2016

Multisensor Concealed Weapon Detection Using the Image Fusion Approach

Tuzhi Xu
University of Windsor

Multisensor Concealed Weapon Detection Using the Image Fusion Approach

By

Tuzhi Xu

A Thesis

Submitted to the Faculty of Graduate Studies

through the Department of Electrical and Computer Engineering

in Partial Fulfillment of the Requirements for

the Degree of Master of Applied Science

at the University of Windsor

Windsor, Ontario, Canada

2016

Multisensor Concealed Weapon Detection Using the Image Fusion Approach

by

Tuzhi Xu

APPROVED BY:

_____

S. Cheng
Civil and Environmental Engineering

_____

R. Rashidzadeh
Electrical & Computer Engineering

_____

J. Wu, Advisor
Electrical & Computer Engineering

May 18, 2016

DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION

I. Co-Authorship Declaration

I hereby declare that this thesis paper incorporates the outcome of a joint research in collaboration with, and under the supervision of, Dr. Q. M. Jonathan Wu, with the review and revision being provided by Dr. Q. M. Jonathan Wu.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publication

This thesis includes one original paper that has been previously published for publication in peer reviewed journals, as follows:

Thesis Chapter	Publication title/full citation	Publication status
Chapter 3	T. Xu and Q. M. J. Wu, "Multisensor concealed weapon detection using the image fusion approach," *6th International Conference on Imaging for Crime Prevention and Detection (ICDP-15)*, July 2015, pp. 1-7.	published

I certify that I have obtained a written permission from the copyright owners to include the above published material in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Detection of concealed weapons is an increasingly important problem for both military and police since global terrorism and crime have grown as threats over the years. This work presents two image fusion algorithms, one at pixel level and another at feature level, for efficient concealed weapon detection application. Both the algorithms presented in this work are based on the double-density dual-tree complex wavelet transform (DDDTCWT). In the pixel level fusion scheme, the fusion of low frequency band coefficients is determined by the local contrast, while the high frequency band fusion rule is developed with consideration of both texture feature of the human visual system (HVS) and local energy basis. In the feature level fusion algorithm, features are exacted using Gaussian Mixture model (GMM) based multiscale segmentation approach and the fusion rules are developed based on region activity measurement. Experiment results demonstrate the robustness and efficiency of the proposed algorithms.

DEDICATION

This thesis is dedicated to my parents for all the love and support I received from them.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Dr. Jonathan Wu, for his devoted guidance, immense support and constant encouragement during my graduate study in University of Windsor. Without his excellent mentoring skills and valuable suggestion, this work would not have been done.

I would like to thank all my friends from the CVSS laboratory for their constant help. I benefited a lot from the discussion with them during my research work.

I would also like to convey my sincere thanks to Dr. Shaohong Cheng and Dr. Rashid Rashidzadeh for being on my advisory committee and the valuable comments on this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS/SYMBOLS

ABBREVIATIONS	DESCRIPTION
CWD	Concealed Weapon Detection
HIS	Intensity-Hue-Saturation
ICA	Independent Component Analysis
PCA	Principal Component Analysis
LCLS	Linearly Constrained Least-Squares
DWFT	Discrete Wavelet frame Transform
SVM	Support Vector Machines
DWT	Discrete Wavelet Transform
HVS	Human Visual System
NMDB	non-multiscale-decomposition-based
MDB	multiscale-decomposition-based
AWA	adaptive weight averaging
MRF	Markov Random Field
NSCT	non-subsampled Contourlet transform
IR	Infrared
MMW	millimeter wave
LPT	Laplacian pyramid transform
EM	expectation-maximization
DDDWT	Double Density Discrete Wavelet Transform
DTCWT	Dual Tree Complex Wavelet Transform
DDDTCWT	Double-density Dual-tree Complex Wavelet transform
1D	1-dimensional
2D	2-dimensional
pdf	probability density function
GMM	Gaussian mixture model
DOG	Difference-of-Gaussian
SD	Standard Deviation
MI	Mutual Information

CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 Concealed Weapon Detection Problem

Detection of concealed weapons has been an increasingly important problem for both military and police since global terrorism and crime have grown as threats over the years. The public's concern about the threats of terrorism and crime using weapons has been rising during the past years. Terrorism has become one of the most important concerns for the public worldwide, especially after the deadly terrorist attacks in Paris and San Bernardino, California which caused hundreds of casualties. The Gallup poll released on Dec. 14, 2015 shows that 16% of Americans identified terrorism as the top challenge facing the country which is the highest percentage in a decade, shown in Fig. 1[1]. According to the Country Reports on Terrorism 2014 provided by the National

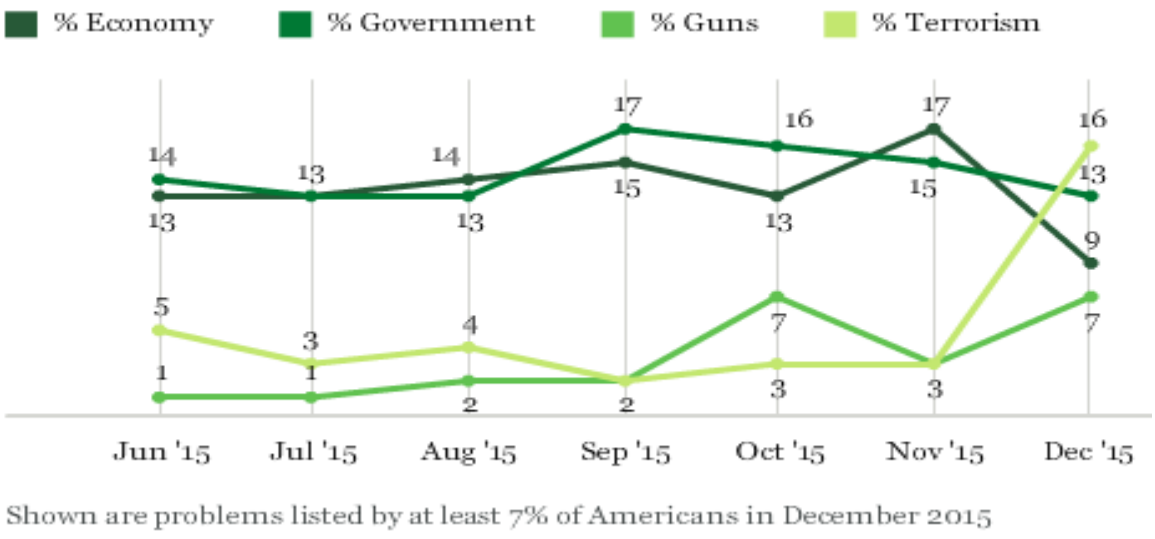

Fig. 1 Most important U.S. problems

Consortium for the Study of Terrorism and Responses to Terrorism [2], the number of terrorist attacks worldwide in 2014 is 13,463, an increase of 35 percentage points since 2013, and the terrorist attacks resulted in more than 67,400 casualties. Another survey released on Apr. 6, 2016 by Gallup indicates that 53% of American adults claimed they worried "a great deal" about crime and violence, hitting the highest point in 15 years, shown in Fig. 2[3]. Nowadays, terrorists and criminals using firearms and explosives have become constant and increasing threats in public areas and are even more harmful comparing with the criminals with white arms like knife. The improving power and the smaller size of weapons bring more challenges to the improvement of the security of the general public as well as the safety of public assets. Therefore, the detection of weapons concealed underneath a person's clothing is one of the greatest and most urgent challenges facing the law enforcement community, crime prevention and anti-terrorism.

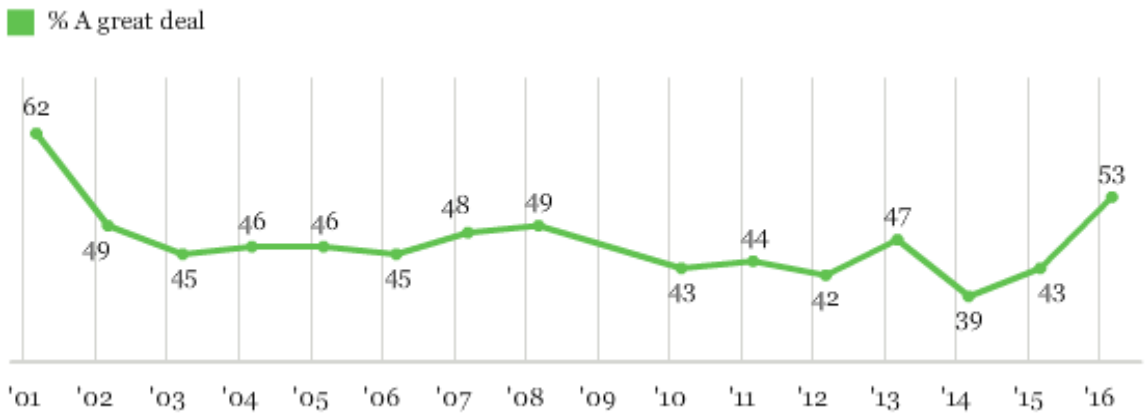

Fig. 2 Americans' level of concern about crime and violence

Easily found in airports and security-sensitive areas, security check points are established to detect concealed weapons like knives, guns and explosives. Traditional devices such as x-ray scanners, hand wands, hand-held metal detector and walk-through metal detectors are commonly used in conjunction with manual screening procedures for the searching of

hidden weapons. The traditional security check procedures have significant limitations despite generally providing great accuracy rate. The disadvantages of the traditional security check procedures involve:

1) Space consuming: devices like x-ray scanners and walk-through metal detectors request huge space for installation, causing reduced feasibility.
2) Time consuming: once a person could be inspected and the manual methods slow the procedures.
3) Labor consuming: huge manpower is needed for the security check.
4) Lack of concealment: the check points are so obvious that the criminals have opportunities to make decision and take action after they discover the existence of the check points.
5) Distance requirement: the effective distances of the traditional devices are limited which could lead to insufficient reaction time for security personnel to safely deal with emergency condition.

Unsafe situation could occur due to the disadvantages. The flow of pedestrian traffic might be impeded in crowded venue like an airport terminal. The crowd trapped in a relatively small area could be a prime target for attack. Also, the safety of security staff is threatened as a result of the insufficient distance. The aforementioned 5 limitations make the traditional methods of Concealed Weapon Detection (CWD) defective and become the motivation of new solutions for CWD problem. A desirable solution should satisfy the following requirements:

1) The hidden weapons are accurately detected.

2) The impacts on pedestrian traffic flow are limited.
3) The CWD procedures are able to take place from a standoff distance.
4) The required space of device installation is small enough to broaden the scope of applications.
5) It is ideally to provide the alarm of suspects carrying weapons without their awareness so the law enforcement is able to take action effectively and safely.
6) The less manpower requirement is preferable.

Thus, image processing methods are considered as efficient and convenient solutions for the CWD problem. Since it is difficult to provide sufficient information with a single sensor in CWD, image fusion has been identified as a key technology to improve CWD algorithms. The image fusion methods for CWD have several merits:

1) The hidden weapons can be clearly detected and recognized with the help of a small amount of personnel.
2) Several targets are able to be inspected simultaneous and the process is automatically carried out by computer program so it is possible to finish the detection procedure rapidly without stopping the pedestrians.
3) The image sensors used enable the detection from a standoff distance and are small enough to be applied in different environment.
4) The image fusion methods are able to provide both the hidden weapon information and the personal identification information so the law enforcement can locate the suspects without drawing their attention.

1.2 Introduction and Literature Review of Image Fusion

Image fusion is the process of combining images acquired using multiple sensors to construct a new image, providing contextual enhancement of the scene being observed. Different fusion algorithms have been proposed for improving spatial and spectral resolutions of the fused images over the decades such as Brovey transform method, Intensity-Hue-Saturation (IHS) method, statistical method, Independent Component Analysis (ICA) method, numerical method and Principal Component Analysis (PCA) method. The algorithms based on a Brovey transform preserve the relative spectral contributions of each pixel but enhance the intensity or brightness component of the image. Each component of the multispectral image bands normalized using a formula is multiplied by a high-resolution co-registered data. The methods based on IHS transformation merge images by preserving most of the spectral information from the H and S components but substituting the intensity image I with histogram-matched high-resolution image. The PCA/ICA method transforms the original images into uncorrelated images and then combines the images by choosing the maximum value among all. PCA is frequently used for fusion as a statistical technique to compact the multivariate data set of inter-correlated variables redundant data into fewer uncorrelated bands.

Image fusion could be performed in four different levels [4], signal level, pixel level, feature level and symbol level, sorted in ascending order of abstraction.

At signal level, image fusion refers to the acquirement of a combined signal of the same type but with greater quality from a group of original sensor signals. Since conclusions are drawn directly from the data and independent of the choice of feature extraction algorithms, Kundur et al. [5] based their blind image restoration and classification techniques on the fusion of individual restorations resulting from single frame algorithms. Xia et al. [6] developed their signal level neural data fusion algorithms based on the Linearly Constrained Least-Squares (LCLS) statistical method.

Pixel level image fusion is a process operated by fusion operators which are decisions depending on individual pixel value or small arbitrary regions of pixels. This kind of techniques makes the fusion decision to increase the information registered in each pixel of the fused images and only takes account of the value of individual pixel or an arbitrary number of surrounding pixels. A feature selection fusion algorithm containing two modes of fusion operation, weighted averaging and selection, of the gradient pyramid representations of the original images has been proposed by Burt et al. [7]. The choice of the modes relies on the local energy match measure which is a local normalized correlation within a neighborhood and is low when more salient feature exists. Li et al. [8] present an image fusion method, based on Discrete Wavelet Frame Transform (DWFT), utilizing Support Vector Machines (SVM) which is superior in finding significant activity at a particular pixel location of the source image. Chandana et al. [9] have suggested a fusion algorithm combining Principal Component Analysis PCA and Discrete Wavelet Transform (DWT) to improve the spatial resolution and spectral content of the fused image. Li [10] have developed a pixel level image fusion method exploiting the

measurement, based on HVS characteristic, of small arbitrary regions of wavelet transform coefficients. This method is designed for remote sensing images.

At feature level fusion, pixels are regarded as making up a feature in an image and features are extracted for image fusion. Segmentation algorithms are applied on the input image first for the extraction of image characteristics and more intelligent semantic fusion rules depending on the activity level measurement of the regions can be developed. Feature level image fusion schemes have the advantages as highlight features and reduced sensitivity to noise. Li et al. [11] have developed their region-based image fusion algorithm using DWT. A fuzzy c-mean clustering is adopted to classify the detail coefficients into regions. A shared multiresolution region representation is then obtained to label all the input images. The fusion process is guided by these multiresolution region representations. Li et al. [12] have proposed a feature based information fusion scheme that spatio-adaptively fuses the complementary information from multiple modalities/channels. Gabor wavelets transform is selected as the multiscale decomposition method for the benefit of feature extraction and characterization measurement. Choose-Max scheme is adopted after applying ICA to the extracted Gabor features for merging the information. A region based approach using Dual Tree Complex Wavelet Transform (DTCWT) for image fusion has been presented by Lewis et al [13]. A combined morphological spectral unsupervised image segmentation algorithm is applied to segment the source images and the region entropy computed using the coefficients of DTCWT high frequency band is calculated as the activity level measurement. The respective regions with maximum entropy are selected to produce the fused image.

The symbol level is the highest abstraction level and features are classified as specific type of symbol. The sets of symbols are then fused to create the fused image. The symbol level image fusion requires classification of feature to make complex decisions and is usually costly. Zhao et al. [14] have proposed a symbol level image fusion approach based on SVM and consensus theory. The classification scheme based on SVM is used to classify the source data and the classification results are fused with consensus theory based fusion rules to obtain the fused image.

Image fusion techniques can also be classified into two group, non-multiscale-decomposition-based (NMDB) fusion methods and multiscale-decomposition-based (MDB) fusion methods [15].

The NMDB techniques, such as adaptive weight averaging (AWA) methods, neural network based methods, Markov Random Field (MRF) based methods and estimation theory based methods, are algorithms without decomposition transform and fusion rules are directly applied to the source images. Lallier et al. [16] have developed an NMDB image fusion scheme for the combination of thermal and visual images based on adaptive weight averaging approach. The weight of a thermal image pixel relies on the divergence of the pixel intensity from the image mean intensity and the local variance in space and time of the visual pixel intensity decide the weight of visual pixel. Xu et al [17] have presented an approach using MRF models to solve the problem of remote sensing image fusion. Both the decision making and true image are modeled as MRF and the Least

Squares technique is used to calculate the fused image. Yang et al. [18] have developed a feature level image fusion algorithm based on estimation theory. The source images are classified into regions by a graph-based image segmentation method and the regions are then analyzed to form a joint region map for the fused image. The region- level expectation-maximization (EM) fusion algorithm is developed to work on the regions modeled using Gaussian mixture distortion to estimate the model parameter and to produce the fused image. Approximate maximum likelihood estimates are produced as the EM algorithm is adopted.

In recent years, multiscale analysis has become one of the most promising methods for image processing and a variety of MDB fusion methods, including pyramid-based methods, DWT-based methods and discrete DWFT-based methods, have been proposed. A MDB fusion method can be generally divided into three steps as shown in Fig. 3.

1) The source images are decomposed into several scale levels using multiscale transform techniques. Low frequency bands containing the approximation coefficients and high frequency bands that consist of detail coefficients are obtained.

2) Different fusion operations are applied to the low/high frequency components to fuse the transform coefficients at each level of the source image.

3) The fused image is produced using the fused coefficients of frequency bands through multiscale reconstruction techniques.

