\alpha^{Xf}(i,j)$$

where $Q_g^{Xf}(i,j)$ and $Q_\alpha^{Xf}(i,j)$ represent the edge strength and orientation values at location (i, j), respectively. The dynamic range for $Q^{XY/f}$ is [0,1] and it should be close to one for better fusion.

The other metrics $L^{XY/f}$, $N^{XY/f}$ and $N_m{}^{XY/f}$ are used to compute the total loss of information and noise or artifacts in fused image which are given in [9,24,29].

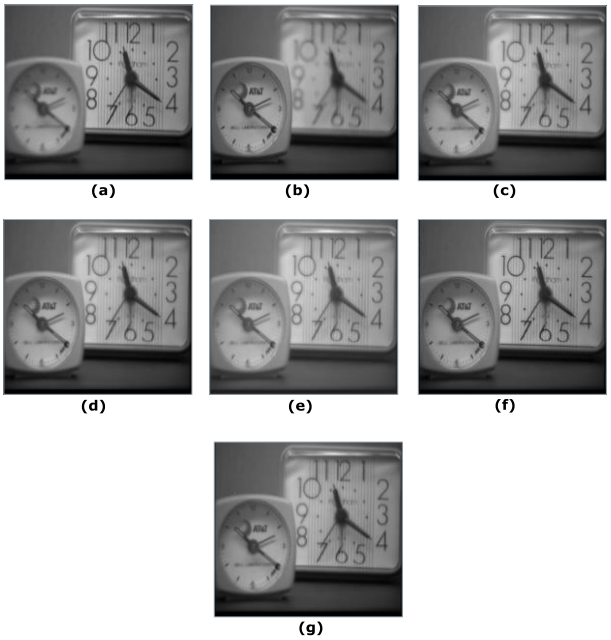

Fig.4. Fusion results for clock image: (a) left part blurred; (b) right part blurred; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image by CBF; (g) fused image by proposed method.

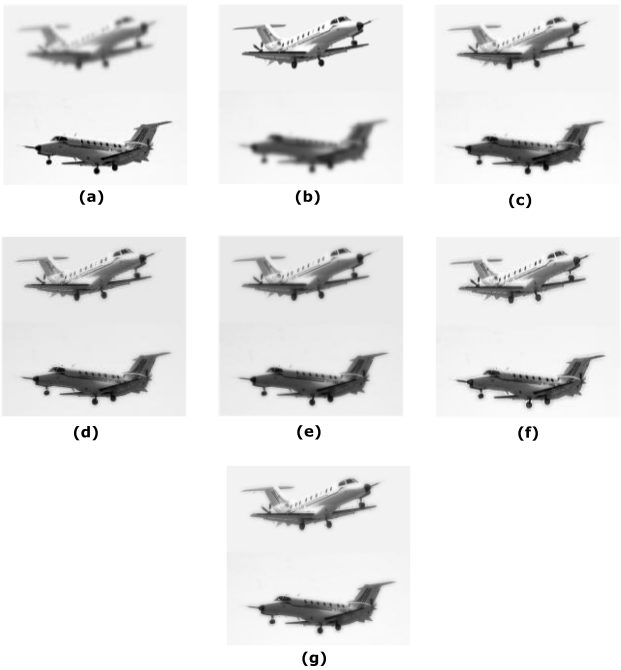

Fig.5. Fusion results for dataset 1: (a) upper part blurred; (b) lower part blurred; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image CBF; (g) fused image by proposed method.

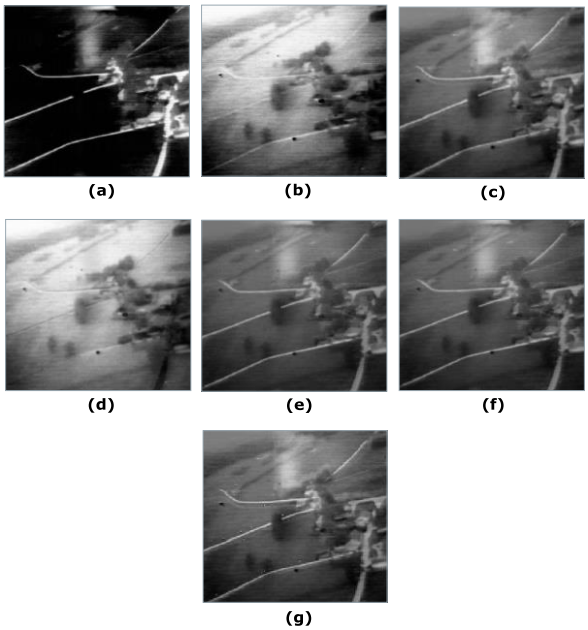

Fig.6. Fusion results for dataset 2: (a) FLIR image; (b) LLTV image; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image by CBF; (g) fused image by proposed method.

