PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

Ministry of Higher Education and Scientific Research

University of Amar Telidji - Laghouat

Faculty of technology
Department of Electronics

Thesis presented for the fulfillment of the degree of

Doctor in Sciences

Field: Electrical Engineering
Speciality: Electronics

Presented by:

LATRECHE Boubakeur

THEME

Development of an image fusion technique based on combined wavelets and DCT: Application to visual sensor networks

Committee members:

Mr. BELKHEIRI Mohammed	**Chair**	Professor	U.A.T Laghouat, Algeria
Mr. SAADI Slami	**Supervisor**	Associate Professor	U.Z.A Djelfa, Algeria
Mr. KIOUS Mecheri	**Co-supervisor**	Associate Professor	U.A.T Laghouat, Algeria
Mr. HAFAIFA Ahmed	**Examiner**	Professor	U.Z.A Djelfa, Algeria
Mr. DJAFER Djelloul	**Examiner**	Research Director	URAER Ghardaia, Algeria
Mrs. CHOUIREB Fatima	**Examiner**	Professor	U.A.T Laghouat, Algeria

October 2018

ملخص : في هذه الأطروحة، اقترحنا تقنية جديدة كحل أمثل لعملية دمج الصور والتي قمنا بتطبيقها في مجال شبكات الاستشعار البصرية. الهدف الأساسي للتقنية المقترحة يشمل تطبيقها على مجموعة الصور الملتقطة لنفس المشهد في شبكات الاستشعار البصرية من أجل دمجها في صورة واحدة ذات بيانات مرئية جد دقيقة، وكذلك غنية بالمعلومات مقارنة بصور الدخل، ويمكن استخدامها بشكل مناسب للإدراك البصري ومعالجة الكمبيوتر. تعتمد هذه التقنية المقترحة على التحويل المنفصل لجيب التمام (DCT) وكذا نظام الرفع الصحيح لتحويل الموجات (ILWT). من أجل تحقيق ذلك، ينقسم هذا العمل إلى ثلاثة أجزاء رئيسية. الجزء الأول يعرض المفاهيم والتطبيقات الأولية لدمج الصورة. في الجزء الثاني، يتم تقديم مراجعة مختصرة للتقنيات الحديثة لدمج الصور. في الجزء الأخير، يتم عرض كل من معلومات تقييم الأداء النوعي والكمي لطرق الدمج المقترحة والمقارنة لإثبات تفوق الأسلوب المقترح على المقارنة.

الكلمات المفتاحية: دمج الصورة; شبكات الاستشعار البصرية; نظام الرفع الصحيح لتحويل الموجات; التحويل المنفصل لجيب التمام.

Abstract: In this thesis, we propose a novel image fusion technique for visual sensor networks (VSNs). The principal goal of this proposed technique is carried out on captured images from VSN to compose a single composite one with more accurate visual data and further information from the input images. It can be suitably used for visual perception and computer processing. This new technique is based on the Discrete Cosine Transform (DCT) and the Integer Lifting Wavelet Transform (ILWT). In order to achieve that, this work is divided into three main parts. The first part presents the preliminary concepts and applications of the image fusion. In the second part, a brief state of the art review of recent image fusion techniques is presented. In the last part, both qualitative and quantitative performance evaluation parameters for the proposed and the compared image fusion methods are presented to prove the outperforming of the proposed method on the compared methods.

Keywords: Image fusion; visual sensor network; integer lifting wavelet transform; discrete cosine transforme.

Résumé : Dans cette thèse, nous proposons une nouvelle technique de fusion d'images pour les Réseaux de Capteurs Visuels (RCV). L'objectif principal de cette technique proposée consiste à utiliser les images capturées à partir du RCV pour composer une seule image composée avec des données visuelles plus précises et des informations supplémentaires que les images d'entrée, et peut-être utilisée pour la perception visuelle et le traitement par ordinateur. Cette nouvelle technique est basée sur la transformée en cosinus discrète (DCT) et le transformée lifting en ondelettes en nombre entier (ILWT). Pour y parvenir, ce travail est divisé en trois parties principales. La première partie présente les concepts préliminaires et les applications de la fusion d'images. Dans la deuxième partie, une brève révision de l'état de l'art sur les techniques de fusion d'images récentes est proposée. Dans la dernière partie, l'évaluation des performances qualitative et quantitative pour les méthodes de fusion d'images (proposée et comparées) sont présentées afin de prouver la surperformance de la méthode proposée par rapport aux méthodes de comparaison.

Mots clé : la fusion d'image ; les réseaux de capteurs visuels ; la transformée lifting en ondelettes en nombre entier ; la transformée en cosinus discrète.

All my praises and all my thanks are to ALLAH the Almighty, who guided me always in my life and made me know that science is the torch that leads to true knowledge of His Majesty.

I would like to express my sincere gratitude to my supervisors**, Dr. Slami SAADI** and **Dr. Mecheri KIOUS** for their generous help, encouragement, flexibility and for their support in hard times and expert guidance were responsible for the completion of this thesis.

Special thanks to **Pr. BELKHEIRI Mohammed** for accepting the chairly of the jury, sincere gratitude also to **Pr. HAFAIFA Ahmed**, **Dr. Djelloul DJAFER** and **Pr. Fatima CHOUIREB** who honored me by examining my thesis and help to riview and improve this work.

I would also like to express my appreciation to my teachers **Pr. Lahcene MITICHE** and **Dr. Ali BENZIANE** for their help, advice and discussions. My thanks go also to all my teachers and My colleagues in the electronics department.

I would also like to thank all those who have contributed in a way or another to this achievement and have made my time enjoyable by providing a pleasant environment.

I would like to dedicate this modest work to

My parents M'hammed and Aicha, ALLAH bless them

My lovely wife Imane and my son Mohammed Hamza

My dear brothers Kouider, Rabeh and Noureddine

My beautiful sisters Fatna and Oumhani

All my precious family members

All my dear friends

Thank you

ABSTRACT…… **i**
AKNOWLIDJEMENTS…… **ii**
DEDICATION ……... **iii**
TABLE OF CONTENTS …….. **iv**
LISTE OF FIGURES …… **vii**
LISTE OF TABLES ……. **viii**
LISTE OF ABREVIATIONS ……... **ix**
CHAPTER 1: GENERAL INTRODUCTION
1.1 Introduction ……... **3**
1.2 Problem statement ……..…… **4**
1.3 State-of-the-art ……... **5**
1.3.1 Fusion methods …… **5**
1.3.1.1 Classical fusion methods …….. **5**
1.3.1.2 Multi-scale fusion methods …….. **6**
1.3.1.3 Miscellaneous fusion methods …….. **7**
1.4 Contributions …….. **8**
1.5 Organisation of thesis …… **9**
CHAPTER 2: AN OVERVIEW ON IMAGE FUSION
2.1 Introduction ……... **12**
2.2 Visual sensor networks …….. **12**
2.2.1 Energy consumption ……... **13**
2.2.2 Real-time performance ……... **13**
2.2.3 Data storage …… **14**
2.2.4 Compression and coding in VSNs …….. **14**
2.3 Concept of image fusion …… **16**
2.4 Single sensor image fusion system …… **17**
2.5 multi-sensor image fusion system …… **18**
2.6 Image fusion processing levels …….. **19**
2.6.1 Pixel-level ……... **19**
2.6.2 Feature-level ……... **20**
2.6.3 Decision-level …… **20**
2.7 Overview on fusion techniques …… **21**

2.7.1 Spatial domain techniques …… **21**
2.7.1.1 Simple fusion ……... **21**
2.7.1.2 Principal component analysis …….. **22**
2.7.1.3 Brovey transform ……. **23**
2.7.1.4 Intensity-Hue-Saturation ……. **23**
2.7.1.5 Hight pass filtering …….. **24**
2.7.2 Transform Domain Techniques …….. **25**
2.7.2.1 The Laplacian pyramid …… **25**
2.7.2.2 The discrete cosine transforms …… **27**
2.7.2.3 The discrete wavelet transforms …….. **27**
2.8 Applications of image fusion techniques ……... **28**
2.8.1 Visual sensor networks …… **28**
2.8.1.1 Surveillance ……. **28**
2.8.1.2 Traffic monitoring ……... **29**
2.8.1.3 Environment monitoring …….. **29**
2.8.2 Remote sensing ……... **29**
2.8.3 Medical imaging ……. **29**
2.9 Fusion performance metrics …….. **30**
2.9.1 Root mean squared error ……. **30**
2.9.2 Peak signal to noise ratio …… **30**
2.9.3 Structural similarity index ……. **30**
2.9.4 Standard deviation …….. **31**
2.9.5 Average gradient ……. **31**
2.9.6 Entropy ……... **31**
2.9.7 Spatial frequency …… **32**
2.9.8 Edge based similarity …….. **32**
2.10 Conclusion ……... **32**

CHAPTER 3: DWT and DCT for Image Fusion: literature review

3.1 Introduction ……...…. **35**
3.2 Discrete Cosine Transform …… **35**
3.2.1 Variance in DCT domain …… **36**
3.3 Wavelet transform ……. **37**
3.3.1 Continuous Wavelet Transforms (CWT) ……... **38**
3.3.2 Discrete Wavelet Transform (DWT) …….. **39**

3.3.3 Filter banks ……. **40**
3.3.4 Two-Dimensional DWT ……. **41**
3.3.5 Integer Lifting Wavelet Transform ILWT …….. **44**
3.4 DCT fusion methods …….. **46**
3.4.1 Contrast-based technique …… **46**
3.4.2 Higher valued AC coefficients ……... **47**
3.4.3 Spatial frequency-based technique ……. **48**
3.5 Wavelet fusion methods …… **49**
3.5.1 Region-based technique …….. **50**
3.5.2 Contrast-based technique …… **50**
3.5.3 DWT-based PCA technique ……... **52**
3.6 Conclusion ……. **53**
3.7 References ……. **64**
CHAPTER 4: THE PROPOSED FUSION ALGORITHM
4.1 Introduction ……...…. **56**
4.2 Proposed ILWT and DCT image fusion method ……... **56**
4.2.1 Block based fusion rule …….. **56**
4.2.2 Pixel significance fusion rule ……. **57**
4.3 Conclusion ……. **61**
CHAPTER 5: SIMULATION RESULTS AND ANALYSIS
5.1 Introduction ……... **63**
5.2 Experimental settings …… **63**
5.3 Results and analysis ……... **65**
5.3.1 Qualitative evaluation ……. **65**
5.3.1.1 First data set ……. **65**
5.3.1.2 Second data set …… **71**
5.3.2 Quantitative evaluation ……... **74**
5.4 Conclusion ……. **78**
Conclusion and future work **80**
References **84**

Figure 2.1: Visual sensor network…… **12**
Figure 2.2: Basic imaging sensor scheme…… **13**
Figure 2.3: Venn diagram representation of the image fusion process…… **16**
Figure 2.4: Single sensor image fusion system……. **17**
Figure 2.5: Multisensory image fusion system……. **18**
Figure 2.6: General processing steps for different fusion levels……... **19**
Figure 2.7: The information flow diagram of PCA……... **22**
Figure 3.1: Different families of wavelets…… **38**
Figure 3.2: Two channel filter bank…… **40**
Figure 3.3: The decomposition and reconstruction filters for Daubechies2 wavelet(db5)…… **41**
Figure 3.4: Filter bank structure of the 2-D DWT Analysis…… **43**
Figure 3.5: 1-level DWT of the 'clock' image…….. **43**
Figure 3.6: Block diagram of inverse 2-D DWT…….. **44**
Figure 3.7: General Flowchart of a forward integer lifting wavelet transform…….. **45**
Figure 3.8: 2-D DCT output block……... **46**
Figure 3.9: General structure of the AC_Max fusion scheme……... **47**
Figure 3.10: Flowchart for SF-based fusion technique……. **48**
Figure 3.11: The block diagram of the image fusion scheme…… **50**
Figure 4.1: General Framework of the proposed image fusion method……... **60**
Figure 5.1: First data set, multi-focus images used for experiments…… **63**
Figure 5.2: Second data set, images used for experiments…… **64**
Figure 5.3: Fusion results of 'Clock' image. ……..…. **65**
Figure 5.4: Edge map of fusion results of 'Clock' image. …… **66**
Figure 5.5: Fusion results of 'Lab' image……. **67**
Figure 5.6: Edge map of fusion results of 'Lab' image. .……. **68**
Figure 5.7: Fusion results of 'Pepsi' image..……. **69**
Figure 5.8: Edge map of fusion results of 'Pepsi' image……..……. **70**
Figure 5.9: Fusion results of 'Cameraman' image ……..…….. **71**
Figure 5.10: Fusion results of 'Traffic' image. ……..…….. **72**
Figure 5.11: Edge map of fusion results of 'Cameraman' image. ……..….. **73**
Figure 5.12: Edge map of fusion results of 'Traffic' image. ……..…….. **74**

Table 2.1: Compression / Coding techniques, In-network processing techniques and Applications of VSN platforms…… **15**

Table 5.1: Quantitative evaluation of different image fusion methods for 'Clock' image…… **75**

Table 5.2: Quantitative evaluation of different image fusion methods for 'Lab' image… **75**

Table 5.3: Quantitative evaluation of different image fusion methods for 'Pepsi' image. **76**

Table 5.4: Quantitative evaluation of different image fusion methods for 'Cameraman' image…… **77**

Table 5.4: Quantitative evaluation of different image fusion methods for 'Traffic' image…… **77**

VSNs	**V**isual **S**ensor **N**etworks
WSNs	**W**ireless **S**ensor **N**etworks
PCA	**P**rincipal **C**omponent **A**nalysis
BT	**B**rovey **T**ransform
IHS	**I**ntensity-**H**ue-**S**aturation
LP	**L**aplacian **P**yramid
DCT	**D**iscrete **C**osine **T**ransforms
CWT	**C**ontinuous **W**avelet **T**ransforms
DWT	**D**iscrete **W**avelet **T**ransforms
SIDWT	**S**hift-**I**nvariant **D**iscrete **W**avelet **T**ransforms
DT-CWT	**D**ual-**T**ree **C**omplex **W**avelet **T**ransforms
BGBSC	**B**ilateral **G**radient **B**ased **S**harpness **C**riterion
AC	**A**lternating **C**urrent
DC	**D**irect **C**urrent
ILWT	**I**nteger **L**ifting **W**avelet **T**ransform
PET	**P**ositron **E**mission **T**omography
CT	**C**omputed **T**omography
MRI	**M**agnetic **R**esonance **I**maging
PAN	**Pan**chromatic
MS	**M**ulti**s**pectral
HPF	**H**igh **P**ass **F**iltering
RGB	**R**ed **G**reen **B**lue
1-D	**One-D**imensional
2-D	**Two-D**imensional
WT	**W**avelet **T**ransforms
FT	**F**ourier **T**ransform
LL	**L**ow-**L**ow
LH	**L**ow-**H**igh
HL	**H**igh-**L**ow
HH	**H**igh-**H**igh
ECT	**E**lectrical **C**apacitance **T**omography
SPECT	**S**ingle **P**hoton **E**mission **C**omputed **T**omography

RMSE	**R**oot **M**ean **S**quared **E**rror
PSNR	**P**eak **S**ignal To **N**oise **R**atio
SSIM	**S**tructural **S**imilarity
STD	**S**tandard **D**eviation
AG	**A**verage **G**radient
He	**E**ntropy
SF	**S**pectral **F**requency
RF	**R**ow **F**requency
CF	**C**olumn **F**requency
CV	**C**onsistency **V**erification

Chapter 1:
General introduction

1.1 Introduction

Wireless Sensor Networks (WSNs) have become a grown technology following a decade of extensive worldwide research and development efforts. WSNs main function is to gather and disseminate critical data that describe the physical phenomena around the sensors [1].. The common examples of the scalar data are pressure, temperature, and humidity. The WSNs are the deployed and ad-hoc networks which link the small devices. These devices are equipped with their private sensing, communication, computation and power resources. These sensors provide a limited amount of information. This information is not enough for many applications even if a big number of such sensors are distributed. This discrepancy is settled by adding camera sensors to WSNs [2].

The networks constituted by the addition of camera sensors to the WSNs are indicated as the Visual Sensor Networks (VSNs). The main difference between scalar sensor nodes and camera sensor nodes is that the scalar sensor nodes collect one-dimensional data in the surrounding area, while the camera sensor nodes collect two-dimensional data in the form of images from the areas which might not be in its neighborhood. The visual data collected by camera sensors offers a huge amount of information. The VSNs provides many valuable applications that include visual surveillance, vehicle traffic monitoring and industrial automation may be enriched with visual data.

The development in the field of image sensor technologies has achieved the accessibility of low-cost image sensors feasible. As a result, the availability of low-cost complementary metal oxide semiconductor (CMOS) cameras has created the opportunities to construct low-cost VSN platforms [3]. VSN platforms are composed of small visual sensor nodes. These nodes consist of a multimedia sensor deployed in large numbers and geographically dispersed at different monitoring points. Each of these sensors (camera) has the ability to collect, process and transmit visual information (images or video sequences) in an autonomous way [4, 5]. Recently, VSNs have received a great deal of interest from many researchers in different valuable applications in various domains such as surveillance [6], remote sensing [7], traffic monitoring [8], environment monitoring [9], military and civil applications [10].

1.2 Problem statement

The challenging issues in VSNs can be caused by a large number of cameras, which provides a huge amount of visual data. Typically, this data are high-resolution images and videos that require complex processing and coding approaches and high bandwidth usage. These requirements imply that visual sensor nodes squander largely higher power than scalar sensor nodes which adversely affects the networking in term of overload information [11] and the power consumption due to the big number of complex operations [12]. However, this leads in addition to the transmission bandwidth in VSNs, to an increasing processing time. Furthermore, sensors employed in transferring data have to be able of buffering large number of packets, this is in terms of quantity.

In terms of quality, a camera rig typically installed on an aerial platform takes images and video frames in multiple modalities. The captured images are subjected to change and may be blurred or contain some ringing; because of: (i) the movement of the frontal objects, (ii) the movement of the platform as it rotates and shifts in respect to the scene. Furthermore, several cameras in VSNs capture partially focused images since the finite depth of field gives incomplete scene [13, 14]. In some images, visual content may be unclear as a result of illumination conditions, noise integrated into the imaging operation, adverse weather conditions as fog and rain…etc.

The main goal is to reduce the large number of images captured by different sensors. Also, to produce images with high visual quality. As a result, to reduce the bandwidth usage, processing time and energy consumption. Vision processing techniques is able to reduce the amount of data by cleverly manipulating the raw data can be used at the cost of earmarking more computation resources. Image fusion techniques have the possibility to solve these problems.

Image fusion could be defined as the process that combines several images of the same scene from different sensors into a smaller set of composite images, usually a single one. Which contains all the important information from the inputs. The goal of image fusion apart from reducing the amount of data is to create new images that can well describe the scene. Also, that are more appropriate for the purpose of visual perception and computer processing [15-17].

1.3 State-of-the-art

In some literature, image fusion techniques may be classified according to their processing level into three different levels which are: pixel (signal) level, feature (segment) level or decision (symbolic) level [18-23]. The pixel-level image fusion based algorithms directly combine the original information exists in the input images into the fused (output) image [15, 24], which are simple and easy to implement and can be analysed by an expert or automatically. These methods are more sensitive to noise. In the feature level, methods require a certain feature extraction from source images using image processing techniques such as segmentation [25], region extraction [26, 27], and morphological measures [13, 28]. Features corresponding to attributes extracted from the source images which are reliance on their environments such as shape, extent and neighborhood. These identical objects from various sources are assigned to each other and then used for further assessment using artificial neural networks or statistical approaches These techniques are less sensitive to noise, but more time and energy consuming because of their complex operations. In decision-level image fusion techniques, the source images are processed individually for information extraction. The acquired information is then combined by applying decision rules to strengthen common interpretation and resolve differences and furnish a better understanding of the observed objects [29, 30].

1.3.1 Fusion methods

1.3.1.1 Classical fusion methods

The simplest image fusion method just takes the pixel-by-pixel average of the input images. This method is able to suppress noises existing in the input images. Regrettably, it suppresses salient features that should be preserved for the resulting image and producing a low contrast. This can be reduced through the selection of “optimal” weights. One common technique to determine such weights is Principal Component Analysis (PCA) [31].

Another simple method is based on the Brovey Transform (BT) [32]. The BT is based on the chromaticity transform. It is a simple technique for combining images from different sensors, with the limitation that only three bands are involved. It is designed to normalize the three multispectral bands used for red, green and blue (RGB) display. Also, to multiply the result by any other desired data to add the intensity or brightness component to the image.

The Intensity-Hue-Saturation (IHS) fusion converts a colour multispectral (MS) image from the RGB space into the IHS colour space. Because the intensity band resembles a panchromatic (PAN) image, it is replaced by a high-resolution PAN image in the fusion. An inverse IHS transform is then conducted on the PAN along with the hue and saturation bands, resulting in an IHS fused image [33, 34].

These classical methods are simple, fast and easy to implement. However, they produce a spectral degradation and suffer from blurring effects which lead to the reduction of the contrast in the fused image [23, 35].

1.3.1.2 Multi-scale fusion methods

So far, Multi-scale or multi-resolution methods have been mostly used for image fusion, such as pyramids [36, 37], discrete wavelets [38-41], complex wavelets [42, 43] and contourlet [44, 45]. For these methods, input images are transformed into another feature domain where the fused representation is obtained by fusing the input images transforms using a specific fusion rule. Finally, the resulting image is obtained after applying the inverse transform on the fused representation [46].

The Laplacian Pyramid (LAP) algorithm which is a famous multi-resolution representation tool. It is widely used in image fusion since the early 1980s. The main idea is to perform a pyramid decomposition on each source image, then integrate all these decompositions to form a composite representation, and finally reconstruct the fused image by performing an inverse pyramid transform. This method has a lack of information direction [36].

Later, an image fusion algorithm based on the discrete cosine transform is proposed in [47], named, DCT+Average. The DCT+Average method performs the fusion by taking the average of all the DCT coefficients of all the input images. This method suffers from blocking artifacts and blurred fused images.

Other most used methods in image fusion are the Discrete Wavelet Transform (DWT) [38, 41] and the Shift-Invariant Discrete Wavelet Transform (SIDWT) [40]. In these methods, input images are transformed from the time domain into the frequency domain. Upon fusion of the wavelet coefficients, the maximum choosing rule is usually used, as large absolute wavelet coefficients mostly correspond to salient features in the images. Unfortunately, the DWT is with lack of information direction. Also, it is not shift-invariant because of the down

sampling used in its computing. These results in ringing artifacts in intraframe and moving artifacts in interframe also the fused image has the pseudo-Gibbs phenomena [15]. While SIDWT method removes the down-sampling at each level, it overcomes the shift-variant problem. But it is much costlier in storage and computation due to the high degree of redundancy, making it impractical for very large data sets. This technique also presents an important challenge in real-time implementation.

To overcome these shortcomings, the Dual-Tree Complex Wavelet Transforms (DT-CWT) is suggested [35, 42, 43]. which is nearly shift-invariant but has lower redundancy than the SIDWT. The DT-CWT has the capability to identify negative and positive frequencies and offers a considerably more compact representation in transform domain which leads to improve the directionality and realize a reduced shift-invariant [48]. The DT-CWT method is more efficient in image fusion compared to DWT and SIDWT.

Authors in [49] suggest an approach operated in the frequency domain. The input images are divided into 8×8 blocks and the DCT coefficients of each block are calculated. Simply, the block with the majority of higher values of 63 Alternating Current (AC) coefficients are selected in the fused image. This method (DCT+Ac_max) leads to low contrast and blocking artifacts fused image due to dividing images into blocks.

1.3.1.3 Miscellaneous fusion methods

A multi-focus image fusion using artificial neural networks is proposed in [50]. In this method the source images are decomposed into blocks. For a Given similar blocks, a neural network is trained to identify which one is clearer. The fusion then proceeds by choosing the clearer block to build the final image. The resulting image has a blocking artefact due to the dividing of the input images into blocks.

In [51], a spatial domain method was proposed based on bilateral gradient and sharpness criterion (BGBSC) for multi-focus image fusion, in which it measures the local sharpness in the blocks of input images. Sharper blocks will make a greater contribution to the fused image. The BGBSC has the ability to best evaluate the local content information, but the fused image has artifacts and block effects that reduce the visual quality.

Another method described in [52], in which the input images are decomposed by the DWT, then the principal component analysis is applied for approximate and detailed coefficients of source images. Next, the mean of the principal components of each source

image constituted weights which are used in the fusion rule. Finally, the output image is constructed by selecting the pixels (spatial domain) through these weights. This method is plagued by some undesired effects such as blurring and reduction in contrast.

In [53], another image fusion technique was developed based on two-scale decomposition and saliency detection. Initially, input images are decomposed into base layers and detail layers containing large-scale differences and small differences respectively. This was done using a simple mean filter. A weight map is constructed using median and mean filters. The fused image is obtained using detail and base layers. Although the algorithm is simple, however, the fused image is with less visual quality due to the drawback of the mean filter. It is weak at preserving image edges.

1.4 Contributions

In order to well contribute to the image fusion field and also to improve the fusion images quality, an efficient and simple image fusion method is proposed for VSNs [54]. This method based on integer lifting wavelet [55-59] and discrete cosine transformers. Firstly, the proposed method decomposes input images using the ILWT, which has principal advantages, since it can easily use integer operations to represent the wavelet transform frequencies (approximation and detail) and the significant features like lines, boundaries and edges. This representation is done in the spatial domain with fixed-point and avoids the shortcomings caused by the floating-point or rounding such as loss of information, computational complexity, time and energy consumption and memory space.