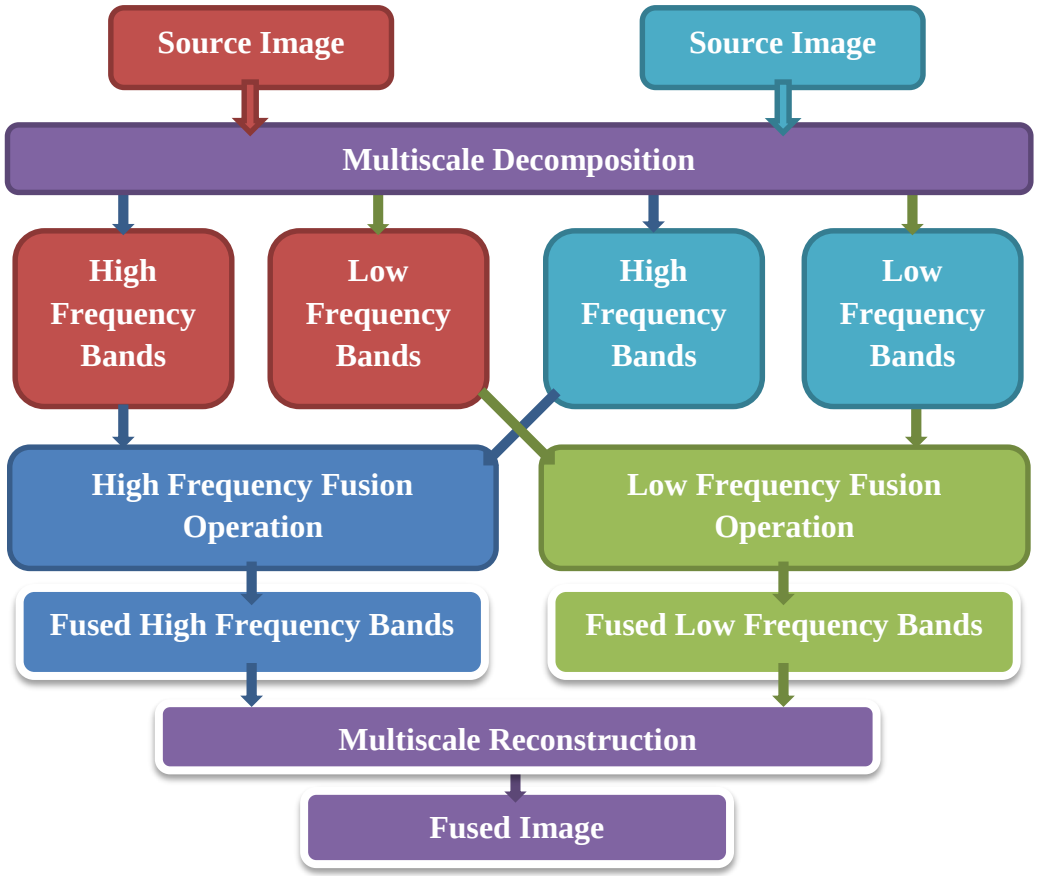

Fig. 3 Block diagram of MDB fusion algorithm

The MDB fusion approaches are able to provide both spatial and frequency domain localization. More intelligent fusion operation can be developed according to the characteristics of the multiscale decomposition representations of the images to achieve much better performance. Burt [19] proposed a multiscale gradient pyramid transform basis which filters and downsamples the input image to produce a sequence of pyramid images representing the input image information at different levels of resolution with the gradient filter for image fusion. Li et al. [20] have suggested a pixel level MDB fusion approach in which the input images are decomposed using DWT. The consistency verification is done along with the activity measurement calculated in a local computing window and maximum selection is adopted as fusion rule. Aiazzi et al. [21] have presented a fusion scheme based on generalized Laplacian pyramid. The method

selectively performs spatial-frequencies spectrum substitution from an image to another and thresholds the local correlation coefficient between the source images utilizing context dependency. Hill et al. [22] have suggested a MDB image fusion framework for complex wavelet transform. The combinations of three fusion operations, maximum selection scheme, weighted average scheme and window based verification scheme, and pre filter complex wavelet transform are implemented to demonstrate the flexibility and efficiency. A MBD fusion algorithm in which the input images are decomposed in multiscale and multi-direction using Non-subsampled Contourlet Transform (NSCT) is proposed by Wang et al. [23]. An accelerated non-negative matrix factorization approach is applied on low frequency bands to generate low frequency fused coefficients and the neighborhood homogeneous measurement rule is employed on the high-frequency components to achieve the fused detail coefficients. Chen et al. [24] have proposed a MDB fusion method based on DDDTCWT. PCA is adopted as the fusion operation of the approximation coefficients of low frequency bands while high frequency band detail coefficients with bigger variance in a window region are selected. Ellmauthaler et al [25] have developed an undecimated wavelet transform based image fusion framework which splits the image decomposition process into two successive filtering operations using spectral factorization of the analysis filters to avoid unwanted side effects like ringing artifacts in the fused reconstruction. Bhatnagar et al. [26] have presented a MDB medical image fusion scheme in which the source images are transformed by NSCT followed by combining low and high frequency bands. Two different fusion operators based on phase congruency and directive contrast are developed and used to fuse low/high-frequency coefficients.

1.3 Image Fusion for CWD

1.3.1 Sensors Choice and Purpose of the CWD Image Fusion Algorithms

For the CWD application, the fusion of a visual image and an image from a special passive sensor indicating concealed weapon is desirable. The term passive refers to the fact that the devices should do not generate or radiate any energy so that it is not harmful for human health. Passive infrared (IR) sensors and millimeter wave (MMW) sensors are preferable among different types of passive imagers since they are able to capture images with clear weapon information from the distance within which the inspection of several persons could be carried out simultaneously without stopping them. Every object with a temperature above absolute zero emits IR radiation which is invisible to the human eye but can be detected by IR sensor proportional to its temperature. A passive IR sensor is an electronic sensor that detects IR light radiation or reflection and converts it into an electronic signal from objects in its field of view. Then the electronic signal is processed to generate a thermal image, where the brightest (warmest) parts of the image are customarily colored white while the dimmest (coolest) parts are colored black. The IR images used for the CWD application are employed with reverse polarity to highlight hidden weapons which appear much darker than the surrounding human body in the original thermal image. The availability of low cost IR technology makes the study of fusing visual and IR images of great interest. A passive MMW sensor also detects the natural thermal energy radiation in environment, much like an IR camera. The thermal energy radiation coming off the body is interrupted by occluding objects which have a

temperature of their own, or reflect other temperatures in the environment, giving rise to contrast in the scene. Compared with IR camera, the most attractive feature of a MMW sensor is that the radiation wavelength of MMW is at W-Band and enables the MMW to penetrate obstacles. However, very sensitive receivers are required to amplify the signal before detection and processing into an image as the total power levels at W-Band are so low.

The purpose of image fusion algorithm for CWD application is to produce a resultant image which allows the viewer to observe:

1) The hidden weapon information from the IR/MMW image. Thus, the interested information from the IR/MMW image is the brighter (white) weapon pixels/regions surrounded by dimmer (black) the human pixels/regions.
2) The appearance information from the visual image. Thus, the target information of visual image is the appearance information of the suspects.

1.3.2 Literature Review of CWD Image Fusion Algorithms

Due to the emerging threats from terrorists and crimes, there is a need to develop an efficient method for heightened security requirements and law enforcement. Thus, a number of multisensor image fusion approaches have been developed to observe objects underneath people's clothing. Zhang et al. [15] have investigated the combination of three multiscale decomposition schems, Laplacian pyramid transform (LPT), DWT and

DWFT, three grouping methods, no-grouping schemes, single-scale grouping scheme and multiscale grouping scheme, two coefficient combining methods, choose-max scheme and weighted average scheme, and four activity level measurement, the coefficient-based activity measures, the weighted average method, the rank filter method and region-based activity measurement. A DWT based fusion algorithm using multiscale grouping scheme and choose max combining method is proposed and the activity level is measured by rank filter method. Varshney et al. [27] have developed an MDB fusion algorithm for CWD application based on DWT. A feature selection scheme using the local energy as the saliency of features has been chosen as the fusion rules. A match measure calculated exploiting the saliency is utilized to decide the two fusion rules, selection and weighted averaging. Yang et al. [28] have proposed a statistical signal processing approach of multisensor CWD image fusion. The source images are first decomposed into multiscale levels using LPT and the multiscale transform coefficients are modeled as the true scene corrupted by additive non-Gaussian distortion modeled using a K-term mixture of Gaussian probability density functions. The EM algorithm is used to estimate the model parameters and produce the fused image. Xue et al. [29] have developed a CWD image fusion algorithm for fusing a color visual image and a corresponding IR image based on DWFT. The RGB visual image is transformed to HSV color space and the V channel representing the intensity of the visual image is used to create two fused images with the IR image and IR image with reverse polarity respectively. The two fused image and the V channel of the visual image are assigned as the R, B and G channel of a new fused RGB color space image which is then transform to LAB color space. The L and A channels of the new fused image are substituted by the L and A channels of the visual image to

produce the final fused color image in LAB color space. Liu et al. [30] have presented a symbol level fusion algorithm based on multiresolution mosaic technique. An unsupervised fuzzy k-means clustering is used to classify the IR/MMW image pixels into regions which are further segmented by a predefined confident threshold. The final detected weapon regions are embedded into the corresponding visual image by using a multiresolution image mosaic technique, which can achieve a seamless boundary between host image and embedded regions. Wang et al. [31] have presented a pixel MDB image fusion approach based on DTCWT for CWD application. The low frequency coefficient measurement relies on the local definition which calculated as the sum of Laplacian operator in the eight directions and the local energy is used to evaluate the high frequency coefficients. The same nonlinear fusion rule depending on the different measure schemes is applied to the low frequency bands and high frequency bands. A pixel level MBD algorithm based on framelet transform also called Double Density Discrete Wavelet Transform (DDDWT) is proposed by Bhatnagar et al. [32]. An AWA fusion strategy based on local energy is used as the low frequency fusion rules and a high frequency fusion operation is developed exploiting the HVS characteristics of image texture. Upadhyaya et al. [33] have proposed a multi-exposed image fusion algorithm for CWD application. The homomorphic filter is applied to the source images to enhance the information content in the images and the source images are divided into small blocks. The Shannon entropy is calculated in each block and the blocks with highest Shannon entropy are selected and blended to produce the fused image.

CHAPTER 2

DOUBLE DENSITY DUAL TREE COMPLEX WAVELET TRANSFORM

2.1 Discrete Wavelet Transform

The DWT is a multiscale transform which has the ability to decompose the input signal into a multiresolution representation using a set of analyzing functions which are dilations and translations of a few wavelet functions and perfectly reconstruct the input signal. The forward and inverse 1-dimensional (1D) DWT are implemented using two-channel analysis filter bank and synthesis filter bank respectively, as shown in Fig. 4. The

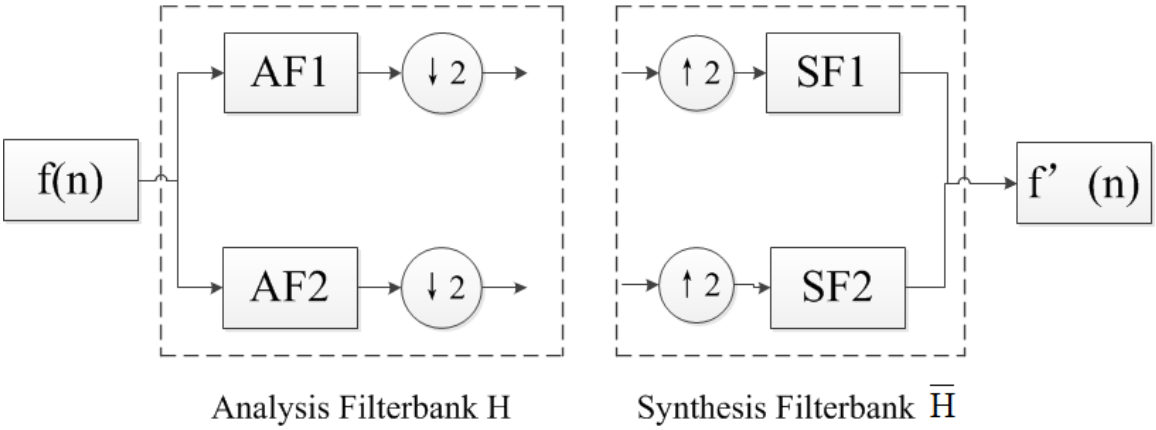

Fig. 4 1D DWT analysis and synthesis filter banks

DWT employs one scaling function $\phi(t)$, and one wavelet $\psi(t)$. The AF1 in Fig.4 denotes the low-pass filter associated with the scaling function and the AF2 is the high-pass filter associated with the wavelet function. AF1 is the quadrature mirror filter of AF2 of which the high-pass amplitude response is a mirror image of the low-pass amplitude response with respect to the middle frequency. Decomposition and

reconstruction based on the wavelet transform consists of recursively applying the 2-channel analysis filter bank on the low-pass output and performing the 2-channel synthesis filter bank on the reconstructed low-pass input and the original high-pass output at the same scale level, shown in Fig. 5.

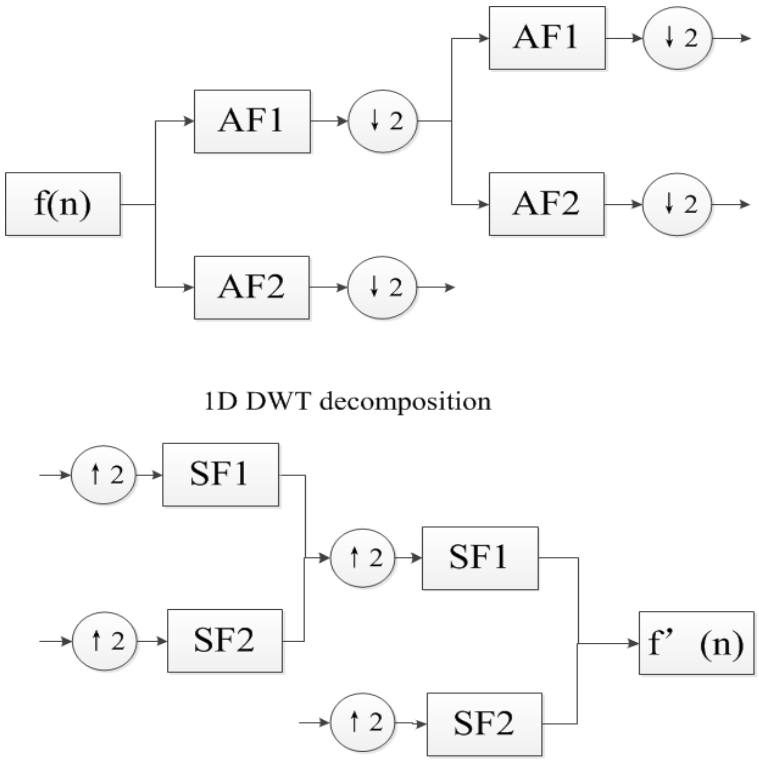

Fig. 5 1D DWT multiscale decomposition and reconstruction

For the applications of image processing, 2-dimensional (2D) DWT is required. The filtering of 2D DWT decomposition is done first by convolving the input image with the digital analysis filter bank in the vertical direction and then downsampling the output along the column. The two output images are further processed along the row followed by down-sampling. The source image is decomposed into four subbands, LL, LH, HL and HH, as shown in Fig.6. The coefficients in LL band which is the low-pass band

contain the approximation information and the coefficients in the other three high

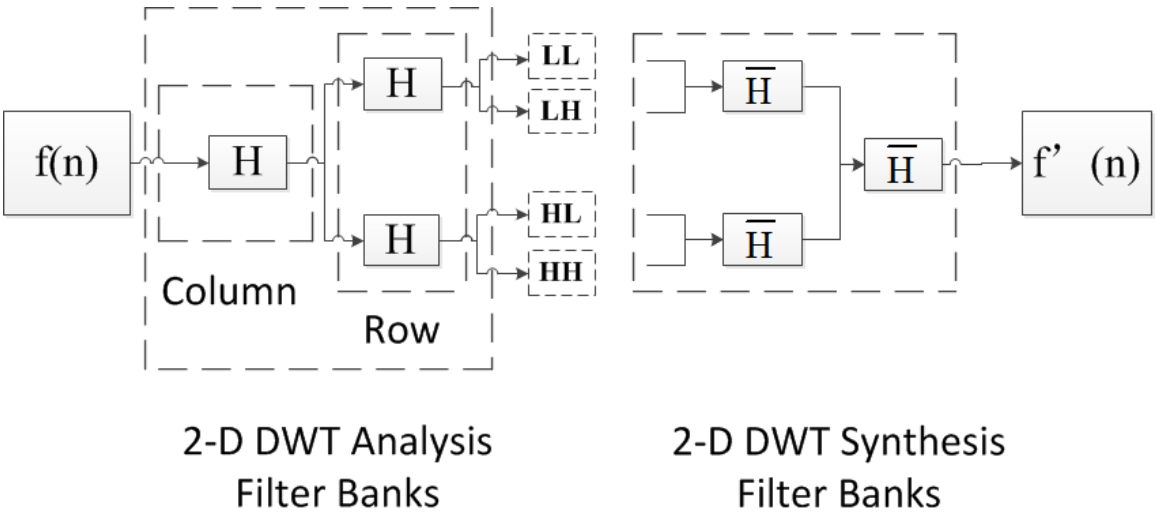

Fig. 6 2D DWT

frequency subbands contain the detail information on different direction. Then the same process can be repeatedly applied to the column and row of the low-pass output bands to produce the multiscale representation of the source image. The 2D DWT is widely used in the image processing applications such as image fusion, noise attenuation, image enhancement and motion detection. The wavelet transform comes in different forms. The critically-sampled form of the wavelet transform provides the most compact representation. However, it has four main limitations:

1) Oscillations: the wavelet coefficients tend to oscillate positive and negative around singularities making singularity extraction and signal modeling challenging as the wavelets are band-pass functions.

2) Shift variance: the coefficient oscillation pattern around singularities is disturbed remarkable by a small shift of the input signal.

3) Aliasing: the repeatedly filtering of the signal with non-ideal low-pass and high-pass filters and the down-sampling operations result in substantial

aliasing which lead to artifacts in the reconstruction. The inverse DWT can only cancel the aliasing when there is no change on the coefficients which is unrealistic for processing.

4) Lack of directionality: the standard tensor product construction of wavelets produces a checkerboard pattern that is simultaneously oriented along several directions. DWT lack the capability of distinguishing orientations in multiple dimensions which is significant in image processing.

2.2 Double Density Discrete Wavelet Transform

Motivated by the recognition of the limitations of DWT, DDDWT [34] have been proposed both of which are able to provide the merits such as nearly shift-invariance. The DDDWT is based on a single scaling functions $\phi(t)$, and two distinct wavelets $\psi_i(t)\ i = 1,2$ where the two wavelets are designed to be offset from one another by one half. Assuming the low-pass and high-pass filters associated with $\phi(t)$, $\psi_1(t)$ and $\psi_2(t)$ are $\mathrm{h}_0(\mathrm{n})$, $\mathrm{h}_1(\mathrm{n})$ and $\mathrm{h}_2(\mathrm{n})$ respectively. The definition of scaling and wavelet functions is given by:

$$\phi(t) = \sqrt{2}\sum_n h_0(n)\phi(2t-n) \tag{2.1}$$

$$\psi_i(t) = \sqrt{2}\sum_n h_i(n)\phi(2t-n)\,,\ i = 1,2 \tag{2.2}$$

The filter banks proposed in Fig. 7 illustrates the implementation of the DDDWT. Three analysis filters make up the analysis filter bank of a 1D DDDWT. The low-pass filter

$h_0(-n)$ associate with the scaling function and the two distinct high-pass filter are denoted by $h_1(-n)$ and $h_2(-n)$. The input signal is decomposed into three subbands which are then down-sampled by 2. The lowpass subband contains the coarser

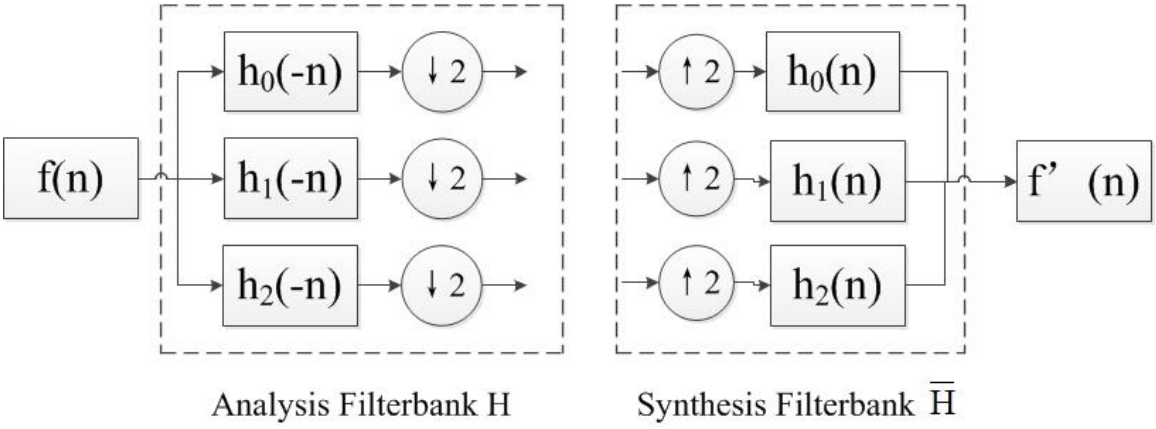

Fig. 7 DDDWT analysis and Synthesis filter banks

approximation information while the high-pass subbands keep the detail information of the input signal.

