SIGNAL PROCESSING ALGORITHMS FOR ENHANCED IMAGE FUSION PERFORMANCE AND ASSESSMENT

ZAID BIN OMAR
M.Sc

A Thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy and Diploma of Imperial College London

Communication and Signal Processing Group
Department of Electrical and Electronic Engineering
Imperial College London
2012

Dedicated to my wife, Manihah.

Thank you for your perseverance.

Abstract

The dissertation presents several signal processing algorithms for image fusion in noisy multimodal conditions. It introduces a novel image fusion method which performs well for image sets heavily corrupted by noise. As opposed to current image fusion schemes, the method has no requirements for *a priori* knowledge of the noise component. The image is decomposed with Chebyshev polynomials (CP) being used as basis functions to perform fusion at feature level. The properties of CP, namely fast convergence and smooth approximation, renders it ideal for heuristic and indiscriminate denoising fusion tasks. Quantitative evaluation using objective fusion assessment methods show favourable performance of the proposed scheme compared to previous efforts on image fusion, notably in heavily corrupted images.

The approach is further improved by incorporating the advantages of CP with a state-of-the-art fusion technique named independent component analysis (ICA), for joint-fusion processing based on region saliency. Whilst CP fusion is robust under severe noise conditions, it is prone to eliminating high frequency information of the images involved, thereby limiting image sharpness. Fusion using ICA, on the other hand, performs well in transferring edges and other salient features of the input images into the composite output. The combination of both methods, coupled with several mathematical morphological operations in an algorithm fusion framework, is considered a viable solution. Again, according to the quantitative metrics the results of our proposed approach are very encouraging as far as joint fusion and denoising are concerned.

Another focus of this dissertation is on a novel metric for image fusion evaluation that is based on texture. The conservation of background textural details is considered important in many fusion applications as they help define the image depth and structure, which may prove crucial in many surveillance and remote sensing applications. Our work aims to evaluate the performance

of image fusion algorithms based on their ability to retain textural details from the fusion process. This is done by utilising the gray-level co-occurrence matrix (GLCM) model to extract second-order statistical features for the derivation of an image textural measure, which is then used to replace the edge-based calculations in an objective-based fusion metric. Performance evaluation on established fusion methods verifies that the proposed metric is viable, especially for multimodal scenarios.

Acknowledgment

All praises be to God for it was with His guidance and mercy I was able to finally complete this doctorate. Embarking on a three-and-a-half years of research work in such a challenging area is a unique endeavour full of trials, disappointments, surprises and wonderful moments – all of which I will cherish for the rest of my life.

My foremost acknowledgements go to Dr Tania Stathaki for her diligent supervision and for guiding me through this research. She has shown tremendous support throughout my progress as a PhD student and I owe much of my knowledge, discipline and passion for research to her.

I would also like to extend my humblest gratitude to fellow colleagues at the Communications & Signal Processing Group – most notably thanks to Dahir Dini, James Pearson, Amanjit Dulai and Dr Bukhari Che Ujang for your excellent company. Further, I must especially thank Dr Roy Jacobs from the Mathematics Department for his help during the early stages of my PhD. I am also grateful to Dr Abdelmalik Bachir, Dr Mohammed Jaward, my sister Khadijah and my father Dr Omar Yaakob for their kind assistance with my research writing. To Hana Aziz and Lila Izhar, thank you too for the enjoyable conversations.

Last but not least, I am indebted to the Ministry of Higher Education (MOHE) and Universiti Teknologi Malaysia (UTM) for their generous funding of my studies.

Contents

Abstract 3

Acknowledgment 5

Contents 6

List of Figures 9

List of Tables 12

Abbreviations 13

Statement of Originality 15

Chapter 1. Introduction 16

1.1 Problem Statements . . . 18

1.1.1 Previous Works on Noisy Fusion . . . 18

1.1.2 Fusion Assessment Methods . . . 21

1.2 Research Aims . . . 23

1.2.1 Motivations for Research . . . 23

1.2.2 Research Objectives . . . 24

1.3 Thesis Outline . . . 25

1.4 Current Publications . . . 25

Chapter 2. The Basics of Image Fusion 27

2.1 Introduction To Image Fusion . . . 27

2.1.1 Benefits of Image Fusion . . . 30

2.2 Abstraction Levels in Fusion . . . 30

2.3 History of Image Fusion . . . 32

2.3.1 Early Fusion Systems . . . 33

2.3.2 Pyramid-based Methods . . . 34

2.3.3 Wavelet-based Methods . . . 35

2.3.4 Data-driven Methods . . . 36

2.3.5 State of the Art in Image Fusion 37
2.4 Image Fusion Applications 37

Chapter 3. Chebyshev Polynomials for Image Fusion 41
3.1 Issues in Fusion for Image Denoising 41
3.2 Univariate Chebyshev Polynomials 42
3.3 Chebyshev Polynomials for Signal Approximation 45
3.3.1 Some Properties of Chebyshev Polynomials 47
3.4 Chebyshev Polynomials as Separable Basis Functions for Image Approximation . 49
3.4.1 Rationale of Approach 49
3.4.2 Derivation of Bivariate Coefficient 49
3.5 A Novel Method of Image Fusion with Chebyshev Polynomials 51
3.5.1 Fusion Rules 51
3.5.2 Fusion Example 52
3.6 Performance Evaluation of CP Fusion 53
3.6.1 Experiment 1: Multimodal Images 53
3.6.2 Experiment 2: Multifocal Clock 53
3.6.3 Experiment 3: Multifocal Computer 65
3.6.4 Discussion 75
3.7 Chapter Summary 76