Since the most used cameras in VSNs are providing images in integer format, so the ILWT is the most appropriate. Secondly, from all approximation sub-bands which represent visual information, we can extract the focused regions using the variance as an activity level measure in DCT domain [60]. The block with the highest activity level is selected as the proper one for the fused approximation sub-band. Then, these blocks are subjected to consistency verification [38] to construct a fused approximation sub-band.

Large variance means that the region is more informative than the others. An important key is that all approximation sub-bands are replaced by their fused one to ensure the transfer of the maximum amount of information into the fused image. Finally, we fuse all sub-bands by taking the optimum weighted average based on pixel significance. Most of the existing approaches in the image fusion field are complex and energy-consuming which are

not suitable for VSN real-time applications such as surveillance. In addition, they are not easy for the purpose of hardware implementation.

Since image is an important part of multimedia data, we propose here a novel image fusion technique which is carried out on the captured images from the VSN (input images) to compose a single one with more precise visual data from the input images, and it can be suitably used for visual perception and computer processing. The proposed method has been examined on different image sets and compared with some algorithms in image fusion literature. The simulation results show the performance of our proposed method. Our main contributions can be summarised as follows:

(1) All operation such as image decomposition, fusion rule and image reconstruction are using integer operations, which lead to solve the shortcoming pointed in some state-of-the-art methods such as computational complexity, energy consumption and loss of information.
(2) An effective primary fusion of approximation sub-bands based on an activity level measurement in the DCT domain, which absorbed the more informative regions.
(3) Additional fusion process to ensure the best fusion of the edges and boundaries. The optimum weights based on pixel significance are adopted as a fusion criterion.
(4) Our proposed method is less sensitive to noise and it is appropriate for real-time image fusion in visual sensor networks.

1.5 Organisation of thesis

The remainder of this thesis is organised as follows:

In chapter 2, we will discuss the basic concepts of image fusion, its main properties and its applications. We will give also an overview on the classification of different image fusion techniques. The image fusion quality metrics are also presented.

In chapter 3, we will present the mathematical explanation of some tools used in image fusion approaches such as the DCT, the DWT and the ILWT transforms and their properties. After that, we will provide a literature review of the existing image fusion methods operating in these transform domains.

In chapter 4, we will present a detailed description of the proposed fusion algorithm. Where the fusion process is done in two phases to ensure a high quality of the fused images.

The first phase is conducted in the DCT domain, where the second phase is performed in the ILWT domain.

chapter 5 will discuss the simulation of the proposed method in comparison with other methods. The simulation results are also given at the end of the chapter.

Finally, we terminate this manuscript by a conclusion that contains an evaluation of our presented work. Suggestions for future research directions will be also discussed.

Chapter 2:

An overview on image fusion

2.1 Introduction

Due to the tremendous development in the fields of image processing, image fusion become a central research area with growing importance and it has received considerable attention from researchers for many applications in different domains such as medical, military, robotics, machine vision, and civil applications. Also, more important applications of image fusion in visual sensor networks such as surveillance, environmental monitoring, smart homes, and Smart meeting rooms [5].

There are many benefits in using image fusion at the integration of disparate and complementary data to increase the reliability of the interpretation as well as to enhance the information apparent in the images. This leads to more accurate data and increased utility. It is also stated that fused data provide for robust operational performance, i.e. improved reliability, reduced ambiguity, increased confidence and improved classification.

2.2 Visual sensor networks

Among the objective differences between the visual sensor networks and the other kinds of sensor networks return to the nature of how the image sensors (cameras) can recognize information from the environment. The majority of sensors provide measurements as one-dimensional data signals. Nevertheless, image sensors consist of a big number of photosensitive cells. A single measurement of image sensor offers a two-dimensional group of data points [3], that we can perceive on Figure 2.1.

Figure 2.1: Visual sensor network.

Furthermore, a camera's sensing model is by nature different from the sensing model of each different type of sensor. Normally, a sensor combines data from its proximity, as defined by its sensing range. On the other hand, cameras are described by a directional sensing model, so that, cameras take images of distant scenes/objects from a particular direction. The two-dimensional sensing area of traditional sensor nodes is - in the case of cameras - substituted by a three-dimensional viewing volume (the field of view) as illustrated in Figure 2.2. Visual sensor networks are in many ways singular and more challenging compared to other kinds of wireless sensor networks [61].

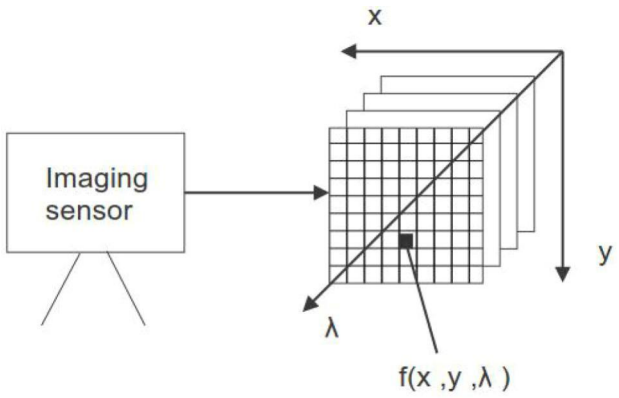

Figure 2.2: Basic imaging sensor scheme.

2.2.1 Energy consumption

The lifetime of the battery-operated camera is restricted by their energy consumption, which is proportional to the energy needed for taking, processing, and transmitting images. Due to a large amount of visual data generated by the cameras, both processing and transmitting image data are very expensive in terms of energy, much more so than for other types of sensor networks. Moreover, visual sensor networks need a large bandwidth for transmitting image data. Consequently, both energy and bandwidth are the most constrained than in other type of sensor networks [62].

2.2.2 Real-time performance

Various applications of visual sensor networks need real-time data from cameras, which impose stringent limits on maximum permissible delays of data from the sources to the user. The real-time performing of a VSN is influenced by the time needed for image data processing and also for the transmission of the processed data via the network. Restricted by limited energy resources and by the speed of embedded processors, most camera nodes have

processors that support only lightweight processing methods. On the network side, the real-time performance of a visual sensor network is restricted by the wireless channel limitations, used wireless standard, and by the current network requirement. As an example, upon detection of an event, the camera nodes can suddenly inject big amounts of data in the network, that could cause data overcrowding and increase data delays [2].

2.2.3 Data storage

A big number of cameras in VSNs generate huge amounts of data over time, which in certain cases should be stored for later analysis. An example is the monitoring of remote areas by a set of cameras, where the frequent transmission of the captured images would rapidly consume the camera's energy resources. Thus, in these cases, the cameras should be provided with memories of higher capacity in order to save the data. To reduce the amount of data that needs storage, the camera node should classify the data according to its importance by using spatiotemporal analysis of image frames, and decide which data should have priority to be stored [63].

In addition, image fusion can reduce the redundancy in the data collected by cameras with overlapped views. It can be achieved via local transmission and processing. This allows the camera nodes to decrease their requirements for storage space by storing only data of fused images.

2.2.4 Compression and coding in VSNs

Pre-storing or transmission steps include image coding and compression, which help further minimize the amount of data stored in memories or sent to its intended. There are two important kinds of coding: the intra-coding and the inter-coding. The intra-coding also was known as transform coding such as wavelet transforms and discrete cosine (JPEG 2000, JPEG), is a technique employed for lossy compression of images and videos. The inter-coding is also known as motion compensated prediction. It can achieve a higher compression rate for images and videos but needs more computations. Traditional techniques, particularly the inter-coding, are not appropriate for VSNs. They offer high-resolution images at the cost of high complexity, memory resources, and power consumption.

Table 2.1: Compression / Coding techniques, In-network processing techniques and Applications of VSN platforms [5].

Platform	Compression / Coding techniques	In-network processing techniques	Targeted applications
Cyclops	JPEG compression	Matrix operation libraries Edge detection (Sobel algorithm) Background subtraction (Object detection) Coordinate conversion	Object detection Hand posture recognition
MeshEye	----	Background subtraction (Object detection) Stereo matching algorithm (Object location) Object acquisition and extraction	Distributed intelligent surveillance
Panoptes	JPEG compression	Motion detection filtering algorithm	Video surveillance
Meerkats	JPEG compression	Object/event detection based on background subtraction or frame differencing	Moving body tracking
FireFly Mosaic	JPEG compression	Frame differencing, Color tracking, Convolution, Edge detection. Connected component analysis, Face detection, etc.	Assisted living application Home activity clustering
MicrelEye	----	Background subtraction Feature extraction Classification algorithms	Image classification People detection
XYZ-ALOHA	JPEG compression	Histogram reconstruction Motion, Edge and Centroid detection	Pattern recognition (hand recognition, hand gesture recognition)
CITRIC	JPEG compression	Background subtraction Frame differencing Object acquisition and extraction	Single target tracking Camera localization using multi-target tracking
Vision Mote	JPEG compression	----	Water conservancy engineering

Throughout the literature review, we have noted that intra-coding (transform-based coding) algorithms are generally preferred over the inter-coding (non-transform-based coding) ones due to their poor compression. The most of existing visual sensor network platforms such as Cyclops [64], Panoptes [65], Meerkats [9], FireFly Mosaic [66], XYZ-ALOHA [67], CITRIC[68] and Vision Mote [69] use quality-scalable JPEG or platform-specific optimized JPEG compression for intra-coding. We summarize the compression algorithms used in some visual sensor network platforms in Table 2.1.

2.3 Concept of image fusion

Image fusion could be defined as the process of combining information from different images into a single composite one, using advanced image processing techniques. The fused image contains more accurate visual data and it contains more information from the input images, and it can be appropriately used for visual perception and computer processing[15]. A general definition can be given as “Image fusion is the combination of two or more different images to form a new image by using a certain algorithm” [70].

The aim of image fusion is to integrate all the extracted useful information from the source images into the fused image. In which, the resulting image will be reliable, and without introducing artifacts or inconsistencies which will be suitable for visual perception and computer processing.

During the process of fusion, input images A and B are combined into a new fused image F by transferring, ideally all of useful information into F. This is illustrated graphically using a simple Venn diagram in Figure 2.3.

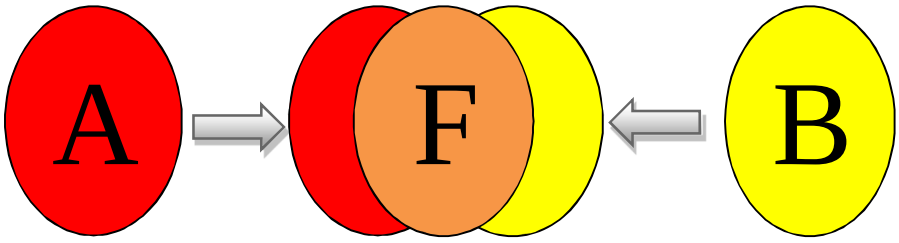

Figure 2.3: Venn diagram representation of the image fusion process.

Occasionally, image fusion techniques can be categorized according to the data set into four categories which are: multimodal, multi-view, multi-focus and multi-temporal. They are [71]:

- **Multi-view fusion**: images of the same modality captured from different places or under different conditions, but at the same time[72].
- **Multimodal fusion**: images captured using various sensors of different modalities: PET, CT, MRI, visible, infrared; ultraviolet…etc[73].
- **Multi-temporal fusion**: images of the same scene taken at different times (of the same modality)[74].
- **Multi-focus fusion**: the original image can be divided into regions such that every region is in focus in at least one channel[75].

2.4 Single sensor image fusion system

An image fusion scheme with a basic single sensor is shown in Figure 2.4. The shown sensor might be a visible-band sensor (digital camera). This sensor can detect and conveys the information that constitutes the real world as a sequence of images. These images are then be fused in a single fused image rich in relevant information and ready to be used by human or computer in many tasks [76]. For example, in object detection, a human operator searches the scene to detect objects such intruders in a security area.

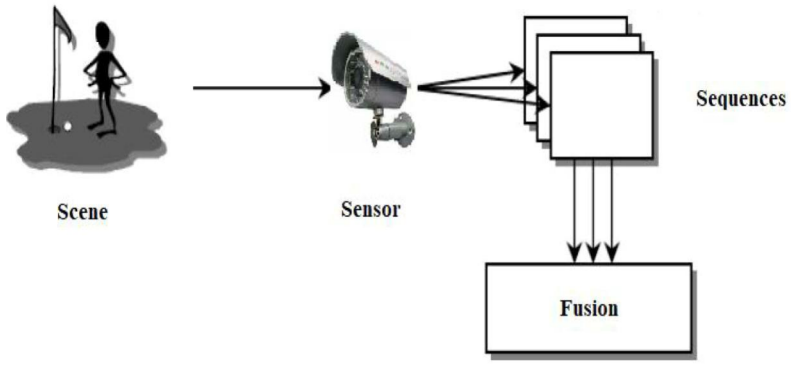

Figure 2.4: Single sensor image fusion system.

The imperfections of this kind of systems underlie the limitation of the used imaging sensor. The conditions whereby the system can operate such as: resolution, the dynamic range...etc. are restricted by the efficiency of the sensor. As an example, a visible-band sensor such as the digital camera is suitable for a brightly illuminated environment such as daylight scenes but is not appropriate for poorly illuminated situations found during night, or under adverse conditions such as in fog or rain.

2.5 Multi-sensor image fusion system

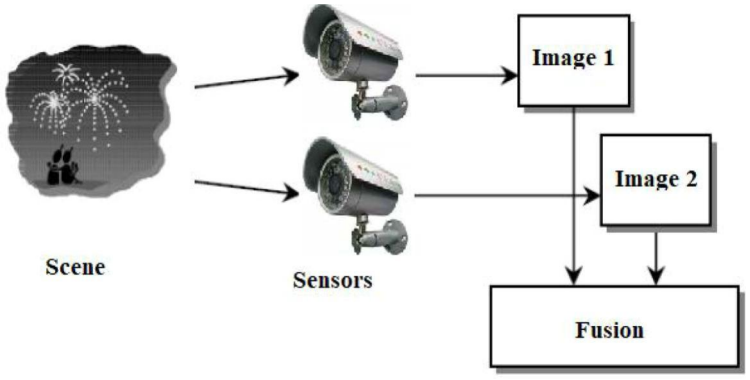

Figure 2.5: Multisensory image fusion system.

A multi-sensor image fusion framework overcomes the shortcomings of a single sensor image fusion by combining all images from all sensors to form a fused composite image. Figure 2.5 represents an illustration of multi-sensor image fusion system. In this picture, the digital camera is accompanied by an infrared camera and their individual images are combined to obtain a fused image. This approach gets over the shortcoming referred before. The infrared camera is suitable for the nightly scenes while the digital camera is appropriate for the daylight scenes.

Some advantages of the multi-sensor image fusion system can be presented here [71]:

- Extended spatial (temporal) coverage: relevant information from sensors that differ in spatial resolution can increase the spatial (temporal) coverage.
- Extended range of operation: numerous sensors that operate under various operating conditions can be deployed to extend the effective range of operation. As an example, different sensors can be employed for day/night operation.
- Powerful system performance: redundancy in multiple measurements can help in systems robustness. In case one or more sensors fail or the performance of a particular sensor deteriorates, the system can depend on the other sensors
- Reduced ambiguity: relevant information from multiple sensors can reduce the ambiguity linked with the decision process.
- Enhanced reliability: The fusion of multiple measurements can reduce noise and therefore enhance the reliability of the measured quantity.

- Compact representation of data: fusion leads to compact representations. For example, in visual sensor networks, instead of storing imagery from several sensors, it is comparatively more efficient to store the fused information.

2.6 Image fusion processing levels

In the literature[23, 77-79], image fusion is usually performed at one of the three different processing levels: pixel-level, feature-level or decision-level. In the context of fusion levels, the terms employed in the literature are not standard. We find that pixel-level also appears as signal or image. Region-level also referred to as regional or object. Other expressions for decision-level are symbolic or information. The more used terms for image fusion levels are:

- Pixel-level.
- Feature-level.
- Decision-level.

Certainly, every individual level needs different processing procedure. General processing steps for different fusion levels are shown in Figure 2.6 below.

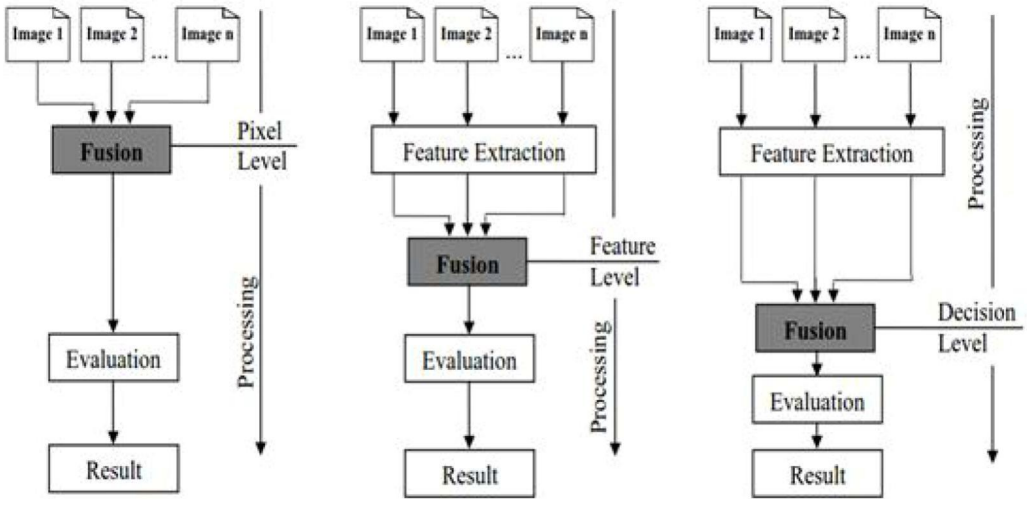

Figure 2.6: General processing steps for different fusion levels.

2.6.1 Pixel-level

Image fusion at pixel-level means that the fusion process is done at the lowest processing level referring to the merging of measured physical parameters, without applying

any image strengthening technique to the input images. Compared with others, pixel-level fusion immediately merges the original information in the input images, which aims at resulting a fused image that is more useful for visual perception and machine processing.

Pixel-level image fusion can be used in many applications that require two or more source images such as surveillance, medical and remote sensing applications. In surveillance applications[6], image fusion can merge the information through the electromagnetic spectrum for the night vision depends on the visible band and the infrared band. In medical imaging applications[80], images from different modalities can be fused jointly for a more precise and credible medical diagnosis. In remote sensing applications[81], the synthesis of a high-resolution panchromatic image (PAN) and a low-resolution multispectral image (MS) is used to result in a fused image bring together the improved spatial resolution with the spectral content of the MS image.

2.6.2 Feature-level

Feature-level image fusion is an intermediate level. This level implies that the source images are processed to extract important characteristics (features) using feature extraction techniques. Feature-level fusion uses a set of image pixels to form adjacent regions and requires extraction of various characteristics from input images. According to [82], “Feature extraction addresses the problem of finding the most compact and informative set of features, to improve the efficiency or data storage and processing”. Features can be edges, pixel intensities, or texture features.

Firstly, related features are extracted from source images with adequate extraction techniques, e.g., using segmentation procedures. Then, the extracted features are assigned to each other and fused with an appropriate fusion method. This preliminary data is transformed and represented as feature vector sets. Different types of features, such as signal amplitude or shape, length or image segments, are contingent on the nature of images and the applying of the fused image. The fused data can also be used for classification, detection or decision based on the fused feature set [83].

2.6.3 Decision-level

Decision fusion is the highest level of data fusion. At this level, fusion happens after the images have been processed for every modality independently. The received (extracted) information from different local classifiers is then fused to obtain the final decision. The result

depends extremely on the design of the decision rules. Decision fusion resolve differences, enhances the common interpretation and offers a better perception of the observed objects [74].

Some decision-level fusion methods are weighting methods, voting, Bayesian inference, Dempster-Shafer, fuzzy decision rule, rank-based, artificial neural networks. According to the different types of classifier outputs, these methods can be divided into three groups: the abstract level, the rank level, and the measurement level [84].

2.7 Overview on fusion Techniques

As it stated before, fusion algorithms are divided into two main categories: spatial domain and transform domain. The following are brief descriptions of the characteristics of each group.

2.7.1 Spatial Domain Techniques

In the spatial domain image fusion methods, the processing is directly done at pixel level of the source images. The pixel values are manipulated to yield desirable outcome. Generally, image fusion methods that processed in the spatial domain such as simple fusion methods, Principal Component Analysis (PCA), Brovey Transform (BT), Intensity-Hue-Saturation (IHS) And High Pass Filtering (HPF) are simple, fast and easy to implement. However, they produce a spatial distortion and suffer from blurring effects in the fused image. Spectral distortion will be a negative impact when we want to go for additional processing such as classification problem [23].

2.7.1.1 Simple fusion

The traditional image fusion methods primarily perform a so fundamental operation such as pixel selection, addition, subtraction or averaging. Although the computational efficiency and the low complexity; these methods are not always effective. The output image suffers from a low contrast and fading of some essential salient features [85].

a. **Simple Maximum:** We assume that we have two input images A and B of the same size $(M \times N)$ and F is the fused image, the output image is obtained by taking the maximum intensity of corresponding pixels from the input images. Which can be defined as:

$$F(i,j) = Max[A(i,j), B(i,j)]. \quad (2.1)$$

b. **Simple Minimum:** The minimum intensity of corresponding pixels from both input images is selected for building the fused image. Which can be given by:

$$F(i,j) = Min[A(i,j), B(i,j)]. \quad (2.2)$$

c. **Simple Average:** In this method, every pixel intensity in the given source image is selected and the mean (average) value is assigned to the corresponding pixel intensity of the output (resultant) image which is defined by equation (2.3). This is recurrent for all pixel values.

$$F(i,j) = Mean\ [A(i,j), B(i,j)]. \quad (2.3)$$

2.7.1.2 Principal Component Analysis

The Principal Component Analysis is a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. It produces a novel set of axes which are orthogonal. The first principal component is chosen to be along the trend with the maximum variance. The second is obliged to lie in the subspace vertical to the first principal component and so on. The aim of principal component analysis is to reduce the number of variables of the dataset while keeping the majority of the original variability in the data[86]. Figure 2.7 represents the information flow diagram of the PCA-based image fusion, whereas the algorithm steps are given below:

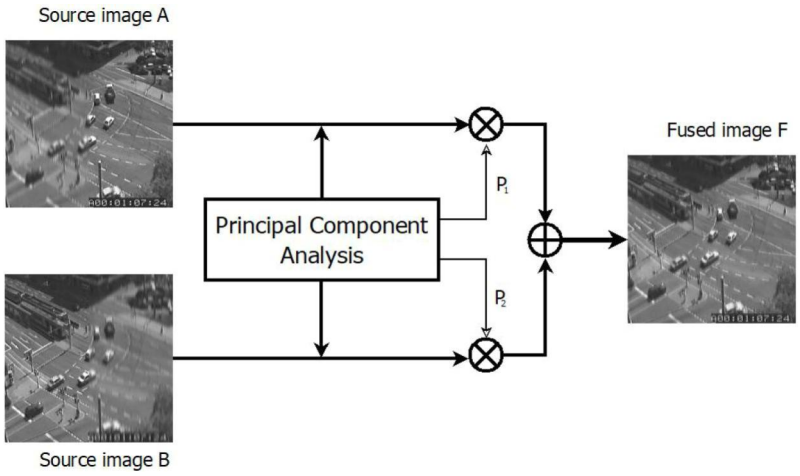

Figure 2.7: The information flow diagram of PCA.

PCA algorithm steps:

1. The input images $(A\ and\ B)$ are arranged into column vectors Z (data matrix).
2. Calculate the empirical mean Me of each column.
3. Subtract the mean vector Me from each column of the data matrix Z to result in a matrix X.
4. Compute the covariance matrix C of the X.
5. Calculate the eigenvectors V and eigenvalue D of the covariance matrix C. Then, sort them by decreasing eigenvalue.
6. The normalized components P_1 and P_2 are computed from the eigenvectors V as:

$$P_1 = \frac{V(1)}{\sum V(1)} \tag{2.4}$$

$$P_2 = \frac{V(2)}{\sum V(2)} \tag{2.5}$$

7. The fused image F is given by:

$$F(i,j) = P_1 * A(i,j) + P_2 * B(i,j) \tag{2.6}$$

2.7.1.3 Brovey Transform

The Brovey Transform was developed by Bob Brovey. It depends on the chromaticity transform also known by the colour normalization transform as it involves the RGB transform method. The BT is a simple technique for merging information from different sensors and it is used to merge multispectral (MS) and panchromatic (PAN) images. The mathematical representation of the BT can be given as a combination of the MS and PAN images, so that each MS image is divided by the sum of the MS images, and multiplied by the PAN image[32]. The fused image is defined by:

$$\begin{bmatrix} Fused\ R \\ Fused\ G \\ Fused\ B \end{bmatrix} = \begin{bmatrix} \frac{R}{R+G+B} \times PAN \\ \frac{G}{R+G+B} \times PAN \\ \frac{B}{R+G+B} \times PAN \end{bmatrix} \tag{2.7}$$

2.7.1.4 Intensity-Hue-Saturation

The Intensity-Hue-Saturation (IHS) method is one of the most frequently used in the image processing field. It has become a standard tool in image fusion to exploit the supplementary nature of multispectral images. It is widely used for colour and feature improvement and enhancement of spatial resolution. The IHS fusion method transforms a colour image from RGB space to IHS space, which effectively separates the colour aspects in intensity (brightness average), hue (dominant wavelength contribution) and saturation (purity). The intensity (I) is replaced by a high-resolution panchromatic image, then transformed back into the RGB space jointly with the same hue (H) and the saturation (S), to produce an IHS fused image [87]. The IHS algorithm can be given by the following steps:

IHS algorithm steps:

1. Step 01: RGB-IHS conversion system.

$$\begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-\sqrt{2}}{6} & \frac{-\sqrt{2}}{6} & \frac{2\sqrt{2}}{6} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2.8}$$

$$H = tan^{-1}\left(\frac{v_2}{v_1}\right) \tag{2.9}$$

$$S = \sqrt{v_1^2 + v_2^2} \tag{2.10}$$

Where v_1 and v_2 are mediator variables.