To use the DDDWT for image processing, the 2D DDDWT can be implemented simply by alternatively applying the 1D DDDWT filter bank structure first to the rows, then to the columns of an image, as shown in Fig. 8. This gives rise to nine subbands, one of which is the low-pass subbands, and the other eight of which make up the eight detail subbands. Afterwards, the same process can be repeated on the lowpass subbands to form the DDDWT multiscale decomposition.

Although the DDDWT exploits more wavelets and has the advantages as nearly shift-invariance and better wavelet smoothness, improvement is required as it lack the ability

of isolating the two diagonal orientations. Indicated in Fig. 9, the eight detail subbands

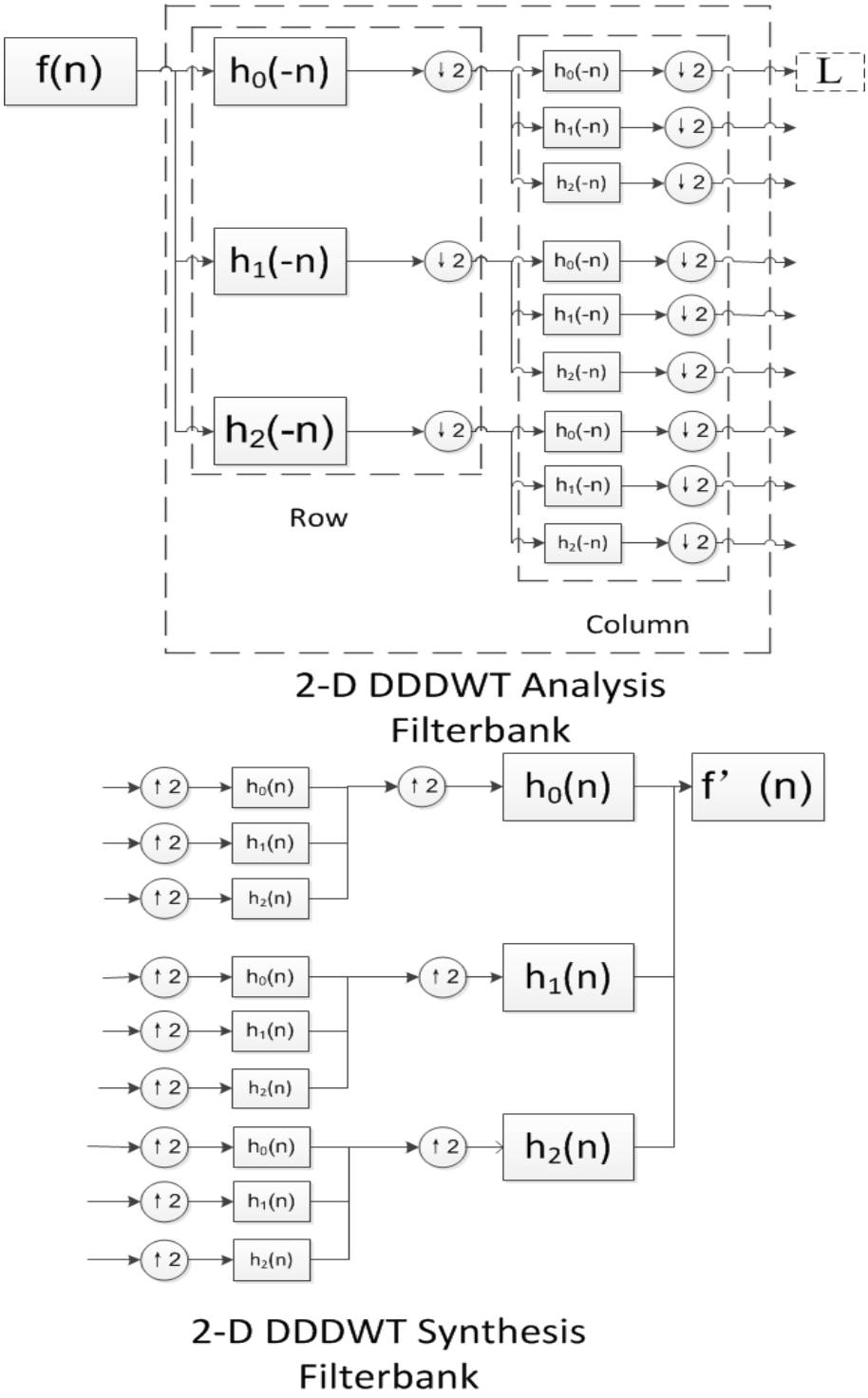

Fig. 8 2D DDDWT filter structure

correspond to eight wavelet components of which two are oriented in the vertical

direction while another two are oriented in the horizontal direction. The rest of the wavelets do not have specific orientations, but rather combine the two diagonal

Fig. 9 The wavelet orientations of DDDWT

orientations, which lead to the checkerboard affect.

2.3 Dual Tree Complex Wavelet Transform

The 1D DTCWT [35] is an expansive wavelet transform and implemented using two critically-sampled DWTs in parallel on the input signal, as shown in Fig. 10. The

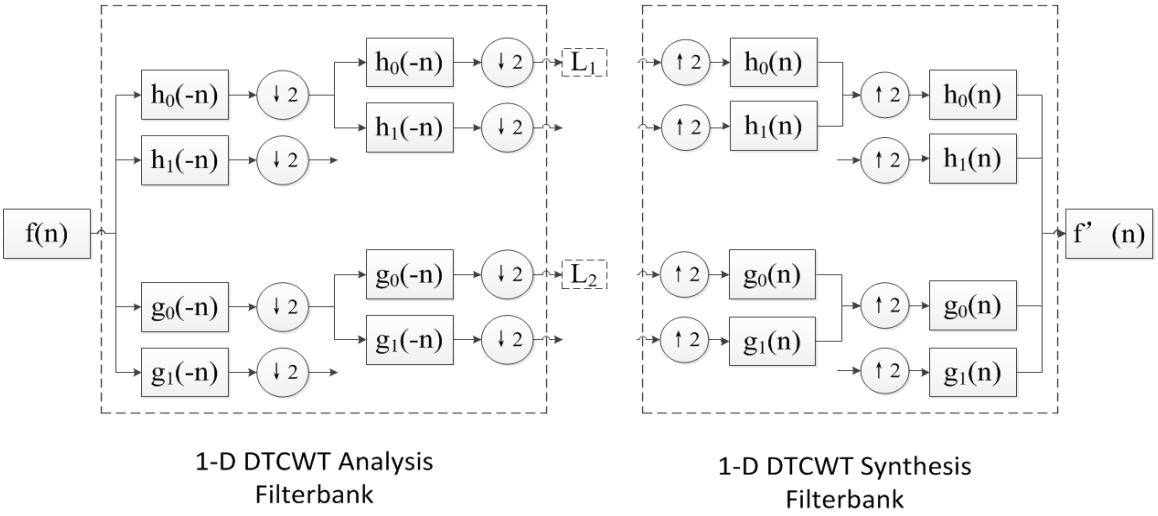

Fig. 10 1D DTCWT

DTCWT has two scaling functions $\phi_h(t)$ and $\phi_g(t)$, and two distinct wavelets $\psi_h(t)$ and $\psi_g(t)$, forming two critically-sampled DWTs. Designed according to specific

standards, the filters of the upper DWT are acting as the real part of a complex wavelet transform while the lower DWT filter can be interpreted as the imaginary part. Equivalently, for specially designed sets of filters, the wavelet associated with the upper DWT can be an approximate Hilbert transform of the wavelet associated with the lower DWT. If the low-pass and high-pass filters associated with $\phi_h(t)$, $\phi_g(t)$, $\psi_h(t)$ and $\psi_g(t)$ are $h_0(n)$, $g_0(n)$, $h_1(n)$ and $g_1(n)$ respectively. The scaling and wavelet functions are defined as:

$$\phi_h(t) = \sqrt{2}\sum_n h_0(n)\phi_h(t) \tag{2.3}$$

$$\phi_g(t) = \sqrt{2}\sum_n g_0(n)\phi_g(t) \tag{2.4}$$

$$\psi_h(t) = \sqrt{2}\sum_n h_1(n)\phi_h(t) \tag{2.5}$$

$$\psi_g(t) = \sqrt{2}\sum_n g_1(n)\phi_g(t) \tag{2.6}$$

$$\psi_g(t) \approx \mathcal{H}\{\psi_h(t)\} \tag{2.7}$$

where $\mathcal{H}$ denotes the Hilbert transform.

The 2D DTCWT is implemented using four critically-sampled separable 2D DWTs in parallel. The recursively applying of the two pairs of DWT analysis and synthesis filter banks, illustrated in Fig. 11, make up the four critically-sampled separable 2D DWTs of which two perform as real part transform and the other act as imaginary part transform. The filter structure of the 2D DTCWT is proposed in Fig. 12. The 2D DTCWT gives rise to 16 frequency bands of which four are low frequency bands including the coarser approximation information and the rest 12 are high frequency bands containing detail

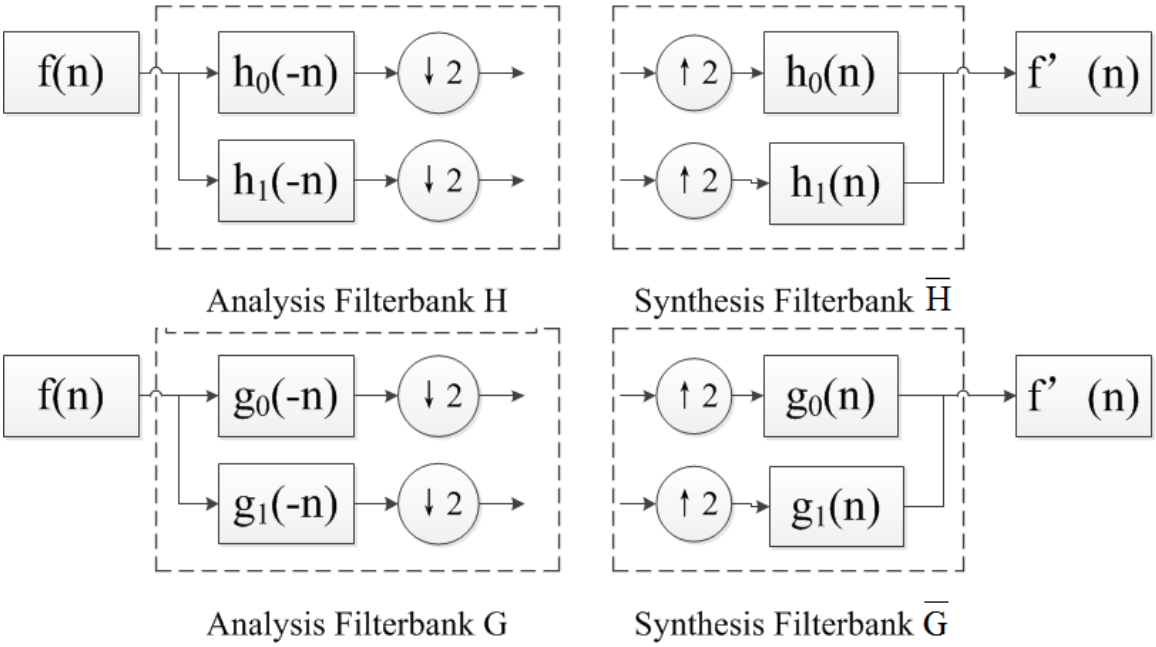

Fig. 11 DWT filter banks for DTCWT

information. The 12 high frequency bands associate with 12 wavelets which are oriented

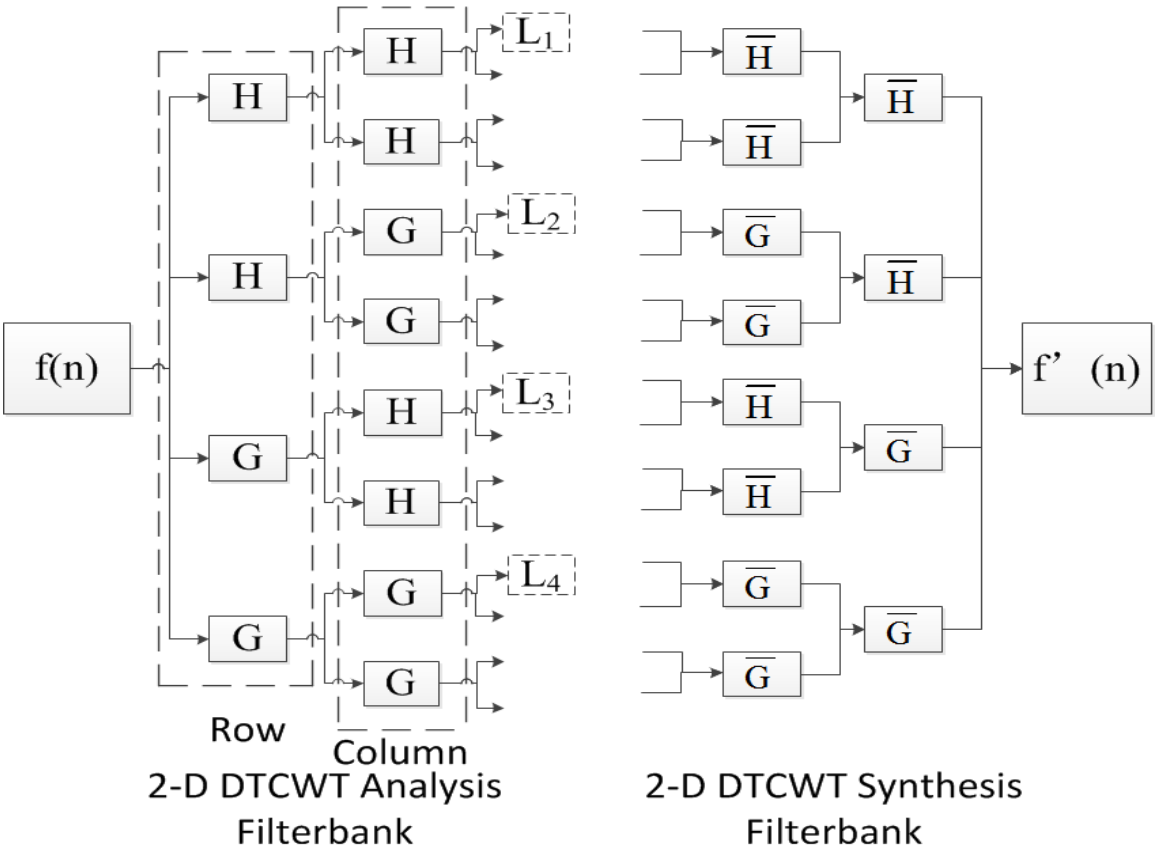

Fig. 12 2D DTCWT filter bank structure

in six directions. Every two wavelets are oriented in the same direction and one of the two wavelets performs as the real part of a complex value 2D wavelet while the other wavelet can be deem as the imaginary part of the complex-valued 2D wavelet, shown in Fig. 13.

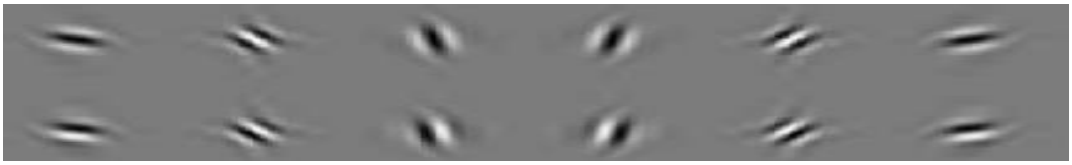

Fig. 13 The wavelet orientations of DTCWT

The 2D DTCWT is nearly shift-invariant and do not have an oscillatory behavior. Moreover, all of the wavelets associated with it are oriented which is desirable for image processing.

2.4 Double Density Dual Tree Complex Wavelet Transform

The DDDTCWT [36] has been developed as complex wavelets transform combining the DTCWT and the DDDWT and possessing their merits and characteristics. The implementation of DDDTCWT is based on two pairs of analysis and synthesis filter banks which form two DDDWT structure respectively, as shown in Fig. 14. The DDDTCWT employs two scaling functions $\phi_h(t)$ and $\phi_g(t)$, and two pairs of distinct wavelets: $\psi_{h,i}(t)$, $\psi_{g,i}(t)$, $i = 1,2$. Let $h_i(n)$ and $g_i(n)$ denote the four high-pass wavelet filters associated with the wavelets, $\psi_{h,i}(t)$, $\psi_{g,i}(t)$, $i = 1,2$. The two low-pass

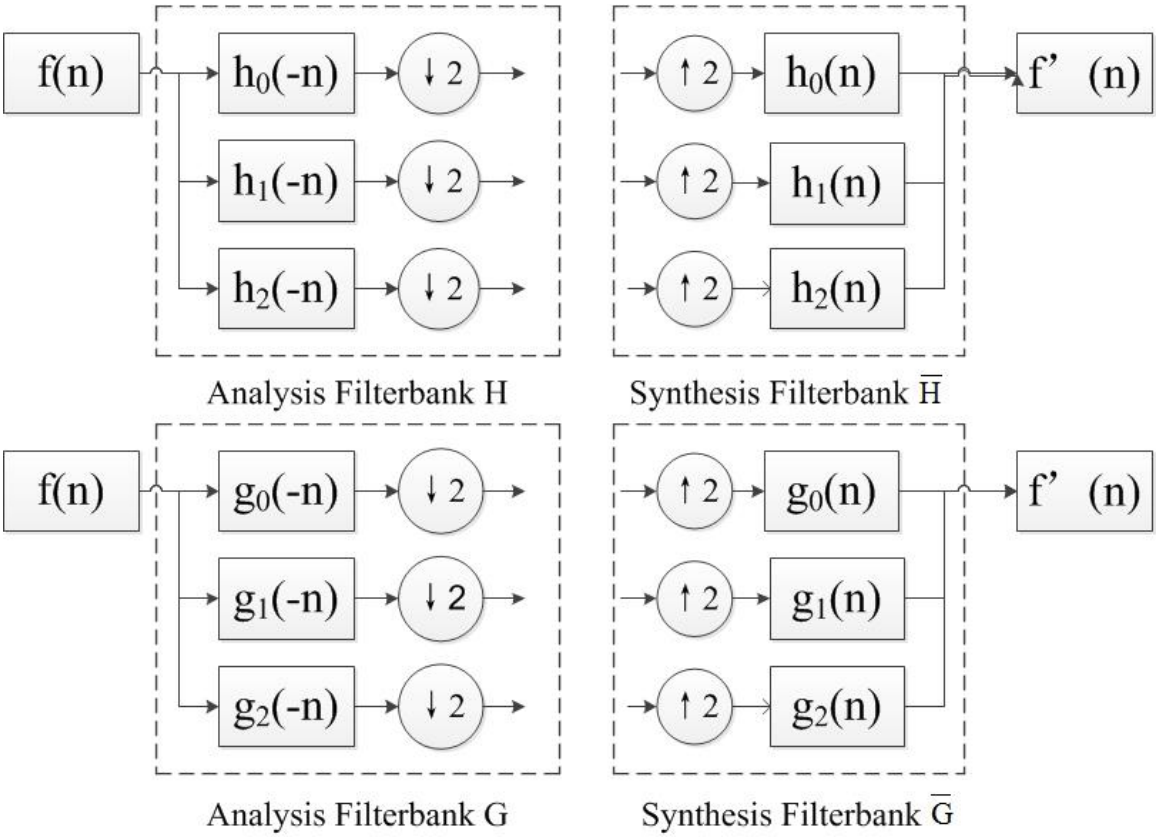

Fig. 14 Basic analysis and synthesis filter banks of DDDTCWT

scaling filters associated with the scaling functions, $\phi_h(t)$ and $\phi_g(t)$, are denoted as $h_0(n)$ and $g_0(n)$.