Chapter 4. Region-based Chebyshev-ICA Fusion Method 78
4.1 Limitations of CP Fusion 78
4.1.1 A Proposed Region-based Approach 79
4.2 Introduction to Independent Component Analysis 79
4.2.1 Comparison of ICA and CP 82
4.3 A Novel Region-based CP-ICA Method for Image Fusion 83
4.3.1 Fusion Rule 83
4.3.2 Region-based Fusion 83
4.3.3 Determining Active and Inactive Regions 84
4.3.4 Mathematical Morphology for Edge Map Enhancement 85
4.4 Performance Evaluation of RBCI 87
4.4.1 Experiment 1: Multifocal Images 88
4.4.2 Experiment 2: Multispectral Images 91
4.5 Chapter Summary 100

Chapter 5. Textural-based Fusion Metric 102
5.1 The Need for a New Fusion Metric 102
5.1.1 Our Definition of Texture 104
5.2 A Review of Image Fusion Assessments Metrics 105

5.2.1 Limitation of Current Metrics 108
5.3 Introduction to Gray-level Co-occurrence Matrix 108
5.4 A Novel Texture Enhancement Feature for GLCM 110
5.5 Textural-based Fusion Metric 113
5.6 Performance Evaluation of Textural Fusion Metric 114
5.7 Chapter Summary 121

Chapter 6. Future Work and Conclusions 123
6.1 Discussion and Future Work 123
6.2 Conclusion 126

Appendix A. Two-dimensional Chebyshev polynomials 128

Appendix B. Discrete Chebyshev polynomials as separable basis functions 131

Bibliography 135

List of Figures

1.1 Example of texture (a) Visual image and (b) NIR image 23
1.2 General framework of thesis 26

2.1 Example of image fusion showing (a) Visual image, (b) NIR image and (c) Fused image 29
2.2 Fusion abstraction levels comprising LLF, ILF and HLF 32

3.1 Chebyshev polynomials, up to $N = 5$ degrees 44
3.2 Comparison of orthogonal polynomials for degrees $n = 1, 2, ..., 5$ 45
3.3 Univariate Chebyshev approximation ($f(x)$ in blue, $\tilde{f}(x)$ in green). R is the ratio of CP order $n = 10$ over data size L. 48
3.4 Multimodal gun example 54
3.5 Occluded jet and cloud example 55
3.6 Multimodal UN Camp example 56
3.7 Multifocus clock example 58
3.8 Fusion performance on images with Gaussian noise, SNR = 21dB 59
3.9 Fusion performance on images with Gaussian noise, SNR = 17dB 60
3.10 Fusion performance on images with Gaussian noise, SNR = 14dB 61
3.11 Fusion performance on images with Gaussian noise, SNR = 11dB 62
3.12 Fusion performance on images with Gaussian noise, SNR = 7dB 63
3.13 Fusion performance on images with Gaussian noise, SNR = 4dB 64
3.14 Comparison of fusion methods for multifocal clock images in noisy conditions using Petrovic and Piella metrics 66
3.15 Multifocus computer example 67
3.16 Fusion performance on computer images with Gaussian noise, SNR = 21dB . . . 68
3.17 Fusion performance on computer images with Gaussian noise, SNR = 17dB . . . 69
3.18 Fusion performance on computer images with Gaussian noise, SNR = 14dB . . . 70
3.19 Fusion performance on computer images with Gaussian noise, SNR = 11dB . . . 71
3.20 Fusion performance on computer images with Gaussian noise, SNR = 7dB 72

3.21 Fusion performance on computer images with Gaussian noise, SNR = 4dB 73
3.22 Comparison of fusion methods for multifocal computer images in noisy conditions using Petrovic and Piella metrics 74

4.1 Framework of region-based combinatory CP-ICA fusion 80
4.2 Edge map derived from selecting all hard and soft edges from both input images (a) UN visual camera, (b) Infrared, (c) Edge map, (d) Enhanced edge map 87
4.3 Multifocal Clock image fusion results for different methods for SNR = 10dB. a) ICA, b) DT-CWT, c) Mean, d) CP, e) RBCI 89
4.4 Multifocal Clock fusion performance for different fusion methods for SNR = 16dB to 10dB using the Max-abs fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon) 92
4.5 Multifocal Clock fusion performance for different fusion methods for SNR = 16dB to 10dB using the Mean fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon) 93
4.6 Multifocal Clock fusion performance for different fusion methods for SNR = 16dB to 10dB using the Weighted Average fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon) 94
4.7 Multispectral UN Camp image fusion results for different methods for SNR = 10dB. a) ICA, b) DT-CWT, c) Mean, d) CP, e) RBCI 96
4.8 Multispectral UN Camp fusion performance for different fusion methods for SNR = 16dB to 10dB using the Max-abs fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon) 97
4.9 Multispectral UN Camp fusion performance for different fusion methods for SNR = 16dB to 10dB using the Mean fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon) 98
4.10 Multispectral UN Camp fusion performance for different fusion methods for SNR = 16dB to 10dB using the Weighted Average fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon) 99

5.1 Brick example of edges versus texture 105
5.2 GLCM 109
5.3 GLCM texture example (a) Bush original image, (b) Bush filtered image, (c) GLCM of original image and (d) GLCM of filtered image 111

5.4 Texture definition in GLCM 112
5.5 Source images and fusion results a) UN Visual, b) UN NIR, c) UN Camp ICA and d) UN Camp CP. 115
5.6 Source images and fusion results a) Gun Visual, b) Gun MMW, c) Gun ICA and d) Gun CP. 116
5.7 Source images and fusion results a) Clock background focus, b) Clock foreground focus, c) Clock ICA and d) Clock CP. 117
5.8 Source images and fusion results a) Field focus, b) Tank focus, c) Tank ICA and d) Tank CP. 118
5.9 Chart depicting relative fusion metric scores a) UN Camp, b) Gun, c) Clock and d) Tank. 119