2. Step 02: The intensity (I) is replaced by the panchromatic image (Pan).
3. Step 03:

$$\begin{bmatrix} Fused\ R \\ Fused\ G \\ Fused\ B \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 1 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} Pan \\ v_1 \\ v_2 \end{bmatrix} \tag{2.11}$$

2.7.1.5 Hight Pass Filtering

The High-Pass Filter can be considered as a method to improve the spatial resolution of MS images. The high spatial resolution image is submitted to a small high-pass filter

(convolution mask), which efficiently reduces the low frequency spectral part and result in the high frequency part of the image which is concerned with the spatial information [77]. Then, filtered result will be added to the MS image and the result will be divided by two to redress the increase in the brightness. The mathematical representation is given below.

$$HPF(i,j)_k = \frac{MS(i,j)_k + FP(i,j)}{2} \quad (2.12)$$

Where HPF represents the result (fused) image, MS represents the multispectral image, FP is the HP filter output and i,j are the pixel index of the band $k(R,G\ or\ B)$.

2.7.2 Transform Domain Techniques

In the transform domain image fusion methods, images are converted from the time domain into the frequency domain by applying some mathematical operations called transforms such as Laplacian pyramid (LAP), the discrete cosine transform (DCT) and the discrete wavelet transform (DWT). In this context, the input images are projected onto local rules which are generally designed to represent the edges and sharpness of an image . Consequently, the resulting coefficients are significant to identify salient features. All the fusion processes are performed in the frequency domain. The fusion is realized by using a feature selection decision method (calculate the average, select the maximum, compute the contrast…etc). The feature selection method chooses the most meaningful pattern from the input image and selects it as a fused coefficient, then the inverse transform (from the frequency domain to the time domain) is performed to obtain the fused image. The transform domain techniques are characterized by the high quality of the spectral content.

2.7.2.1 The Laplacian pyramid

The Laplacian Pyramid (LAP) is one efficient structure adopted to represent an image with multi-resolution. The LAP introduced by Burt and Adelson in 1983[88], is used to decompose the input image into several low-frequency levels. These levels similar to the structure of the pyramid where any level is generated iteratively by filtering images applying a low pass filter. Laplacian pyramids are achieved by first compressing the input image to yield the next level then subtracting the input image from the low passed image. This method handles with the pixels of the image at all levels. Consequently, it is more able to identify features of the image efficiently and effectively[89]. So, the Laplacian pyramid

decomposition is divided into two steps: the Gaussian pyramid decomposition, and the Laplacian pyramid.

A. Gaussian Pyramid Decomposition:

Suppose the zero level of the pyramid G_0 is equal to the input image, G_0 is on the bottom of the pyramid, and the $l-th$ level of Gaussian pyramid which indicated as G_l is resulted by these steps:

Firstly, the convolution is performed between the $l-1-th$ level image G_{l-1} and the window function $\omega(m,n)$ which has low-pass characteristics. Then ,convolution results were separated out in the down sampling, which can be defined as:

$$G_l(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} \omega(m,n)\, G_{l-1}(2i+m, 2j+n), \tag{2.13}$$

With: $1 \le l \le N, 0 \le i < R_b, 0 \le j < C_l$.

Where N is the maximal level of pyramid, C_l and R_b represent the column and row number of the $l-th$ level pyramid, respectively. $\omega(m,n)$ is called weighting function or generating kernel, which is a 2-D separable 5×5 window function defined by:

$$\omega = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}. \tag{2.14}$$

It is convenient to consider this operation as a standard reduce process, and simply write :

$$G_l = Reduce(G_{l-1}). \tag{2.15}$$

Then, the Gaussian pyramid is constituted by $G_0, G_1, \dots, G_N$, where G_0 is the bottom and G_N is the top of the pyramid, and the total number of Gaussian pyramid layers is $N+1$.

B. Laplacian Pyramid Decomposition:

In order to reduce the large number of repeated information from Gaussian pyramid, it is required to find the difference between the adjacent two images and get the band-pass

filtered images, this set is the Laplacian Pyramid. The specific algorithm is as follows. Let G_l^* be the image obtained by expanding G_l , then G_l^* has the same size with G_{l-1} . So, the amplification operator Expand can be used, namely :

$$G_l^* = Expand(G_l). \tag{2.16}$$

According to (2.15), the Expand operator is defined as:

$$G_l^*(i,j) = 4\sum_{m=-2}^{2}\sum_{n=-2}^{2}\omega(m,n)\,G_{l-1}\left(\frac{i+m}{2},\frac{j+n}{2}\right), \tag{2.13}$$

With: $1 \le l \le N, 0 \le i < R_b, 0 \le j < C_l$.

Where:

$$G_l^*\left(\tfrac{i+m}{2},\tfrac{j+n}{2}\right) = \begin{cases} G_l\left(\tfrac{i+m}{2},\tfrac{j+n}{2}\right); & \tfrac{i+m}{2},\tfrac{j+n}{2}\ are\ integers, \\ 0 \quad ; & others\ . \end{cases} \tag{2.14}$$

Set:
$$\begin{cases} LP_l = G_l - G_{l+1}^*; 1 \le l < N, \\ LP_l = G_N \quad ; \quad l = N. \end{cases} \tag{2.15}$$

Where N is the number of Laplacian pyramid levels, LP_l is the $l-th$ level image decomposed from Laplacian pyramid, and Expand operator is the inverse of Reduce operator. Now, we can get the Laplacian pyramid composed of $LP_0, LP_1, \dots, LP_N$. Each of them is the difference of its Gaussian pyramid image itself and the last level's which has been interpolated and enlarged, the course just like band-pass filtering.

2.7.2.2 The discrete cosine transforms

The Discrete Cosine Transform is an important transform widely used in image processing such as fusion, watermarking and compression. It is a limited sequence of data points from the point of a sum of cosine function oscillating at various frequencies. The DCT based image fusion methods are very effective in JPEG-coded images for archiving or transmission. The input images are divided into 8×8 blocks, then the two-dimensional 2-D DCT is performed on each block, this produces 64 DCT coefficients for each block which are quantized to decrease their magnitude. Then, the fusion process is performed at the coefficient level. Finally, the fused image is obtained by applying the inverse 2-D DCT transform [47].

2.7.2.3 The discrete wavelet transforms

Mallat has introduced the original theory and the principal concept of wavelet-based multiresolution analysis [90]. The wavelet transforms (WT) has been largely used in image processing such as feature detection, data compression and image fusion. It can be used to decompose images (2D signals) by applying the two-dimensional discrete wavelet transform (2-D DWT). It is realized by performing at first the 1-D DWT on the rows and the columns of an image by separately filtering and down sampling. This generates a single sub-band of approximation coefficients and three sub-bands of detail coefficients. This are low-low (LL), low-high *(LH)*, high-low *(HL)*, and high-high *(HH)* sub-bands (sub-images) [91].

The LL sub-band represents the half resolution of the original image which includes smooth spatial data with high spatial correlation. While the LH, HL, and HH sub-bands corresponds to three different directional-orientations (Horizontal, Vertical and Diagonal) respectively. The 2-D DWT can be iterated in a similar manner on the LL sub-band to provide a multilevel decomposition.

The basic concept of image fusion based on DWT is to perform the decomposition on all source images. The result in coefficients in all sub-bands (low-frequency and high-frequency) are then performed with a specific fusion rule. Finally, the fused image is obtained by performing the inverse 2-D DWT on the selected coefficients.

2.8 Applications of image fusion techniques

Recently, image fusion has been attracting a considerable attention in a large range of applications, such as in visual sensor networks, remote sensing and medical imaging.

2.8.1 Visual sensor networks

A huge number of images can be captured in Visual Sensor Networks (VSNs). These images are used in several applications, such as surveillance, traffic monitoring and environment monitoring.

2.8.1.1 Surveillance

The most popular examples of VSNs applications in surveillance contain the habitat, home security, intrusion detection and environmental monitoring. The captured images of the surroundings are processed for the identification, recognition and classification of the required

objects. According to the results, some control is applied or some preventive measures are taken. Due to the bandwidth limitations of the VSNs, these applications require a lowest possible processing operation on images and the least amount of data for transmission or archiving. For this purpose, we can use image fusion techniques [92].

2.8.1.2 Traffic monitoring

Image fusion in VSNs could be employed in traffic monitoring applications. The traffic monitoring systems can provide us with an important information about vehicles such as the number, speeds, classifications, deceleration or acceleration and license plate recognition. Also, accidents, traffic offences...etc. the main advantage of image fusion in this application is the ability to cover all roads and accurately vehicles recognition[8].

2.8.1.3 Environment monitoring

This application is very helpful to solve nature conservation questions so that it can gather data of animals and the environment to answer several science issues. It can be used to survey the variety and availability of the bird communities and terrestrial mammal day and night as a result of the fusion of visible images and infra-red images [93].

2.8.2 Remote sensing

Lately, Remote sensing image fusion is a new field falls within the overall framework of data fusion. It has become a common tool for analysing various features offered on satellite images. Generally, satellite images are two image datasets which are the multispectral (MS) images and the panchromatic (PAN) images. The MS image is with low-resolution and having the best spectral resolution than PAN image, while this last is with high-resolution and it is having the best spatial resolution than the MS image. Because of this trade-off between the MS and the PAN image resolutions, it might be difficult to maintain spatial and spectral resolution in a one image without using a fusion technique [81].

2.8.3 Medical imaging

Medical image fusion is employed to extract helpful information from multi-modal medical images. Image fusion has a board application field of medicinal diagnostic. Medical images have various kinds, such as computerized tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), electrical capacitance tomography

(ECT) and Single Photon Emission Computed Tomography (SPECT). The anatomical image CT and MRI include high-spatial resolution, while the functional images ECT and SPECT includes low-spatial resolution. Medical fusion image is to merge the anatomical image and the functional image jointly into a single one. The fused image can offer plentiful information to diagnose diseases [94].

2.9 Fusion performance metrics

In order to evaluate image fusion methods, extensive experiments are carried out to clearly demonstrate the performance of the proposed method. Here, the current image fusion performance metrics are classified into two classes. In the first class we found metrics that requires a referenced image, while the other does not. In this section, we chose nine performance metrics: The Root Mean Squared Error (RMSE) [95], Peak Signal To Noise Ratio (PSNR) [96] and the Structural Similarity (SSIM) index [97], used when there is a reference image. Otherwise, for non-reference image we adopt the Standard Deviation (STD) [98], Average Gradient (AG) [99], entropy (He) [100], Spectral Frequency (SF) [101], edge based similarity metric ($Q^{\frac{AB}{F}}$) [102, 103] and the total loss of information ($L^{\frac{AB}{F}}$) [78].

2.9.1 Root mean squared error

RMSE is widely used to calculate the variation between the fused and the reference images [95] and can be defined by Equation 2.16.

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{m=1}^{M}\sum_{n=1}^{N}[\boldsymbol{R}(\boldsymbol{m},\boldsymbol{n}) - \boldsymbol{F}(\boldsymbol{m},\boldsymbol{n})]^2}{\boldsymbol{M} \times \boldsymbol{N}}}, \qquad (\mathbf{2.16})$$

Here, R and F are the referenced and the fused image respectively. When the RMSE is closest to zero, the fused image is more similar to the reference image.

2.9.2 Peak signal to noise ratio

In assessing image fusion methods, the PSNR is commonly used to evaluate the fused image quality [96]. Higher value demonstrates a better fusion. It is given by equation 2.17:

$$PSNR = 10\log_{10}\left(\frac{L^2}{\frac{1}{M\times N}\sum_{m=1}^{M}\sum_{n=1}^{N}[R(m,n)-F(m,n)]^2}\right), \qquad (2.17)$$

With L is the maximum possible pixel value of the image.

2.9.3 Structural similarity index

In objective assessment and for comparing the pixel intensities of local patterns between the fused and source images [97], the SSIM is used. It is defined by equation 2.18:

$$SSIM(R,F) = \frac{2\mu_R\mu_F + C_1}{\mu_R^2 + \mu_F^2 + C_1} \times \frac{2\sigma_{RF} + C_2}{\sigma_R^2 + \sigma_F^2 + C_2}, \quad (2.18)$$

With μ_R, σ_R^2, μ_F and σ_F^2 are the mean and the variance of the referenced (R) and the fused (F) images respectively, σ_{RF} is the covariance between R and F. C_1 and C_2 are constants with $C_1 \ll 1$ and $C_2 \ll 1$. SSIM takes values between -1 to 1 and the fused image is similar to reference image if SSIM equals 1 [104].

2.9.4 Standard deviation

The standard deviation measures the contrast of an image, that reflects the dispersion in the data [98]. It is given by equation 2.19.

$$STD = \sqrt{\frac{\sum_{m=1}^{M}\sum_{n=2}^{N}[F(m,n) - \bar{F}]^2}{M \times N}}, \quad (2.19)$$

With $F(m,n)$ and $\bar{F}$ are the localised and averaged pixels of the fused image. For high contrast, STD must be with high value.

2.9.5 Average gradient

To compute the sharpness and the clarity in the fused image, the average gradient is the best metric. It is defined [78] by equation 2.120 as indicated below:

$$AG = \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{\sqrt{[F(m,n) - F(m+1,n)]^2 + [F(m,n) - F(m,n+1)]^2}}{M \times N}. \quad (2.20)$$

High value of AG indicates more clarity and sharpness image.

2.9.6 Entropy

As a key indicator of information quality, entropy (He) is used to appreciate the amount of information content in the fused image [100]. The large value in He means image more information. The entropy is given by equation 2.21:

$$He = -\sum_{k=0}^{255} p_k \log_2(p_k), \tag{2.21}$$

Where p_k is the probability associated with grey level k.

2.9.7 Spatial frequency

The study of the human visual system leads to the emergence of the spatial frequency [101]. It is calculated by computing the raw and column frequencies and defined by equation 2.22:

$$SF = \sqrt{(RF)^2 + (CF)^2}, \tag{2.22}$$

With RF and CF are the row frequency and column frequency respectively and they are given by:

$$RF = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} \sqrt{[F(m,n) - F(m,n-1)]^2}. \tag{2.23}$$

$$CF = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} \sqrt{[F(m,n) - F(m-1,n)]^2}. \tag{2.24}$$

2.9.8 Edge based similarity

The edge-based similarity metric $(Q^{\frac{AB}{F}})$ is widely used to assess the amount of the edge information transmitted from the source images into the fused image [105]. The $(Q^{\frac{AB}{F}})$ value varies in the range 0 to 1, where 1 indicates the optimal fusion [106]. It is given by equation 2.25:

$$Q_F^{AB} = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} [Q^{AF}(m,n) w^A(m,n) + Q^{BF}(m,n) w^B(m,n)]}{\sum_{m=1}^{M} \sum_{n=1}^{N} [w^A(m,n) + w^B(m,n)]}, \tag{2.25}$$

Where A and B are the source images. F represents the fused image. $w^A(m,n)$ and $w^B(m,n)$ are the weighted factors for Q^{AF} and Q^{BF} respectively. While these last two are the edge preservation values [102, 103].

2.10 Conclusion

In this chapter, we have presented the preliminary concepts and applications of the image fusion technique. In the first part, we defined the concept of image fusion and discussed the image fusion system. After that, we discussed the different image fusion processing levels and the different processing procedures. In the second part, we addressed an overview of image fusion techniques in both the spatial domain and transform domain. Also, we broached to some image fusion applications in different domains. We ended this chapter by some fusion quality metrics.

In the next chapter a brief state of the art review of recent image fusion techniques is presented. We will focus on image fusion methods based on DCT and DWT transforms. The proposed design system is also given.

Chapter 3:
DWT and DCT for Image Fusion: Literature review

3.1 Introduction

In chapter 2, we have addressed briefly some traditional image fusion techniques, which operate in spatial domain. These techniques are very simple to implement. But, they produce a spatial distortion and suffer from spectral degradation. The transform domain image fusion schemes are exploited employing various frequency domain analysis. The most popular used transforms in image fusion methods are the DCT transform and the DWT transform. The main purpose of using transform domain is the working in a framework where the important features of images are more obviously described than spatial domain.

Here, we present a brief literature review of some recent image fusion methods used in the transform domain. We are just going to focus on image fusion methods based on the DCT domain and the DWT domain.

3.2 Discrete Cosine Transform

The discrete cosine transform is mostly used in image fusion techniques. The 2-D DCT transform F of an $N \times N$ image block $b(i,j)$is given by [47, 107]:

$$F(k,l) = \frac{2}{N} c(k)c(l) \times \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b(i,j) \times \cos\left[\frac{(2i+1)k\pi}{2N}\right] \times \cos\left[\frac{(2j+1)l\pi}{2N}\right], \quad (3.1)$$

Where $k,l = 0,1,\dots,N-1$ and:

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0, \\ 1, & k \neq 0. \end{cases} \quad (3.2)$$

The original image can be recovered from the DCT coefficients by applying the inverse 2-D DCT transformer as follows:

$$b(i,j) = \frac{2}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} c(k)c(l) \times F(k,l) \times \cos\left[\frac{(2k+1)i\pi}{2N}\right] \times \cos\left[\frac{(2l+1)j\pi}{2N}\right], \quad (3.3)$$

Where $i,j = 0,1,\dots,N-1$.

In Equation (3.1), $F(0,0)$ represents the direct current (DC) coefficient and all the rest $F(k,l)$ are the alternating current (AC) coefficients.

3.2.1 Variance in DCT domain

The normalised DCT coefficients are given as [108]:

$$\hat{F}(k,l) = \frac{F(k,l)}{N}. \tag{3.4}$$

As it is known, in the spatial domain we can calculate the mean μ value and the variance σ^2 of an $N \times N$ block of an image as below:

$$\mu = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b(i,j). \tag{3.5}$$

and

$$\sigma^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [b^2(i,j) - \mu^2]. \tag{3.6}$$

The variance in the DCT domain can be defined by [60]:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b^2(i,j) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b(i,j) \times b(i,j)$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b(i,j) \left(\sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \frac{2\,c(k)c(l)}{N} \times F(k,l) \times \cos\left[\frac{(2k+1)i\pi}{2N}\right] \times \cos\left[\frac{(2l+1)j\pi}{2N}\right] \right)$$

$$= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \frac{2\,c(k)c(l)}{N} \times F(k,l) \left(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b(i,j) \times \cos\left[\frac{(2k+1)i\pi}{2N}\right] \times \cos\left[\frac{(2l+1)j\pi}{2N}\right] \right)$$

$$= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \frac{2\,c(k)c(l)}{N} \times F(k,l) \times \frac{N \times F(k,l)}{2\,c(k)c(l)}$$

$$= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k,l) \times F(k,l)$$

$$\Rightarrow \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} b^2(i,j) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} F^2(k,l). \quad (3.7)$$

As we know, the mean value μ of a block is given by $\hat{F}(0,0)$, and from Equations (3.6) and (3.7) we have:

$$\sigma^2 = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}[b^2(i,j) - \mu^2],$$

$$\Rightarrow \sigma^2 = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\frac{F^2(k,l)}{N^2} - \hat{F}^2(0,0). \quad (3.8)$$

Finally, the variance of an $N \times N$ block of an image in the DCT domain can be defined as the sum of squared normalized alternating current (AC) coefficients.

3.3 Wavelet transform

As a great mathematical tool, the wavelet transforms (WT) was developed in the field of signal and image processing. More recently, wavelets have begun playing an extremely crucial role in image fusion. WT is a more appropriate technique compared to Fourier transform (FT). The key difference is that the Wavelet Transform is a multi-resolution transform, that is, it allows a form of time–frequency analysis (or translation–scale in wavelet speak). When using the Fourier Transform the result is a very precise analysis of the frequencies contained in the signal, but no information on when those frequencies occurred. Also, it can't observe frequencies varying over time because the FT produces a function independent of time[109, 110].

Unlike the FT, the wavelet transform decomposes the input signal in terms of space and scale by a basis wavelet function which is a varying frequency in limited duration. The use of functions localized in time eliminates the windowing requirements generally found in Fourier techniques [111].

3.3.1 Continuous Wavelet Transforms (CWT)

Starting with *mother wavelet* ψ, a family member $\psi_{\tau,s}$ of *wavelet daughters* can be resulted by simply *translating* and *scaling* the mother wavelet ψ . Mathematically, can be defined as :

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right), \tag{3.9}$$

Were τ and s are location and scaling parameters respectively. Translating a wavelet simply means shifting its location in time, while *scaling* it simply means compressing it if $|s| < 1$ or stretching it if $|s| > 1$.

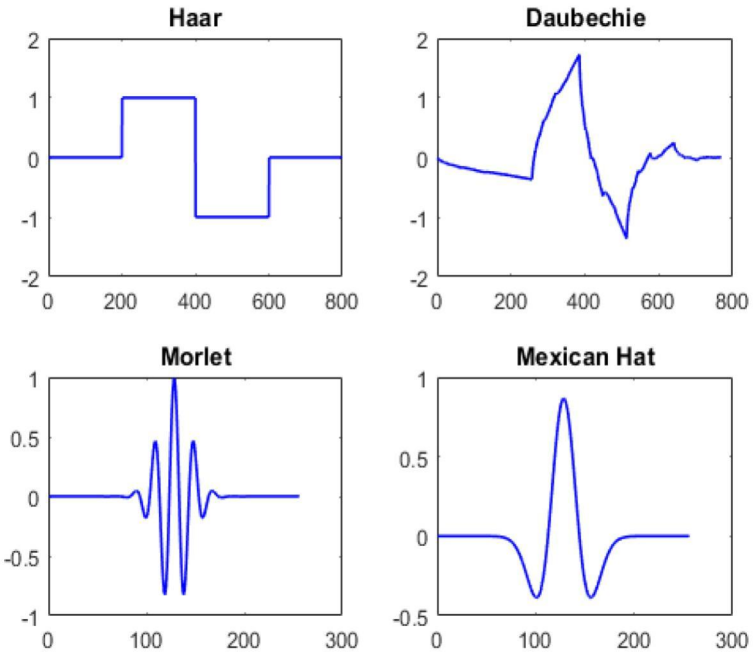

Figure 3.1: Different families of wavelets.

The mother wavelet (wavelet function) $\psi(t)$ must satisfy the conditions described below [112, 113]:

- First, a wavelet must have finite energy:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < +\infty. \tag{3.10}$$

- Second, a wavelet must verify the admissibility condition:

$$0 < C_\psi = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|}{|\omega|} d\omega < +\infty, \tag{3.11}$$

Were $\Psi(\omega)$ is the Fourier transform of the wavelet $\psi(t)$ and defined by:

$$\Psi(\omega) = \int_{-\infty}^{\infty} \psi(t) e^{-jt\omega} dt. \tag{3.12}$$

The admissibility condition implies that the wavelet has no zero-frequency component $(\Psi(0) = 0)$, i.e. the mean of the wavelet $\psi(t)$ must equal zero.

We can now define the continuous wavelet transform of a signal $x(t)$ by the function of two variables τ and s, which are given by:

$$X(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-\tau}{s}\right) dt \tag{3.13}$$

Where ψ^* denotes the complex conjugate of the mother ψ. The signal energy is normalized at different scales by dividing the wavelet coefficients by $\frac{1}{\sqrt{|s|}}$. which ensures that the wavelets have the same energy at different scales.

As the Fourier transform, the wavelet transform also has an inverse transformation defined by:

$$x(t) = \frac{1}{{C_\psi}^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(\tau, s) \frac{1}{s^2} \psi \left(\frac{t-\tau}{s}\right) d\tau ds \tag{3.14}$$

3.3.2 Discrete Wavelet Transform (DWT)

Since coming to the digital computing machines, it becomes necessary to express the discrete version of the wavelet transform. That it is possible by representing the continuous translation and scaling variables τ and s in terms of integer values. The most common way to descritise them is by using Equations (3.15) and (3.16) given below [114]:

$$a = a_0^m, \tag{3.15}$$

$$b = n b_0 a_0^m, \tag{3.16}$$

Where m and n are two integer parameters. By replacement the new formula of τ and s in Equation (3.9), the discrete wavelet can be given by [115]:

$$\psi_{m,n}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0), \tag{3.17}$$

Based on the dyadic sampling method, a_0 and b_0 values are chosen as 2 and 1 respectively. Therefore, $a = 2^m$ and $b = n \times 2^m$. By using these values, the discrete wavelet can be represented by:

$$\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n), \tag{3.18}$$

Generally, for dyadic decomposition, the wavelet coefficients of any function $x(t)$ are given by:

$$c_{m,n}(\omega) = 2^{-m/2}\int x(t)\,\psi(2^{-m}t - n)\,dt\,. \tag{3.19}$$

from the discrete wavelet coefficients, we can reconstruct the signal $x(t)$ as:

$$x(t) = \sum_{m=-\infty}^{+\infty}\sum_{n=-\infty}^{+\infty} c_{m,n}(\omega)\,\psi_{m,n}(t). \tag{3.20}$$

3.3.3 Filter banks

The filter bank is composed of two filter banks, which separate any signal into frequency bands, one each for the forward transform (analysis) and the inverse transform (synthesis) [116, 117]. An example of the two-channel filter banks is shown in Figure (3.2).