To keep both of the properties of the DDDWT and the DTCWT, the two wavelets from the same DDDWT are designed to be offset from one another by one half, as are the filters designed for DDDWT:

$$\psi_{h,1}(t) \approx \psi_{h,2}(t - 0.5) \tag{2.8}$$

$$\psi_{g,1}(t) \approx \psi_{g,2}(t - 0.5) \tag{2.9}$$

The two corresponding wavelets from different DDDWTs form approximate Hilbert transform pairs, as do the filters required for DTCWT:

$$\psi_{g,1}(t) \approx \mathcal{H}\{\psi_{h,1}(t)\} \tag{2.10}$$

$$\psi_{g,2}(t) \approx \mathcal{H}\{\psi_{h,2}(t)\} \tag{2.11}$$

The design of the filters of DDDTCWT implementation enables the complex and directional wavelet transforms. The scaling functions and wavelets are defined through the dilation and wavelet equations:

$$\phi_h(t) = \sqrt{2}\sum_n h_0(n)\phi_h(2t-n) \tag{2.12}$$

$$\phi_g(t) = \sqrt{2}\sum_n g_0(n)\phi_g(2t-n) \tag{2.13}$$

$$\psi_{h,i}(t) = \sqrt{2}\sum_n h_i(n)\phi_h(2t-n),\ i = 1,2 \tag{2.14}$$

$$\psi_{g,i}(t) = \sqrt{2}\sum_n g_i(n)\phi_g(2t-n),\ i = 1,2 \tag{2.15}$$

As shown in Fig.15, the 1D DDDTCWT filter bank structure is based primarily on concatenating two oversampled iterated DDDWT filter banks in parallel on the input signal. The process gives rise to six subbands of which two are low-pass subbands containing approximation information, indicated as L_1 and L_2 in Fig. 15. The multiscale representation of 1D DDDTCWT is obtained by repeatedly applying the two DDDWT filter banks on the two low-pass outputs at each scale level.

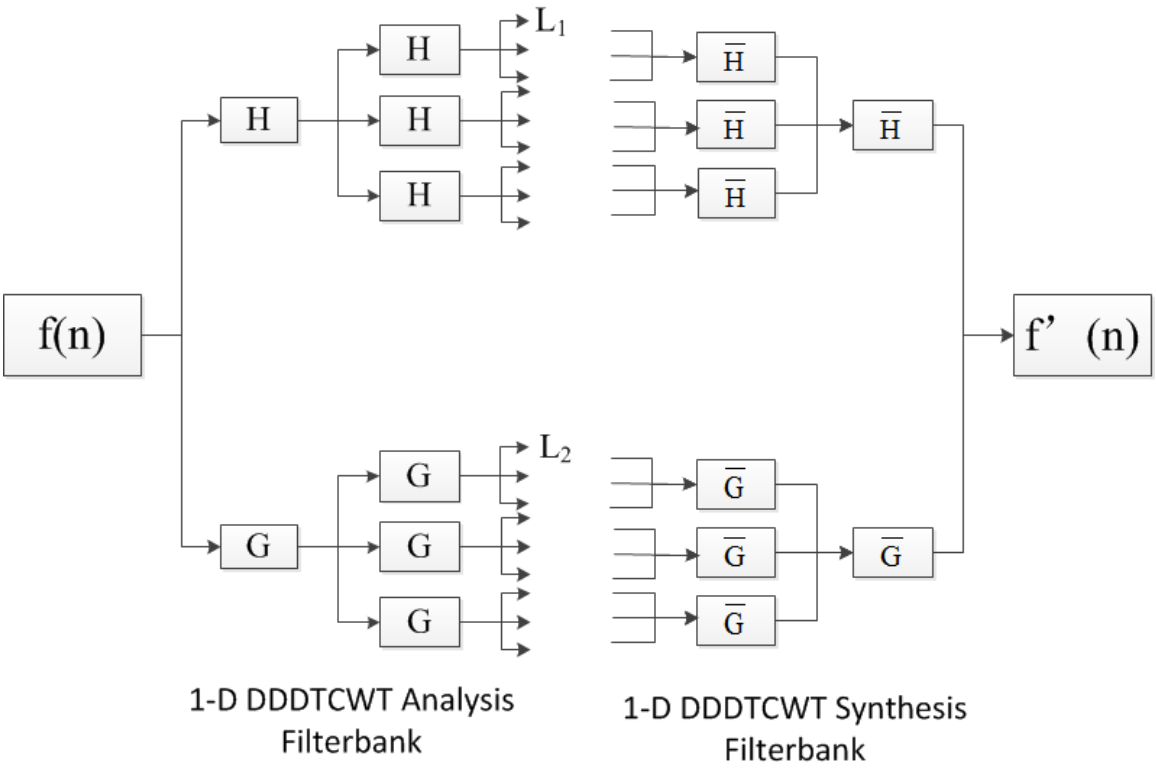

Fig. 15 1D DDDTCWT filter bank structure

The 2D DDDTCWT is implemented by employing four oversampled 2D DDDWT in parallel to the same input image. The 2D DDDTCWT analysis and synthesis filter bank structure is shown in Fig. 16.

The implementation of the four oversampled 2D DDDWTs is realized using the combinations of the basic filter banks shown in Fig. 14. The filter banks are applied to the rows and columns of the image data alternately in the DDDWTs. Indicated in Fig. 14, the two oversampled analysis filter banks, denoted by H and G, and the two synthesis filter banks, denoted by $\overline{H}$ and $\overline{G}$, are constructed and the synthesis filters are the time-reversed versions of the analysis filters. The 2D DDDTCWT gives rise to 36 frequency bands in

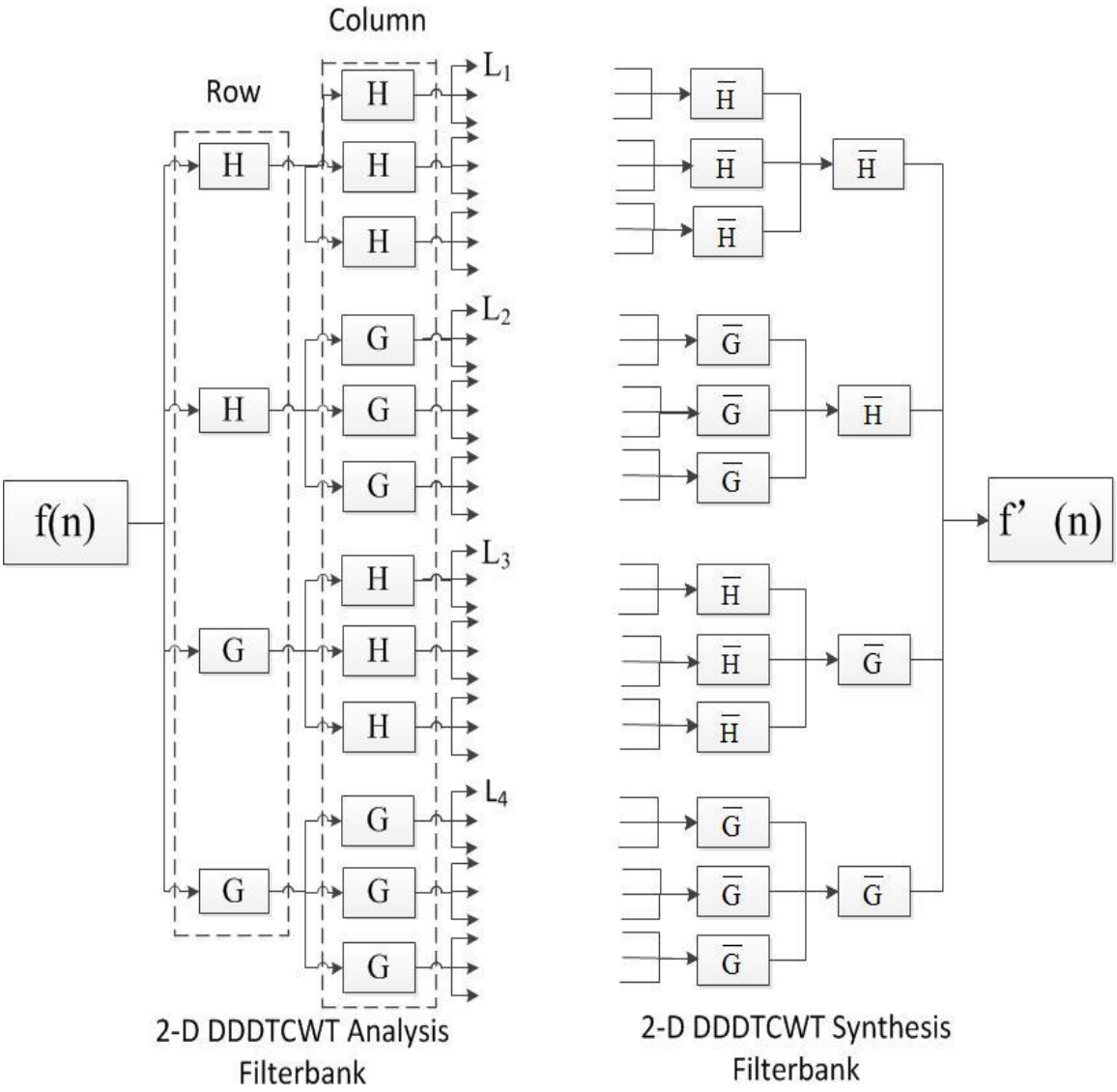

Fig. 16 2D DDDTCWT filter bank structure

each level. 4 of the frequency bands are the low frequency bands containing approximation information and rest make up the 32 high frequency bands with oriented edge and texture information. The 32 wavelets associated with the 32 high frequency bands are oriented in 16 distinct directions and there are 2 wavelets oriented in each direction, as shown in Fig. 17. For each direction, one wavelet acts as the real part of a complex-valued 2D wavelet function, while the other can be interpreted as the imaginary

part. Therefore, the DDDTCWT possesses better directional selectivity than both the DDDWT and the DTCWT.

Fig. 17 The wavelet orientations of DDDTCWT

CHAPTER 3

PIXEL LEVEL IMAGE FUSION FOR CWD

3.1 Procedure of the Pixel Level Image Fusion Algorithm

The pixel level image fusion methods are popular since the cost of the schemes is commonly low and the algorithms are easy to implement. One pixel level image fusion algorithm has been developed for the application of CWD. The block diagram of the algorithm is illustrated in Fig. 18.

The algorithm is implemented in 3 steps:

1) The source images from the visual sensor and IR/WWM sensor are decomposed using the DDDTCWT. The multiscale representation of the source images in the form of low frequency bands and high frequency bands are obtained.

2) Two different fusion operations are applied to the low frequency bands and high frequency bands respectively and the fused low and high frequency bands are produced. The low frequency fusion rule is a feature selection scheme based on local contrast and the high frequency fusion rule is also a feature selection scheme with the measurement based on the HVS characteristics and the characteristics of the DDDTCWT.

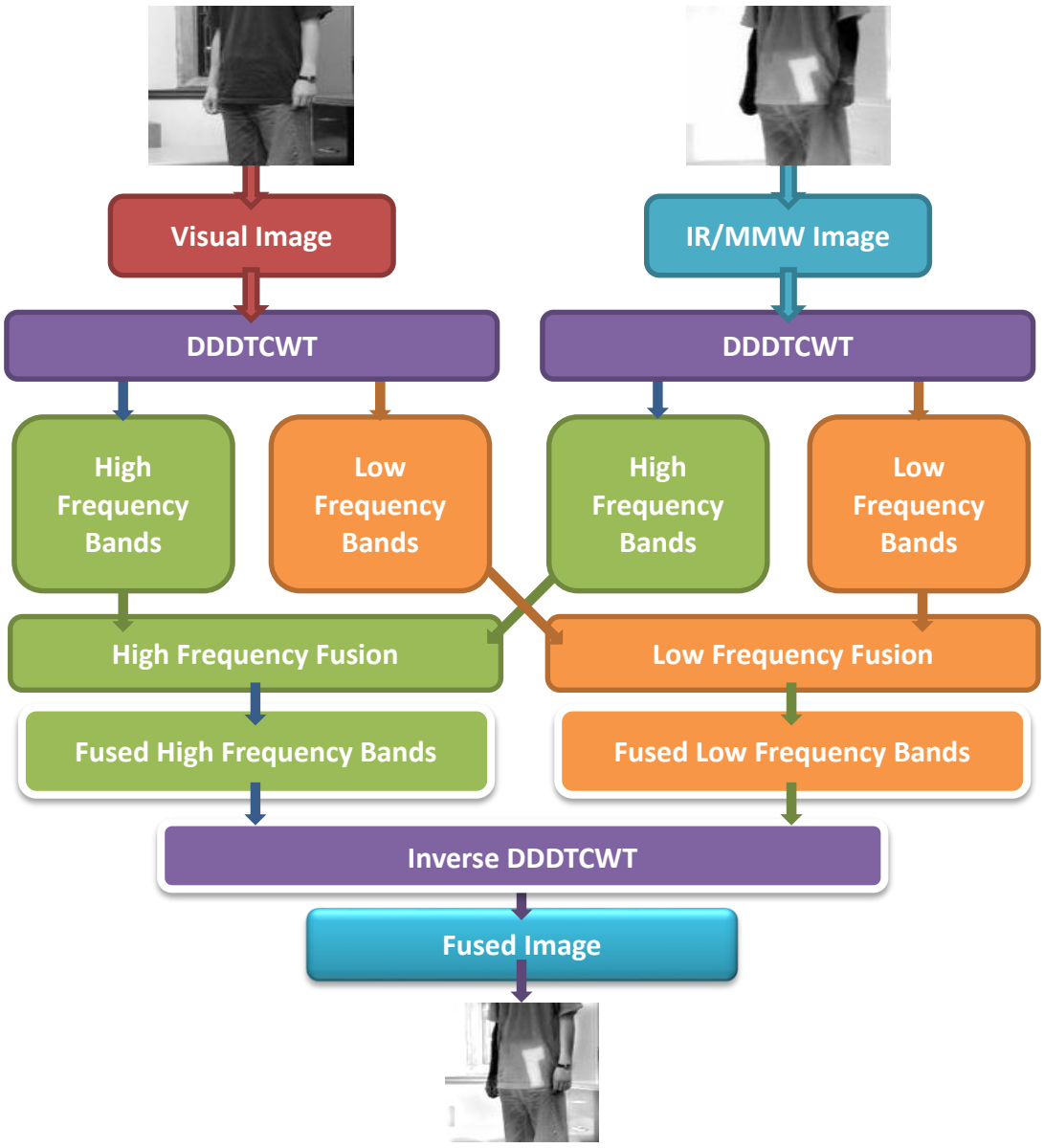

Fig. 18 Block diagram of pixel level image fusion algorithm

3) The fused image is created through the inverse DDDTCWT using the fused low and high frequency bands.

Experiments and comparisons demonstrate the robustness and efficiency of the proposed approach and indicate that the fusion rules can be applied to different multiscale

transformations. Also, it shows that the fusion result using the proposed fusion rules on DDDTCWT is superior to other combinations as well as previously proposed approaches in literature.

3.2 The Low Frequency Fusion Rule

The low frequency bands of the DDDTCWT of an image reflect the coarser approximation of the original image. The low frequency bands are similar to a compressed version of the source image with blurred texture and edge. Averaging is a convenient method to fuse the approximation coefficients of the low frequency bands and maintain the reasonable mean intensity for the fused image. However, averaging results in missing important information and reduced contrast. Thus, a low frequency fusion rule is developed based on the local contrast of the corresponding low frequency coefficients.

The core aim of the algorithm is to produce a high quality fused image that displays both the hidden weapons and the face and figure characteristics of the person carrying the weapon for the convenience of observers. The preferable information is the concealed weapon information in the IR/MMW image and the personal identification information in the visual image. In the IR/WWM low frequency subbands, the more bright regions consisting of the larger value approximation coefficients indicate the concealed weapons and correspond to the cloth region commonly made up of the approximation coefficients with smaller value (darker) in the visual images. In the visual image low frequency

subbands, the interested coefficients containing the appearance information of the suspects are stored in the area of which the corresponding IR/MMW regions consist of the lower value (darker) coefficients indicating human body. Therefore, the preferable regions are those with larger brightness difference.

For the multiscale DDDTCWT representation of an image, only the four low frequency bands at the highest decomposition level $\boldsymbol{l}$ are taken into consideration as the low frequency bands at other levels are produced using the four bands in the inverse DDDTCWT. The corresponding low frequency bands of the visual image and IR/WWM image are denoted by f_{1,L_i}^{l} and f_{2,L_i}^{l}, for $i = 1,2,3,4$ representing the four low frequency bands. To preserve the pertinent regions of the IR/WWM subbands, the normalized approximation coefficients $f_{1,L_i}^{l,N}$ and $f_{2,L_i}^{l,N}$ are used for measurement calculation:

$$f_{k,L_i}^{l,N}(m,n) = \frac{f_{k,L_i}^{l}(m,n) - min\left(f_{k,L_i}^{l}\right)}{max\left(f_{k,L_i}^{l}\right) - min\left(f_{k,L_i}^{l}\right)}, k = 1,2 \tag{3.1}$$

where (m,n) is the position of the current coefficient.

The local contrast of the approximation coefficients used as the low frequency band fusion measurement is calculated as:

$$C_i(m,n) = \frac{S_{1,i}(m,n) - S_{2,i}(m,n)}{S_{1,i(m,n)} + S_{2,i(m,n)}} \tag{3.2}$$

where $S_{k,i}, k = 1,2$ representing the sum of local normalized coefficients are used to eliminate block effect. $S_{k,i}, k = 1,2$ are calculated in a calculation window centered at (m,n):

$$S_{k,i}(m,n) = \sum_{a\in S_p}\sum_{b\in S_p} f_{k,L_i}^{l,N}(m+a,n+b) \tag{3.3}$$

The size of the local calculation window is $S_p \times S_p$.

The fusion rule of the low frequency band is a feature selection scheme, based on the local contrast measure $C_i(m,n)$, which has two modes. When the contrast of the two corresponding coefficients is larger than a threshold T, a maximum chosen scheme is adopted and the coefficient with larger value is selected. In this case, the selected coefficient is high possible to contain the preferable information. If the contrast of the two corresponding coefficients is smaller than the threshold T, an averaging scheme combining the value of the coefficients is used to maintain the reasonable mean intensity for the fused image.

The fusion rule of the low frequency band is defined as:

$$f_{f,L_i}^{l}(m,n) = k_1^i \times f_{1,L_i}^{l}(m,n) + k_2^i \times f_{2,L_i}^{l}(m,n) \tag{3.4}$$

Where k_1^i and k_2^i are the weights following the rule:

$$k_1^i + k_2^i = 1 \tag{3.5}$$

k_1^i is defined as:

$$k_1^i = \begin{cases} 1 & if\ C_i(m,n) \geq T \\ 0 & if\ C_i(m,n) \leq -T \\ 0.5 & if -T < C_i(m,n) < T \end{cases} \tag{3.6}$$

there T is a positive threshold.