Fig.7. Fusion results for hoed image: (a) middle part blurred; (b) boundary part blurred; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image CBF; (g) fused image by proposed method.

Table 2. Image fusion performance measure

		Fusion metrics											
Input Images	Fusion methods	*SD*	*CC*	*API*	*AG*	*EN*	*MI*	*FS*	*SF*	$Q^{XY/f}$	$L^{XY/f}$	$N^{XY/f}$	$N_m{}^{XY/f}$
Clock	[8]	49.315	0.9721	97.038	3.879	3.8869	5.728	1.8904	6.2841	0.8341	0.1659	**0.0017**	0.0016
	[26]	49.315	0.9754	97.038	3.878	4.3445	5.875	1.8722	6.2817	0.8338	0.1662	0.0021	0.0013
	[12]	49.408	0.9872	97.038	4.471	4.4882	6.492	1.9021	7.4981	0.8700	0.1299	0.0118	**0.0010**
	[24]	**49.892**	0.9869	96.548	**5.5265**	7.2755	**7.3415**	1.9600	**10.142**	**0.8982**	**0.0995**	0.0114	0.0011
	Proposed	49.417	**0.9890**	**97.080**	4.4044	**7.2761**	6.8667	**1.9934**	8.0545	0.8583	0.1406	0.0155	0.0023
Dataset1	[8]	51.872	0.9746	221.43	3.1869	2.5333	5.0198	1.9587	10.2225	0.7958	0.2042	0.0296	0.0072
	[26]	51.873	0.9798	221.43	3.1871	2.8173	5.0368	1.9590	10.2232	0.7959	0.2041	0.0279	0.0079
	[12]	52.172	0.9876	221.43	4.2307	3.5672	5.3640	1.9526	13.5788	0.8698	0.1293	**0.0036**	**0.0052**
	[24]	52.561	0.9886	220.58	4.9215	4.2667	**5.4154**	1.9635	16.9262	**0.9284**	**0.0635**	0.0345	0.0081
	Proposed	**55.242**	**0.9904**	**221.45**	**5.8073**	**4.7166**	5.0230	**1.9766**	**18.6560**	0.8451	0.1473	0.0285	0.0076
Dataset2	[8]	40.394	0.5703	84.631	7.0495	1.9383	3.0799	1.7832	10.6120	0.6249	0.3751	0.0173	0.0036
	[26]	41.220	0.5770	84.631	9.7021	2.0432	3.0765	1.7704	**21.9084**	0.4599	0.5132	0.1035	0.1182
	[12]	41.087	0.5689	84.631	9.3246	2.0486	3.0832	1.7543	14.0195	0.7226	0.2753	**0.0078**	0.0065
	[24]	41.358	0.5317	84.665	**9.7106**	7.3465	3.0842	1.7735	16.8049	0.7305	0.2651	0.0184	0.0044
	Proposed	**41.429**	**0.5815**	**86.642**	7.7882	**7.8537**	**3.4738**	**1.8188**	17.4319	**0.8310**	**0.1026**	0.2124	**0.0021**
Hoed	[8]	56.559	0.9521	95.951	13.2480	1.9248	4.5982	1.8954	17.4967	0.7168	0.2832	0.0276	0.0028
	[26]	56.568	0.9460	95.951	13.2807	2.0431	4.7601	1.8643	17.5939	0.7188	0.2812	1.0231	**0.0020**
	[12]	57.947	0.9479	95.951	**28.0801**	7.6941	4.7803	1.8092	30.7982	0.8957	0.1007	0.0254	0.0036
	[24]	58.097	**0.9628**	95.958	22.8185	7.6988	4.6434	**1.9987**	30.9057	0.8775	0.1152	0.0626	0.0072
	Proposed	**61.849**	0.9566	**96.173**	25.2017	**7.7399**	**6.9666**	1.9952	**33.9649**	**0.9478**	**0.0483**	**0.0226**	0.0039
CT and MRI	[8]	34.883	0.6502	32.082	5.5005	1.6771	3.4601	1.6204	10.2717	0.6440	0.3560	0.0657	0.0110
	[26]	35.155	0.6521	32.082	6.6750	1.8607	3.4028	1.6082	13.1607	0.6140	0.3850	**0.0052**	**0.0011**
	[12]	35.107	0.6246	32.082	6.1676	2.0452	3.4122	1.6091	11.3466	0.6914	0.3086	1.5826	0.0021
	[24]	35.785	**0.6905**	32.166	7.2431	5.9698	3.4311	**1.6554**	14.7174	0.7143	0.2843	0.0065	0.0014
	Proposed	**37.862**	0.6545	**41.414**	**11.5408**	**6.7736**	**5.5999**	1.6122	**21.0958**	**0.9116**	**0.0758**	0.0873	0.0126