A.1 Graphs of 2D Chebyshev polynomials 129

B.1 Discrete Chebyshev polynomials, with $N = 33$ points 132
B.2 Comparison of the first five discrete versus classical Chebyshev polynomials for data length $L = 11, 15, 21$ 133

List of Tables

3.1 Fusion metric evaluation for clock images – ICA, CP and DT-CWT 65
3.2 Fusion metric evaluation for computer images – ICA, CP and DT-CWT 75

4.1 Average Petrovic score for multifocal Clock example for noise levels 16dB to 10dB. 91
4.2 Average Petrovic score for multispectral UN Camp example for noise levels 16dB to 10dB. 100

5.1 Analysis of objective fusion metric scores for ICA and CP results of various image sets 120

Abbreviations

The most significant mathematical notations and abbreviations used in this document:

CP:	Chebyshev polynomials
ICA:	Independent component analysis
MRA:	Multi-resolution analysis
AWGN:	Additive white Gaussian noise
LPF:	Low-pass filter
NIR:	Near infrared
EMD:	Empirical mode decomposition
DT-CWT:	Dual-tree complex wavelet transform
IMF:	Intrinsic mode function
RS:	Remote sensing
MS:	Multi-spectral
PAN:	Panchromatic
IHS:	Intensity-hue-saturation
RGB:	Red-green-blue
PCA:	Principal component analysis
DCT:	Discrete cosine transform
SCS:	Sparse code shrinkage
RBCI:	Region-based Chebyshev-ICA method
SNR:	Signal-to-noise ratio
MSE:	Mean square error
HVS:	Human visual system

MI: Mutual information

GLCM: Gray-level co-occurrence matrix

$I(x, y)$: Pixel intensity

$T_n(x)$: Chebyshev polynomial of degree n

$\mu(x)$: Weight function of CP

$\tilde{f}(x)$: Approximated function of $f(x)$

a_n: n-th coefficient of CP

P: Power spectrum of image

$Q^{\text{AB/F}}$: Petrovic metric

$Q(a, b, f)$: Piella metric

$C_{m,n}$: GLCM magnitude at coordinate (m, n)

G_{max}: Maximum gradient

T: Measure of texturedness

$Q_{\text{T}}^{\text{AB/F}}$: Textural-based fusion metric

Statement of Originality

To the extent of the author's knowledge, the following aspects of the thesis are believed to be original contributions:

1. The use of classical Chebyshev polynomials as basis functions for image approximation, by means of the derivation of the orthogonality relation for separable approximation of bivariate signals.

2. A first application of the Chebyshev approximation above to perform a transform domain fusion of multiple grayscale images. Our scope is heavily focused on multimodal images of varying levels of noise.

3. Development of a hybrid scheme that merges the results of two incumbent image fusion techniques, namely independent component analysis (ICA) and Chebyshev polynomials (CP) through the selection of active and non-active regions. This is based on the Canny edge detector coupled with morphological operations. The approach is primarily intended for multimodal scenarios with medium to high levels of noise.

4. Derivation of a novel textural-based metric to measure the performance of image fusion algorithms, emphasising on texture preservation. The method utilises the gray level co-occurrence matrix (GLCM) to elicit meaningful features and develop a measure of texturedness, according to a novel interpretation of texture.

5. Modification of the well known objective fusion evaluation measure to incorporate parameters that prioritise textural preservation for fusion assessment.

Chapter 1

Introduction

WHAT is a good image? Numerous criteria that define the characteristics of a high quality image have been proposed and discussed extensively in image processing literature. Obviously the notion of image quality has as much to do with user applications and requirements as it does with perceived visual quality in general. Almost always, the definition of a good image relates directly to the user's ability to distinguish an object or several objects of interest from non-important regions. This task however becomes problematic in cases involving noise. The presence of noise is undesirable as it introduces unwanted artefacts in the final image, betraying its true quality.

The prevalence of a wide variety of sensor modalities for imaging applications, particularly for commercial and industrial purposes, has seen the widespread use of low to medium cost camera sensors. A significant drawback of this, apart from the high compression rate, is the prospect of low quality input images [1]. This owes to the natural characteristics of photoelectric sensors as well as their working environments [2], which will almost certainly render the acquired images susceptible to packet losses within a transmitted network [3], corruption and subsequently, noise. Low cost cameras offering narrow bandwidth tend to limit the total amount of light, causing low signal-to-noise ratio (SNR) and are thereby more likely to suffer from noise corruption [4].

Charge-coupled device (CCD) cameras make up the majority of image acquisition tools in practice. CCD sensors work by measuring the charge induced by incident photons against a mesh of electro-optical elements. The number of photons reflected on a particular element is tantamount

to the amount of charge it stores, and therefore the brightness of the resulting pixel. There are two main sources of sensor noise: dark (leakage) current noise, which is sensitive to temperature, and photon noise, which relates to the number of incident photons. In low to medium cost sensors to which category most visible and infrared sensors in practice belong, both noise components are combined where at room temperature the noise can generally be modelled as Gaussian [5].