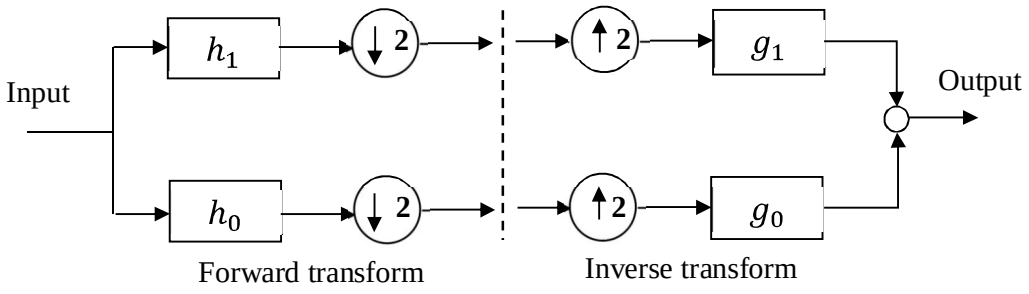

Figure 3.2: Two channel filter bank.

In this filter bank, the input signal is passed on the analysis bank via two separate channels of digital filters. It is filtered by a high-pass filter (h_1) and a low-pass filter (h_0), supported by a down sampling process by a factor of 2, in every channel. The high-passed data (H) contains the high-frequency content (approximation information) of the input signal while the low passed data (L) retains the low frequency content (detail information) of the input signal.

The synthesis filter bank used to reconstruct the original signal, the frequencies are first up sampled with a factor of 2 and then passed via the synthesis filter banks (g_1) and (g_0).

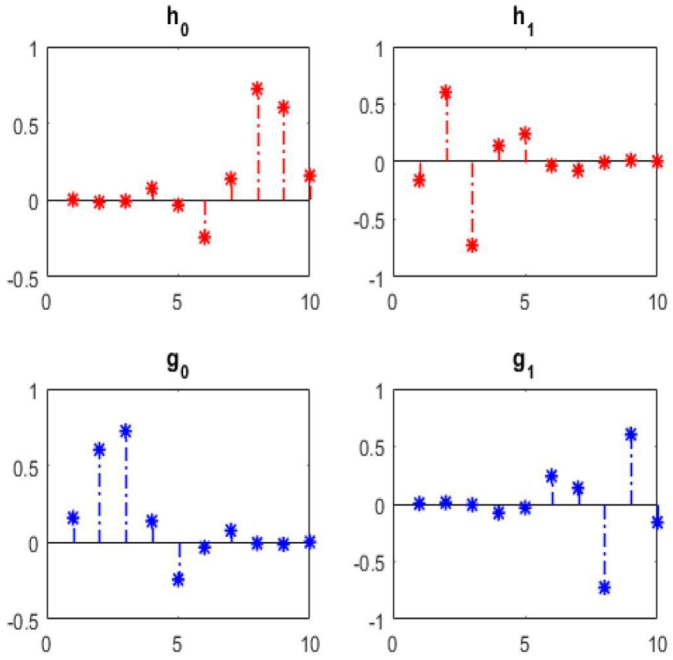

Figure 3.3: The decomposition and reconstruction filters for Daubechies2 wavelet(db5).

The synthesis filters (g_1) and (g_0) must be particularly adjusted to the analysis filters (h_1) and (h_0) in order to obtain an absolute reconstruction. By looking at the Z-transfer function of the two-channel filter bank, it is simple to get the relationship that those filters need to meet**.**

After the Forward transforms, the two sub-bands (high “Hsb” and low “Lsb”) can be given by:

$$Hsb = \frac{1}{2}\left[H_1\left(z^{1/2}\right)X\left(z^{1/2}\right) + H_1\left(-z^{1/2}\right)X\left(-z^{1/2}\right)\right]. \quad (3.21)$$

$$Lsb = \frac{1}{2}\left[H_0\left(z^{1/2}\right)X\left(z^{1/2}\right) + H_0\left(-z^{1/2}\right)X\left(-z^{1/2}\right)\right]. \quad (3.22)$$

the relationship between the input signal $X(z)$ and the output signal $Y(z)$ of the filter banks is defined by [118]:

$$Y(z) = \frac{1}{2}[G_0(z)H_0(z) + G_1(z)H_1(z)]X(z) + \cdots$$

$$\ldots + \frac{1}{2}[G_0(z)H_0(-z) + G_1(z)H_1(-z)]X(-z). \quad (3.23)$$

In order to remove the shortcomings of distortion and aliasing, the below conditions must be met:

$$G_0(z) = H_1(-z). \quad (3.24)$$

And

$$G_1(z) = -H_0(-z). \quad (3.25)$$

The final filtering equation is presented as:

$$Y(z) = \frac{1}{2}z^{-N}[H_0(z)H_0(z^{-1}) + H_0(z)H_0(-z^{-1})]X(z). \quad (3.26)$$

3.3.4 Two-Dimensional DWT

The two-dimensional extension of DWT (2-D DWT) is necessary for transformation 2-D signals (images). An image can be mathematically represented as a matrix (2-D array) $X[M,N]$ with M rows and N columns, where M and N are positive integers. The 2-D DWT implementation can be achieved by performing the 1-D DWT row-wise to generate an intermediate result and then perform the same 1-D DWT column-wise on this mediator result to generate the final result [114, 119]. This is presented in Figure (3.4).

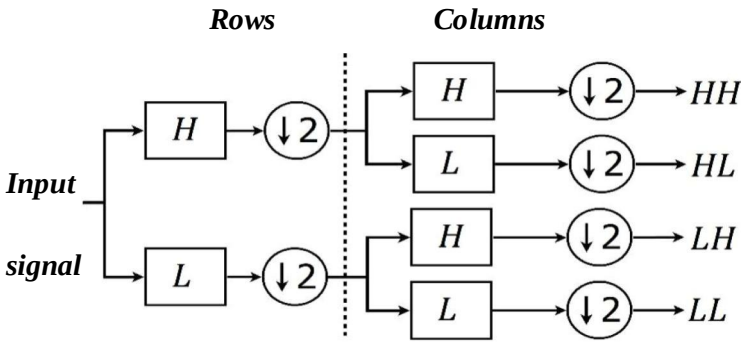

Figure 3.4: Filter bank structure of the 2-D DWT Analysis.

As mentioned in chapter 2, applying a 2-D DWT on an $(M \times N)$ image, we produce four sub-bands LL, LH, HL, and HH of the same reduced size $(\frac{M}{2}, \frac{N}{2})$. The 2-D DWT can be repeated in the same way on the LL sub-band to provide a multilevel decomposition. Figure (3.5) shown an example of 1 level 2-D wavelet decomposition of the 'clock' image using 'Haar' wavelet.

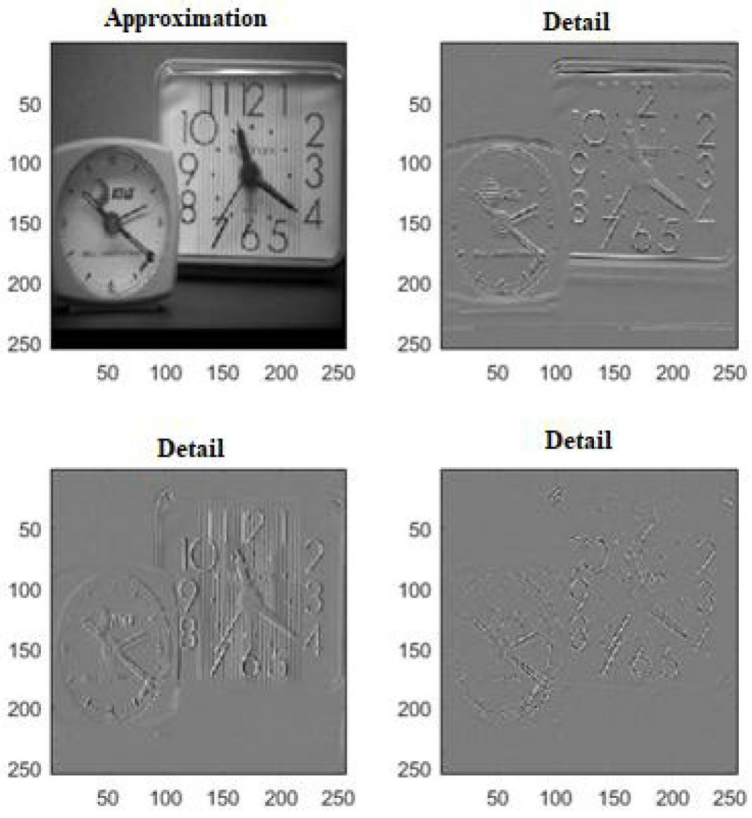

Figure 3.5: 1-level DWT of the 'clock' image.

A 2-D scaling function $\varphi(x,y)$, and three 2-D wavelets $\psi^H(x,y)$, $\psi^V(x,y)$ and $\psi^D(x,y)$ are crucial components for 2-D wavelet transform. These directional wavelets and scaling function are consisted of the product of a 1-D scaling function φ and corresponding wavelet ψ which are proved below [120]:

$$\begin{cases} \varphi(x,y) = \varphi(x)\varphi(y) \longrightarrow LL \\ \psi^H(x,y) = \varphi(x)\psi(y) \longrightarrow LH \\ \psi^V(x,y) = \varphi(y)\psi(x) \longrightarrow HL \\ \psi^D(x,y) = \psi(x)\psi(y) \longrightarrow HH \end{cases} \tag{3.27}$$

In order to reconstruct the output signal, all sub-bands are firstly up-sampled by a factor of 2. Then, they are further passed via the same set of low pass and high pass filters in both columns and rows. All reconstruction steps are shown in Figure (3.6).

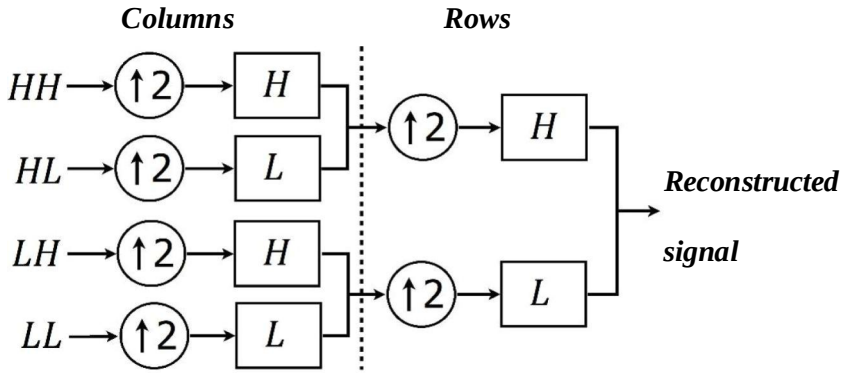

Figure 3.6: Block diagram of inverse 2-D DWT.

3.3.5 Integer Lifting Wavelet Transform ILWT

As known, the traditional wavelet based on the Fourier transformer in which the input signal (image) is separately filtered by low-pass and high-pass filters, then the output streams are sub-sampled to produce the approximation and detail sub-bands respectively. This process is computationally complex, requires more time for processing and requires a considerable memory space. To overcome these shortcomings, the lifting scheme is proposed in [121]. It is a powerful tool in signal processing, in which can easily calculate the wavelet transform without having the Fourier transformer, but the output signal of LWT is represented by floating-point numbers which usually hamper accurately reconstruct the signal.

Researchers in [122, 123] solve this problem and proposed the integer lifting scheme of the wavelet transform. The realisation of the lifting wavelet consists of three steps as follows [56, 57] :

(1) Split: this step generates a simple wavelet referred to as the lazy wavelet which divided the input signal into even and odd sequences.

$$split(x_i) = (s_i^0, d_i^0). \tag{3.28}$$

(2) Predict: in this step, the event sequence is used to predict the odd sequence based on the correlation in the input signal. The output of this step is the same as the high-pass output of DWT filter.

$$d_i^n = d_i^{n-1} - P(s_i^{n-1}). \tag{3.29}$$

(3) Update: here, the modified odd sequence is added to the even sequence to form the updated event sequence. The output of this step represents the low-pass output dwt filter.

$$\boldsymbol{s_i^n = s_i^{n-1} - U(d_i^n)}. \tag{3.30}$$

An important feature of the lifting scheme is that it can immediately make the inverse transform, also it offers the suitable context to construct the integer-to-integer DWT filters. To do this, quantizers are placed promptly after computing the predicting and the updating steps, also before the adjustment of the even and odd sequence values [124, 125]. For more understanding, all these steps are summarised in Figure (3.7).

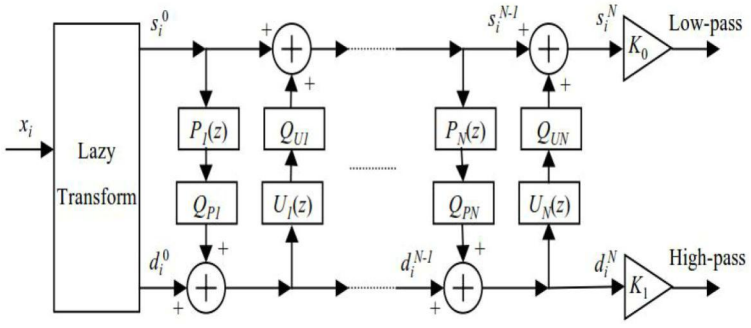

Figure 3.7: General Flowchart of a forward integer lifting wavelet transform.

In figure 3.7, $P(z), U(z)$ and Q represent the predicting step, the updating step and the quantizer, respectively.

3.4 DCT fusion methods

In literature, several image fusion methods based on DCT domain have been presented. In this part, we will focus on some of theme [47, 49, 101, 126, 127].

3.4.1 Contrast-based technique

Due to the sensitivity of the human visual system to the local contrast, the fusion criterion is based on the contrast measure in DCT domain [47]. First, the input images are divided into 8×8 non-overlapping blocks, then the 2-D DCT transformation is calculated for any block as in Equation (3.1). Figure (3.8) represents the DCT output block. Then the DCT coefficients are grouped into 15 $(k = i + j - 1)$ different frequency band.

$$D = \begin{bmatrix} d_{00} & d_{01} & d_{02} & d_{03} & d_{04} & d_{05} & d_{06} & d_{07} \\ d_{10} & d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ d_{20} & d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} & d_{27} \\ d_{30} & d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} & d_{37} \\ d_{40} & d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} & d_{47} \\ d_{50} & d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} & d_{57} \\ d_{60} & d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} & d_{67} \\ d_{70} & d_{71} & d_{72} & d_{73} & d_{74} & d_{75} & d_{76} & d_{77} \end{bmatrix}$$

(E_1 and E_4 mark the first and fourth diagonal bands)

Figure 3.8: 2-D DCT output block.

The contrast at each coefficient is defined by:

$$C_{u,v} = \frac{d_{u,v}}{\sum_{k=0}^{n-1} E_k}, \tag{3.31}$$

Where $d(u, v)$ is the DCT coefficient, E_k is the average amplitude in a frequency band is computed as:

$$E_k = \frac{\sum_{u+v=k} |d_{u,v}|}{N}, \tag{3.32}$$

And

$$N = \begin{cases} k + 1; & k < 8, \\ 14 - k + 1; & k \geq 8. \end{cases} \tag{3.33}$$

In order to obtain the fused DCT coefficients, the authors in [47] compare the contrast of the similar DCT coefficients of all source images, then choose the DCT coefficient with the largest contrast value as the DCT coefficient of the fused image.

3.4.2 Higher valued AC coefficients

The fusion process of the input images has been based on higher valued AC coefficients calculated in the DCT domain [49]. The fusion process can be given by the following steps:

1. Divide the source images A and B into 8×8 blocks (JPEG Standard), (A_n, B_n) where n represents the n^{th} block.
2. Perform the 2-D DCT transformation on each block (D_n^A, D_n^B).
3. Extract the AC coefficients $AC_{n[1\ldots63]}^A, AC_{n[1\ldots63\}}^B$ from the DCT representation of each block (D_n^A, D_n^B).
4. Initialise $(C_n^A = C_n^B = 0)$ where C_n^A and C_n^B represent the number of maximum valued AC coefficients found in the n^{th} block of the input images (D_n^A, D_n^B).
5. Compare each AC coefficient of D_n^A with the corresponding AC coefficient of D_n^B and increment C_n^A / C_n^B whose AC coefficient value is higher.
6. select the block with the majority AC coefficient value is higher as the fused DC block.

$$D_n^F = \begin{cases} D_n^A; & D_n^A \geq D_n^B, \\ D_n^B; & D_n^A < D_n^B. \end{cases} \tag{3.34}$$

7. Subject the fused DCT blocks D^F to consistency verification (CV).
8. Perform the inverse 2-D DCT transformation to reconstruct the fused image F.

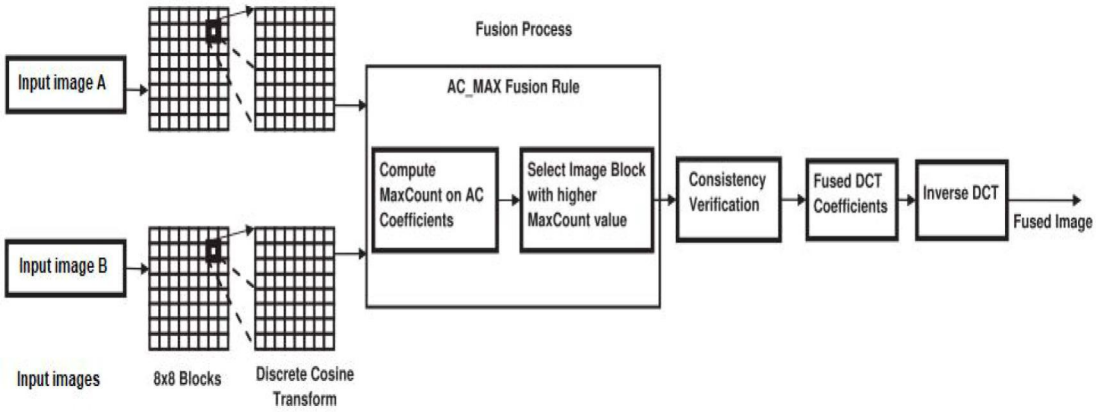

Figure 3.9: General structure of the AC_Max fusion scheme.

The Figure (3.9) the general framework of the higher valued AC coefficients image fusion scheme (AC_Max).

3.4.3 Spatial frequency-based technique

It is well known that the spatial frequency can be given in the spatial domain by Equation 2.19. In this part, a general image fusion technique based on the spatial frequency (SF) in the DCT domain is proposed by Cao et al [101]. The mathematical expression of the SF of an 8×8 block image is given by:

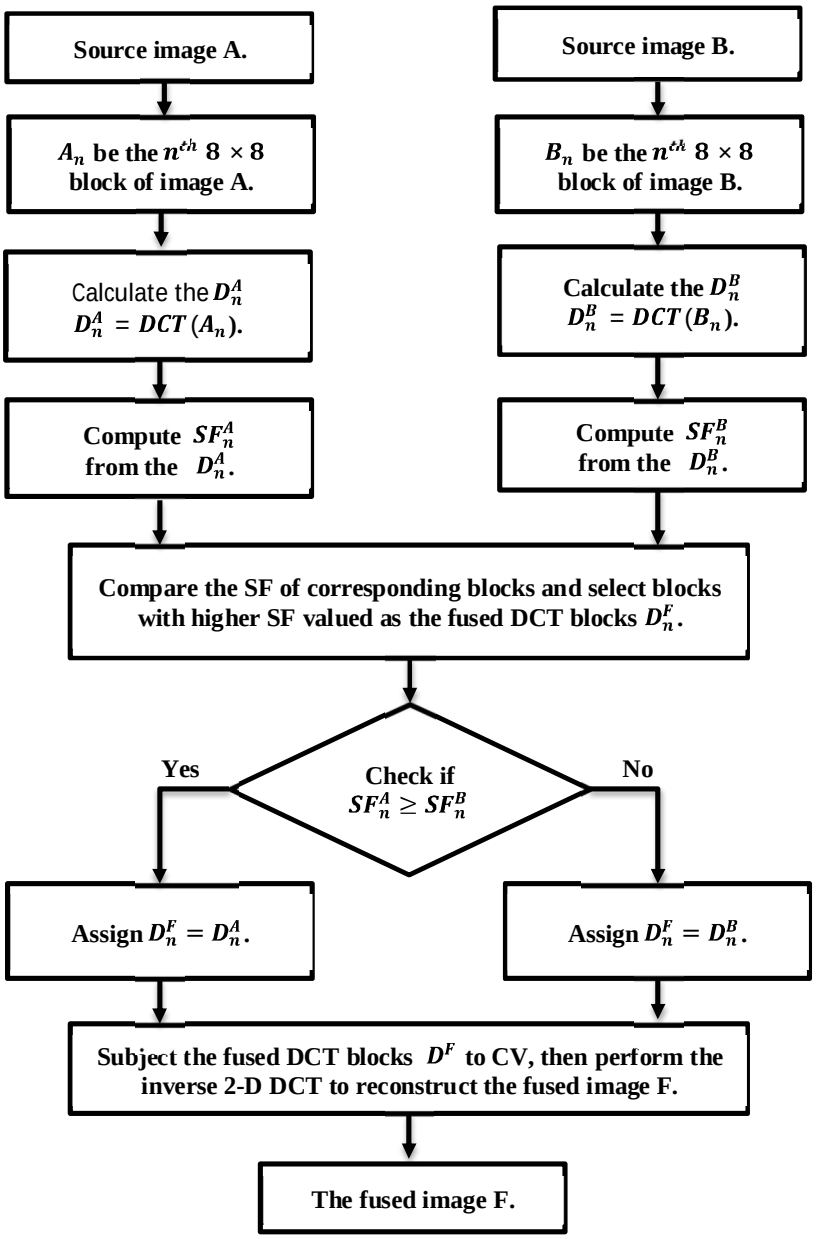

Figure 3.10: Flowchart for SF-based fusion technique.

$$SF^2 = RF^2 + CF^2 = \frac{1}{8 \times 8}\sum_{u=0}^{7}\sum_{v=0}^{7}\left[\left(D(u,u) + D(v,v)\right) \times F^2(u,v)\right]$$

$$= \frac{1}{8 \times 8}\sum_{u=0}^{7}\sum_{v=0}^{7}[E(u,v) \times F^2(u,v)], \quad (3.35)$$

where $E = D(u,u) + D(v,v)$, $D = B \times B^t$ and B is the DCT representation of b where [101]

$$b = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & 0.1522 & 0.5858 & 1.2346 & 2.0000 & 2.7654 & 3.4142 & 3.8478 \\ 0.1522 & 0.3045 & 0.7380 & 1.3869 & 2.1522 & 2.9176 & 3.5665 & 4.0000 \\ 0.5858 & 0.7380 & 1.1716 & 1.8204 & 2.5858 & 3.3512 & 4.0000 & 4.4335 \\ 1.2346 & 1.3869 & 1.8204 & 2.4693 & 3.2346 & 4.0000 & 4.6488 & 5.0824 \\ 2.0000 & 2.1522 & 2.5858 & 3.2346 & 4.0000 & 4.7654 & 5.4142 & 5.8478 \\ 2.7654 & 2.9176 & 3.3512 & 4.0000 & 4.7654 & 5.5307 & 6.1796 & 6.6131 \\ 3.4142 & 3.5665 & 4.0000 & 4.6488 & 5.4142 & 6.1796 & 6.8284 & 7.2620 \\ 3.8478 & 4.0000 & 4.4335 & 5.8478 & 5.8478 & 6.6131 & 7.2620 & 7.6955 \end{bmatrix}$$

In summary, we can note that the SF of a block of pixels can be exactly computed by the weighted sum of squares of alternative current (AC) coefficients in the DCT domain [101]. Figure (3.10) summarises the fusion algorithm.

3.5 Wavelet fusion methods

The DWT-based methods are suitable for performing image fusion application. They are multiscale procedures entirely appropriate to manage the various image resolutions. The image decomposition using DWT allows preserving the image information. In this section, some image fusion techniques are presented [38, 52, 128].

3.5.1 Region-based technique

A very simple method for multi sensor image fusion in DWT domain was developed by Hui et al in [38]. In this method, the authors have chosen the larger absolute coefficients in DWT domain to form the fused image because it corresponds to the sharper brightness changes. Consequently, this contributes to define the salient features in the local area.

After the decomposition of the input image using DWT, the fusion process takes place in all detail and approximation sub-bands. The coefficients are selected by choosing the maximum absolute value inside the window as an activity measure relevant with the centre pixel.

A decision map equal in size with the DWT transform is then created to save the selection results. This map is subject to CV. Specifically in the transform domain, if the centre pixel value from the image B while the majority of surrounding pixel values come from image A, the centre pixel value is switched to that of image A, and vice versa. Finally, the fused image is obtained by performing the inverse DWT on the selected coefficients [38].