3.3 The High Frequency Fusion Rule

3.3.1 Noise Visibility Function

HVS research provides mathematical models regarding how human eyes respond to the visual stimulus of the world. Lots of research works on HVS have been proposed and applied in image processing applications as the HVS is an excellent image processor capable of detecting and recognizing image information. HVS-based model [37] was first introduced in image compression algorithms. Nowadays, different HVS models have been widely adopted in various research areas such as image enhancement, digital watermarking, image segmentation, motion estimation and image fusion. Noise Visibility Function (NVF), based on the noise visibility of an image, is one of the popular models and modelled by Voloshynovskiy [38]. NVF characterizes the local properties and indicates edge, texture and flat regions. Modelling the original image as a random variable with stationary generalized Gassain probability density function (pdf), the NVF at each pixel position can be calculated as:

$$NVF(m,n) = \frac{w(m,n)}{w(m,n)+\theta\delta^2} \tag{3.7}$$

where (m,n) indicates the pixel location and δ^2 is the global variance of the image. θ is a tuning parameter which must be chosen for every particular image and computed as:

$$\theta = \frac{D}{{\delta^2}_{max}} \tag{3.8}$$

where D is an experimentally determined parameter and chosen as 60. ${\delta^2}_{max}$ is the maximum value of the local variance ${\delta^2}_l$ calculated in an $L \times L$ window as:

$${\delta^2}_l(m,n) = \frac{1}{(2L+1)^2}\sum_{k=-L}^{L}\sum_{l=-L}^{L}\big(x(m+k,n+l) - \bar{x}(m,n)\big)^2 \tag{3.9}$$

With $\bar{x}(m,n)$ is the local mean of pixel:

$$\bar{x}(m,n) = \frac{1}{(2L+1)^2}\sum_{k=-L}^{L}\sum_{l=-L}^{L} x(m+k,n+l) \tag{3.10}$$

$w(m,n)$ is computed as:

$$w(m,n) = \frac{\gamma[\eta(\gamma)]}{\|r(m,n)\|^{2-\gamma}} \tag{3.11}$$

where γ is the shape parameter of the Gaussian distribution and is typically in the range from 0.3 to 1 for real images. The parameter γ depends on the decomposition level and can be estimated using a maximum likelihood estimate method [39].

$\eta(\gamma)$ and $r(m,n)$ are calculated as:

$$r(m,n) = \frac{x(m,n) - \overline{x}(m,n)}{\delta_x} \tag{3.12}$$

$$\eta(\gamma) = \sqrt{\frac{\Gamma(3/\gamma)}{\Gamma(1/\gamma)}} \tag{3.13}$$

$\Gamma(.)$ is the standard gamma function and defined as:

$$\Gamma(t) = \int_0^\infty e^{-u}u^{t-1}du \tag{3.14}$$

It is clear by definition that NVF characterizes the local image properties and identifies texture, edges and flat regions. The smaller NVF values indicate texture and edge areas while the higher NVF values represent the flat regions.

3.3.2 The High Frequency Fusion Rule

The DDDTCWT high frequency band coefficients of an image contain most of the edge and texture information. The higher value of a coefficient represents a more significant change of the pixel values and indicates the edge or texture information.

The purposed of the high frequency band fusion is to better preserve the edge and texture information. As NVF is a measure of the edge and texture region based on local window calculation, the local energy of the detail coefficients can act as an additional factor to further improve the performance. A high frequency fusion rule is developed base on the combination measure of NVF and local energy. The core idea is to use the local energy weighted NVF as the texture masking function to extract all texture and edges from all high-frequency bands and select most prominent texture and edges for fused images.

For the DDDTCWT of an image, 32 high frequency bands are produce at each decomposition level. The high frequency bands from all decomposition levels are fused under the same rule. Thus, given a detail subband f_{1,H_i}^{θ} and its corresponding detail subband f_{2,H_i}^{θ} at decomposition level θ and $i = 1,2,\dots,32$ indicating different subbands of the same level, the measurement of a high frequency band coefficient located in (m,n) is defined as:

$$D_{f_{k,H_i}^{\theta}}(m,n) = E_{f_{k,H_i}^{\theta}}(m,n) \times \left(1 - NVF_{f_{k,H_i}^{\theta}}(m,n)\right) \tag{3.15}$$

where $NVF_x(m,n)$ represents the NVF value of the coefficient calculated in a $S_p \times S_p$window centred at (m,n) in the frequency band x and $E_x(m,n)$ is the local energy weight computed in the same window of the subband x:

$$E_x(m,n) = \sum_{a \in S_p} \sum_{b \in S_p} [x(m+a, n+b)]^2 \tag{3.16}$$

The larger energy of a detail coefficient represents more detail information.

The fusion rule of the high frequency band is also a feature selection scheme, based on the local energy weighted NVF value, which has two modes. The texture and edge information from whichever image is desirable. When the measurement value of the two corresponding coefficients is different, a maximum chosen scheme is adopted and the coefficient with larger measurement value is selected. This means the coefficient with more texture and edge information is chosen. Only if the measurement value of the two corresponding coefficients is equal, an averaging scheme combining the value of the coefficients is applied. The fusion rule of detail coefficients can be interpreted as:

$$f^{\theta}_{f,H_i}(m,n) = k_3^i \times f^{\theta}_{1,H_i}(m,n) + k_4^i \times f^{\theta}_{2,H_i}(m,n) \tag{3.17}$$

Where k_3^i and k_4^i are the weights following the rule:

$$k_3^i + k_4^i = 1 \tag{3.18}$$

k_3^i is defined as:

$$k_3^i = \begin{cases} 1 & if\ D_{f^{\theta}_{1,H_i}}(m,n) > D_{f^{\theta}_{2,H_i}}(m,n) \\ 0 & if\ D_{f^{\theta}_{1,H_i}}(m,n) < D_{f^{\theta}_{2,H_i}}(m,n) \\ 0.5 & if\ D_{f^{\theta}_{1,H_i}}(m,n) = D_{f^{\theta}_{2,H_i}}(m,n) \end{cases} \tag{3.19}$$

CHAPTER 4

FEATURE LEVEL IMAGE FUSION FOR CWD

4.1 Procedure of the Feature Level Image Fusion Algorithm

The feature level image fusion methods are less sensitive to noise and the utilization of the feature characteristics leads to more intelligent fusion schemes improving the performance of the algorithms. A feature level image fusion approach has been developed for the application of CWD. The block diagram of the algorithm is illustrated in Fig. 19.

The algorithm includes 4 steps:

1) The source images from the visual sensor and IR/WWM sensor are decomposed using the DDDTCWT. The low frequency bands and high frequency bands of the source images are produced.

2) Multiscale image segmentation scheme based on GMM is applied to the multiscale representation of the source images to create the label matrix for visual and IR/WWM image respectively.

3) Two different fusion operations based on the region saliency according to the label matrixes are developed to fuse the low frequency bands and high frequency bands.

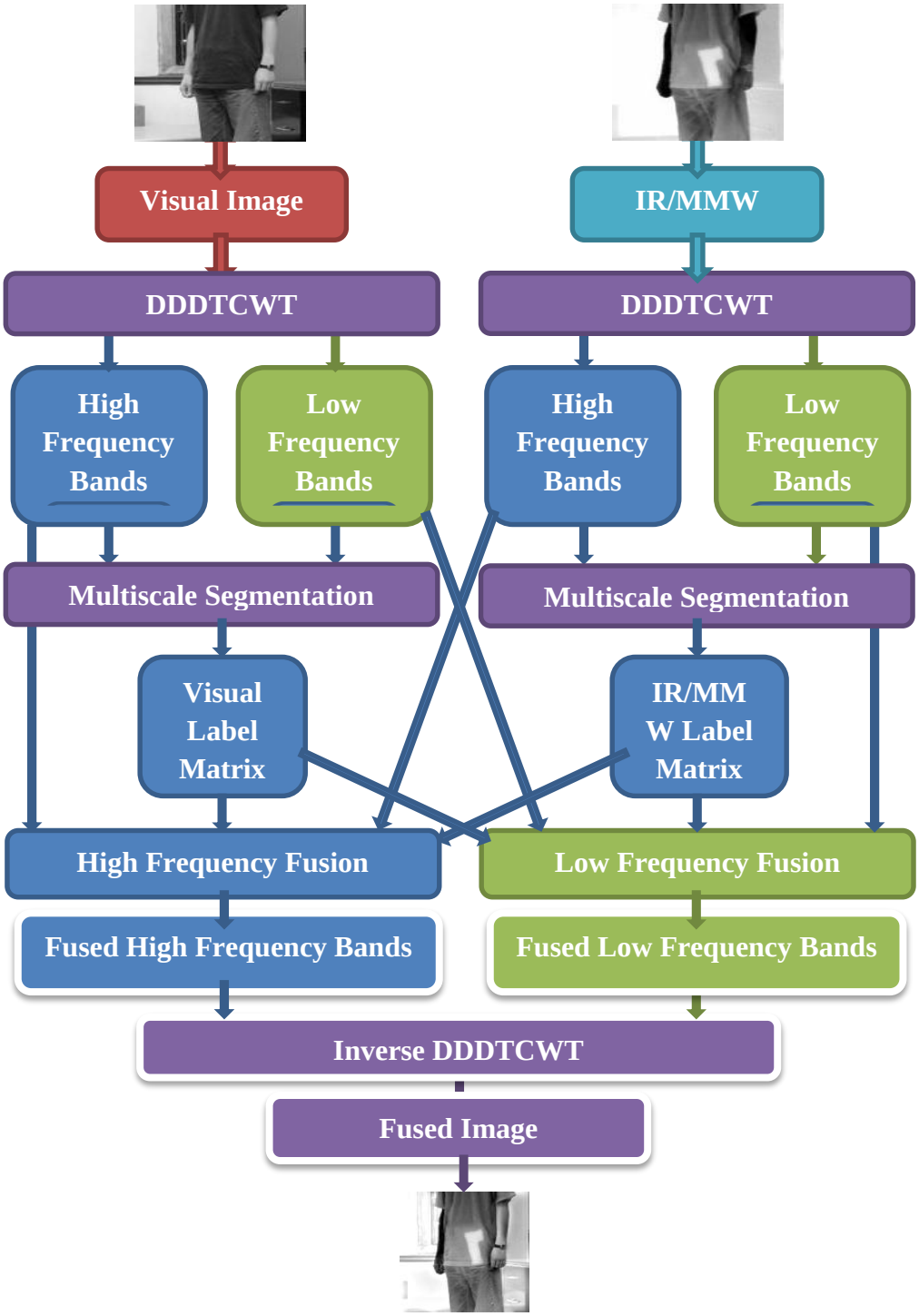

Fig. 19 Block diagram of feature level image fusion algorithm

4) The fused image is created through the inverse DDDTCWT using the fused low and high frequency bands.

Experiments and comparisons demonstrate the efficiency of the developed approach and indicate that the feature level image fusion owns advantages on feature preservation.

4.2 Image Segmentation Based on GMM

GMMs are flexible and powerful statistical models that assume all the data points are characterized by a mixture of a finite number of Gaussian distributions with unknown parameters. The GMM pdf is represented as a weighted sum of component Gaussian densities, as shown in Fig. 20. Each component density of GMM is a Gaussian distribution.

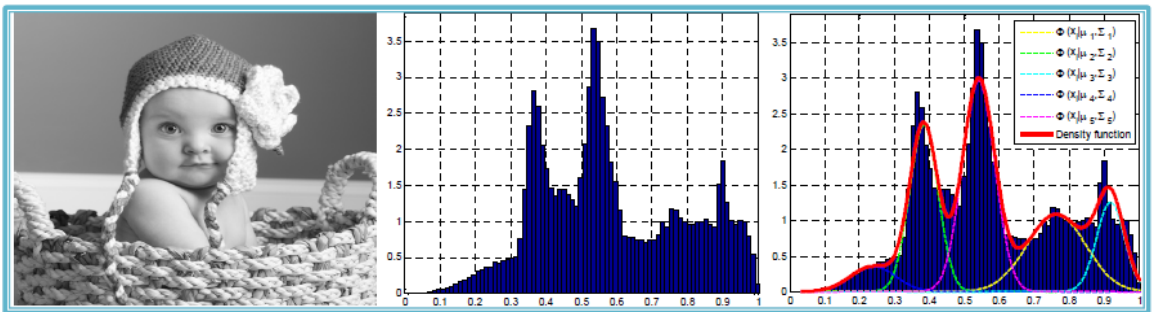

Fig. 20 GMM pdf of a image

The GMM is a well-known probabilistic model which has been widely used in applications, such as image fusion, image sequence analysis, image compressions and image segmentation, due to its simplicity and ease of implementation. In the feature level image fusion algorithm, a GMM segmentation algorithm based on MRF [40] is used to classify the DDDTCWT coefficients into regions and produce the label matrixes. The segmentation algorithm incorporates spatial relationships among neighboring pixels in a

simpler metric to make the scheme fast and easy to implement. An EM algorithm is directly applied to optimize the parameters.

The main objective of the GMM based segmentation approach is to classify the image pixels into *K* labels. Let $x_i\ i == (1,2,\ldots,N)$, present an observation at the i^{th} pixel of an image with dimension D. The neighborhood of the i^{th} pixel is denoted by δ_i. Consider there are *K* random sources, labeled as $\Omega_j, j = (1,2,\ldots,K)$, of which each characterized by a Gaussian pdf:

$$\Phi(x_i|\Theta) = \frac{1}{2\pi^{D/2}|\Sigma_j|^{1/2}} exp\left\{-\frac{1}{2}\left(x_i - \mu_j\right)^T {\Sigma_j}^{-1}\left(x_i - \mu_j\right)\right\} \quad (4.1)$$

where Σ_j denotes the $D \times D$ covariance matrix and $|\Sigma_j|$ is the determinant of Σ_j. μ_j denotes the *D*-dimension mean vector and $\Theta_j = \{\mu_j, \Sigma_j\}$. The density function at an observation x_i is given by:

$$f(x_i|\Pi, \Theta) = \sum_{j=1}^{K} \pi_{ij} \Phi(x_i|\Theta) \quad (4.2)$$

where $\Pi = \pi_{ij}, i = (1,2,\ldots,N), j = (1,2,\ldots,K)$ is the set of prior distributions modeling the probability that pixel x_i belongs to the label Ω_j, which satisfies the constraints:

$$0 \leq \pi_{ij} \leq 1\ and\ \sum_{j=1}^{K} \pi_{ij} = 1 \quad (4.3)$$

In order to classify an image consisting of *N* pixels into *K* labels, the standard finite GMM assumes that each observation x_i is modeled as statistically independent. The joint conditional density of the data set $X = (x_1,\ x_2\ \ldots\ x_N)$ is given by:

$$p(X|\Pi,\Theta) = \prod_{i=1}^{N} f(x_i|\Pi,\Theta) = \prod_{i=1}^{N}\sum_{j=1}^{K}\pi_{ij}\Phi(x_i|\Theta) \quad (4.4)$$

Taking the spatial correlation between the neighboring pixels into consideration, the MRF distortion is applied to reduce the sensitivity of noise and illumination. The posterior probability density function can be written as:

$$p(\Pi,\Theta|X) \propto p(X|\Pi,\Theta)p(\Pi) \quad (4.5)$$

where $p(\Pi)$ is the MRF distortion:

$$p(\Pi) = Z^{-1}exp\left\{-\frac{1}{T}U(\Pi)\right\} \quad (4.6)$$

Z denotes a normalizing constant and T is a temperature constant. $U(\Pi)$ is the smoothing function:

$$U(\Pi) = -\sum_{i=1}^{K}\sum_{j=1}^{K} G_{ij}^{(t)} \log {\pi_{ij}}^{(t+1)} \quad (4.7)$$

where $G_{ij}^{(t)}$ is a factor G_{ij} based on the posterior probability z_{ij} and π_{ij} at the tth iteration step and calculated as:

$$G_{ij}^{(t)} = \exp\left\{\frac{\beta}{2N_i}\sum_{m\in\delta_i}\left(z_{mj}^{(t)} + \pi_{mj}^{(t)}\right)\right\} \quad (4.8)$$

where β is the temperature value controlling the smoothing prior and set to 12. δ_i is a square 5×5 window and N_i denoting the number of pixel in δ_i is equal to 25. The log-likelihood function can be written as:

$$L(\Pi,\Theta|X) = \log p(\Pi,\Theta|X)$$

$$= \sum_{i=1}^{N}\log\left\{\sum_{j=1}^{K}{\pi_{ij}}^{(t+1)}\Phi\left(x_i\middle|\Theta^{(t+1)}\right)\right\} - \log Z + \frac{1}{T}\sum_{i=1}^{K}\sum_{j=1}^{K} G_{ij}^{(t)} \log {\pi_{ij}}^{(t+1)} \quad (4.9)$$

where Z and T are set to 1. Maximizing the log-likelihood function will lead to an increase in the value of the objective function:

$$J(\Pi, \Theta|X) = \sum_{i=1}^{N}\sum_{j=1}^{K} z_{ij}^{(t)}\left\{\log {\pi_{ij}}^{(t+1)} + \log \Phi\left(x_i\middle|\Theta^{(t+1)}\right)\right\} +$$

$$\sum_{i=1}^{K}\sum_{j=1}^{K} G_{ij}^{(t)} \log {\pi_{ij}}^{(t+1)} \qquad (4.10)$$

The segmentation is then realized by applying the EM scheme to optimize the parameters in order to maximize the objective function $J(\Pi, \Theta|X)$. When the parameter-learning phase is complete, the determination of the classification of each pixel x_i is made based on the optimized posterior probability:

$$x_i \in \Omega_j : if\ z_{ij} \geq z_{ik},\ k = (1,2,\dots,K) \qquad (4.11)$$

Each pixel is assigned to the label with the largest posterior probability z_{ij}.

The EM algorithm procedure is summarized as follows and shown in Fig. 21:

1) Initialize the parameters$\{\Pi, \Theta\}$ using a k-mean algorithm: the prior distributions π_{ij}, the mean μ_j and the covariance Σ_j.
2) E step: Evaluate the value z_{ij} and update the factor G_{ij}.
3) M step: update the parameters $\{\Pi, \Theta\}$.
4) Estimate the log-likelihood and check the convergence of either the log-likelihood function or the parameter value. If the convergence criterion is not satisfied, then go to step 2. If the convergence criterion is satisfied, then the parameter-learning phase is complete.

The DDDTCWT representation of an image at every scale level can be deemed as an image with 36 dimensions. Thus the GMM segmentation algorithm is applied to the subbands of the source image at each level to create the label matrixes. In order to separate the weapon coefficients from the background coefficients, then the region map of the IR/MMW subbands is relabeled to ensure that the regions are continuous.