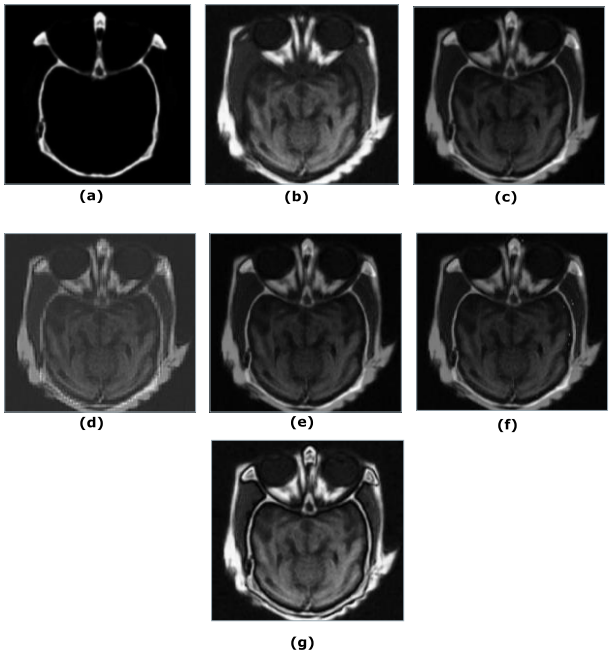

Fig.8. Fusion results for CT and MRI image: (a) CT image; (b) MRI image; (c) fused image by DWT; (d) fused image by PCA; (e) fused image by SWT; (f) fused image CBF; (g) fused image by proposed method.

V. CONCLUSIONS

In this paper, a fusion method of two multi-focus or multisensor images based on DWT and CBF algorithm is presented. In the proposed framework, the average and sharp details of the source images are integrated very cleverly into a single fused image. The performance of the proposed fusion method is tested against other fusion methods using several pairs of test images. Fusion performance is evaluated using many quantitative measurement criterions. Through the experimental results, it is found that the proposed method preserves more sharp information while eliminating artifacts and has shown better performance than other existing fusion methods visually as well as quantitatively.

ACKNOWLEDGMENTS

The authors would gratefully acknowledge the support of Dr. V.P.S. Naidu for providing us dataset1 and dataset2 for the test purpose.