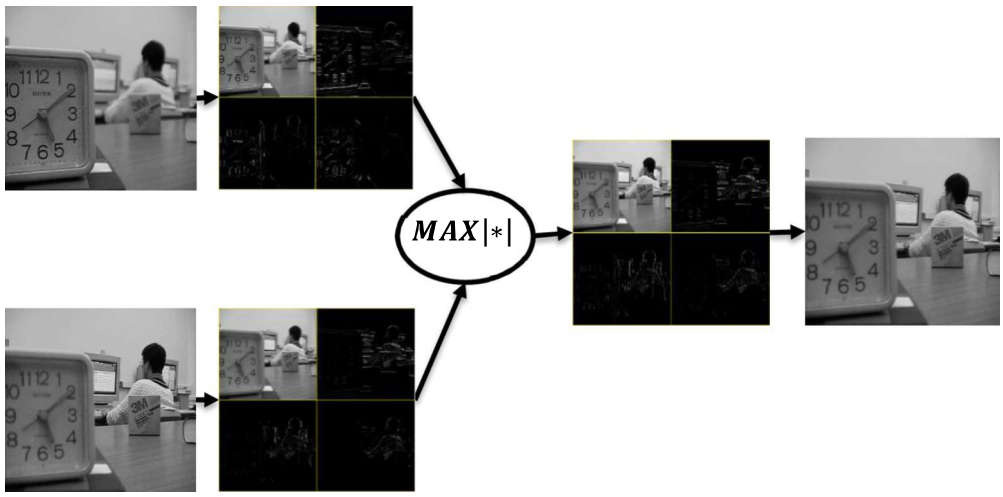

Figure 3.11: The block diagram of the image fusion scheme [38].

3.5.2 Contrast-based technique

Generally, contrast-based image fusion methods produce fused images with better visual quality, because they consider the properties of the human visual system which is more

sensitive to local luminance contrast. The fusion methods in [128, 129] are contrast based image fusion in the DWT domain.

The contrast of an image in the DWT domain at a given decomposition level can be defined as the absolute value of the ratio of the detail coefficients to the approximate coefficients [128]. A series of directive contrast is given as:

$$\begin{cases} C_k^H = \dfrac{LH_k}{LL_k} \\ C_k^V = \dfrac{HL_k}{LL_k} \\ C_k^D = \dfrac{HH_k}{LL_k} \end{cases}, \tag{3.36}$$

Where C_k^H, C_k^V and C_k^D are the horizontal, vertical and diagonal contrast in the k-level, respectively.

The fusion process can be given by the following steps:

1. Perform the DWT decomposition for any source image with a certain level (k).
2. For each level, a series of directive contrast is calculated using Equation (3.36).
3. The high frequencies of the fused image (LH_k^F, HL_k^F and HH_k^F) can be selected by comparing the corresponding frequencies. As an example, suppose we have two source images A and B, and F denotes to the fused image, we have:

$$LH_k^F(i,j) = \begin{cases} LH_k^A(i,j); \left|C_{A,k}^H(i,j)\right| \geq \left|C_{B,k}^H(i,j)\right| \\ LH_k^B(i,j); \left|C_{A,k}^H(i,j)\right| < \left|C_{B,k}^H(i,j)\right| \end{cases}. \tag{3.37}$$

$$HL_k^F(i,j) = \begin{cases} HL_k^A(i,j); \left|C_{A,k}^V(i,j)\right| \geq \left|C_{B,k}^V(i,j)\right| \\ HL_k^B(i,j); \left|C_{A,k}^V(i,j)\right| < \left|C_{B,k}^V(i,j)\right| \end{cases}. \tag{3.38}$$

$$HH_k^F(i,j) = \begin{cases} HH_k^A(i,j); \left|C_{A,k}^D(i,j)\right| \geq \left|C_{B,k}^D(i,j)\right| \\ HH_k^B(i,j); \left|C_{A,k}^D(i,j)\right| < \left|C_{B,k}^D(i,j)\right| \end{cases}. \tag{3.39}$$

4. The low frequencies of the fused image (LL_k^F) can be selected by taking the average of the corresponding frequencies.

$$LL_k^F(i,j) = \frac{LL_k^A(i,j) + LL_k^B(i,j)}{2}. \quad (3.40)$$

5. Finally, the fused image can be reconstructed by performing the inverse DWT.

3.5.3 DWT-based PCA technique

In the previous techniques (sections: 3.5.1 and 3.5.2), the fusion rule is based on the selection of suitable coefficients (with ignoring the unselected coefficients). In contrast, the fusion procedure of the algorithm proposed in [52] occurs in the DWT domain where all coefficients are contributing with specified weights to construct the fused image. The principal components for the DWT coefficients are evaluated, then the weights are constituted by taking the averaging of these principal components.

For simplification, we consider that we have two images A and B. Firstly, each input image is decomposed into approximation (LL) and detail (LH, HL, HH) coefficients. The two approximation coefficients LL^A and LL^B are considered as input for the principal component analysis. As a result, the largest principal components of LL coefficients are given as p_1 and p_2. Likewise, detail coefficients are processed to assess the principal components.

Let x_i^A and x_i^B are a column vectors expression of LL^A and LL^B respectively.

$$x_i^A = LL_k^A = \begin{bmatrix} x_1^A \\ x_2^A \\ \vdots \\ x_n^A \end{bmatrix}, \quad (3.41)$$

$$x_i^B = LL_k^B = \begin{bmatrix} x_1^B \\ x_2^B \\ \vdots \\ x_n^B \end{bmatrix}, \quad (3.42)$$

Where $i = 1, 2, \dots, n$; n is the number of the vector elements (coefficients), and k is the level of decompositions.

The covariance between x_i^A and x_i^B is given by:

$$cov(x_i^A, x_i^B) = E\left[(x_i^A - \mu_{x_i}^A)(x_i^B - \mu_{x_i}^B)\right], \quad (3.43)$$

where $\mu_{x_i}^A$ and $\mu_{x_i}^B$ represent means of x_i^A and x_i^B.

$$\mu_{x_i}^{A} = \frac{1}{n}\sum x_i^{A}. \tag{3.44}$$

$$\mu_{x_i}^{B} = \frac{1}{n}\sum x_i^{B}. \tag{3.45}$$

The eigenvalues matrix D and the corresponding eigenvectors matrix V are calculated. The normalized components p_1 and p_2 are computed as below:

$$p_1\left(LL_k^{A,B}\right) = \begin{cases} \dfrac{V(1,1)}{V(1,1)+V(2,1)}, & D(1,1) \geq D(2,2) \\ \dfrac{V(1,2)}{V(1,2)+V(2,2)}, & D(1,1) < D(2,2) \end{cases}. \tag{3.46}$$

$$p_2\left(LL_k^{A,B}\right) = \begin{cases} \dfrac{V(2,1)}{V(1,1)+V(2,1)}, & D(1,1) \geq D(2,2) \\ \dfrac{V(2,2)}{V(1,2)+V(2,2)}, & D(1,1) < D(2,2) \end{cases}. \tag{3.47}$$

Likewise, p_1 and p_2 for the detailed coefficients are too assessed. The averaging of all these normalized components (p_1 and p_2) give us the weights $p_{1(av)}$ and $p_{2(av)}$ for the fusion rule.

$$p_{1(av)} = \frac{p_1\left(LL_k^{A,B}\right) + p_1\left(LH_k^{A,B}\right) + p_1\left(HL_k^{A,B}\right) + p_1\left(HH_k^{A,B}\right)}{K}, \tag{3.48}$$

$$p_{2(av)} = \frac{p_2\left(LL_k^{A,B}\right) + p_2\left(LH_k^{A,B}\right) + p_2\left(HL_k^{A,B}\right) + p_2\left(HH_k^{A,B}\right)}{K}, \tag{3.49}$$

Where K is the number of the decomposed sub-bands.

Finally, the fused image F can be given by:

$$F = p_{1(av)} \times A + p_{2(av)} \times B. \tag{3.50}$$

3.6 Conclusion

In this chapter, we present a brief literature review of some transform domain image fusion techniques. The most commonly used methods are those based on DCT transform and DWT transform. This is due to the fact that significant features of images in transform domain

are more clearly represented than the spatial domain. In the next chapter, the proposed image fusion algorithm will be presented.

Chapter 4:

The proposed image fusion algorithm

4.1 Introduction

In this section, we propose a novel image fusion method for visual sensor networks based on the integer lifting wavelet transform and the discrete cosine transform [54]. To improve the fused image quality, the fusion process is done in two phases. The source images are decomposed by 2-D ILWT into approximation and detail sub-bands. In the first fusion process, the approximation sub-bands of all input images are fused in the DCT domain by selecting the variance as an activity level measure. Then, these sub-bands are updated by their fused ones. The second fusion process is performed at pixel-level. This will exploit the correlation between the wavelet coefficients across all the levels to calculate the weight of every source coefficient giving more weightage to the coefficient with sharper neighborhood.

4.2 Proposed ILWT and DCT image fusion method

For simplicity, we assume that we want to fuse two input images A and B of the same size in which they are formerly aligned to obtain a fused image F. First, we perform the ILWT on both input images A and B to obtain the four bands (LL_1, HL_1, LH_1, HH_1) in the first level decomposition. The HL_1, LH_1 and HH_1 bands represent the detail information in horizontal, vertical and diagonal directions respectively. Whereas, LL_1 is a scaled down copy of the input image itself (approximation).

Always, the approximation (LL_1) is used to perform the decomposition process. At the third level, we obtain ten sub-bands for any input image, where (HL_1, HL_2, HL_3), (LH_1, LH_2, LH_3) and (HH_1, HH_2, HH_3) represent horizontal, vertical and diagonal sub-bands respectively. LL_3 represents the approximation sub-band, which contains the average information of the input image.

4.2.1 Block based fusion rule

As mentioned, the focused regions are which are more informative and give more clarity. As known, the variance value is widely used as a contrast criterion in image processing fields [11]. Similarly, we can use the variance as the activity level criterion to identify the focused blocks in an approximation sub-band of an image [60]. Firstly, we divide the approximation sub-bands (LL_3^A, LL_3^B) of the input images (A and B) into 8×8 blocks and we compute the DCT coefficients for every block. From Equation (3.8), we can easily compute the variance value of the comparable blocks from input images. Then, the block with the greatest value in

variance is absorbed to the fused approximation sub-band (LL_3^F). Finally, the fused approximation sub-band is subjected to consistency verification (CV) [38] with 3×3 neighbourhood window. The main objective of the CV is to verifying if the centre block comes from image A and the majority of surrounding blocks are coming from image B, the centre block is switched to the corresponding block in source image B.

4.2.2 Pixel significance fusion rule

Some image fusion algorithms use weighted average as a fusion criterion [130-132], in which all pixels of the clear and blurred regions will contribute in calculating these weights. This caused undesirable effects like reduction in image contrast. To avoid this problem, we replace the approximation sub-bands of all input images by their fused once, and then we calculate the optimum weighted average. To ensure that, only clear regions (blocks) contribute in the fused image.

In wavelet decompositions, at the first decomposition level, the signal is passed through the high pass and lowpass filters, followed by down sampling of 2. The output of the high pass (low pass) filter has the half resolution and so for other levels. As a result, any coefficient in the third level is correlated with 2×2 windows of children coefficients in the second level, which is also correlated with 4×4 window of grand children in the first level, whereas coefficients in level 2 has only children. Also, level 3 coefficients don't have any children or grandchildren, so the weight can be calculated by taking the average energy of a 3×3 neighbourhood window [130, 132].

To simplify, we denote the sub-bands by SB_L, in which SB represents the detail sub-bands (HL, LH, HH) and L indicates the decomposition level. The weight factor of any detail sub-band coefficient (WT_{SB_L}) can be calculated as follows:

$$WT_{SB3}(i,j) = |SB_3(i,j)| + \sum_{k=0}^{1}\sum_{l=0}^{1}|SB_2(2i-k, 2j-l)| + \cdots$$

$$\ldots + \sum_{m=0}^{3}\sum_{n=0}^{3}|SB_1(4i-m, 4j-n)|. \quad (4.1)$$

$$WT_{SB2}(i,j) = |SB_2(i,j)| + \sum_{k=0}^{1}\sum_{l=0}^{1}|SB_1(2i-k, 2j-l)|. \quad (4.2)$$

$$WT_{SB1}(i,j) = \frac{1}{3^2} \times \sum_{m=-1}^{1} \sum_{n=-1}^{1} [SB_1(i-m, j-n)]^2 . \quad (4.3)$$

Here, each coefficient in the approximation sub-band (LL_3) is more correlated with the corresponding coefficient in detail sub-bands (HL_3, LH_3, HH_3), and the weight factor is given by equation (4.4):

$$WT_{LL3}(i,j) = WT_{HL3}(i,j) + WT_{LH3}(i,j) + WT_{HH3}(i,j). \quad (4.4)$$

These weights (pixel significance) are calculated for all sub-bands in both input images A and B. If WT_A and WT_B are the weights of the corresponding pixels $A(i,j)$ and $B(i,j)$ from the input images A and B respectively. The fused coefficient $f(i,j)$ can be computed as the optimum weighted average of both weights (WT_A, WT_B), which is given by equation (4.5):

$$f(i,j) = \frac{WT_A \times A(i,j) + WT_B \times B(i,j)}{WT_A + WT_B}. \quad (4.5)$$

Finally, the fused image can be constructed by performing the backward ILWT on the fused coefficients.

Figure (4.1) shows the general framework of the proposed image fusion method, that can be summarized in the following steps. For simplicity, we assume that we want to fuse two input images A and B of the same size (512×512) in which they are formerly aligned, in order to obtain the fused image F:

Step 01: Perform the 3-level 2-D ILWT of the input images A and B. This produces the approximation sub-band matrix (LL) and nine detail sub-bands $(HL, LH,$ and $HH)$ for any input image.

$(LL_3^I, HL_k^I, LH_k^I, HH_k^I; I = A, B; k = 1,2,3)$.

First fusion phase:

Step 02: Divide the approximation sub-band into 8×8 blocks.

$(LL_{3,b}^A, LL_{3,b}^B;\ b(number\ of\ blocks) = 1, \dots, 64)$

and compute the DCT coefficients for every block.

$X_b^I = dct\left(LL_{3,b}^I\right); I = A, B; b(number\ of\ blocks) = 1, \dots, 64.$

Step 03: From Equation (3.8), compute the variance value (σ^2) of the comparable blocks (X_b^A, X_b^B). The block with the greatest value in variance is fed to the fused approximation sub-band.

$$LL_{3,b}^F = \begin{cases} LL_{3,b}^A, & \sigma^2(X_b^A) \geq \sigma^2(X_b^B), \\ LL_{3,b}^B, & \sigma^2(X_b^A) < \sigma^2(X_b^B). \end{cases}$$

The fused approximation sub-band (LL_3^F) is subjected to consistency verification (CV) [38].

Second fusion phase:

Step 04: update the approximation sub-bands ($LL_3^A\ and\ LL_3^B$) with the fused one LL_3^F.

$LL_3^A = LL_3^B = LL_3^F.$

Step 05: calculate the weight factor of any sub-band coefficient (WT_{SB_L}) from Equation (4.1) to Equation (4.3).

Step 06: calculate the final fused coefficients $f(i,j)$ from Equation (4.4). Finally, the fused image (F) can be constructed by performing the backward 2-D ILWT on the fused coefficients.

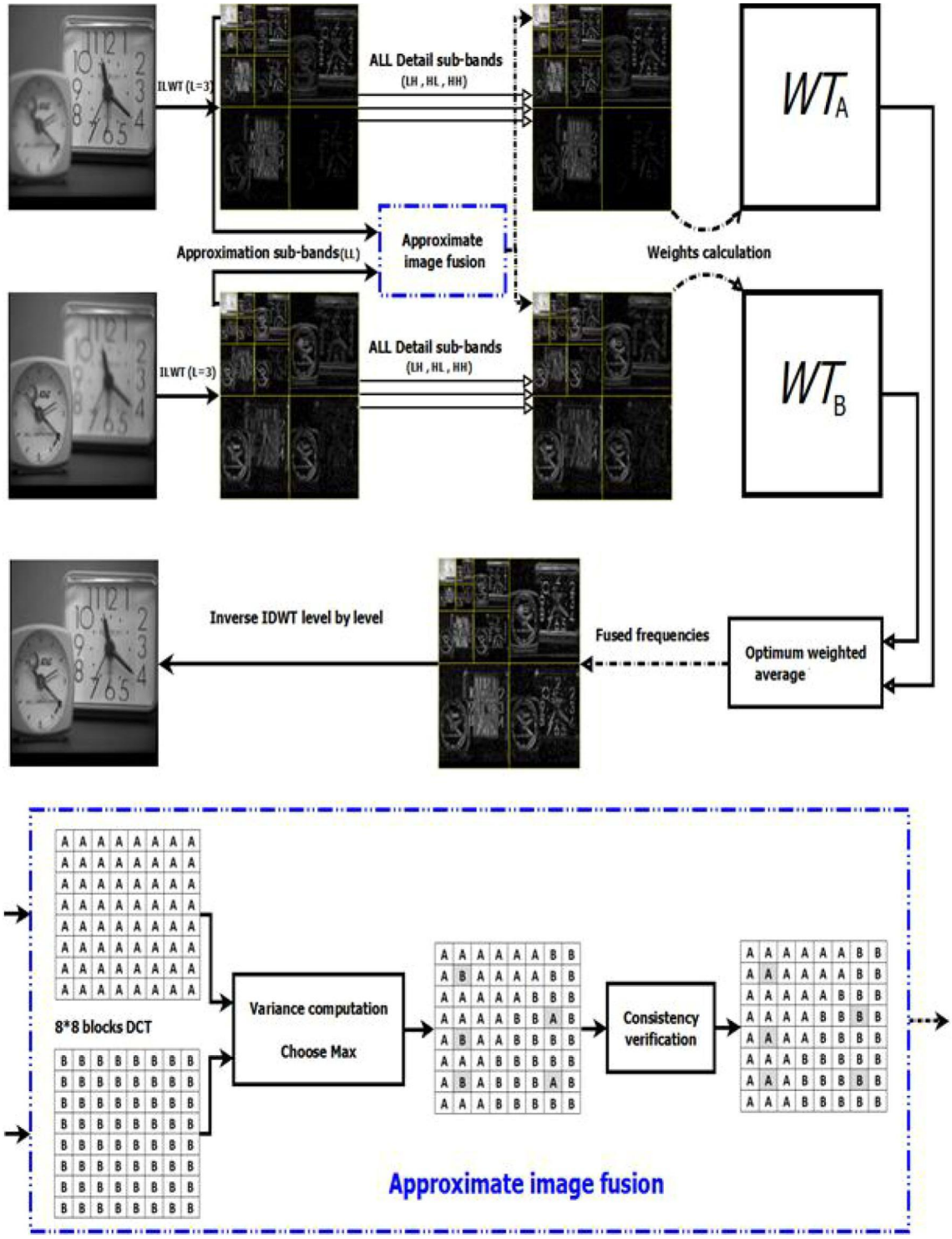

Figure 4.1: General Framework of the proposed image fusion method.

4.3 Conclusion

In this chapter, we have presented a brief description of the proposed image fusion algorithm. In this proposed algorithm the fusion process is done in two phases to ensure a high quality of the fused images. The first phase is conducted in the DCT domain by taking the variance measure of the approximate coefficients as a contrast criterion to select the focused regions. The second phase is performed in the ILWT domain by taking the weighted average of the detail coefficients to establish an edge map which is more appropriate to the chosen regions.

In order to evaluate the design system, the proposed method will be tested examined using different image sets and compared with some state-of-the-art methods quantitatively and qualitatively, in the next chapter.

Chapter 5:
Simulation results and analysis

5.1 Introduction

Image fusion technique produces some distortions in the fused image. Therefore, image quality evaluation is an important issue. The quality of the fused images can be evaluated in terms of qualitative (subjective) measurement and quantitative (objective) measurement. Qualitative evaluation is time-consuming and cannot objectively discriminate image qualities, therefore the fusion images have also been evaluated through the quantitative methods. In the following subsections, the evaluation methods are firstly introduced and then the evaluation results are discussed.

5.2 Experimental settings

In this part, to verify the performance of the fusion algorithms, we use two image data sets, which are available in [133, 134].

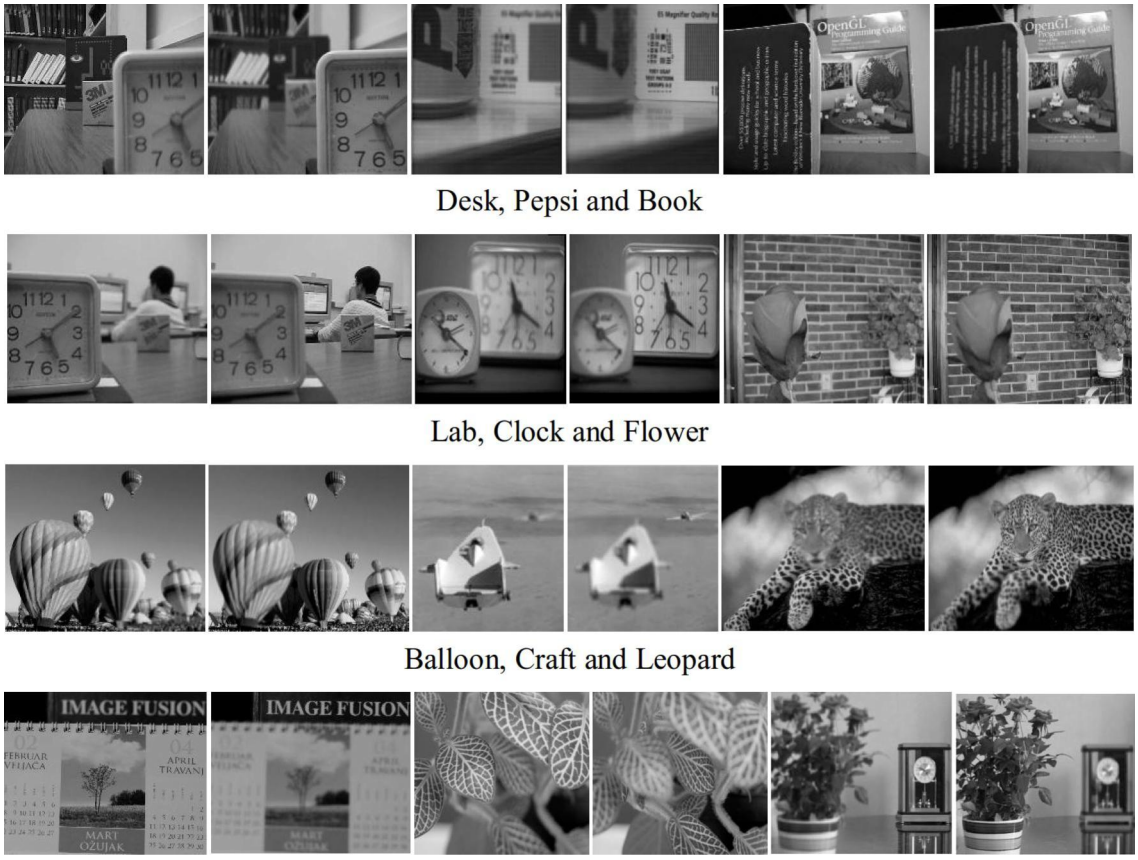

Figure 5.1: First data set, multi-focus images used for experiments.

The first data set contains a twelve pairs non-referenced multi-focus images (Figure5.1), with 256-level grey scale, with different characteristics, different focused zone

and different sizes. The right and the left focused images are put in the right and the left position respectively in each pair of ‘Desk’, ‘Pepsi’, ‘Book’, ‘Lab’, ‘Clock’, ‘Balloon’, ‘Leopard’ and ‘Corner’. The in-focus images in pairs ‘Flower’, ‘Craft’, ‘Calendar’ and ‘leaf’ are put in the left, unlike the out-focus of the same images are put in right.

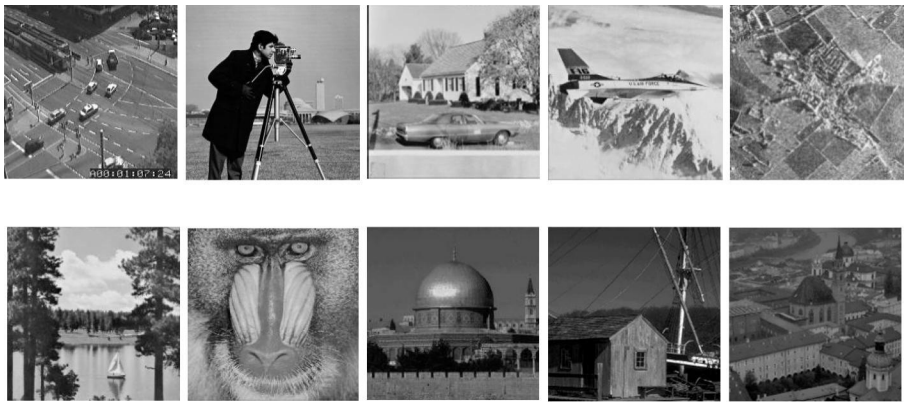

Figure 5.2: Second data set, images used for experiments.

The second data set contains ten pairs (with reference image) of artificially generated input images by filtering ten standard original images (Figure 5.2), with a 2-D Gaussian smoothing kernel with standard deviation (sigma = 5). The filter is centred at the left half and the right half of each image, to obtain a left blurred image and right blurred image respectively.