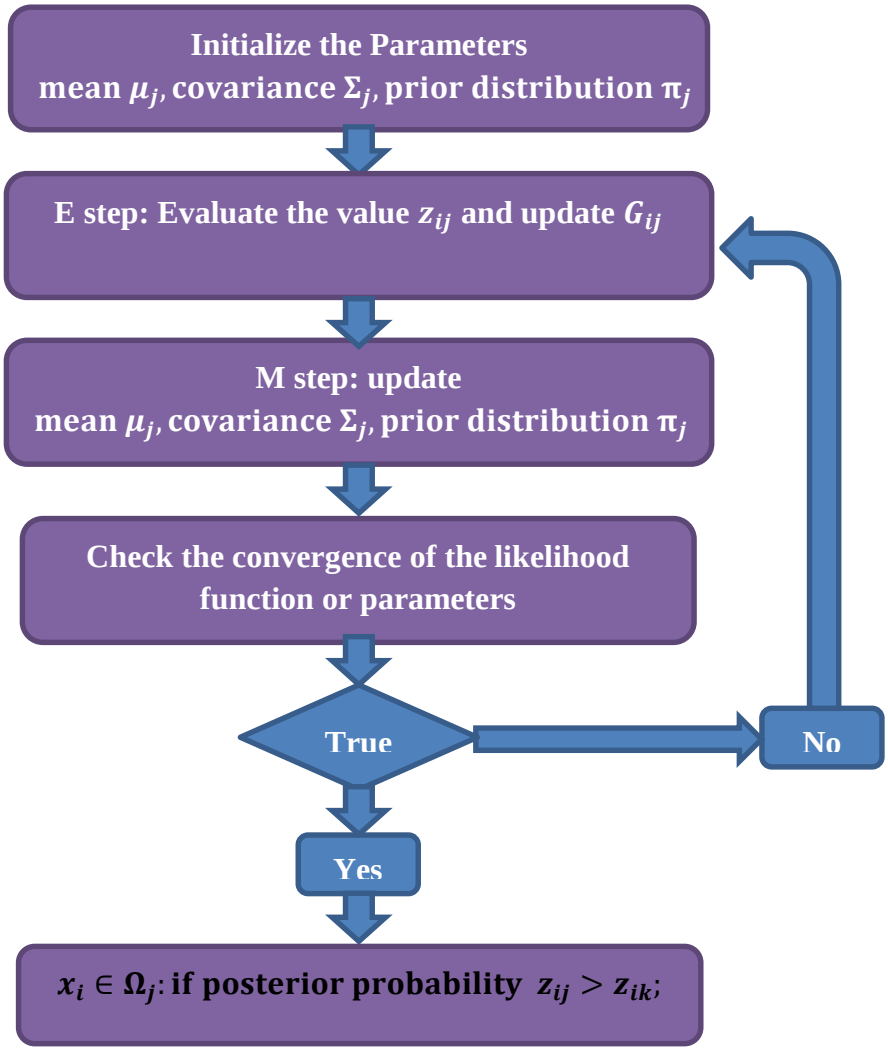

Fig. 21 the block diagram of EM algorithm

4.3 Feature Level Low Frequency Fusion Rule

The preferable regions for the CWD application are the region containing concealed weapon information from the IR/MMW image and the areas that the personal identification information is stored from the visual image. In the feature level low frequency image fusion scheme, the region saliency of low frequency band is used as measurement.

Humans have the ability of locating the most information part of an image. This ability of HVS is called visual attention. Computational methods for modeling the visual attention provide the saliency map which has the capability to indicate the regions or objects standing out from the rest of the image. Thus, saliency map is a measurement of the informative degree of an image. Saliency detection methods can be roughly categorized into bottom-up and top-down approaches [41]. The top-down schemes are task-dependent methods based on knowledge about the scenes, objects, face and their interrelations. By contrast, bottom-up approaches are data-driven and easy to implement and understand. Thus, bottom-up algorithms have a wider application prospect and have been extensively adopted. A bottom-up region saliency detection method is developed as the measurement of the feature level image fusion approach.

Let $\Omega_i^v, i = (1,2,\ldots,N)$ and $\Omega_j^s, j = (1,2,\ldots,K)$ denote the segmentation labels of the highest level visual image subbands and IR/MMW image subbands, respectively. The region color mean of a low frequency band can be calculated as:

$$sc_n^k(l) = \frac{1}{N_l}\sum_{x_n^k(r,c)\in\Omega_l^k} x_n^k(r,c) \qquad k \in \{v,s\}, l \in \{i,j\} \tag{4.12}$$

where $k \in \{v,s\}$ distinguishes the source image and $l \in \{i,j\}$ indicates the region labels. $x_n^k(r,c)$ is an approximation coefficient located at (r,c) and N_l is the total number of the coefficients belong to the label Ω_l^k. $n \in \{1,2,3,4\}$ denotes the 4 low frequency bands of a source image at the highest level. The spatial position $sp_n^k(l)$ of a region labeled as Ω_l^k is defined as the center location of the region. One factor of the saliency measure is created based on the color contrast cue that the salient stimulus should be distinct from its neighborhood. Thus, the color contrast factor is designed to be proportional to the color contrast of the region and its neighbor regions:

$$cons_n^k(a)\_o = \alpha\sum_{a\neq b} exp\left\{-d\left(sp_n^k(a), sp_n^k(b)\right)\right\} \times d\left(sc_n^k(a), sc_n^k(b)\right) \tag{4.13}$$

where $d(x,y)$ is the normalized Euclidean distance between x and y. α is a weight to better preserve the weapon information and calculated as:

$$\alpha = \begin{cases} \frac{sc_n^k(a)}{\mu_n^k} & k = s \\ 1 & k = v \end{cases} \tag{4.14}$$

Where μ_n^k is the intensity mean of the low frequency band. Another saliency factor is designed according to the color distribution cue that the salient objects are commonly centered and compact:

$$ndiss_n^k(a)\_o = \beta\sum_{a\neq b} d\left(sp_n^k(a), sp_n^k(b)\right) \tag{4.15}$$

where β is a parameter to measure the size of the region:

$$\beta = \frac{N_a}{N_n^k} \tag{4.16}$$

N_a is the total coefficients of the region and N_n^k is the number of the total coefficients of the subband. The two factor $cons_n^k(a)\_o$ and $ndiss_n^k(a)\_o$ are then normalized to [0, 1], denoted by $cons_n^k(a)$ and $ndiss_n^k(a)$. The region saliency is computed as:

$$sal_n^k(a) = cons_n^k(a) \times \left(1 - ndiss_n^k(a)\right) \tag{4.17}$$

A saliency map S_n^k can be created as the coefficients belong to the same region own the same saliency value. Therefore, the fusion rule of the low frequency bands:

$$f_n^f(r,c) = k_1^i \times f_n^s(r,c) + k_2^i \times f_n^v(r,c) \tag{4.18}$$

where $f_n^s(r,c)$, $f_n^v(r,c)$ denote the corresponding approximation coefficients from the IR/MMW and visual image. k_1^i and k_2^i are the weights following the rule:

$$k_1^i + k_2^i = 1 \tag{4.19}$$

k_1^i is defined as:

$$k_1^i = \begin{cases} 1 & if\ S_n^s(r,c) - S_n^v(r,c) \geq T \\ 0 & if\ S_n^s(r,c) - S_n^v(r,c) \leq -T \\ 0.5 & if\ -T < S_n^s(r,c) - S_n^v(r,c) < T \end{cases} \tag{4.20}$$

where T is a positive threshold.

4.4 Feature Level High Frequency Fusion Rule

The DDDTCWT high frequency band coefficients of an image contain most of the edge and texture information. The better detail information preservation is desirable for the high frequency band fusion.

For the feature level fusion of high frequency bands, a region detail information measure scheme has been developed. The frequency tuned saliency algorithm [42] is used to generate the saliency map for the high frequency bands. A Difference-of-Gaussian (DOG) filter is applied to the high frequency bands to extract the low level features. The variance of the DOG filter is specifically designed to preserve the edge and texture information. Let $\Omega_{v,m}^{L,(i)}\ i = (1,2,\dots,N)$ and $\Omega_{s,m}^{L,(j)}\ j = (1,2,\dots,K)$ denote the labels of the visual image subbands and IR/MMW subbands, respectively. L is the decomposition level and $m = (1,2,\dots,32)$ represents the 32 high frequency bands. A saliency map for a high frequency band $f_{k,m}^{L}$ at level L is obtained as:

$$sal_{k,m}^{L} = DOG(f_{k,m}^{L}) \tag{4.21}$$

where $k \in \{v,s\}$ distinguishes the source images. The region saliency of a high frequency band region labeled by $\Omega_{k,m}^{L,(l)}$ is given by:

$$SAL_{k,m}^{L}(l) = \frac{1}{N_{k,m}^{L,(l)}} \sum_{x_{k,m}^{L}(r,c) \in \Omega_{k,m}^{L,(l)}} sal_{k,m}^{L}(r,c) \tag{4.22}$$

where $l \in \{i,j\}$ indicate the label and $N_{k,m}^{L,(l)}$ is the number of the coefficients labeled by $\Omega_{k,m}^{L,(l)}$. $x_{k,m}^{L}(r,c)$ is a detail coefficient. The region detail information measurement is defined as:

$$M_{k,m}^{L}(l) = E_{k,m}^{L}(l) \times SAL_{k,m}^{L}(l) \tag{4.23}$$

where $E_{k,m}^{L}(l)$ is the average energy of the subband region:

$$E_{k,m}^{L}(l) = \frac{1}{N_{k,m}^{L,(l)}} \sum_{x_{k,m}^{L}(r,c) \in \Omega_{k,m}^{L,(l)}} x_{k,m}^{L}(r,c)^2 \tag{4.24}$$

The coefficients of the same region have the same measurement value, leading to a measurement map $MP_{k,m}^{L}$. The high frequency fusion rule is defined as:

$$f_{f,m}^{L}(r,c) = k_3^i \times f_{s,m}^{L}(r,c) + k_4^i \times f_{v,m}^{L}(r,c) \tag{4.25}$$

where k_3^i and k_4^i are the weights following the rule:

$$k_3^i + k_4^i = 1 \tag{4.26}$$

k_3^i is defined as:

$$k_3^i = MP_{s,m}^{L}(r,c) \Big/ \Big(MP_{s,m}^{L}(r,c) + MP_{v,m}^{L}(r,c) \Big) \tag{4.27}$$

CHAPTER 5

EXPERIMENTAL RESULTS AND ANALYSIS

5.1 Quality Metrics

Better image fusion algorithms should preserve complimentary features from source images and should not introduce artifacts and inconsistencies. Therefore, several objective quality metrics have been used to estimate the quality of the images. Higher metrics values indicate better image quality.

1) Standard Deviation (SD): SD is able to characterize the dispersion of the pixel values of an image:

$$SD = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}|x(i,j) - \overline{x}|/MN} \tag{5.1}$$

where $x(i,j)$ is an image pixel, M and N are the size of the image and $\overline{x}$ is the mean value of the image pixels.

2) Entropy: The entropy measures the information presented in an image and is defined as follows:

$$H = -\sum_{i=0}^{L} p(i) log_2 p(i) \tag{5.2}$$

where L is the number of grey levels, $p(i)$ is the corresponding probability of grey level i of the image.

3) Mutual Information (MI): MI is a measurement indicating the degree of dependence of two images. When two images are independent, the value of MI is zero.

$$MI_{AB} = \sum_{a=0}^{L-1} \sum_{b=0}^{L-1} h_{AB}(a,b) log_2 \frac{h_{AB}(a,b)}{h_A(a) h_B(b)} \quad (5.3)$$

where $h_{AB}(a,b)$ is the normalized joint grey level histogram of images A and B. $h_A(a)$ and $h_B(b)$ are the normalized marginal histograms of the two images, and L is the number of grey levels.

MI between the fused image F and the two source images A and B is:

$$MI_F^{AB} = (MI_{FA} + MI_{FB})/2 \quad (5.4)$$

4) Objective edge based quality $Q^{AB/F}$ [43]: Objective edge based quality, based on Sobel's operator, indicates the amount of edge information that has been transferred:

$$Q^{AB/F} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [Q^{AF}(i,j) w^A(i,j) + Q^{BF}(i,j) w^B(i,j)]}{\sum_{i=1}^{M} \sum_{j=1}^{N} [w^A(i,j) + w^B(i,j)]} \quad (5.5)$$

where A, B and F represent source images and the fused image, respectively. Q^{AF} and Q^{BF} are edge information preservation values and w^A and w^B are weights.

5.2 Experiment Background

In our experiments, four datasets of visual and IR/MMW images are used as source images. 5 existing algorithms, the gradient pyramid-based method [7], DWT-based

method [15], DTCWT-based method [31], DDDWT-based method [32] and an existing DDDTCWT-based algorithm [24], are used in the comparison.

The source images are depicted as Fig. 22.

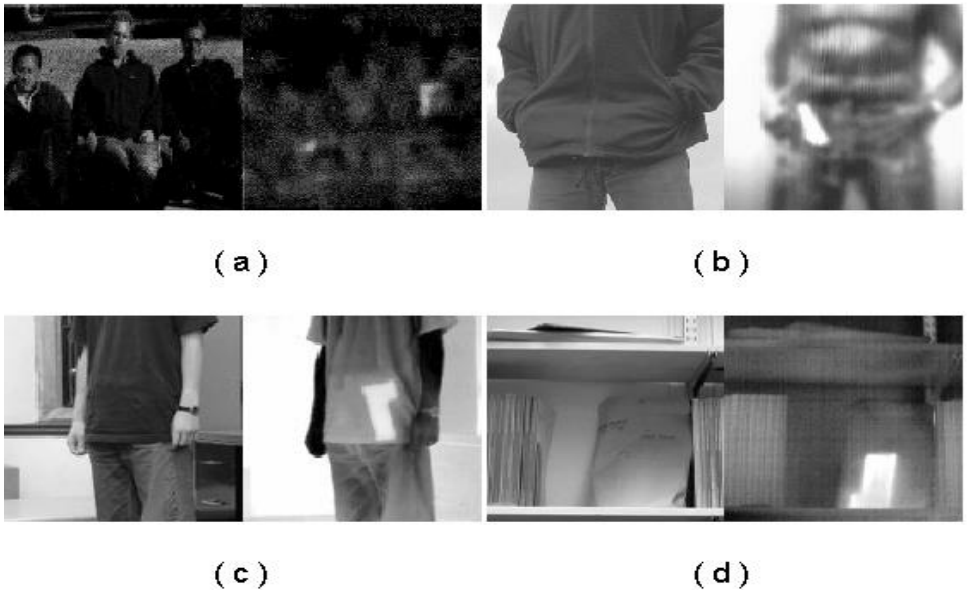

Fig. 22 Four pairs of source images (a) Dataset 1 (b) Dataset 2 (c) Dataset 3 (d) Dataset 4

5.3 Experimental Result and Analysis for Pixel Level Image Fusion

The performance of the proposed pixel level fusion algorithm is demonstrated through the experiments on different image datasets and the comparisons with other methods. In the experiments, the calculation window is chosen as 3×3 when computing the local contrast, NVF, and local energy. The highest decomposition level l is selected as 3.

The fused images using different fusion algorithms are shown in Fig. 23, Fig. 24, Fig. 25 and Fig. 26. The comparison of the statistical performance for the algorithms can be found in Table 1 and Table 2. According to the results of our experiments, the proposed algorithm has preserved the pertinent information from the source images and improved the visual effect of the fused image. Some existing algorithms, such as those based on gradient pyramid and DWT, suffer from the problem of reduced contrast. The existing DTCWT based algorithm performs well on some performance statistics but introduces inconsistencies that can be easily seen in the figures. The existing algorithm based on DDDTCWT performs poorly for several datasets because it uses PCA which has no feature selectivity as the low frequency fusion operation. In most cases, the performance improves when the existing fusion rules are replaced by our proposed fusion rules, proving the effectiveness of the rules. The proposed pixel level fusion algorithm performs best on objective edge based quality, mainly because of the better directional selectivity of DDDTCWT and the improved high frequency fusion rule, which leads to better preservation of edge information. Due to the advantages of DDDTCWT and the fusion rules utilizing HVS characteristics, the proposed pixel level fusion algorithm performs better than the other existing algorithms.

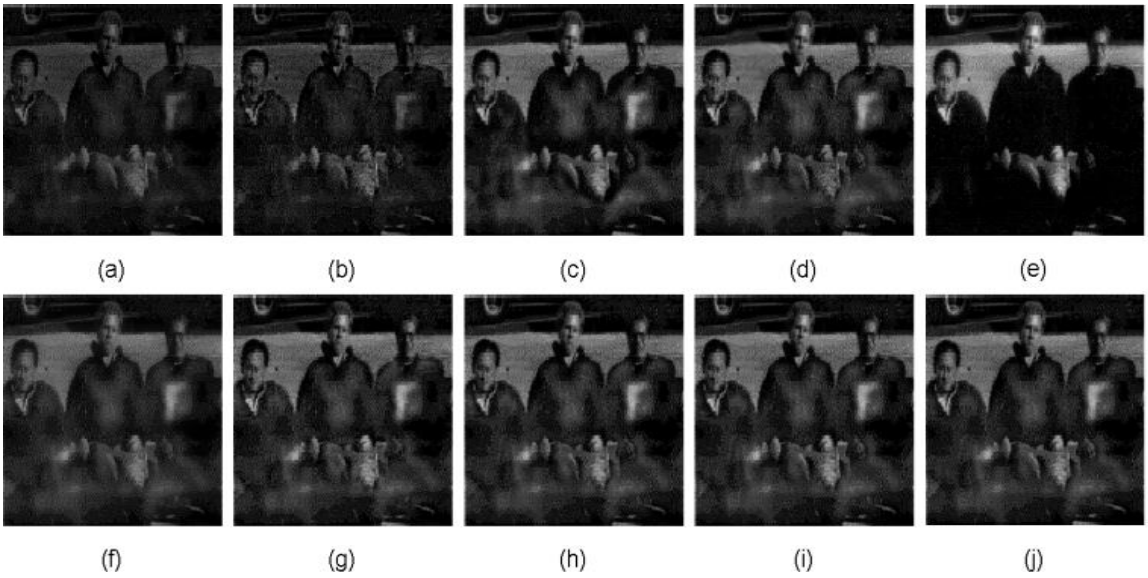

Fig. 24 Fused images of Dataset 1 (a) GP-based existing method (b) DWT-based existing method (c) DTCWT-based existing method (d) DDDWT-based existing method (e) DDDTCWT-based existing method (f) GP with proposed fusion rules (g) DWT with proposed fusion rules (h) DTCWT with proposed fusion rules (i) DDDWT with proposed fusion rules (j) Proposed Method

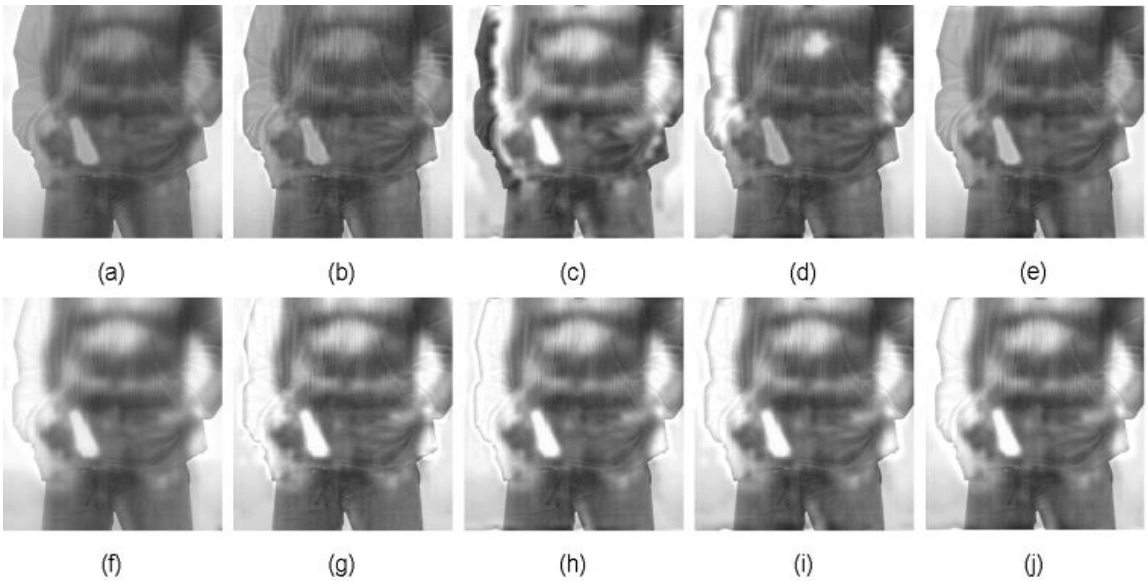

Fig. 23 Fused images of Dataset 2 (a) GP-based existing method (b) DWT-based existing method (c) DTCWT-based existing method (d) DDDWT-based existing method (e) DDDTCWT-based existing method (f) GP with proposed fusion rules (g) DWT with proposed fusion rules (h) DTCWT with proposed fusion rules (i) DDDWT with proposed fusion rules (j) Proposed Method