Research into the phenomenon of image noise and efforts on denoising have been vast and extensive throughout image processing literature. Notable sources for an introduction into noise and denoising theories can be found in [6, 7], and it is not our intention to further analyse the theories of image noise. Interestingly though, recent works have pointed out the plausibility of using fusion for denoising purposes. As image fusion is concerned with the combination of useful information whereas denoising is essentially removing noise corrupted information, the merging of both processes in an integrated approach of fusion and noise suppression is essential. The rationale behind this is twofold, as issues concerning noise in practice are largely common:

- A single corrupted image may be treated by multiple denoising filters, each of which tends to specific features of the scene. For instance the spatial median filter (SMF) [8] is regarded as a uniform smoothing algorithm that can maintain shape edges of round objects within the image. Likewise, the Canny edge detector [9] incorporates a noise reduction tool and uses a multi-stage filter to detect edges of various degrees. Fusion is then applied between the results of the respective filters [10]. This has been the case with signal-independent impulse noise [10, 11], which are common in digital images. Hence pixel-based fusion is regarded strictly as a tool with which to perform denoising of completely redundant image inputs. A related example is in [12] which proposes the fusion of a non-noisy long-exposed image with non-blurry short-exposed images to overcome their respective weaknesses.

- The second motivation lies with the fact that fusion applications themselves, as noted above, are prone to noise. The performance of a surveillance system involving fusion of visual and infrared images for example, may be degraded by the presence of sensor noise attributed to these sensors, which owes to the physical limitations of recording such observations [1]. For instance one might acquire images that are corrupted in different locations of the scenery, and a fused image may be developed by only selecting the uncorrupted pixels from various

image parts [11]. Noise may be caused by faulty memory addresses or noisy transmission channels in which only some pixel intensities are altered [13]. Therefore a robust fusion algorithm with the inherent ability for noise detection, suppression and removal is desired.

1.1 Problem Statements

1.1.1 Previous Works on Noisy Fusion

A major limitation of transform- and gradient-based fusion algorithms is the fact that the gradient operator is heavily influenced by noise elements prevalent in the data. This occurs as some noise particles in the high frequency sub-band are treated as edge information, and therefore transferred along in the fusion process [14]. Of late, significant efforts in research have been directed towards the development of fusion in noisy image scenarios. Noise-based fusion is a sub-field in which extensions or modifications are often applied to popular techniques to account for noise corrupted signals.

Examples of this include sparse coding of independent component analysis (ICA) coefficients [15], which shrinks the magnitude of coefficients towards zero so as to minimise the amount of noise transferred. The amount of shrinkage invariably depends on the strength of noise – a knowledge of noise variance is therefore essential. The procedure necessitates the estimations of a density model, which consists of normal and Laplacian density, and a non-linear shrinkage function. Similar shrinkages of noisy wavelet and contourlet coefficients were performed in [2, 16]. Elsewhere the contourlet transform [2], due to its directionality, requires a selection of shrinkage value suitable for use in its high frequency sub-bands, where the noise variance σ^2 is modelled using the Monte Carlo method.

Another application of sparse coding on fusion is found in the study by Liao and Liu [17], which makes use of orthogonal matching pursuit. The output quality is fundamentally controlled by configuring the value of approximation error, and the fusion result is found to have good robustness to noise whilst uncovering underlying image structures. Furthermore in [18] a layered approach is taken by Rahman et al. whereby after the denoising of source images, signal dominated coefficients are modified towards maximum, while noise dominated coefficients are min-

imised. This is an adept approach except that denoising source images prior to fusion is largely undesirable as it causes loss to edge details, leading to blurry inputs for fusion.

Other promising approaches to noisy fusion include bias removal techniques [19] and a robust Gaussian moments-based ICA algorithm [20] applied onto ICA data that enables the removal of signal bias due to noise. In [14], Xin et al. introduces a solution using multi-scale products for the wavelet transform. The notion behind this is that in contrast to signals, noise coefficients tend to decrease with a gain in scale. A product of wavelet coefficients on adjacent scales can therefore enhance edges while suppressing noise.

An effective fusion method for denoising has also been proposed by Indu and Ramesh in [11], that involves feedback of noise information to its basic restoring median filter and works well in high density noise areas. However the method is limited to only impulse noise and requires a large number of n input images to enable the selection of uncorrupted pixels. Another simple yet efficient method uses background subtraction denoising [21], by which the histogram of the estimated noisy background is subtracted from the histogram of the fused image. For signal dependent noise solutions, the work by Kumar and Miller in [22] used a local affine model followed by a total variation (TV) regularisation which is generally robust and does not require *a priori* knowledge. Though similar to [11], multiple captures of the input image are necessary.

Autonomous Denoising for Fusion in High Density Noise

There are two notable drawbacks of these schemes: first is their significant reliance on the knowledge of the input data at hand. Sparse code ICA for instance requires an estimation of the noise variance beforehand, whilst wavelet shrinkage for hyperspectral fusion is scale-specific and requires approximate knowledge of the resolution difference between two inputs [23]. For other algorithms the calculation of the shrinkage value is necessary [15]. To accommodate these requirements, a number of noise models have been proposed [3].

The second weakness, notwithstanding their effectiveness on low to moderately noised images, is that the approaches have only so far been restricted to tests on low density (15-20dB SNR) noise scenarios, barring in [11], with no real emphasis on highly corrupted images of $\leq$10dB SNR. The dearth in studies involving fusion of highly corrupted images is presumably due to the

notion that the performance of these algorithms tends to diminish as noise levels increase. The ability to measure denoising capacity is highly desirable in a fusion system, as acute noise and corrupted image scenarios are found to be common occurences in real world phenomena. A good fusion algorithm must therefore be adept at capturing these elements.

For noise models, such an approach is understandable; as in most cases estimating information regarding noise involves the proposition of several asssumptions that have largely been proven correct on certain datasets. To mimic real world phenomena, it has been desirable to model as accurately as possible noise and the impairment behaviour of scenarios causing noise [1, 3]. Indeed, algorithms such as multi-resolution analysis (MRA) and ICA which have been adapted for noisy signals perform well compared to non-denoising fusion techniques.