To demonstrate the superiority of the proposed algorithm, it has been compared with some existing image fusion algorithms, such as DWT with DBSS (2,2) [38], shift invariant DWT (SIDWT with Haar basis) [40], PCA based algorithm [135], LAP based algorithm [36], and recently proposed algorithms, such a bilateral gradient-based sharpness criterion (BGBSC) [51], DCT+Ac_Max [49], DWT+Pca+av [52], and saliency detection based algorithm (SC) [53].

In this chapter, all algorithms (the proposed and the compared) are implemented in MATLAB environment (R2016a) and all the experiments are carried out on a computer with intel processor i5-2430M (2.40GHz) and 8192MB (665.1MHz) memory. In addition, DWT, SIDWT, PCA based and LAP based image fusion toolbox can be downloaded from [136]. Bilateral gradient-based sharpness criterion source code found in the website [137]. DWT+Pca+av and saliency detection based algorithm can be downloaded from the websites [138, 139] respectively.

5.3 Results and analysis

In this subsection, a quantitative and qualitative evaluation has been used to demonstrate the performance and effectiveness of the proposed method.

5.3.1 Qualitative evaluation

Here, we demonstrate the visual quality (qualitative) of the fused images generated by our method compared with methods cited in section 5.2.

5.3.1.1 First data set

For the lack of memory space, we choose three images from the first data set of the multi focus images, as an example; which are the 'Clock', 'Lab' and 'Pepsi' images.

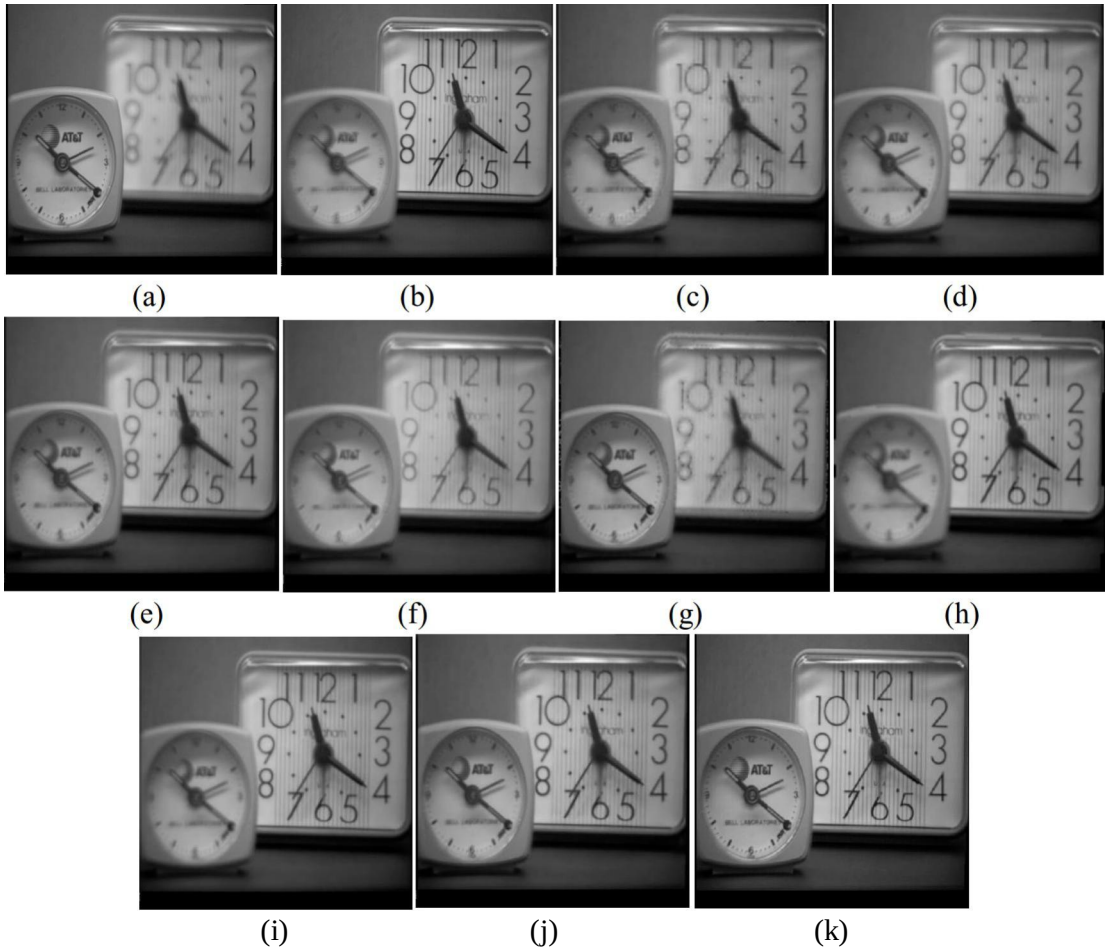

Figure 5.3: Fusion results of 'Clock' image. (a),(b) Source images. fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method.

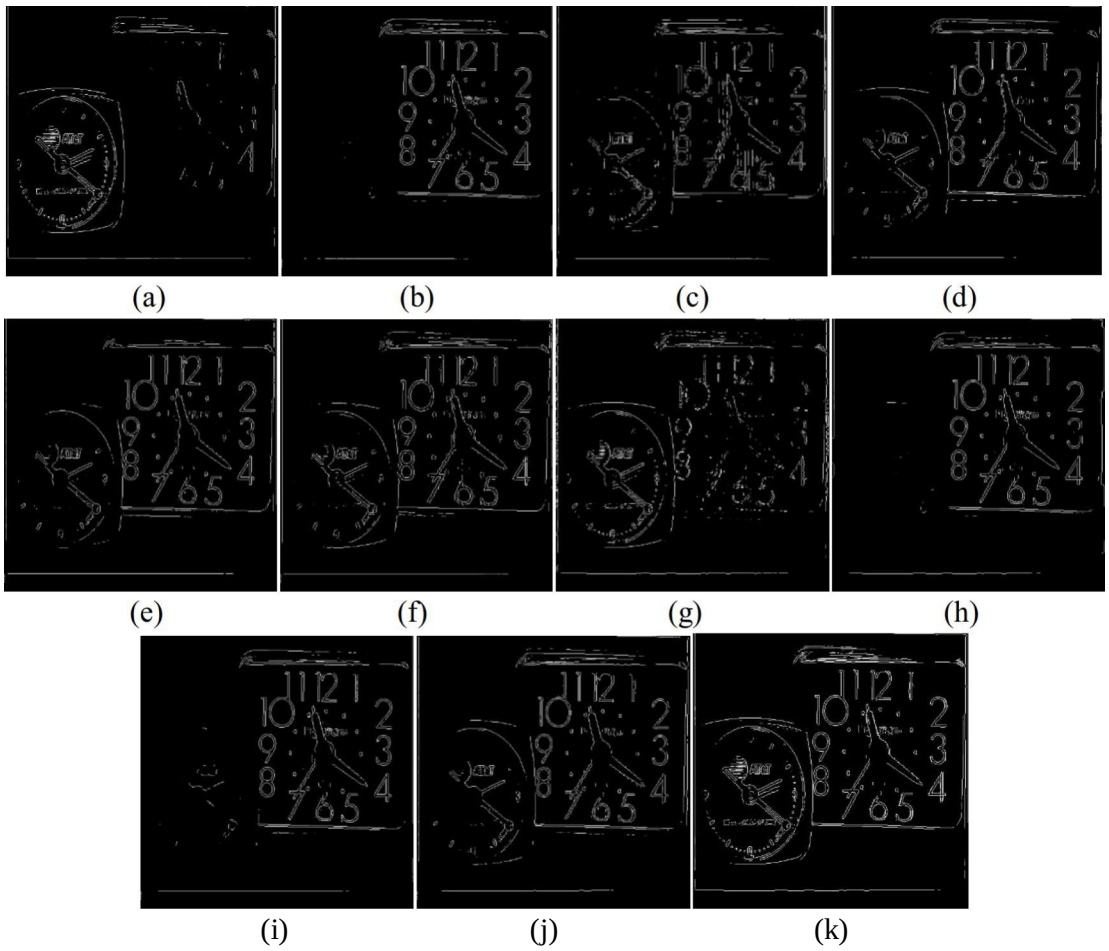

Figure 5.4: Edge map of fusion results of 'Clock' image. (a),(b) Source images. Fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method.

As shown, Figure 5.3 presents the visual display of the 'Clock' fused image. The source image 5.3(a) is focused on the left part in which the small clock is completely focused while the big clock is unfocused. On the contrary, the source image represented by Figure 5.3(b) is focused on the right part in which the small clock is completely unfocused while the big clock is focused. It is clear that the fused image with DWT based method (Figure 5.3(c)), SIDWT based method (Figure 5.3(d)), PCA based method (Figure 5.3(e)) and DWT+Pca+av method (Figure 5.3(i)) suffer from blurring effects, which have a negative impact on the fused image quality. LAP based method results image with undesirable extra illumination and blurring effects as shown in Figure 5.3(f). BGBSC method leads to artifacts at the boundaries between regions (Figure 5.3(g)). DCT+Ac_max method gives blocking artifacts in the fused image (Figure 5.3(h)), since input images are divided into 8×8 blocks. The saliency detection-

based algorithm SC is little better than other comparable methods, but it suffers from a blurring (right side of image) and ringing effects (top right the big clock) (Figure 5.3(j)). For our reformulate this sentence. The fused image shows the better visual quality than other method (Figure 5.3(k)).

For careful observation, the illustrated images in Figure 5.4 present the edges of the given images in figure 5.3 using the Prewitt edge detector technique. One can note that the proposed method is able to produce a clean and precise edge map by extracting the principal edge features from the input images. Thereby this leads to the best interpretation of image contents (objects/features), that it is one of the main objectives in computer vision specifically in image fusion.

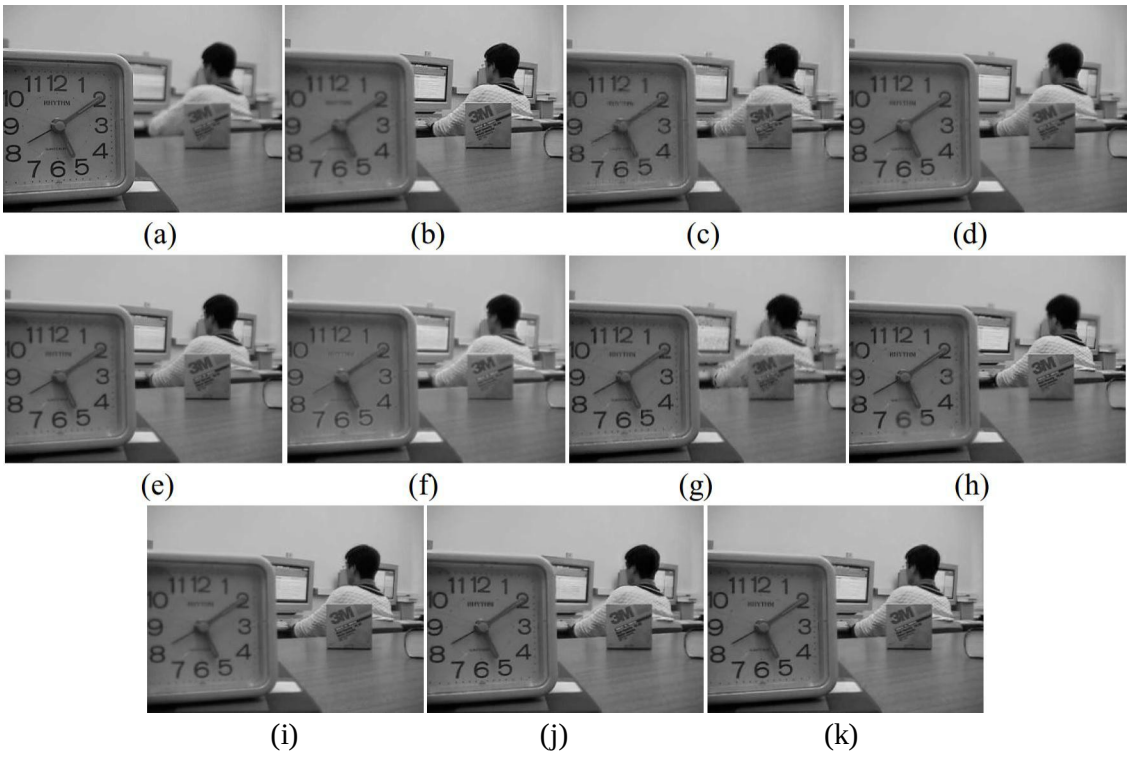

Figure 5.5: Fusion results of 'Lab' image. (a),(b) Source images. fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method.

Figure 5.5 Shows the fusion result of input images called 'Lab'. Figures 5.5(a) and(b) indicate the left and right focused images respectively. Both DWT and BGBSC methods give out a distorted image as illustrated in Figure 5.5(c) and (g) respectively. The same problem for DCT+Ac_max method that produces blocking artifacts (see numbers 3, 5, 6, 8, 9 and 10 inside the clock in Figure 5.5(h)). SIDWT, LAP, PCA and DWT+Pca+av methods always

give a blurred fused image as observed in Figure 5.5(d), (e), (f) and (i) respectively. The SC method provides a fused image with some ringing around the man head area and blurring inside the clock as illustrated in Figure 5.5(j) generally, better than the previous methods.

Whereas, our proposed method is the best, because it could well absorb the focused regions and the important features from the source images into the fused image, and keeps them in good consistency. Therefore, the fused images are produced with the best visual quality, sharper edges, clear numbers and without blurring and ringing as presented in Figure 5.5(k).

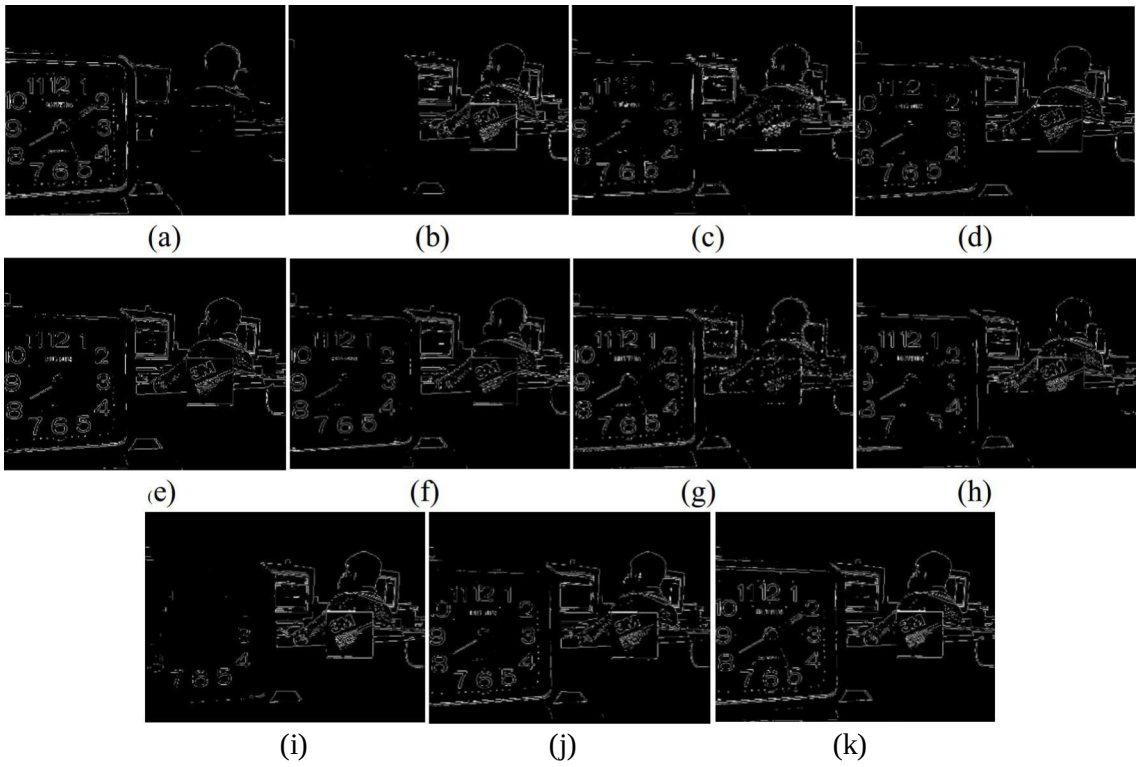

Figure 5.6: Edge map of fusion results of 'Lab' image. (a),(b) Source images. fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method.

The same observation as Figure 5.4. The edges of the fused 'Lab' image using different methods displayed in Figure 5.5 is presented here (Figure 5.6). The comparative methods couldn't properly produce fused images with clear edge map except for the SC method. We can observe that it is succeeded to an extent to present the edges as shown in Figure 5.6(j). It is absolutely clear that the proposed method possessed the preference (Figure 5.6(k)) compared with other methods results (Figure 5.6(c)-(j)).

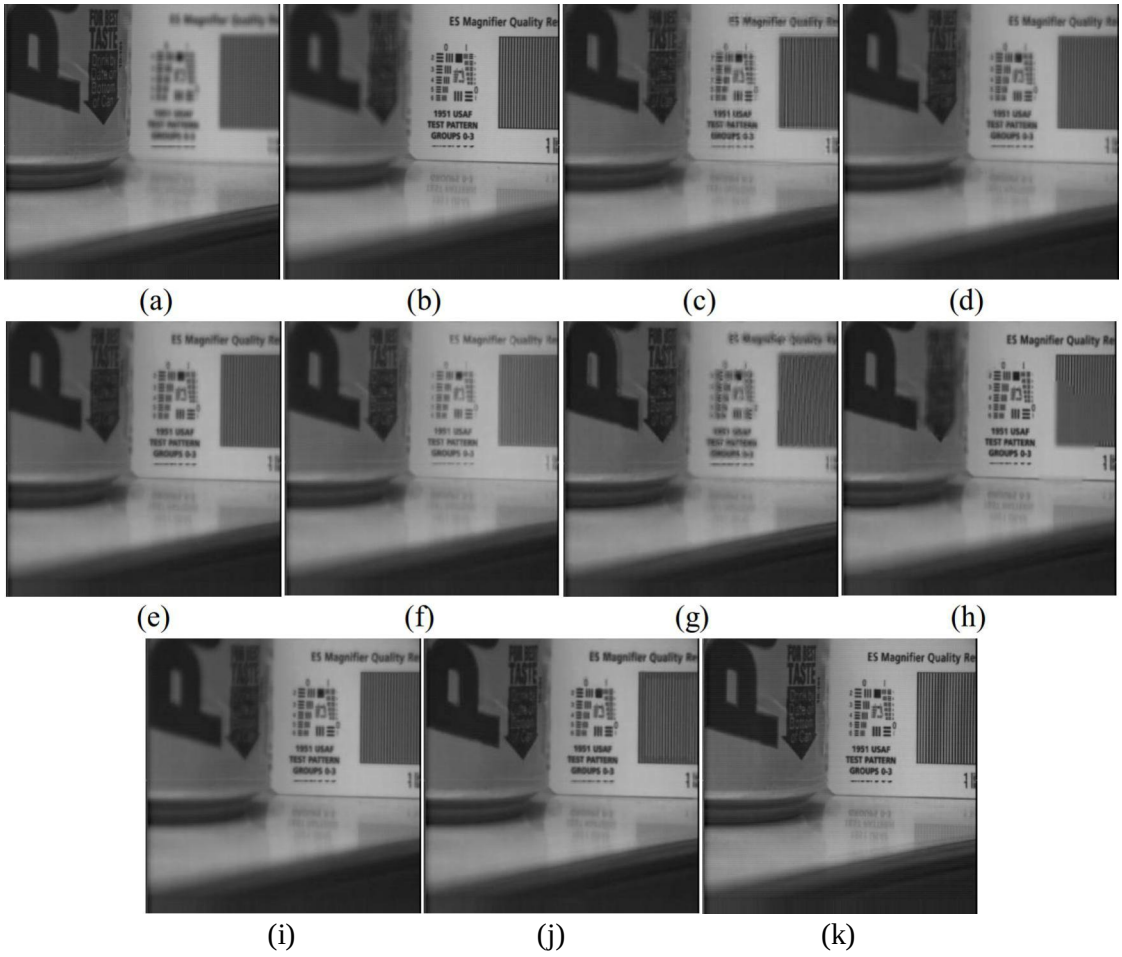

Figure 5.7: Fusion results of 'Pepsi' image. (a),(b) Source images. fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method.

As predicted, the proposed method produces a fused image as shown in Figure 5.7(j) with more information, more clarity text regions and with careless blurring. Unlike that, DWT, BGBSC, DCT+Ac_max and SC methods are not able to extract clear text regions (as seen in the top right of images) and the bar-code from the source images as shown in Figure 5.7(c), (g), (h), (j) respectively. The other compared methods, SIDWT, LAP, PCA and DWT+Pca+av (Figure 4.7(d), (e), (f), (i) respectively), as usual, suffer from blurring effects which are generally a big problem in image processing and especially in image fusion.

Images in Figure 5.8 fully confirm what we observe about the fused images illustrated in Figure 5.7. While the proposed method is the single method that it could present a complete and clear edge map as shown in Figure 5.8(k).

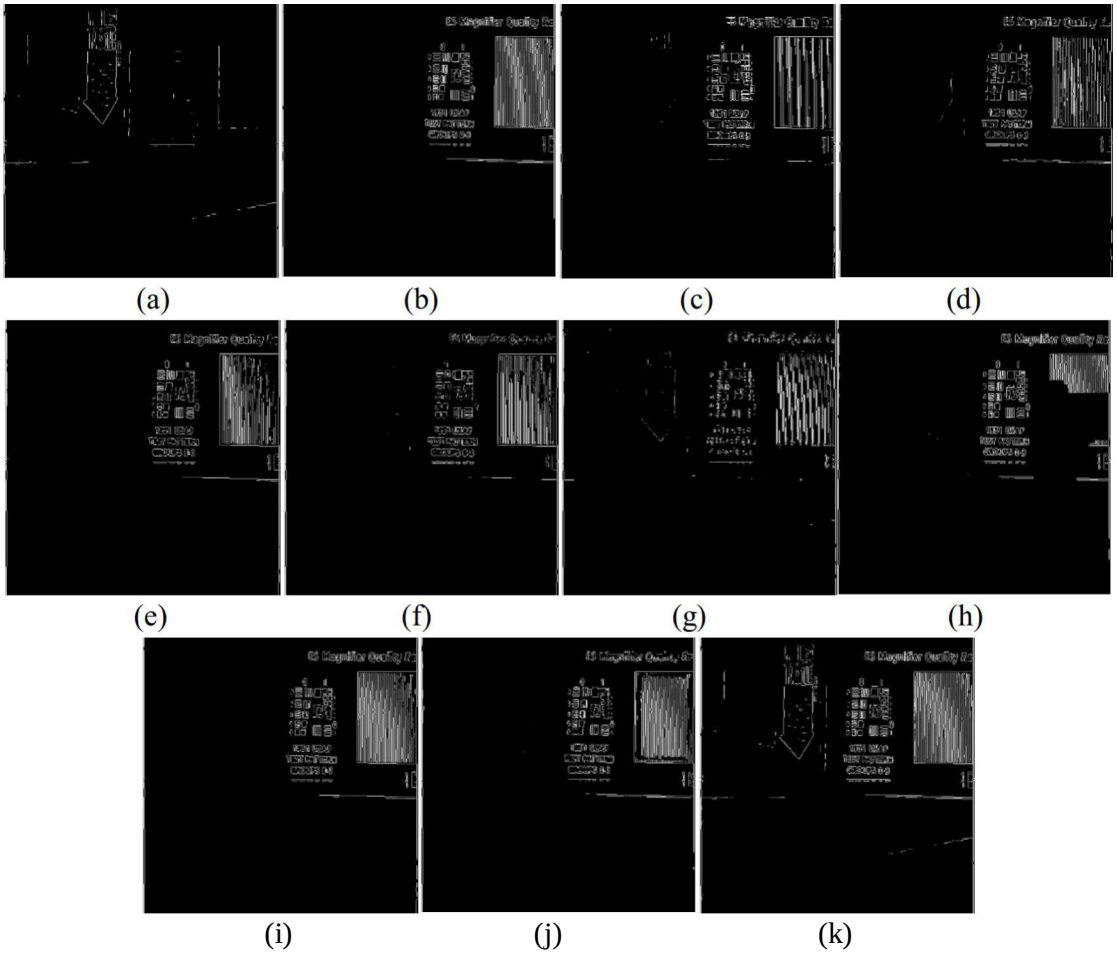

Figure 5.8: Edge map of fusion results of 'Pepsi' image. (a),(b) Source images. fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method.

5.3.1.2 Second data set

Now, we select two images from the second data set of the artificially blurred images, 'Cameraman' and 'Traffic' taken as truth images. As usual, in Figure 5.9 and Figure 5.10, images (a) and (b) are the source artificially left and right blurred images respectively. Images (c)-(k) are the fused images using DWT, SIDWT, PCA, LAP, BGBSC, DCT+Ac_Max, DWT+Pca+av, SC and the proposed method. While the image (l) represent the original (truth) image.