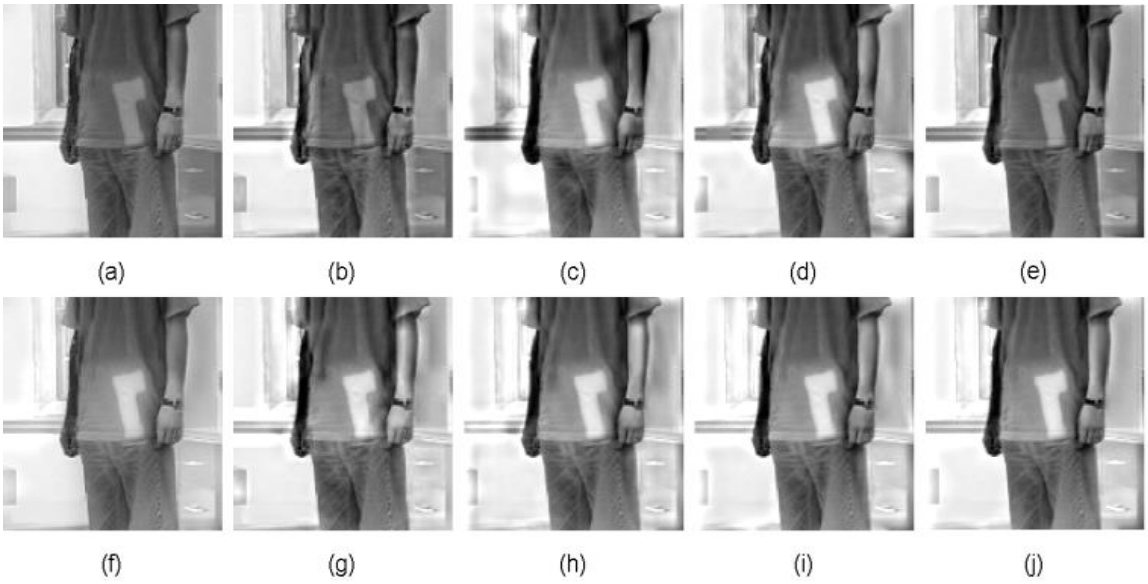

Fig. 26 Fused images of Dataset 3 (a) GP-based existing method (b) DWT-based existing method (c) DTCWT-based existing method (d) DDDWT-based existing method (e) DDDTCWT-based existing method (f) GP with proposed fusion rules (g) DWT with proposed fusion rules (h) DTCWT with proposed fusion rules (i) DDDWT with proposed fusion rules (j) Proposed Method

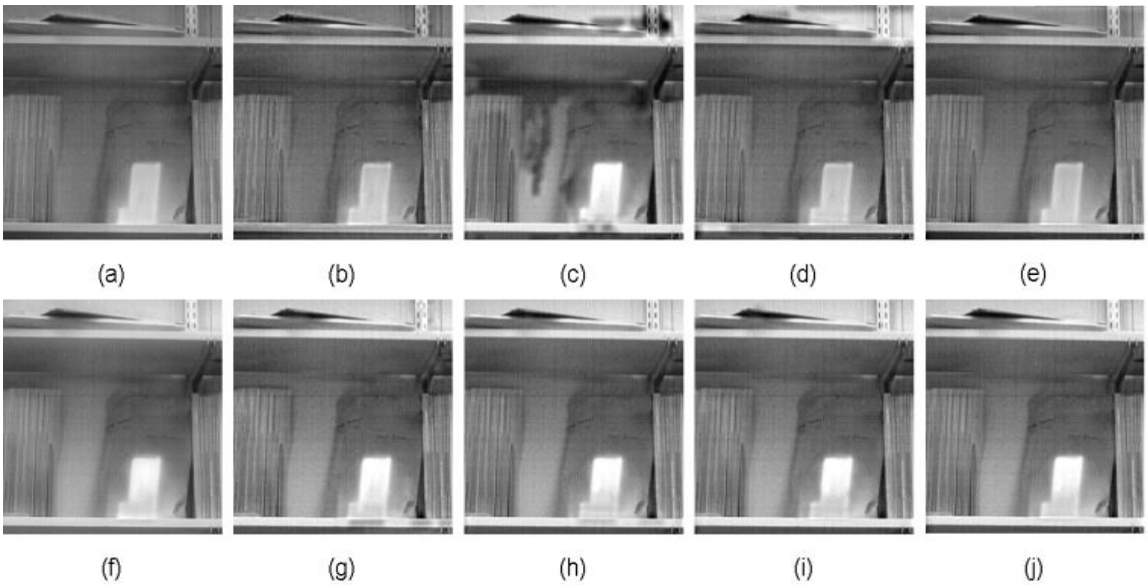

Fig. 25 Fused images of Dataset 4 (a) GP-based existing method (b) DWT-based existing method (c) DTCWT-based existing method (d) DDDWT-based existing method (e) DDDTCWT-based existing method (f) GP with proposed fusion rules (g) DWT with proposed fusion rules (h) DTCWT with proposed fusion rules (i) DDDWT with proposed fusion rules (j) Proposed Method

Table 1 Pixel level fusion statistical performance for Dataset 1 and Dataset 2

		Dataset 1		**Dataset 2**	
Transform	**Quality Matrics**	existing fusion rules	proposed fusion rules	existing fusion rules	proposed fusion rules
Gradient Pyramid	SD	25.087	31.3951	51.1916	60.1233
	Entropy	5.9586	6.6081	7.2698	7.4431
	MI	0.8655	0.9306	1.8768	2.0685
	Q	0.5994	0.6193	0.547	0.5899
Wavelet	SD	27.363	36.8852	51.4419	60.8988
	Entropy	5.8929	6.6095	7.2939	7.4406
	MI	0.7675	1.4626	1.7871	2.1912
	Q	0.5785	0.6266	0.5322	0.5727
DTCWT	SD	38.992	36.7567	60.2957	61.1402
	Entropy	6.424	6.6094	7.4357	7.4432
	MI	1.2586	1.4584	1.6809	**2.1937**
	Q	0.6312	0.6484	0.566	0.575
DDDWT	SD	36.603	37.2823	61.0801	61.1904
	Entropy	6.6302	**6.645**	7.3004	7.439
	MI	1.1107	1.4455	1.6102	2.0936
	Q	0.5635	0.6522	0.5212	0.5835
DDDTCWT (Proposed)	SD	**55.262**	37.3287	53.1471	**61.2846**
	Entropy	4.7276	6.6118	7.3406	**7.4447**
	MI	0.8219	**1.5255**	1.8066	2.13
	Q	0.5849	**0.6707**	0.5727	**0.6025**

Table 2 Pixel level fusion statistical performance for Dataset 3 and Dataset 4

		Dataset 3		**Dataset 4**	
Transform	**Quality Matrics**	existing fusion rules	proposed fusion rules	existing fusion rules	proposed fusion rules
Gradient Pyramid	SD	64.3899	64.6956	35.9215	47.1954
	Entropy	7.5351	7.6041	7.1173	7.4136
	MI	**2.6028**	2.5994	1.4398	1.7538
	Q	0.5955	0.5965	0.5616	0.5669
Wavelet	SD	67.8323	68.923	37.6591	51.9953
	Entropy	7.6064	7.6648	7.1847	7.4773
	MI	2.4998	2.5229	1.3152	1.9084
	Q	0.5894	0.5837	0.5491	0.5478
DTCWT	SD	73.5217	**73.6173**	**54.9537**	52.2417
	Entropy	7.6261	7.275	**7.6263**	7.4987
	MI	2.4901	2.5187	1.9119	**2.0108**
	Q	0.5744	0.5822	0.5554	0.5586
DDDWT	SD	68.3498	69.7896	45.5946	52.3291
	Entropy	7.662	7.6772	7.3781	7.5027
	MI	2.4367	2.4861	1.4419	1.9823
	Q	0.5498	0.5904	0.5182	0.5618
DDDTCWT (Proposed)	SD	70.9037	70.9415	41.3382	52.4251
	Entropy	7.5974	**7.7051**	7.333	7.5148
	MI	2.5353	2.5852	1.4535	1.9649
	Q	0.6099	**0.6168**	0.5692	**0.5763**

The rating of the fused images produced by different algorithms using different datasets is illustrated in Fig. 27, Fig. 28, Fig. 29 and Fig. 30. It is clearly indicated that the rating goes up when the fusion rules are replaced by the proposed fusion rules most of time and the proposed pixel fusion method outperform the other algorithms.

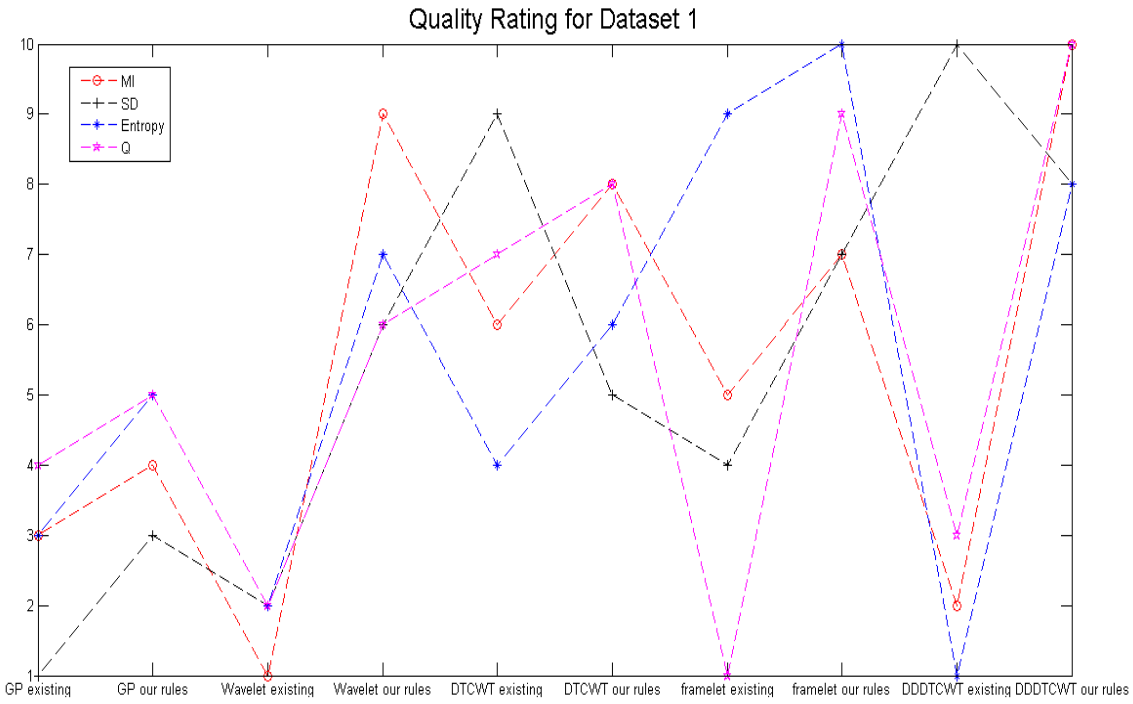

Fig. 27 Quality rating for Dataset 1

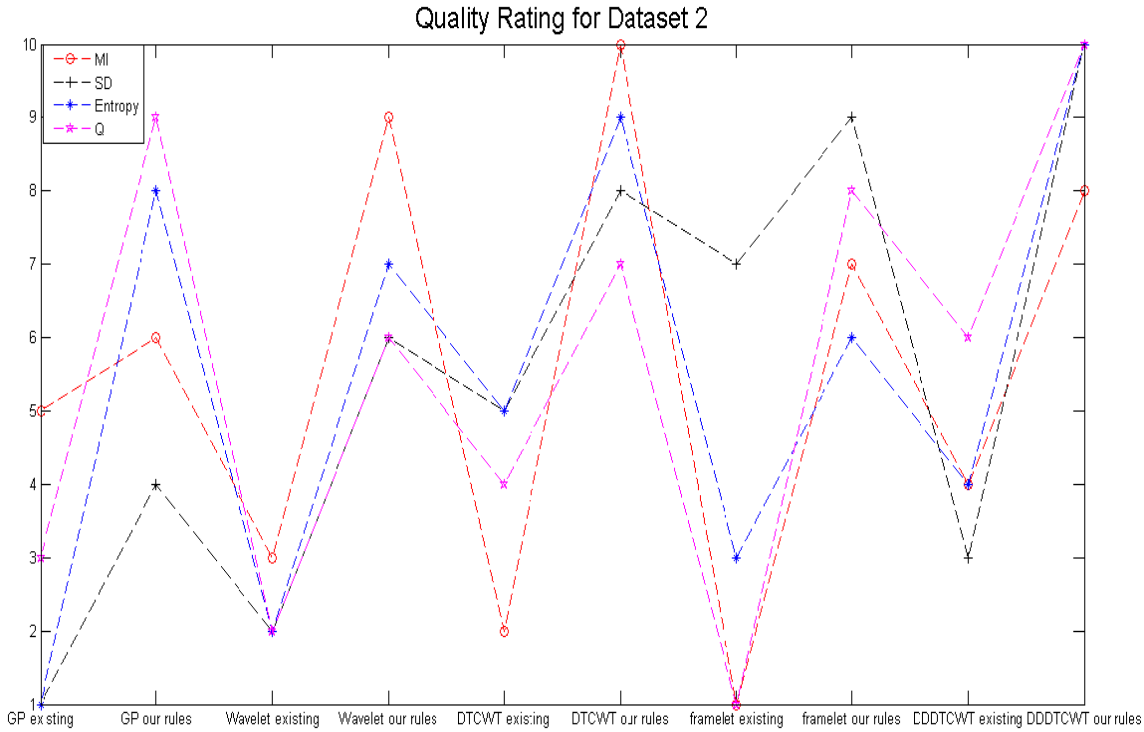

Fig. 28 Quality rating for Dataset 2

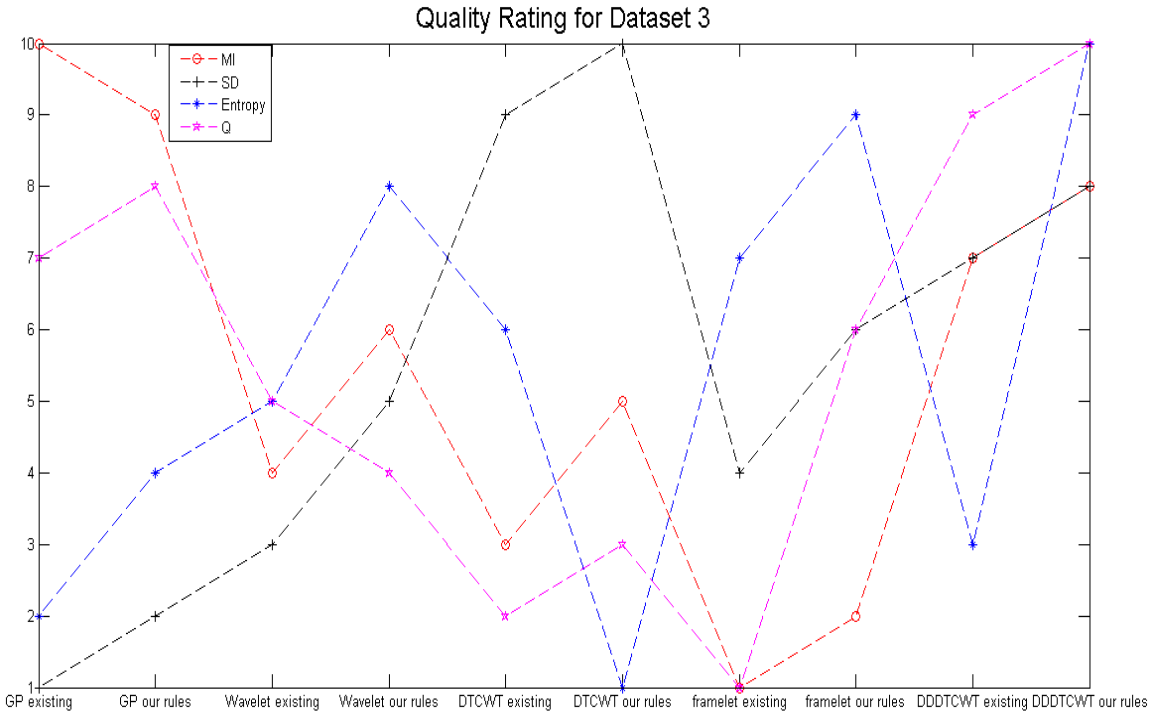

Fig. 29 Quality rating for Dataset 3

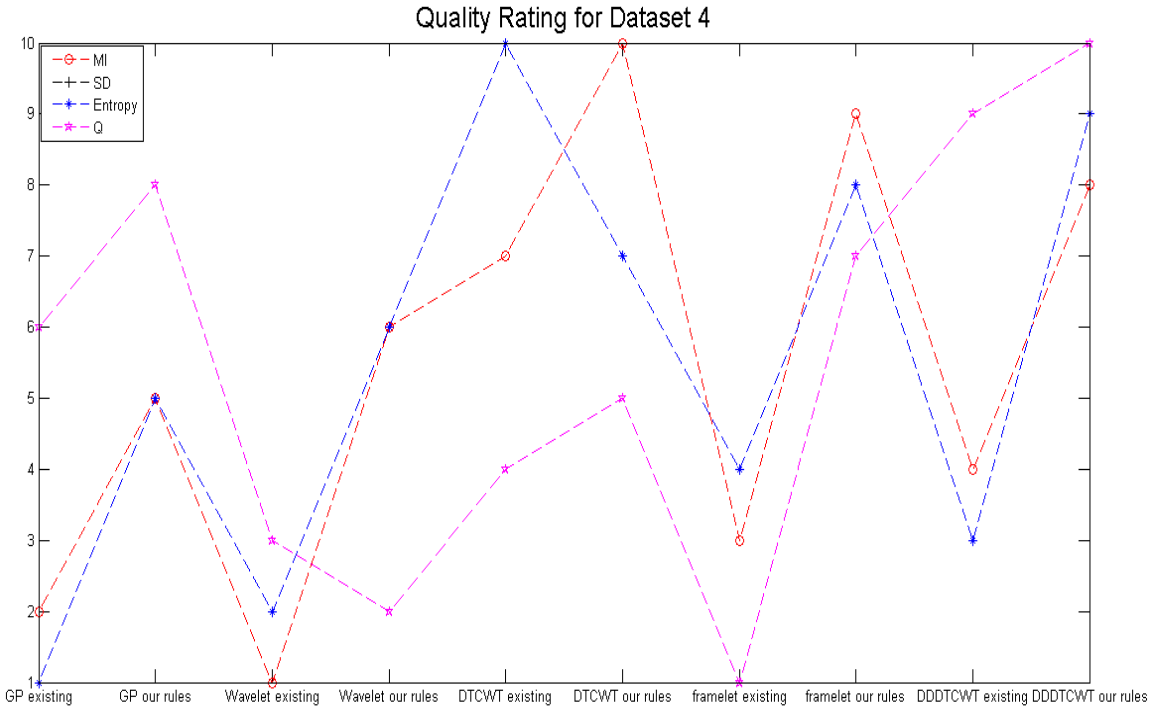

Fig. 30 Quality rating for Dataset 4

5.4 Experimental Result and Analysis for Feature Level Image Fusion

The experiments of the feature level image fusion algorithm have been done based on the 4 datasets. The GMM based multiscale segmentation results are show in the Fig. 31. The comparison of the feature level fusion with the pixel level fusion image is carried out to demonstrate the efficiency of the methods. The highest decomposition level l is selected as 3.