REFERENCES

[1] Blum R S, Liu Z. Multi-sensor Image Fusion and Its Applications. CRC Press, Taylor & Francis Group, 2005.
[2] Hill D, Edwards P, Hawkes D. Review of pixel-level image fusion. Fusing medical images [J]. Image Processing, 1994, 6(2): 22-24.
[3] Qu G H, Zhang D L, Yan P E. Medical image fusion by wavelet transform modulus maxima [J]. Opt Express, 2001, 9(4): 184-190.
[4] Smith M I, Ball A N, Hooper D. Real-time image fusion: A vision aid for helicopter pilotage [J]. Proc SPIE, 2002, 4713: 83-94.
[5] Slamani M A, Ramac L C, Uner M K, Varshney P K, Weiner D D, Alford M G, Ferris D D, Vannicola V C. Enhancement and fusion of data for concealed weapons detection [J]. Proc SPIE, 1997, 3068: 8-19.
[6] Daniel M M, Willsky A S. A multiresolution methodology for signal-level fusion and data assimilation with applications to remote sensing [J]. Proc IEEE, 1997, 85(1): 164-180.
[7] Mitianoudis N, Stathaki T. Pixel-based and region-based image fusion schemes using ICA bases. Inf. Fusion, 2007, 8 (2): 131-42.
[8] Li H, Manjunath B S, Mitra S K. Multisensor image fusion using the wavelet transform [J]. Graph Models Image Process, 1995, 57(3): 235-245.
[9] Petrovic V. Multisesor pixel-level image fusion. PhD Thesis, Department of Imaging Science and Biomedical Engineering, Manchester School of Engineering, United Kingdom, 2001.
[10] Shreyamsha Kumar B K, Swamy M N S, Omair Ahmad M. Multiresolution DCT decomposition for multifocus image fusion. Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, Canada, 2013, 1-4.
[11] Goshtasby A A, Nikolov S. Image fusion: advances in the state of the art. Inf. Fusion, 2007, 8(2): 114-118.
[12] Huafeng Li, Shanbi W, Chai Yi. Multifocus image fusion scheme based on feature contrast in the lifting stationary wavelet domain. EURASIP Journal on Advances in Signal Processing. 2012, 39: 1-16.
[13] Selesnick I W, Baraniuk R G, Kingsbury N C. The dual-tree complex wavelet transform. IEEE Signal Process, Mag, 2005, 22 (6): 123-151.
[14] Ali F E, El-Dokany I M, Saad A A, Abd El-Samie F E. A curvelet transform approach for the fusion of MR and CT images. Journal of Modern Optics, Taylor & Francis. 2010, 57 (4): 273-286.
[15] Do M N, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation. IEEE Trans. on Image Processing, 2005, 14(12): 2091–2106.
[16] Zhang Q, Guo B L. Multifocus image fusion using the nonsubsampled contourlet transform. Signal Processing, 2009, 89(7): 1334-1346.
[17] Naidu V P S. Image fusion technique using multiresolution singular value decomposition. Defence Sci. J., 2011, 61 (5): 479-484.
[18] Shutao Li, Xudong Kang, Jianwen Hu, Bin Yang. Image matting for fusion of multi-focus images in dynamic scenes. Inf. Fusion, 2013, 14: 147-162.
[19] Naidu V P S. Discrete cosine transform-based image fusion, Defence Sci. J., 2010, 60 (1): 48-54.
[20] Sonam Gautam, Manoj Kumar. An Effective Image Fusion Technique based on Multiresolution Singular Value Decomposition. INFOCOMP Journal of Computer Science, 2015, 14 (2): 31-43.
[21] Tomasi C, Manduchi R. Bilateral filtering for gray and color images. Proceedings of the Sixth International Conference on Computer Vision, 1998: 839–846.
[22] Manoj Diwakar, Sonam, Manoj Kumar. CT image denoising based on complex wavelet transform using local adaptive thresholding and Bilateral filtering. ACM, Proceedings of the Third International Symposium on Women in Computing and Informatics (WCI), 2015, 297-302.
[23] Jianwen Hu, Shutao Li. The multiscale directional bilateral filter and its application to multisensor image fusion. Inf. Fusion, 2012, 13(3): 196-206.
[24] Shreyamsha Kumar B K. Image fusion based on pixel

significance using cross bilateral filter. SIVIP, 2013.
[25] Petschnigg G, Agrawala M, Hoppe H, Szeliski R, Cohen M,Toyama K. Digital photography with flash and no-flash image pairs. ACM Trans. Gr. 2004, 23(3): 664–672.
[26] Naidu V P S, Raol J R. Pixel-level image fusion using wavelets and principle component analysis. Defence Sci. J., 2008, 58 (3): 338-352.
[27] Li S, Kwok J T, Wang Y. Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images. Inf. Fusion, 2002, 3(1): 17–23.
[28] Wang Z, Ziou D, Armenakis C, Li D, Li Q. A Comparative Analysis of Image Fusion Methods. IEEE Trans. on Geoscience and Remote Sensing, 2005, 43 (6): 1391 - 1402.
[29] Petrovic V, Xydeas C. Objective image fusion performance characterization. In: Proceedings of the International Conference on Computer Vision (ICCV), 2005, 2: 1866–1871.
[30] Eleyan Alaa. Enhanced face recognition using data fusion. I. J. Intelligent systems and applications, 2013, 01: 98-103.
[31] Kaur S, Dadhwal H S. Biorthogonal wavelet transform using bilateral filter and adaptive histogram equalization. I. J. Intelligent systems and applications, 2015, 03: 37-43.

Authors' Profiles

Sonam received the M.C.A. degree in computer science from Babasaheb Bhimrao Ambedkar University, Lucknow, India, in 2012. She is currently working toward the Ph.D. degree at Babasaheb Bhimrao Ambedkar University, Lucknow, India. Her research interest focuses on image fusion and image enhancement.

Manoj Kumar received the Ph.D. degree in Computer Vision and Image Processing from Indian Institute of Technology Roorkee, India, in 2011 and M.C.A. degree in computer science from Jawaharlal Nehru University, New Delhi, India, in 2003. He is currently working as Assistant Professor in the Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow, India. His general research interest include the areas of computer graphics, vision & image processing, medical imaging, image compression, security and watermarking. He has published various research papers in international and national journals and conferences.

How to cite this paper: Sonam, Manoj Kumar,"Discrete Wavelet Transform and Cross Bilateral Filter based Image Fusion", International Journal of Intelligent Systems and Applications(IJISA), Vol.9, No.1, pp.37-45, 2017. DOI: 10.5815/ijisa.2017.01.04