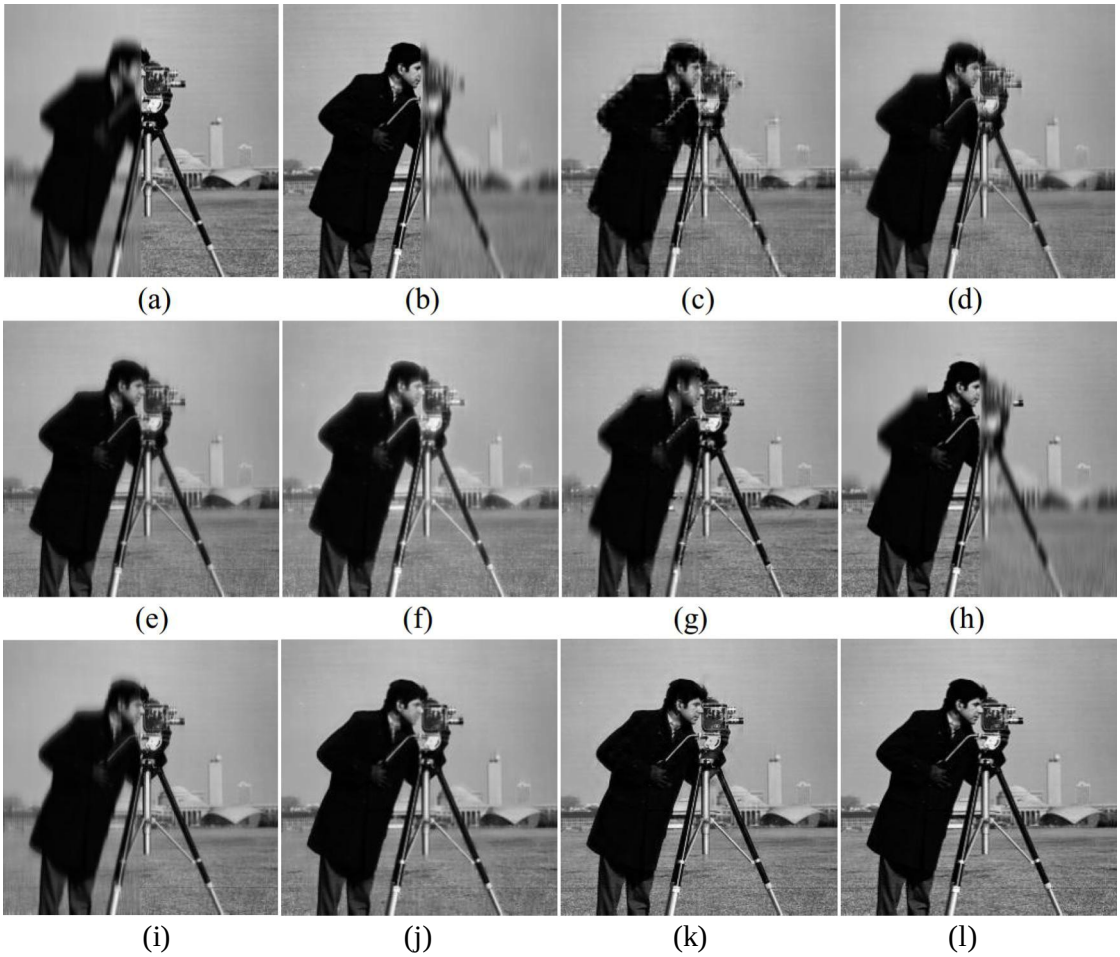

Figure 5.9: Fusion results of 'Cameraman' image. (a) left blurred image. (b) right blurred image. Fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method. (l) Original image.

The proposed method produces the best fused images with very similar quality to the reference images (see images (k) and (l) in Figure 5.9 and Figure 5.10). We can observe in Figure 5.9(k) that the proposed method preserves the same features of 'cameraman' face and

without ringing around the head area. Also, in 'Traffic' image (show Figure 5.10(k)), it is able to extract all objects (cars) and produces images with sharper lines and sharper edges.

The comparative methods (except SC method) don't live up to the acceptable quality (Figure 5.9(c)-(i)). SC method somewhat is better than others, it produces some blurring, like the ones on the face in Figure 5.9(j)). For the second image 'Traffic' as presented in Figure 5.10, most comparative methods are not capable to achieve informative and desired fused image. They produce blurred images with ringing effects around edges and can't present the important features such as cars.

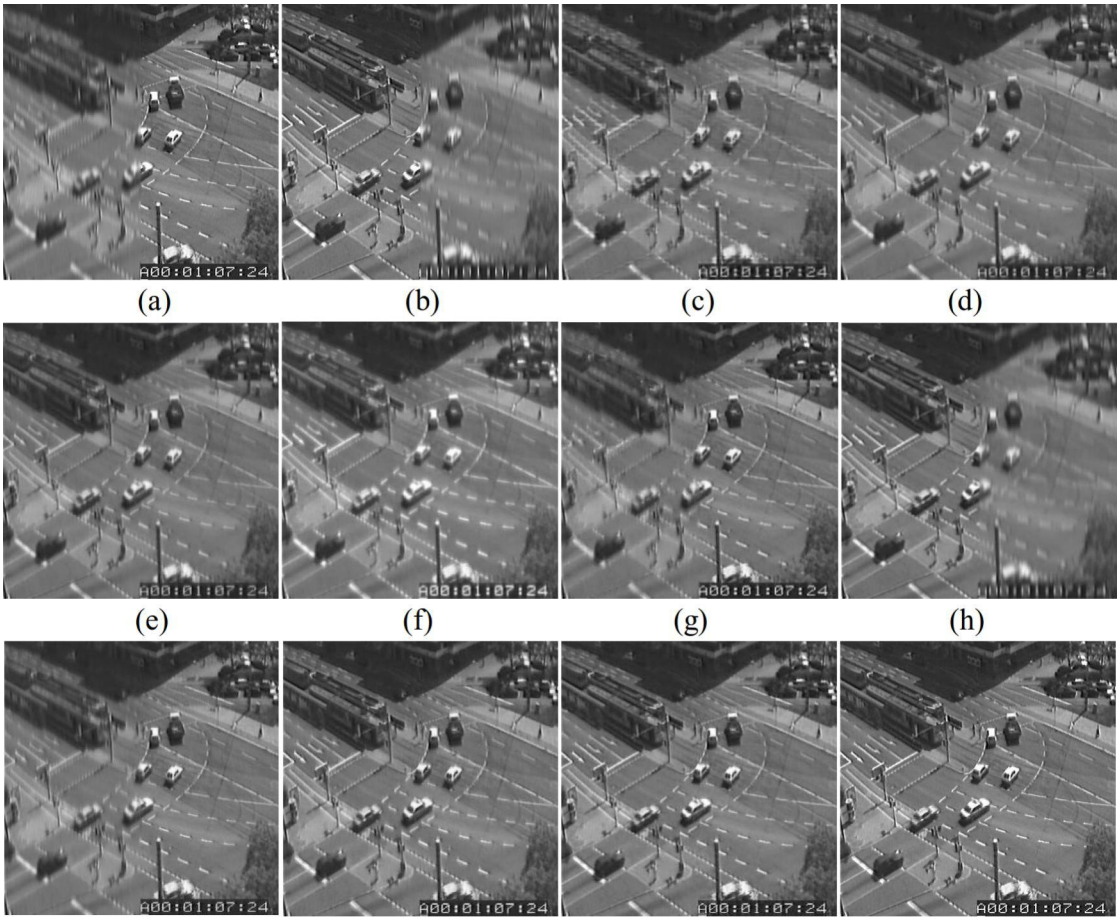

Figure 5.10: Fusion results of 'Traffic' image. (a) left Blurred image. (b) right blurred image. Fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method. (l) Original image.

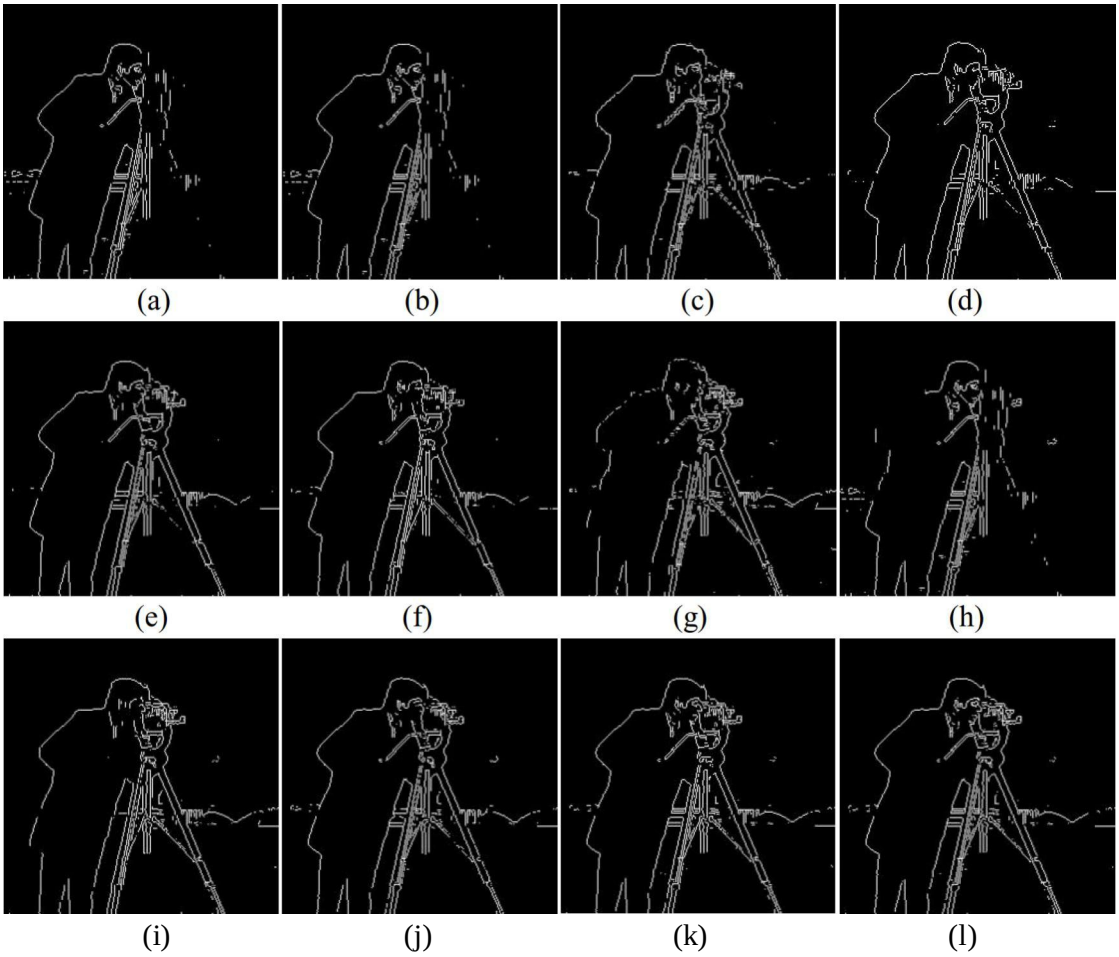

Figure 5.11: Edge map of fusion results of 'Cameraman' image. (a) left blurred image. (b) right blurred image. Fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method. (l) Original image.

The illustrated images in Figure 5.11 and Figure 5.12 present the edge map of the fused images in Figure 5.9 and Figure 5.10 respectively, for all comparative and the proposed methods, in addition, the edge map of the original image.

It is observed that some methods could give a clear edge map for the general context of the fused image such as the SIDWT, PCA, LAP and SC methods. While they could not provide a fine edge map to the small and overlapping features (the features of the face in Figure 5.11 and the continuity of lines in Figure 5.12).

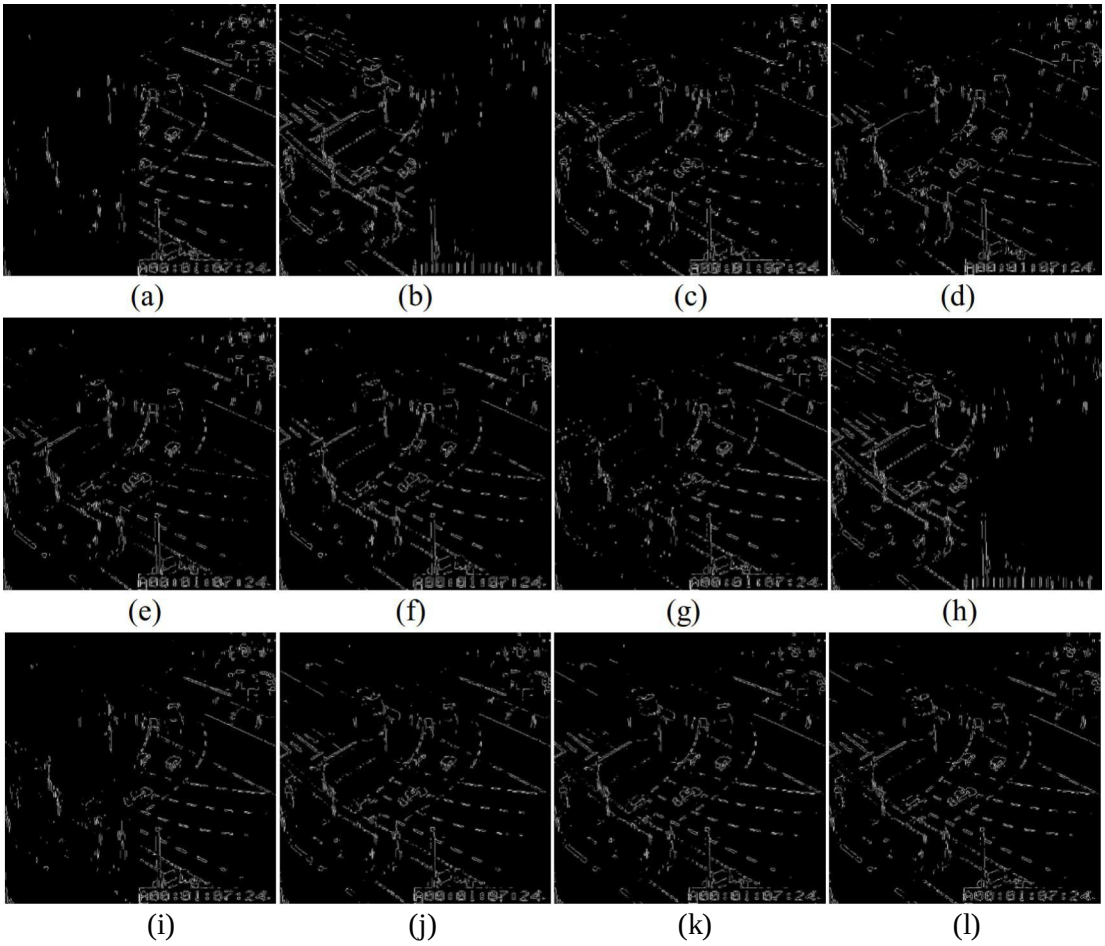

Figure 5.12: Edge map of fusion results of 'Traffic' image. (a) left blurred. (b) right blurred. Fused images using: (c) DWT. (d) SIDWT. (e) PCA. (f) LAP. (g) BGBSC. (h) DCT+Ac_Max. (i) DWT+Pca+av. (j) SC. (k) Proposed method. (l) Original image.

5.3.2 Quantitative evaluation

For the scientific honesty, it is not fair that any fusion method can be judged by visual display only. For this reason, a quantitative evaluation is conducted to evaluate the proposed method by applying some fusion quality metrics (RMSE, PSNR, SSIM, SD, AG, Entropy, SF, $Q^{AB/F}$ and $L^{AB/F}$) on the results of all comparative and proposed methods presented in the qualitative evaluation.

Table 5.1: Quantitative evaluation of different image fusion methods for 'Clock' image.

methods	measures					
	SD	**AG**	**He**	**SF**	$\mathbf{Q^{AB/F}}$	$\mathbf{L^{AB/F}}$
DWT	50.6792	4.8495	7.3575	8.3281	0.7283	0.2583
SIDWT	50.5017	4.2727	7.345	7.142	0.7708	0.2269
PCA	50.5657	3.9554	7.2934	6.4488	0.8277	0.1723
LAP	52.0906	3.9577	7.3843	6.5932	0.6914	0.3073
BGBSC	51.2888	4.9012	7.2715	**10.1011**	0.5995	0.3822
DCT+Ac_max	51.8873	5.1166	7.3103	9.3062	0.7979	0.2011
DWT+Pca+av	50.7362	4.6322	7.3102	7.9319	0.8069	0.1931
SC	52.3085	5.0679	7.3611	9.4894	0.8381	0.1456
The proposed	**52.4074**	**5.6536**	**7.4317**	9.5031	**0.8812**	**0.1126**

Table 5.2: Quantitative evaluation of different image fusion methods for 'Lab' image.

methods	measures					
	SD	**AG**	**He**	**SF**	$\mathbf{Q^{AB/F}}$	$\mathbf{L^{AB/F}}$
DWT	44.1064	5.6946	7.0803	11.4193	0.7432	0.2454
SIDWT	43.7745	5.1907	7.0790	10.0246	0.7536	0.2441
PCA	43.9980	4.9762	7.0740	10.2924	0.8195	0.1805
LAP	44.4390	5.0028	7.0971	10.5303	0.7361	0.2633
BGBSC	44.2860	5.9240	7.1061	12.3155	0.6222	0.3629
DCT+Ac_max	44.8486	5.5999	7.0970	11.9756	0.7333	0.2657
DWT+Pca+av	44.2265	5.4376	7.0808	11.1359	0.9297	0.0703
SC	45.0994	6.2902	7.0968	12.6025	0.8997	0.0927
The proposed	**45.1708**	**6.8227**	**7.1197**	**12.8860**	**0.9444**	**0.0522**

Table 5.3: Quantitative evaluation of different image fusion methods for 'Pepsi' image.

methods	measures					
	SD	**AG**	**He**	**SF**	$\mathbf{Q^{AB/F}}$	$\mathbf{L^{AB/F}}$
DWT	45.6182	5.3361	6.9907	9.7130	0.6854	0.2935
SIDWT	45.5587	4.4810	6.9666	8.5319	0.7541	0.2441
PCA	45.4459	4.1445	6.9556	7.7613	0.8065	0.1935
LAP	44.5197	4.4739	6.9584	8.4751	0.7470	0.2504
BGBSC	45.8629	5.6550	6.8826	11.6776	0.6414	0.3393
DCT+Ac_max	46.6487	5.5097	6.9100	11.3255	0.7480	0.2496
DWT+Pca+av	45.9385	4.4524	6.9124	9.1259	0.7636	0.2364
SC	47.0845	5.5278	6.9963	11.3336	0.8615	0.1287
The proposed	**47.4137**	**6.4005**	**7.1075**	**11.7791**	**0.8962**	**0.1004**

The quantitative performance comparisons are presented in tables 5.1, 5.2 and Table 5.3 that give the performance metric of the non-referenced images (Clock, Lab and Pepsi). Theoretically, the higher value of SD, AG, He, SF and $Q^{AB/F}$ parameters mean better quality of the fused image unlike the lower value of the $L^{AB/F}$ parameter that meanbetter quality. Values of these quantities confirm the subjective results and prove the superiority of the proposed method.

From tables 5.1, 5.2 and 5.3, it is clearly seen that the proposed method performs better than other algorithms in most cases. This is because the proposed algorithm always gives best metrics values. The fused image has the higher value of AG and SD, which means that it has more sharpness, more clarity and with high contrast. Usually, the proposed method proves here effectiveness in term of entropy measures (He). We can see that it has the biggest value in the three tables (5.1, 5.2 and 5.3). This indicates that our method is able to transfer a big amount of information content into the fused image.

Also, in term of SF, the proposed method can yield fused images with more visual information. The edge-based similarity metric $Q^{AB/F}$ is usually with higher values in case of the proposed method and it gives a lower values of $L^{AB/F}$. This indicates the amount of edge information transferred from the source images into the fused image as shown in Figures that present the edge map, and with neglected loss of edge information. This also indicates that the

proposed algorithm could well preserve the prominent texture information from the source images.

Table 5.4: Quantitative evaluation of different image fusion methods for 'Cameraman' image.

method	measures								
	RMSE	**PSNR(dB)**	**SSIM**	**SD**	**AG**	**He**	**SF**	**$Q^{AB/F}$**	**$L^{AB/F}$**
DWT	12.7638	26.0112	0.8769	59.5704	11.5403	7.1063	22.3391	0.8039	0.1757
SIDWT	10.7429	27.5084	0.9199	58.8701	9.4524	7.0352	18.8734	0.8434	0.1550
PCA	9.7561	28.3452	0.9311	58.9127	9.0698	7.0212	18.3595	0.8824	0.1176
LAP	15.0961	24.5535	0.9031	58.3059	10.0854	7.0986	20.7851	0.8566	0.1401
BGBSC	12.3579	26.2919	0.9040	59.9414	12.352	7.1133	24.9779	0.7960	0.1940
DCT+Ac_max	14.9405	24.6435	0.8281	59.2755	9.2320	7.0387	21.3941	0.6978	0.3017
DWT+Pca+av	10.6074	27.6186	0.9218	59.0171	9.6430	7.0428	19.3885	0.8359	0.1641
SC	6.5693	31.7804	0.9564	61.0124	11.2574	7.0560	24.2608	0.8777	0.1169
The proposed	**4.8263**	**34.4586**	**0.9764**	**61.2854**	**13.5830**	**7.1187**	**26.8804**	**0.9021**	**0.0915**

Table 5.5: Quantitative evaluation of different image fusion methods for 'Traffic' image.

method	measures								
	RMSE	**PSNR(dB)**	**SSIM**	**SD**	**AG**	**He**	**SF**	**$Q^{AB/F}$**	**$L^{AB/F}$**
DWT	11.9313	26.5970	0.8552	36.3610	10.4073	6.9682	18.4998	0.791	0.1852
SIDWT	10.3107	27.8651	0.9097	35.1547	8.4722	6.8955	15.7085	0.8272	0.1711
PCA	9.5232	28.5552	0.9215	35.0079	8.1388	6.8904	14.8993	0.8731	0.1269
LAP	14.1174	25.1357	0.8941	38.1845	9.1129	6.9816	16.8163	0.8708	0.1244
BGBSC	11.0445	27.2679	0.8632	37.2929	10.9748	6.9646	22.3881	0.772	0.2154
DCT+Ac_max	14.9496	24.6382	0.8299	35.6226	8.7645	6.9253	15.9241	0.6704	0.3296
DWT+Pca+av	9.8005	28.3058	0.8952	35.7210	8.5243	6.9172	17.2391	0.8051	0.1949
SC	6.5858	31.7587	0.9477	38.5476	10.6377	7.0078	20.4194	0.8998	0.0907
The proposed	**3.7581**	**36.6314**	**0.9747**	**38.8058**	**12.1336**	**7.0161**	**22.5349**	**0.9183**	**0.0752**

To demonstrate the efficiency of our proposed work, some mathematical measures such as RMSE, PSNR, SSIM, SD, AG, Entropy, SF, $Q^{AB/F}$ and $L^{AB/F}$ are also used for the referenced images (Cameraman and Traffic) and the corresponding results are shown in Tables 5.4 and 5.5. Higher values of fusion quality metrics PSNR, SSIM, SD, AG, Entropy, SF and $Q^{AB/F}$ signify the better quality of fused image. Opposite to RMSE and $L^{AB/F}$, lower value means the best quality.

From the results shown in table 5.4 and 5.5, we can note that the proposed method gives the best performance in terms of RMSE, PSNR and SSIM. The fused image is very similar to the reference image with low variation between the fused and the reference image (the lower RMSE). The proposed method can reconstruct the fused images with enhanced visual quality because it gives higher PSNR value. According to the same tables which contains SSIM values, the superiority of the proposed method is evident.

It is clearly seen that the proposed method is more efficient than other methods used for comparison because it always gives best metrics values. The fused image has the higher value of AG and SD, which means that, it is with more clarity and with high contrast. In addition, our method gives a larger value of $Q^{AB/F}$, entropy and SF and it gives lower value of $L^{AB/F}$. This indicates that our method performs better than other methods. It is able to transfer more information from input images to the fused image.

5.4 Conclusion

In this chapter, extensive simulations are carried out to demonstrate that the proposed method outperforms some literature image fusion techniques in the qualitative and quantitative assessment. To do this, two image data sets of multi-focus images (without reference image) and artificially blurred images (with reference image) are used.

These analysis and results make it clear that the performance of the proposed method is better compared to other fusion methods. One can note that the proposed method can lead to the best visual qualities on both multi-focus images and artificially blurred images. Our proposed method is able to achieve a fused image better visual quality and with more information.

The proposed algorithm yields the best results in all the cases. This is because it can well combine the independent information from each source image to form one single image. Here, the independent information is referred as the focused regions detected from each source image.

Conclusion and future work

Conclusion and future work

In this thesis, we have introduced a novel image fusion technique for visual sensor networks based on the combined integer lifting wavelet transform (ILWT) and the discrete cosine Transform (DCT). The proposed method is based on the variance in the DCT domain as a contrast measure in the focused regions and the correlation between the ILWT coefficients through all levels.

We started this work with a general overview of the concepts of the image fusion process. It contains some basics of sensors of image fusion system, fusion processing levels, and a short overview about the traditional image fusion methods in both the spatial domain and the transform domain. Also some applications and some performance metrics of image fusion techniques are presented.

Also, we have reviewed a brief state of the art of some recent image fusion techniques in the transform domain and compared with the proposed method. We have focused on the image fusion methods based on:

- The DCT domain: contrast-based method, the method based on higher valued of the alternative current (AC) coefficients and the spatial frequency-based technique.
- The DWT domain: region-based technique, contrast-based technique and DWT-based PCA technique.