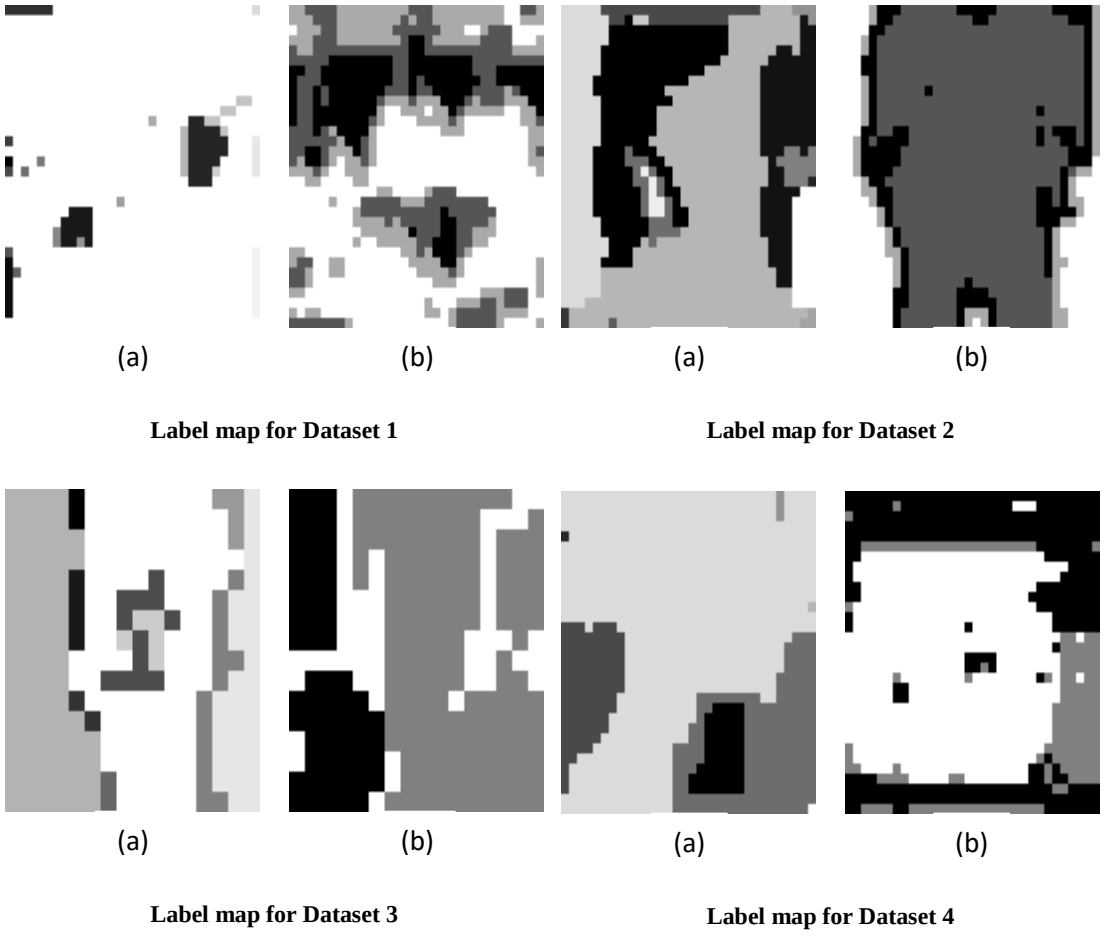

Fig. 31 Multiscale segmentation results for (a)IR/MMW (b)visual

The fused images using the feature level fusion algorithms are shown in Fig. 32. The comparison of the statistical performance between the feature level and pixel algorithms can be found in Table 3.

Table 3 Pixel level and Feature level fusion statistical performance

Dataset	Quality Metrics	Pixel	Feature
Dataset1	SD	37.329	**37.47**
	Entropy	6.6118	**6.6266**
	MI	**1.5255**	1.4576
	Q	**0.6707**	0.6674
Dataset2	SD	61.285	**64.133**
	Entropy	7.4447	**7.4487**
	MI	2.13	**2.1817**
	Q	**0.6025**	0.5873
Dataset3	SD	70.942	**73.872**
	Entropy	**7.7051**	7.6438
	MI	2.5852	**2.6094**
	Q	**0.6168**	0.6154
Dataset4	SD	52.425	**52.454**
	Entropy	7.5148	**7.5256**
	MI	**1.9649**	1.8942
	Q	0.5763	**0.5806**

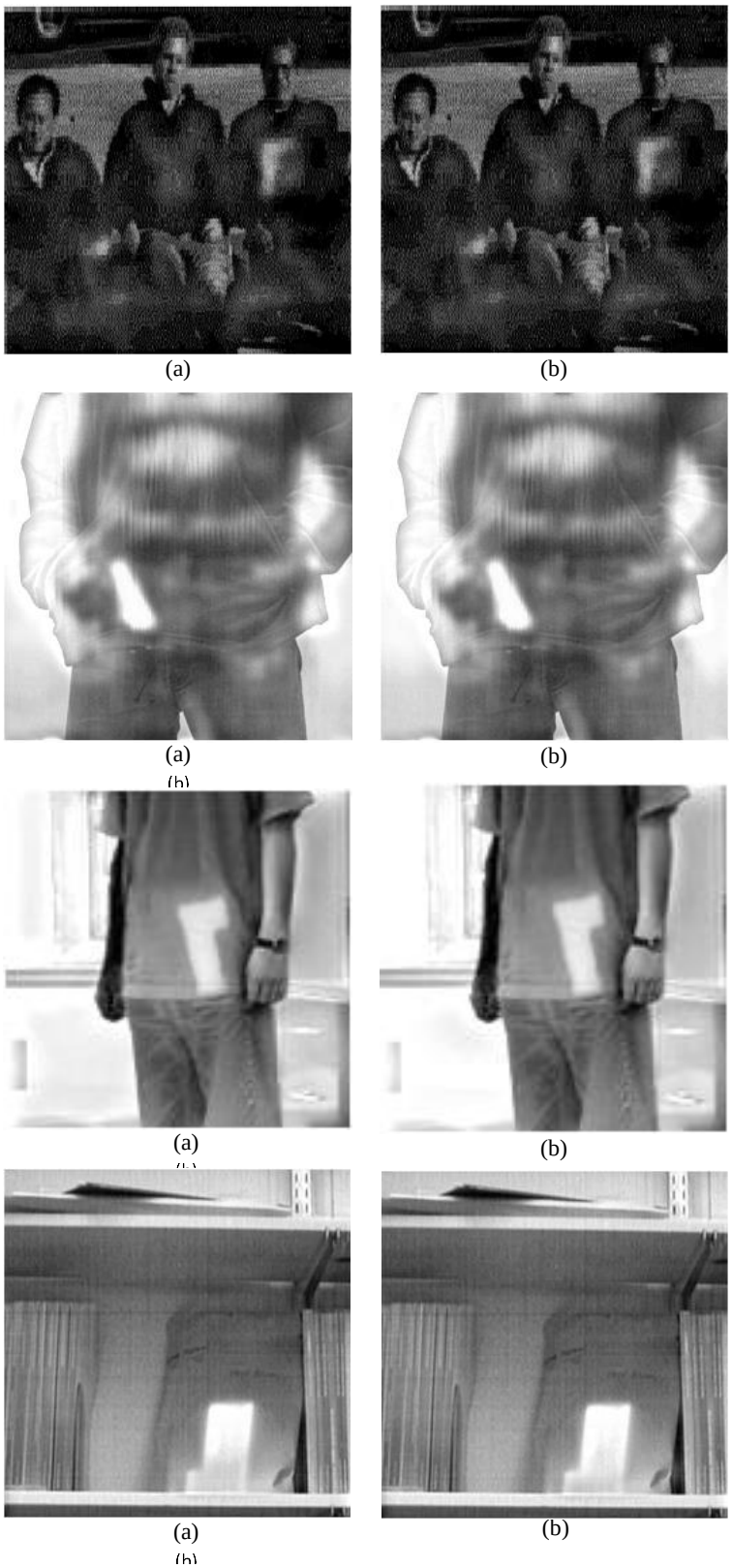

Fig. 32 Fused images of (a) pixel level fusion (b) feature level fusion

According to the results of our experiments, the proposed feature level algorithm has preserved the pertinent information from the source images and improved the visual effect of the fused image. As the fusion rules of feature fusion algorithm are based on region saliency, preferable features are better preserve in the fused images. In most cases, the proposed pixel level fusion approach performs better than the feature level fusion algorithm on objective edge based quality, mainly because the fusion based on regions eliminates part of the edge and texture information . However, the feature level image fusion is less sensitive to noise leading to higher entropy performance. According to the experimental results, the pixel level and feature level algorithms have their own advantages.

CHAPTER 6

CONCLUSION AND FUTURE WORK

In this thesis, two MDB image fusion algorithms, one at pixel level and another at feature level, for CWD application are introduced. The goal of the algorithms is to produce fused images with a clear representation of both the identity information of the suspect and the concealed weapon. For this purpose, images from multiple sensors containing complementary information are fused to provide a detailed description of the person and the weapon hidden beneath the suspect's clothing. DDDTCWT is adopted as the decomposition scheme for both of the algorithms as DDDTCWT possesses the advantages of both DDDWT and DTCWT and has the best directional selectivity. The fusion strategies for the pixel level fusion approach are mainly based on the pixel values of the subband coefficients of the multiscale representation of the source images, leading to better preservation of preferable information from coefficients. The pixel level low frequency fusion rule is a two-mode selection method based on the local contrast measurement and the high frequency fusion operation is developed with consideration of the characteristics of both HVS and DDDTCWT. The fusion rules have been applied with five different decomposition schemes and the experimental results have been compared with five existing pixel level MDB algorithms. The experimental results show that the proposed pixel level method can better preserve the information and improve the quality of the fused images compared to the existing algorithms. Also, the experiment results demonstrate that the proposed pixel level fusion rules can be applied to a variety of multiscale decomposition schemes and that it performs well in the application of CWD.

The DDDTCWT is superior to many other multiscale decomposition schemes, especially because of its better directional selectivity. The feature level fusion algorithm employs GMM based multiscale segmentation scheme to the multiscale representation of the source images to extract features and classify the coefficients into regions. A region based saliency measurement is exploited to create the feature level low frequency fusion operator and the high frequency subbands are fused according to the region based detail information measure. The experimental results show the efficiency and robustness of the proposed feature level fusion algorithm. The pixel level fusion algorithm has better preservation of edge and texture information since the feature level fusion algorithm eliminates part of the texture and edge information from undesired regions. However, the feature level fusion can better preserve the preferable features and is less sensitive to noise, leading to better quality of the fused images.

In the future work, more intelligent fusion rules can be developed to improve the performance of the feature level fusion algorithm. Besides that, more research on effective and accurate evaluation methods for the fusion of thermal images and visual images is required.

REFERENCES

[1] Gallup, "Americans name terrorism as No. 1 U.S. problem," 14 December 2015. [Online]. Available: http://www.gallup.com/poll/187655/americans-name-terrorism-no-problem.aspx. [Accessed 30 March 2016].

[2] Department of State, "Re statistical information on terrorism in 2014," Country Reports on Terrorism 2014, 2014. [Online]. Available: http://www.state.gov/j/ct/rls/crt/2014/239416.htm. [Accessed 30 March 2016].

[3] Gallup, "In U.S., concern about crime climbs to 15-year high," 6 April 2016. [Online]. Available: http://www.gallup.com/poll/190475/americans-concern-crime-climbs-year-high.aspx?g_g_medium=search&g_campaign=tiles [Accessed 30 March 2016].

[4] M. A. Abidi and R. C. Gonzalez, Data Fusion in Robotics and Machine Intelligence, Academic Press, San Diego, 1992.

[5] D. Kundur, D. Hatzinakos and H. Leung, "Robust classification of blurred imagery," *IEEE Trans. Image Process*., vol. 9, no. 2, pp. 243–255, Feb. 2000.

[6] Y. Xia, H. Leung and E. Bossé, "Neural data fusion algorithms based on a linearly constrained least square method," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 320–329, Apr. 2002.

[7] P.J. Burt and R.J. Kolczynski, "Enhanced image capture through fusion", *Fourth Int. Conf. Computer Vision*, 1993, pp. 173-182.

[8] S. Li, J. T. Kwok, I. W. Tsang and Y. Wang, "Fusing images with different focuses using support vector machines," *IEEE Trans. Neural Networks,* vol. 15, pp. 1555-1563, 2004.

[9] M. Chandana, S. Amutha and N. Kumar, "A hybrid multifocus medical image fusion based on wavelet transform". *Int. J. Res. Rev. in Comput. Sci. (IJRRCS),* vol. 2, no. 4, pp. 948–953, 2011.

[10] H. Li, "Wavelet-based weighted average and human vision system image fusion," *Int. J. Wavelets, Multiresolut. Inf. Process.*, vol. 4, no. 1, pp. 97–103, 2006.

[11] Z. H. Li, Z. L. Jing, G. Liu, S. Y. Sun and H. Leung, "A Region-based image fusion algorithm using multiresolution segmentation," *IEEE International Conference on Intelligent Transportation Systems,* 2003, pp. 96-101.

[12] Y. Li and R. Verma, "Multichannel image registration by feature-based information fusion," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 707–720, 2011.

[13] 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," *Inf. Fusion,* vol. 8, no. 2, pp. 119–130, 2007.

[14] S. Zhao, X. Chen, S. Wang, J. Li and W. Yang, "A new method of remote sensing image decision-level fusion based on support vector machine," *in Proc. Int. Conf. Recent Advances in Space Technologies,* 2003, pp. 91–96.

[15] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proc. IEEE,* vol. 87, no. 8, pp. 1315– 1326, Aug. 1999.

[16] E. Lallier and M. Farooq, "A real time pixel-level based image fusion via adaptive weight averaging," *Proc. 3rd Int. Conf. Inf. Fusion*, 2000, pp. WEC3/3-WEC313.
[17] M. Xu, H. Chen and P. K. Varshney, "An image fusion approach based on Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5116–5127, Dec. 2011.
[18] J. Yang and R. S. Blum, "A region-based image fusion method using the expectation maximization algorithm," in *Proc. Conf. Information Sciences and Systems*, Mar. 2006, pp. 468–473.
[19] P. J. Burt, "A gradient pyramid basis for pattern-selective image fusion," *Proc. the Society for Information Display*, pp. 467-470, 1992.
[20] H. Li, B. S. Manjunath and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.
[21] B. Aiazzi, L. Alparone, S. Baronti and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.
[22] P. R. Hill, D. R. Bull, and C. N. Canagarajah, "Image fusion using a new framework for complex wavelet transforms," in Proc. Int. Conf. Image Processing, 2005, pp. 1338–1341.
[23] J. Wang, S. Lai and M. Li, "Improved image fusion method based on NSCT and accelerated NMF," *Sensors*, vol. 12, pp. 5872- 5887, 2012.
[24] G. Chen, and Y. Gao, "Multisource image fusion based on double density dual-tree complex wavelet transform," in *Int. Conf. on Fuzzy Systems and Knowledge Discovery*, 2012, pp. 1864-1868.
[25] A. Ellmauthaler, C. L. Pagliari and E. A. B. da Silva, "Multiscale image fusion using the undecimated wavelet transform with spectral factorization and nonorthogonal filter banks," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1005–1017, March 2013.
[26] G. Bhatnagar, Q. M. J. Wu and Z. Liu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1014–1024, Aug. 2013.
[27] P. K. Varshney, H. Chen, L. C. Ramac and M. Uner, "Registration and fusion of infrared and millimeter wave images for concealed weapon detection", *Proc. of Int. Conf. on Image Processing, Japan*, 1999, vol. 3, pp. 532–536.
[28] Y. Yang and R. S. Blum, "A statistical signal processing approach to image fusion for concealed weapon detection,'' *in Proc. 2002 Int. Conf. Image Processing*, Rochester, NY, Sept. 2002, vol. 1, pp. I-513–516.
[29] Z. Xue and R.S. Blum, "Concealed weapon detection using color image fusion," in *Proc. Int. Conf. on Image Fusion*, Queensland, Australia, 2003, pp. 622–627.
[30] Z. Liu, Z. Xue, R. S. Blum, and R. Laganiere, "Concealed weapon detection and visualization in a synthesized image," *Pattern Analysis and Applications*, vol. 8, no. 4, pp. 375-389, 2006.
[31] Y. Wang and M. Lu, "Image fusion based concealed weapon detection," in *Proc. Int. Conf. on Computational Intelligence and Software Engineering*, 2009, pp. 1-4.
[32] G. Bhatnagar and Q. M. J. Wu, "Human visual system based framework for concealed weapon detection," in *Canadian Conf. on Computer and Robot Vision*, St. Johns, NL, 2011, pp. 250-256.

[33] E. M. Upadhyaya and M. K. Rana, “Exposure fusion for concealed weapon detection,” *2nd International Conference on Devices, Circuits and Systems*, 2014, pp. 1-6.
[34] I. W. Selesnick, “The double density DWT,” In *Wavelets in Signal and Image Analysis: From Theory to Practice*, A. Petrosian, F. G. Meyer. Eds. Bost, MA: Kluwer, 2001.
[35] N. G. Kingsbury, “The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters,” in *Proc. 8th IEEE DSP Workshop*, Citeseer, 1998, vol. 8, p. 86.
[36] I. W. Selesnick, “The double-density dual-tree discrete wavelet transform,” *IEEE Trans. Signal Processing,* vol. 52, no. 5, pp. 1304–1314, May 2004.
[37] D. Levicky and P. Foris, “Human visual system models in digital watermarking”, *Radioengineering*, vol. 13, no. 4, pp. 38–43, 2004.
[38] S. Voloshynovskiy, A. Herrigel, N. Baumgartner and T. Pun, “A stochastic apporach to content adaptive digital image watermarking”, *Proc. Int. Workshop on Information Hiding*, Dresden, Germany, pp. 211–236, 1999.
[39] M. N. Do and M. Vetterli, “Wavelet-based texture retrieval using generalized Gaussian density and Kullbackcleibler distance,” *IEEE Trans.Image Process.,* vol. 11, no. 2, pp. 146–158, Feb. 2002.
[40] T. M. Nguyen and Q. M. J. Wu, “Fast and robust spatially constrained Gaussian mixture model for image segmentation,” *IEEE Trans. Circuits Syst. Video Technol.,* vol.23, no. 4, pp. 621-635, Apr. 2013.
[41] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
[42] R. Achanta, S. Hemami, F. Estrada and S. Susstrunk, “Frequency-tuned salient region detection,” *in Proc. IEEE Conf. Comput. Vis. Pattern Recogn.,* 2009, pp. 1597–1604.
[43] C. S. Xydeas and V. Petrovid, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, pp. 308-309, 2000

VITA AUCTORIS

NAME:	Tuzhi Xu
PLACE OF BIRTH:	Fengxin, Jiangxi, China
YEAR OF BIRTH:	1985
EDUCATION:	Civil Aviation University of China, B.Eng., Tianjin, China, 2002-2006
	University of Windsor, M.Sc., Windsor, ON, 2013-2016