However we deem the aims of real world fusion applications are better served by fast yet simple processing algorithms, consisting of minimal parametric inputs. The techniques mentioned above generally employ an approximate estimation of noise data, along with training procedures, which considerably add to the processing complexity. Whilst effective, these models are based on the presumption of separating signal from noise [1]:-

$$A_n = A + k_n N \tag{1.1}$$

Whereby a noise signal N, scaled by a factor k_n in which n denotes different SNR levels i.e. $n = 30, 20, 10, 3, 1$, is added to an image signal A to yield the resulting corrupted image, A_n. Most noise models work on the basis of measuring N according to several noise characteristics – including independent average white Gaussian noise (AWGN), Poisson noise that produces signal outliers, and noise of the signal-dependence variety [22]. The problem expands in fusion, where instead of one, numerous noise estimations for multiple input images have to be considered.

Based on the preceding theories, we find modelling noise as Gaussian with zero mean and unit variance [1] to be a legitimate and feasible standard measure of image corruption in this work. Assuming an intermediate level fusion (ILF) framework, a faster approach to noisy fusion would therefore be to incorporate the denoising process directly into the transformation step. This can be achieved by a low-pass filtered (LPF) approximation of signals, the concept of which is

elaborated in the Chapter 3, to develop a novel fusion algorithm using orthogonal polynomials as the basis functions. Polynomial approximation for fusion enables a robust fusion of images and an automated, indiscriminate removal of noise without the need for *a priori* knowledge. Moreover, enhanced performances by joint-fusion of incumbent algorithms are also explored in Chapter 4.

1.1.2 Fusion Assessment Methods

Another important consideration in image fusion is formal methods of assessing fusion performance. Fusion metrics evaluate the quality of a fusion result in terms of information transferred i.e. the fidelity of preserving salient features from its source images. The notion however becomes difficult given that image fusion is a highly application-specific field. The focus of content preservation in medical imaging fusion, for example, is vastly different from that in multifocal fusion for photography purposes. A common goal of fusion is to preserve salient image features but in multispectral scenarios, for instance, it is difficult to decide which type of information from the inputs should feature in the fused image. Ultimately, the desired content will depend on the user requirements. Also, the lack of ground truth inherently means that image quality assessment (IQA) methods alone are inadequate for fusion assessment.

As it is, there is often a notable discrepancy between the scores of fusion metric themselves that is theoretically due to the different measures of saliency utilised by each metric. Fusion techniques vary on the type of salient features that are given priority – several methods focus on signal energy, some on image contrast, whilst others incorporate edge and textural information into consideration when performing fusion. The metrics, in turn, tend to reflect this diversity.

A comprehensive review into the study of fusion measures can be found in Chapter 5. In essence, there exist numerous metrics for evaluating the effectiveness of fusion methods. They distinguish how much valuable information have been transferred from the source images to the output image. The information in question tends to vary between metrics. Petrovic [24] and Piella [25] metrics, for example, measure the edge details transferred whereas Qu [26] metric calculates mutual information between the images. In certain cases, IQA methods have proven a useful basis to develop a fusion metric, as the works by Piella have shown [25]. This is an interesting premise as IQA and fusion assessment inevitably share many similarities in the measurement

of visual quality. As such, fusion metrics may do well to exploit the plethora of IQA methods in existence.

Nevertheless, there is yet to be a formal way of determining which metric is best. Indeed, fusion tests of a random scenario by different metrics may probably yield vastly different scores. Each fusion metric tends to have its own advantages and is best suited for certain applications.

And therein lies the problem. The general uncertainty regarding the definition of fusion quality has instigated the need for a generic list of criteria that represents a good fusion performance measure. For example, a good fusion metric has to have strong correlations with subjective, human-based evaluations. This means that images which we perceive to be visually pleasing will directly translate into a high score in the fusion metric. This enables us to provide a standard on which engineers and developers of image fusion metrics may base their algorithm.

Emphasis on Texture for Natural Imagery

It has been noted that in critical surveillance, military and remote sensing applications, the detection and preservation of natural imagery are vital. These include grass, rocks, bushes, pebbles, earth and other entities which make up the background scenery. While their textural structure may be insignificant compared to the primary objects of interest such as human presence, it is important to preserve these details to ensure quality for further processing. The presence of texture adds depth and structure which allows us define a better description of an image or scenario. An example is given in the UN Camp image pair in Figure 1.1.

The two images depict the same scenario using different image acquisition methods. The visual camera features very rich texture, in addition to edges and other salient features. The near-infrared (NIR) image has almost zero texture in a sense that the main objects are displayed as plain and monotonous. Intra-object intensity variation is almost zero, though edges are intact and, more crucially, human presence is detected. The objective therefore is to fuse the two images so that the respective object textures are preserved, but the human is also highlighted. A theoretically ideal fusion would be able to extract just the human outline from NIR and ‘paste’ it onto the detailed background of the visual camera so as to fully preserve the image texture.

Existing fusion metrics do not possess as strong an emphasis on textural representations,

(a) (b)

Figure 1.1: Example of texture (a) Visual image and (b) NIR image

but are more focused on edges. Due to this, the performance of current metrics on natural imagery is prone to inconsistency. It has been shown, for instance, that the Piella metric tends to give low scores to grass features. We therefore aim to establish a fusion quality measure, based on textural representations, that is suited to work in outdoor scenes.

1.2 Research Aims

1.2.1 Motivations for Research

The review of current literature in the previous section can be simplified into several underlying motivations that make up the basis of this research:

- Most of the literature on noisy fusion applications revolve around the estimation and *a priori* knowledge of noise signals. This introduces an extra step in processing and makes use of several statistical assumptions, which may or may not apply across a broad range of image datasets.