A comparative study is performed to demonstrate the performance of the proposed method in comparison with other methods methods. Simulations was done under the same condition such as multi-focus input images and artificially blurred input images. Also, various image fusion quality metrics are used. These extensive simulations clearly demonstrate that our proposed method provides the best visual quality. It also often gives best metrics values compared to the other methods.

Generally, Image fusion techniques in the DCT domain resulted blocking artifacts and blurring in the fused images. This is due to the that fact that images are divided into blocks and wavelet analysis can't effectively represent the line and plane singularities of the images and thus can't represent the edge directions in images correctly.

Our proposed method is able to solve the precedent shortcomings, by using two fusion phases to accurately adjust the fused blocks in the DCT domain with appropriate edges in the wavelet domain. Since the ILWT, quite easily uses integer operations to represent the wavelet coefficients of images, it is very simple and effective for real-time image fusion in the visual sensor networks. This method is able to generate fused images with high visual quality (as shown).

As a future work, we will try to:

- Include a large dataset for other applications (medical, remote sensing…etc.)
- Embed implementation in reconfigurable programmable hardware FPGA or SoC.
- Extend this proposed algorithm to be used in video fusion fields with its embedded implementation.

References

References

[1] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE,* vol. 91, no. 8, pp. 1247-1256, 2003.

[2] Y. Charfi, N. Wakamiya, and M. Murata, "Challenging issues in visual sensor networks," *IEEE Wireless Communications,* vol. 16, no. 2, 2009.

[3] S. Soro and W. Heinzelman, "A survey of visual sensor networks," *Advances in multimedia,* vol. 2009, 2009.

[4] Y. Yang, S. Tong, S. Huang, P. Lin, and Y. Fang, "A Hybrid Method for Multi-focus Image Fusion Based on Fast Discrete Curvelet Transform," *IEEE Access,* 2017.

[5] B. Tavli, K. Bicakci, R. Zilan, and J. M. Barcelo-Ordinas, "A survey of visual sensor network platforms," *Multimedia Tools and Applications,* vol. 60, no. 3, pp. 689-726, 2012.

[6] N. Paramanandham and K. Rajendiran, "Multi sensor image fusion for surveillance applications using hybrid image fusion algorithm," *Multimedia Tools and Applications,* pp. 1-32, 2017.

[7] J. Yang, Y. Li, J. C.-W. Chan, and Q. Shen, "Image Fusion for Spatial Enhancement of Hyperspectral Image via Pixel Group Based Non-Local Sparse Representation," *Remote Sensing,* vol. 9, no. 1, p. 53, 2017.

[8] A. Koutsia, T. Semertzidis, K. Dimitropoulos, N. Grammalidis, and K. Georgouleas, "Intelligent traffic monitoring and surveillance with multiple cameras," in *Content-Based Multimedia Indexing, 2008. CBMI 2008. International Workshop on*, 2008, pp. 125-132: IEEE.

[9] J. Boice *et al.*, "Meerkats: A power-aware, self-managing wireless camera network for wide area monitoring," in *Proc. Workshop on Distributed Smart Cameras*, 2006, pp. 393-422: Citeseer.

[10] I. Pavlidis, V. Morellas, P. Tsiamyrtzis, and S. Harp, "Urban surveillance systems: from the laboratory to the commercial world," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1478-1497, 2001.

[11] M. Abdipour and M. Nooshyar, "Multi-focus image fusion using sharpness criteria for visual sensor networks in wavelet domain," *Computers & Electrical Engineering*, vol. 51, pp. 74-88, 2016.

[12] A. Redondi, M. Cesana, M. Tagliasacchi, I. Filippini, G. Dán, and V. Fodor, "Cooperative image analysis in visual sensor networks," *Ad Hoc Networks,* vol. 28, pp. 38-51, 2015.

[13] Y. Zhang, X. Bai, and T. Wang, "Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure," *Information Fusion,* vol. 35, pp. 81-101, 2017.

[14] B. Zhang, X. Lu, H. Pei, Y. Liu, W. Zhou, and D. Jiao, "Multi-focus image fusion based on sparse decomposition and background detection," *Digital Signal Processing,* vol. 58, pp. 50-63, 2016.

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

[16] Z. Liu, Y. Chai, H. Yin, J. Zhou, and Z. Zhu, "A novel multi-focus image fusion approach based on image decomposition," *Information Fusion,* vol. 35, pp. 102-116, 2017.

[17] Q. Zhang and X. Maldague, "An adaptive fusion approach for infrared and visible images based on NSCT and compressed sensing," *Infrared Physics & Technology,* vol. 74, pp. 11-20, 2016.

[18] D. E. Nirmala and V. Vaidehi, "Comparison of Pixel-level and feature level image fusion methods," in *Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on*, 2015, pp. 743-748: IEEE.

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

[20] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Information fusion,* vol. 13, no. 1, pp. 10-19, 2012.

[21] G. Siddalingesh, A. Mallikarjun, H. Sanjeevkumar, and S. Kotresh, "Feature-Level Image Fusion Using DWT, SWT, and DT-CWT," in *Emerging Research in Electronics, Computer Science and Technology*: Springer, 2014, pp. 183-194.

[22] B. Yang, Z.-l. Jing, and H.-t. Zhao, "Review of pixel-level image fusion," *Journal of Shanghai Jiaotong University (Science),* vol. 15, pp. 6-12, 2010.

[23] H. Ghassemian, "A review of remote sensing image fusion methods," *Information Fusion,* vol. 32, pp. 75-89, 2016.

[24] Z. Chen and S. Muramatsu, "Multi-focus pixel-based image fusion in dual domain," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 1736-1740: IEEE.

[25] Y. Zuo, J. Liu, G. Bai, X. Wang, and M. Sun, "Airborne Infrared and Visible Image Fusion Combined with Region Segmentation," *Sensors,* vol. 17, no. 5, p. 1127, 2017.

[26] X. Xia, S. Fang, and Y. Xiao, "High resolution image fusion algorithm based on multi-focused region extraction," *Pattern Recognition Letters,* vol. 45, pp. 115-120, 2014.

[27] B. Zhang, X. Lu, H. Pei, H. Liu, Y. Zhao, and W. Zhou, "Multi-focus image fusion algorithm based on focused region extraction," *Neurocomputing*, vol. 174, pp. 733-748, 2016.

[28] S. Bickelhaupt *et al.*, "Independent value of image fusion in unenhanced breast MRI using diffusion-weighted and morphological T2-weighted images for lesion characterization in patients with recently detected BI-RADS 4/5 x-ray mammography findings," *European radiology*, vol. 27, no. 2, pp. 562-569, 2017.

[29] W. Ouerghemmi, A. Le Bris, N. Chehata, and C. Mallet, "A TWO-STEP DECISION FUSION STRATEGY: APPLICATION TO HYPERSPECTRAL AND MULTISPECTRAL IMAGES FOR URBAN CLASSIFICATION," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, 2017.

[30] D. Zhang, F. Song, Y. Xu, and Z. Liang, "Decision level fusion," in *Advanced Pattern Recognition Technologies with Applications to Biometrics*: IGI Global, 2009, pp. 328-348.

[31] H. R. Shahdoosti and H. Ghassemian, "Combining the spectral PCA and spatial PCA fusion methods by an optimal filter," *Information Fusion*, vol. 27, pp. 150-160, 2016.

[32] R. Gharbia, A. H. El Baz, A. E. Hassanien, and M. F. Tolba, "Remote sensing image fusion approach based on Brovey and wavelets transforms," in *Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014*, 2014, pp. 311-321: Springer.

[33] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE transactions on geoscience and remote sensing*, vol. 49, no. 1, pp. 295-309, 2011.

[34] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 746-750, 2010.

[35] Y. Yang, S. Tong, S. Huang, and P. Lin, "Dual-tree complex wavelet transform and image block residual-based multi-focus image fusion in visual sensor networks," *Sensors*, vol. 14, no. 12, pp. 22408-22430, 2014.

[36] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on communications*, vol. 31, no. 4, pp. 532-540, 1983.

[37] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," *RCA engineer*, vol. 29, no. 6, pp. 33-41, 1984.

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

[39] A. Ben Hamza, Y. He, H. Krim, and A. Willsky, "A multiscale approach to pixel-level image fusion," *Integrated Computer-Aided Engineering,* vol. 12, no. 2, pp. 135-146, 2005.

[40] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *Image Processing, 1997. Proceedings., International Conference on*, 1997, vol. 3, pp. 288-291: IEEE.

[41] L. Ming and W. Shunjun, "A new image fusion algorithm based on wavelet transform," in *Computational Intelligence and Multimedia Applications, 2003. ICCIMA 2003. Proceedings. Fifth International Conference on*, 2003, pp. 154-159: IEEE.

[42] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE signal processing magazine,* vol. 22, no. 6, pp. 123-151, 2005.

[43] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Information fusion,* vol. 8, no. 2, pp. 119-130, 2007.

[44] V. Bhateja, H. Patel, A. Krishn, A. Sahu, and A. Lay-Ekuakille, "Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains," *IEEE Sensors Journal,* vol. 15, no. 12, pp. 6783-6790, 2015.

[45] P. Mangalraj and A. Agrawal, "Fusion of Multi-Sensor Satellite Images Using Non-Subsampled Contourlet Transform," *Procedia Computer Science,* vol. 54, pp. 713-720, 2015.

[46] M. Nejati, S. Samavi, and S. Shirani, "Multi-focus image fusion using dictionary-based sparse representation," *Information Fusion,* vol. 25, pp. 72-84, 2015.

[47] J. Tang, "A contrast based image fusion technique in the DCT domain," *Digital Signal Processing,* vol. 14, no. 3, pp. 218-226, 2004.

[48] P. Hill, M. E. Al-Mualla, and D. Bull, "Perceptual Image Fusion Using Wavelets," *IEEE Transactions on Image Processing,* vol. 26, no. 3, pp. 1076-1088, 2017.

[49] Y. A. V. Phamila and R. Amutha, "Discrete Cosine Transform based fusion of multi-focus images for visual sensor networks," *Signal Processing,* vol. 95, pp. 161-170, 2014.

[50] S. Li, J. T. Kwok, and Y. Wang, "Multifocus image fusion using artificial neural networks," *Pattern Recognition Letters,* vol. 23, no. 8, pp. 985-997, 2002.

[51] J. Tian, L. Chen, L. Ma, and W. Yu, "Multi-focus image fusion using a bilateral gradient-based sharpness criterion," *Optics communications,* vol. 284, no. 1, pp. 80-87, 2011.

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

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

[54] B. Latreche, S. Saadi, M. Kious, and A. Benziane, "A novel hybrid image fusion method based on integer lifting wavelet and discrete cosine transformer for visual sensor networks," *Multimedia Tools and Applications,* pp. 1-23, 2018.

[55] D. Chen, Y. Li, H. Zhang, and W. Gao, "Invertible update-then-predict integer lifting wavelet for lossless image compression," *EURASIP Journal on Advances in Signal Processing,* vol. 2017, no. 1, p. 8, 2017.

[56] G. Hu, Y. Zheng, and X.-q. Qin, "Image Fusion based on integer lifting wavelet transform," in *Image Fusion and Its Applications*: InTech, 2011.

[57] Z. Wang, X. Yu, and L. Zhang, "A remote sensing image fusion algorithm based on integer wavelet transform," in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, 2008, pp. 5950-5954: IEEE.

[58] M. Grangetto, E. Magli, M. Martina, and G. Olmo, "Optimization and implementation of the integer wavelet transform for image coding," *IEEE transactions on image processing,* vol. 11, no. 6, pp. 596-604, 2002.

[59] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Lossless image compression using integer to integer wavelet transforms," in *Image Processing, 1997. Proceedings., International Conference on*, 1997, vol. 1, pp. 596-599: IEEE.

[60] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "Multi-focus image fusion for visual sensor networks in DCT domain," *Computers & Electrical Engineering,* vol. 37, no. 5, pp. 789-797, 2011.

[61] M. Al Najjar, M. Ghantous, and M. Bayoumi, "Visual sensor nodes," in *Video Surveillance for Sensor Platforms*: Springer, 2014, pp. 17-35.

[62] C. B. Margi, V. Petkov, K. Obraczka, and R. Manduchi, "Characterizing energy consumption in a visual sensor network testbed," in *Testbeds and Research Infrastructures for the Development of Networks and Communities, 2006. TRIDENTCOM 2006. 2nd International Conference on*, 2006, pp. 8 pp.-339: IEEE.

[63] D. Ganesan, B. Greenstein, D. Estrin, J. Heidemann, and R. Govindan, "Multiresolution storage and search in sensor networks," *ACM Transactions on Storage (TOS),* vol. 1, no. 3, pp. 277-315, 2005.

[64] M. Rahimi *et al.*, "Cyclops: in situ image sensing and interpretation in wireless sensor networks," in *Proceedings of the 3rd international conference on Embedded networked sensor systems*, 2005, pp. 192-204: ACM.

[65] W.-C. Feng, E. Kaiser, W. C. Feng, and M. L. Baillif, "Panoptes: scalable low-power video sensor networking technologies," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM),* vol. 1, no. 2, pp. 151-167, 2005.

[66] A. Rowe, D. Goel, and R. Rajkumar, "Firefly mosaic: A vision-enabled wireless sensor networking system," in *Real-time systems symposium, 2007. RTSS 2007. 28th IEEE international*, 2007, pp. 459-468: IEEE.

[67] T. Teixeira, E. Culurciello, J. H. Park, D. Lymberopoulos, A. Barton-Sweeney, and A. Savvides, "Address-event imagers for sensor networks: evaluation and modeling," in *Information Processing in Sensor Networks, 2006. IPSN 2006. The Fifth International Conference on*, 2006, pp. 458-466: IEEE.

[68] P. Chen *et al.*, "CITRIC: A low-bandwidth wireless camera network platform," in *Distributed smart cameras, 2008. ICDSC 2008. Second ACM/IEEE international conference on*, 2008, pp. 1-10: IEEE.

[69] M. Zhang and W. Cai, "Vision mesh: A novel video sensor networks platform for water conservancy engineering," in *Computer science and information technology (ICCSIT), 2010 3rd IEEE international conference on*, 2010, vol. 4, pp. 106-109: IEEE.

[70] J. Van Genderen and C. Pohl, "Image fusion: Issues, techniques and applications," 1994.

[71] C. Pohl and J. Van Genderen, *Remote sensing image fusion: A practical guide*. Crc Press, 2016.

[72] J. L. Rubio-Guivernau *et al.*, "Wavelet-based image fusion in multi-view three-dimensional microscopy," *Bioinformatics,* vol. 28, no. 2, pp. 238-245, 2011.

[73] M. Kim, D. K. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Information fusion,* vol. 27, pp. 198-214, 2016.

[74] P. Du, S. Liu, J. Xia, and Y. Zhao, "Information fusion techniques for change detection from multi-temporal remote sensing images," *Information Fusion,* vol. 14, no. 1, pp. 19-27, 2013.

[75] Q. Li, X. Yang, W. Wu, K. Liu, and G. Jeon, "Multi-Focus Image Fusion Method for Vision Sensor Systems via Dictionary Learning with Guided Filter," *Sensors*, vol. 18, no. 7, p. 2143, 2018.

[76] I. Bloch, *Information fusion in signal and image processing: major probabilistic and non-probabilistic numerical approaches*. John Wiley & Sons, 2013.

[77] C. Pohl and J. L. Van Genderen, "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *International journal of remote sensing*, vol. 19, no. 5, pp. 823-854, 1998.

[78] B. S. Kumar, "Image fusion based on pixel significance using cross bilateral filter," *Signal, image and video processing*, vol. 9, no. 5, pp. 1193-1204, 2015.

[79] Y. Yang, S. Tong, S. Huang, P. Lin, and Y. Fang, "A Hybrid Method for Multi-Focus Image Fusion Based on Fast Discrete Curvelet Transform," *IEEE Access*, vol. 5, pp. 14898-14913, 2017.

[80] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, pp. 4-19, 2014.

[81] V. Solanky and S. Katiyar, "Pixel-level image fusion techniques in remote sensing: a review," *Spatial Information Research*, vol. 24, no. 4, pp. 475-483, 2016.

[82] I. Guyon and A. Elisseeff, "An introduction to feature extraction," in *Feature extraction*: Springer, 2006, pp. 1-25.

[83] S. Kor and U. Tiwary, "Feature level fusion of multimodal medical images in lifting wavelet transform domain," in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, 2004, vol. 1, pp. 1479-1482: IEEE.

[84] Y. Gu, M. Yang, Z. Shi, and Z. Wu, "The Applications of Decision-Level Data Fusion Techniques in the Field of Multiuser Detection for DS-UWB Systems," *Sensors*, vol. 15, no. 10, pp. 24771-24790, 2015.

[85] K. Rani and R. Sharma, "Study of different image fusion algorithm," *International journal of Emerging Technology and advanced Engineering*, vol. 3, no. 5, pp. 288-291, 2013.

[86] M. R. Metwalli, A. H. Nasr, O. S. F. Allah, and S. El-Rabaie, "Image fusion based on principal component analysis and high-pass filter," in *Computer Engineering & Systems, 2009. ICCES 2009. International Conference on*, 2009, pp. 63-70: IEEE.

[87] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Transactions on Geoscience and Remote sensing,* vol. 44, no. 6, pp. 1672-1682, 2006.

[88] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics (TOG),* vol. 2, no. 4, pp. 217-236, 1983.

[89] A. Sahu, V. Bhateja, and A. Krishn, "Medical image fusion with Laplacian pyramids," in *Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 2014 International Conference on*, 2014, pp. 448-453: IEEE.

[90] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence,* vol. 11, no. 7, pp. 674-693, 1989.

[91] Y. Yang, D. S. Park, S. Huang, and N. Rao, "Medical image fusion via an effective wavelet-based approach," *EURASIP Journal on Advances in Signal Processing,* vol. 2010, p. 44, 2010.

[92] M. Chitnis, Y. Liang, J. Y. Zheng, P. Pagano, and G. Lipari, "Wireless line sensor network for distributed visual surveillance," in *Proceedings of the 6th ACM symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks,* 2009, pp. 71-78: ACM.

[93] R. Kays *et al.*, "Camera traps as sensor networks for monitoring animal communities," in *Local Computer Networks, 2009. LCN 2009. IEEE 34th Conference on*, 2009, pp. 811-818: IEEE.

[94] K. Parmar, R. K. Kher, and F. N. Thakkar, "Analysis of CT and MRI image fusion using wavelet transform," in *Communication Systems and Network Technologies (CSNT), 2012 International Conference on*, 2012, pp. 124-127: IEEE.

[95] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature," *Geoscientific model development,* vol. 7, no. 3, pp. 1247-1250, 2014.

[96] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics letters,* vol. 44, no. 13, pp. 800-801, 2008.

[97] M. G. Albanesi, R. Amadeo, S. Bertoluzza, and G. Maggi, "A New Class of Wavelet-Based Metrics for Image Similarity Assessment," *Journal of Mathematical Imaging and Vision,* vol. 60, no. 1, pp. 109-127, 2018.

[98] P. Jagalingam and A. V. Hegde, "A review of quality metrics for fused image," *Aquatic Procedia,* vol. 4, pp. 133-142, 2015.

[99] J. Wu, H. Huang, Y. Qiu, H. Wu, J. Tian, and J. Liu, "Remote sensing image fusion based on average gradient of wavelet transform," in *Mechatronics and Automation, 2005 IEEE International Conference*, 2005, vol. 4, pp. 1817-1821: IEEE.

[100] W. Xu, M. Li, and X. Wang, "Information fusion based on information entropy in fuzzy multi-source incomplete information system," *International Journal of Fuzzy Systems*, vol. 19, no. 4, pp. 1200-1216, 2017.

[101] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang, and Y. Zhang, "Multi-focus image fusion based on spatial frequency in discrete cosine transform domain," *IEEE signal processing letters*, vol. 22, no. 2, pp. 220-224, 2015.

[102] C. a. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics letters*, vol. 36, no. 4, pp. 308-309, 2000.

[103] V. Petrovic and T. Cootes, "Objectively optimised multisensor image fusion," in *Information Fusion, 2006 9th International Conference on*, 2006, pp. 1-7: IEEE.

[104] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600-612, 2004.

[105] R. Singh and A. Khare, "Fusion of multimodal medical images using Daubechies complex wavelet transform–A multiresolution approach," *Information Fusion*, vol. 19, pp. 49-60, 2014.

[106] Z. Zhou, S. Li, and B. Wang, "Multi-scale weighted gradient-based fusion for multi-focus images," *Information Fusion*, vol. 20, pp. 60-72, 2014.

[107] K. R. Rao and P. Yip, *Discrete cosine transform: algorithms, advantages, applications*. Academic press, 2014.

[108] M. A. Naji and A. Aghagolzadeh, "A new multi-focus image fusion technique based on variance in DCT domain," in *Knowledge-Based Engineering and Innovation (KBEI), 2015 2nd International Conference on*, 2015, pp. 478-484: IEEE.

[109] L. Prasad and S. S. Iyengar, *Wavelet analysis with applications to image processing*. CRC press, 1997.

[110] C. K. Chui, *Wavelets: a mathematical tool for signal analysis*. Siam, 1997.

[111] C. M. Leavey, M. N. James, J. Summerscales, and R. Sutton, "An introduction to wavelet transforms: a tutorial approach," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 45, no. 5, pp. 344-353, 2003.

[112] I. Daubechies, *Ten lectures on wavelets*. Siam, 1992.

[113] G. Kaiser, "Continuous Wavelet Transforms," in *A Friendly Guide to Wavelets*: Springer, 2011, pp. 60-77.

[114] T. Acharya and A. K. Ray, *Image processing: principles and applications*. John Wiley & Sons, 2005.

[115] T. Y. Yan, *Wavelet theory approach to pattern recognition*. World Scientific, 2009.

[116] G. Strang and T. Nguyen, *Wavelets and filter banks*. SIAM, 1996.

[117] S. Basu and B. C. Levy, *Multidimensional Filter Banks and Wavelets: Research Developments and Applications*. Springer Science & Business Media, 2013.

[118] K. K. Shukla and A. K. Tiwari, *Efficient algorithms for discrete wavelet transform: with applications to denoising and fuzzy inference systems*. Springer Science & Business Media, 2013.

[119] A. Jensen and A. la Cour-Harbo, *Ripples in mathematics: the discrete wavelet transform*. Springer Science & Business Media, 2001.

[120] R. C. Gonzalez and R. E. Woods, "Digital image processing," ed: Prentice hall New Jersey, 2002.

[121] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM journal on mathematical analysis*, vol. 29, no. 2, pp. 511-546, 1998.

[122] A. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," *Applied and computational harmonic analysis*, vol. 5, no. 3, pp. 332-369, 1998.

[123] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier analysis and applications*, vol. 4, no. 3, pp. 247-269, 1998.

[124] G. Dillen, B. Georis, J.-D. Legat, and O. Cantineau, "Combined line-based architecture for the 5-3 and 9-7 wavelet transform of JPEG2000," *IEEE Transactions on Circuits and Systems for video technology*, vol. 13, no. 9, pp. 944-950, 2003.

[125] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," *Signal processing: Image communication*, vol. 17, no. 1, pp. 3-48, 2002.

[126] V. Naidu and B. Elias, "A novel image fusion technique using DCT based Laplacian pyramid," *International Journal of Inventive Engineering and Sciences (IJIES) ISSN*, pp. 2319-9598, 2013.

[127] B. S. Kumar, M. Swamy, and M. O. Ahmad, "Multiresolution DCT decomposition for multifocus image fusion," in *Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference on*, 2013, pp. 1-4: IEEE.

[128] T. Pu and G. Ni, "Contrast-based image fusion using the discrete wavelet transform," *Optical Engineering,* vol. 39, no. 8, pp. 2075-2083, 2000.

[129] S. M. Rahman, M. O. Ahmad, and M. Swamy, "Contrast-based fusion of noisy images using discrete wavelet transform," *IET image Processing*, vol. 4, no. 5, pp. 374-384, 2010.

[130] P. Shah, T. Srikanth, S. Merchant, and U. B. Desai, "A novel multifocus image fusion scheme based on pixel significance using wavelet transform," in *IVMSP Workshop, 2011 IEEE 10th*, 2011, pp. 54-59: IEEE.

[131] P. Shah, S. N. Merchant, and U. B. Desai, "Multifocus and multispectral image fusion based on pixel significance using multiresolution decomposition," *Signal, Image and Video Processing*, pp. 1-15, 2013.

[132] B. S. Kumar, "Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform," *Signal, Image and Video Processing*, vol. 7, no. 6, pp. 1125-1143, 2013.

[133] "https://sites.google.com/site/uze1989/."

[134] "http://decsai.ugr.es/cvg/CG/base.htm//."

[135] P. Kwarteng and A. Chavez, "Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis," *Photogramm. Eng. Remote Sens,* vol. 55, pp. 339-348, 1989.

[136] *http://www.metapix.de/toolbox.htm*.

[137] "https://www.mathworks.com/matlabcentral/profile/authors/869694-kanchi."

[138] "https://www.mathworks.com/matlabcentral/profile/authors/1477144-r-vijayarajan."

[139] "https://www.mathworks.com/matlabcentral/profile/authors/4989349-durga-prasad-bavirisetti."