- Most works encompass fusion of images carrying low to moderate noise densities. There is a notable lack of fusion literature that test their approach on medium to high density noise conditions, owing to extreme circumstances that tend to occur in fusion applications. The lack of performance assessment in severe noise levels is based on the presumption that such

conditions are wholly unfeasible in which to perform fusion.

- Performance evaluation of fusion methods is a relatively new area within image fusion. Comprehensive and accurate portrayal of image fusion efficiency remains a challenging ordeal, in spite of numerous state-of-the-art fusion metrics in existence. This is mainly due to three reasons: the significant absence of ground truth images, the wide-ranging nature of multimodal image datasets especially for remote sensing and surveillance applications, and a lack of well-defined properties of saliency that constitute the aim of fusion. For these metrics, measures of image saliency are represented generally by gradients. Whilst true, objective evaluation of fusion methods may still benefit from other image analysis techniques. This includes using a novel measure of texturedness for image quality, which may be essential in particular for multimodal scenarios involving outdoor scenery and natural imagery.

1.2.2 Research Objectives

With the above issues in mind, we have set out to achieve the following objectives:-

1. To create a basal yet robust fusion algorithm with inherent noise removal properties, that maintains strong performance even in medium to high noise density scenarios. The approach shall require no prior knowledge of noise components.

2. To explore the possibility of merging multiple fusion algorithms; and to develop an image fusion approach that is able to combine the benefits of conventional methods within an algorithm fusion framework, for enhanced performance in noisy conditions.

3. To define a generic set of requirements that all objective fusion measures can adhere to.

4. To develop a new measure of image fusion performance that emphasises alternative means of feature saliency, mainly for the benefit of fusion applications that contain natural scenery.

Although a readily established branch of image processing, image fusion may still be open to ideas for exploration and development. The works on noisy fusion and the use of alternative

bases as a measure of saliency provide an incentive to further expand on this field and contribute for the benefit of future researchers. This has been the aim of the thesis.

1.3 Thesis Outline

This section highlights the outline of the thesis. Chapter 1 presents the problem background, more specifically issues surrounding fusion in noise corrupted images, its limitations and possibilities for improvements. It also deals with issues concerning performance measure for image fusion. Chapter 2 introduces the concept of image fusion. The chapter reviews the history of its inception, analyses the various techniques related to fusion and briefly describes the applications attributed to the area. This is followed by a thorough discourse in Chapter 3 on the expansion of Chebyshev polynomials for bivariate data and its use in image fusion, particularly in noisy scenarios. An enhanced method for noisy fusion is proposed in Chapter 4, that performs an algorithm fusion involving CP and ICA results which are demarcated according to active defined regions. Chapter 5 turns the attention to evaluation of image fusion performance, where a novel approach of utilising texture as a saliency measure for fusion assessment is proposed. Finally, conclusions and suggestions for future work are presented in Chapter 6.

The whole framework of the thesis can be incorporated into a flow chart, as shown in Figure 1.2.

1.4 Current Publications

Conference Papers

1. Two-dimensional Chebyshev polynomials for image fusion: Z. Omar, T. Stathaki and N. Mitianoudis, *Proc. Picture Coding Symposium*, PCSJ, Nagoya, Japan (2010)

2. Region-based image fusion using a combinatory Chebyshev-ICA method: Z. Omar, T. Stathaki and N. Mitianoudis, *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, IEEE, Prague, Czech Republic (2011)

3. GLCM-based metric for image fusion assessment: Z. Omar and T. Stathaki, *Proc. Int. Conf. Information Fusion*, ISIF, Singapore (2012)

Journal Paper

1. Region-based combinatory fusion for images in noisy conditions: Z. Omar, T. Stathaki and N. Mitianoudis, *Information Fusion Journal*, Elsevier, (2012) *(Currently under review)*

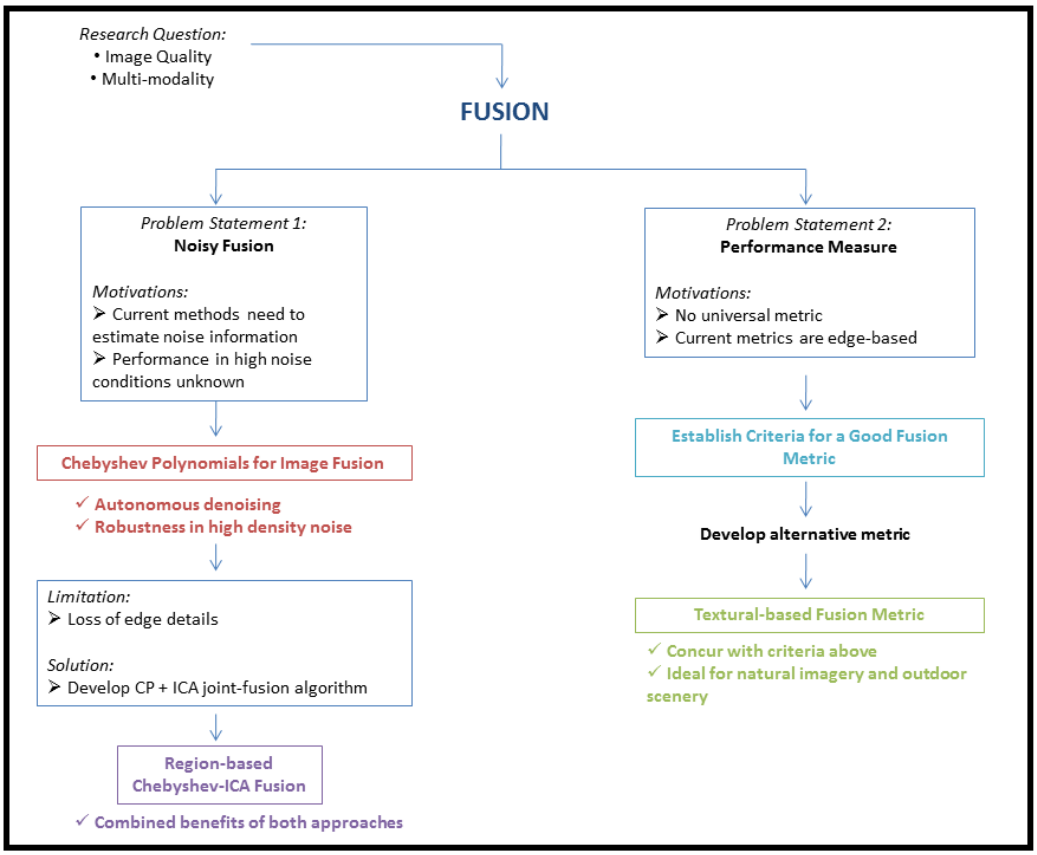


Figure 1.2: General framework of thesis

Chapter 2

The Basics of Image Fusion

2.1 Introduction To Image Fusion

In real world applications where various optical sensors are used for image acquisition, it is often difficult to obtain a good quality image from a single sensor alone. Decisions pertaining to system conditions are very rarely made upon the measurement of a single parameter. This condition remains true across many branches of modern technology be it medicine, geography or the military. The term 'good quality' itself tends to encompass various elements of the image scenery - illumination, sharpness, noise and contrast among others [28]. There exists a multitude of sensor tools which include optical camera, millimeter wave (MMW) camera, infrared (IR) and near-infrared (NIR), X-ray, radar, magnetic resonance imaging (MRI) among others [29], each of which tends to emphasise a different aspect of a captured image. In addition to sensor modalities, the pluralistic nature of input images is also necessary due to many other factors – occlusion of objects of interest due to smoke, fog and other unwanted objects, changing illumination in a scenery for photography applications e.g. daylight exposure at different times of day, and adjustable parameters within the sensors themselves, such as focal length [30, 31].

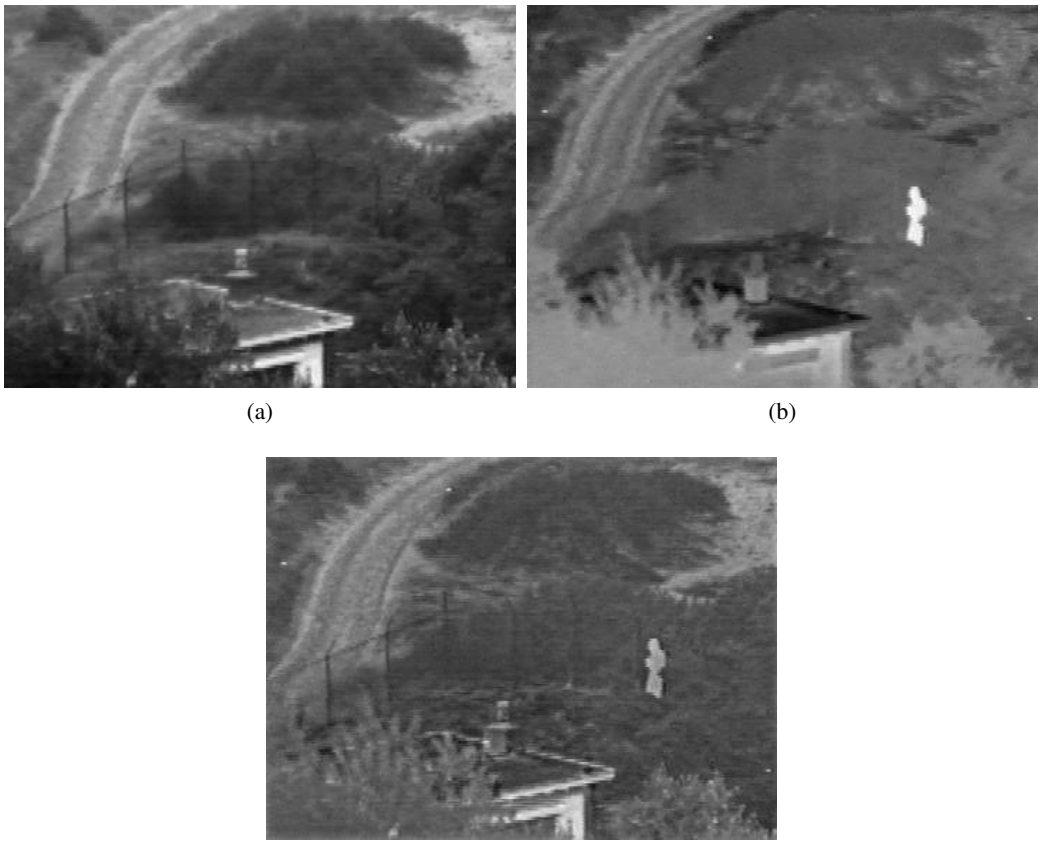
Notwithstanding the above, large amounts of data tend to contribute to problems common in signal processing including storage requirements, computation time, limited bandwidth and inconsistency with decisions pertaining sensor systems [32], not to mention the lack of a standard assessment criteria to measure sensors belonging to different modalities. It therefore makes sense

to reduce these multi-dimensional data into just a compact single image that preserves relevant information and whose quality exceeds any of its inputs. The successful merging of contrasting, but complementary, features from multiple sensors should therefore be the goal of the image acquisition system.

A simple example would be of the UN Camp sequence set [1] in Figure 2.1. A landscape is captured during night time using two image acquisition techniques. An NIR camera that detects strong thermal presence (within the 1.5-15μm spectrum) such as humans comprise one input. In general though, NIR sensors suffer from lower image resolution, prevalent image noise and the lack of availability of data sets which render them unsuitable for solitary use [33]. The second input is a standard image of the same scene, which is taken by a visual camera that captures strong textural background details (in the 0.45-0.7μm spectrum) but is severely limited in sparse illumination conditions. In this case the purpose is to enhance the lighting conditions in the scene and improve image qualities, so as to facilitate the detection of various moving objects and isolate pre-specified objects of interest i.e. tracking. Depending on user requirements, a good image must possess the ability to detect the human figure against the detailed backdrop of terrain and forestry. The formation of a good quality image is crucial as it enables us to have a proper understanding of the scenery context, which may prove decisive in real world surveillance and target recognition systems.

Therein lies the concept of fusion. As part of the grand challenge in image processing, image fusion aims to merge the salient aspects of two (or more) source images from these sensors to produce a singular output image, that contains all pertinent image features [34] and has a higher visual and numerical quality than either input, which may be essential in critical applications such as military surveillance. In this case, a simple correlation of pixels will not suffice due to the extremely diverse modalities employed for image acquisition. Fusion, also known by terms such as merging, integration, combination and synergy [35] exploits the synergetic relationship between two or more variables and makes better use of the available data at hand [36]. In modern times the definition of fusion often goes beyond the scope of algorithm alone, but rather is better regarded as a general framework that encompasses the combination of processing of images using algorithms,

[1] Image provided by TNO Human Factors Research Institute



(a) (b)

(c)

Figure 2.1: Example of image fusion showing (a) Visual image, (b) NIR image and (c) Fused image

means and tools to form a new image of greater quality [35]. The result may subsequently be evaluated using fusion metrics.

2.1.1 Benefits of Image Fusion

The purpose of fusion transcends the output image quality alone. It essentially enables users to visualise different sets of data under one scene. An important factor why fusion has been so successful is that engineers, developers and users are able to save costs in utilising signal processing techniques in lieu of designing an expensive system for image acquisition. Fusion reduces data dimensionality while preserving salient information content, thus reducing storage costs [30].

Sensory systems [33] focus mainly on how information can be extracted from sensory data. Fusion is therefore desired so as to improve visual accuracy and imply specific inferences that could not be achieved by a single sensor [36]. The fusion framework resembles the way humans locate their surroundings by using various cues from multiple modatilies. For instance, humans utilise binocular vision whereby they combine visual content from both the left and right eyes for visual processing. The advantages of this are several: firstly it gives a wider field of view, a maximum of 200 degrees horizontally of which 120 degrees makes up the binocular field, adjacent to 40 degrees each to the left and right uniocular fields [37]. Secondly, the parallax in vision created due to each eye's different relative position enables humans to generate depth perception and an enhanced ability to detect faint and far objects [38].

Fusion improves system reliability by reducing uncertainty in variables, thereby increasing accuracy [30] and aiding situational awareness [33]. It enables better decision making, localisation and discrimination of objects of interest, due to the potentiality of more complete information [33, 36]. Most importantly, it enhances and enables us to distinguish complementary information for the purpose of detection and segmentation in many security applications.

2.2 Abstraction Levels in Fusion

Image and data fusion literature generally categorises the framework into low, intermediate and high level fusion, or LLF, ILF and HLF respectively [36, 39–42].

Low-level Fusion

LLF, otherwise known as data level fusion merges raw real world data directly from the sources comprising pixels and signals. An example of this would be the pixel averaging method, where the mean values of the source images are taken pixel-by-pixel as the fused image.

Intermediate-level Fusion

Most approaches are categorised under ILF [43], also called fusion of features as the process involves extracting relevant features from the image using techniques such as multi-resolution analysis (MRA) and signal decomposition [30]. In [44], Nikolov et al. has classified image fusion algorithms into spatial and transform domain methods. Almost all fusion algorithms have since been based on a specific type of the transform domain, whereby a transform is performed on each input image and the transform coefficients undergo a fusion step. The resulting composite image is obtained by applying the inverse transform of the coefficients [15]. The central idea of transform-based fusion methods is to modify the magnitude of the source images' coefficients, so that edges and gradients are maximised.

Identifying notable and discreet salient image features such as segmented regions of interest and edge maps often incorporates a transformation process like pyramid [45] and wavelet-based [46] methods, blind signal processing approaches like independent component analysis (ICA) [47] and a non-parametric, data-driven analysis tool in empirical mode decomposition (EMD) [48]. The transform process decomposes an image into several parts and can illustrate hidden properties such as structural relationships between data, which in turn may enhance the information content and omit redundancies.

High-level Fusion

HLF, also called symbol level fusion is the highest level of abstraction and deals with fusion at the decision-making stage. Various confidence measures are taken into consideration and combined to draw a conclusive decision [36]. High-level fusion algorithms may also be in the form of image descriptions such as relational graphs [49]. A brief example is given in [30], where abstract

information rather than imagery data are used as inputs in the fusion process, and also in [50].

Figure 2.2 illustrates the levels of abstraction in fusion. Three procedures are involved in the system: input, process and output. At the input, a target of surveillance is captured on two sensor cameras, which convert it to digital images $I_1(x, y)$ and $I_2(x, y)$ respectively. During processing, LLF makes a direct comparison of input pixels whereas ILF performs a transform to elicit meaningful features from the images before fusing them. HLF merges information from the two input images regarding the possibility of the target being an object of interest.

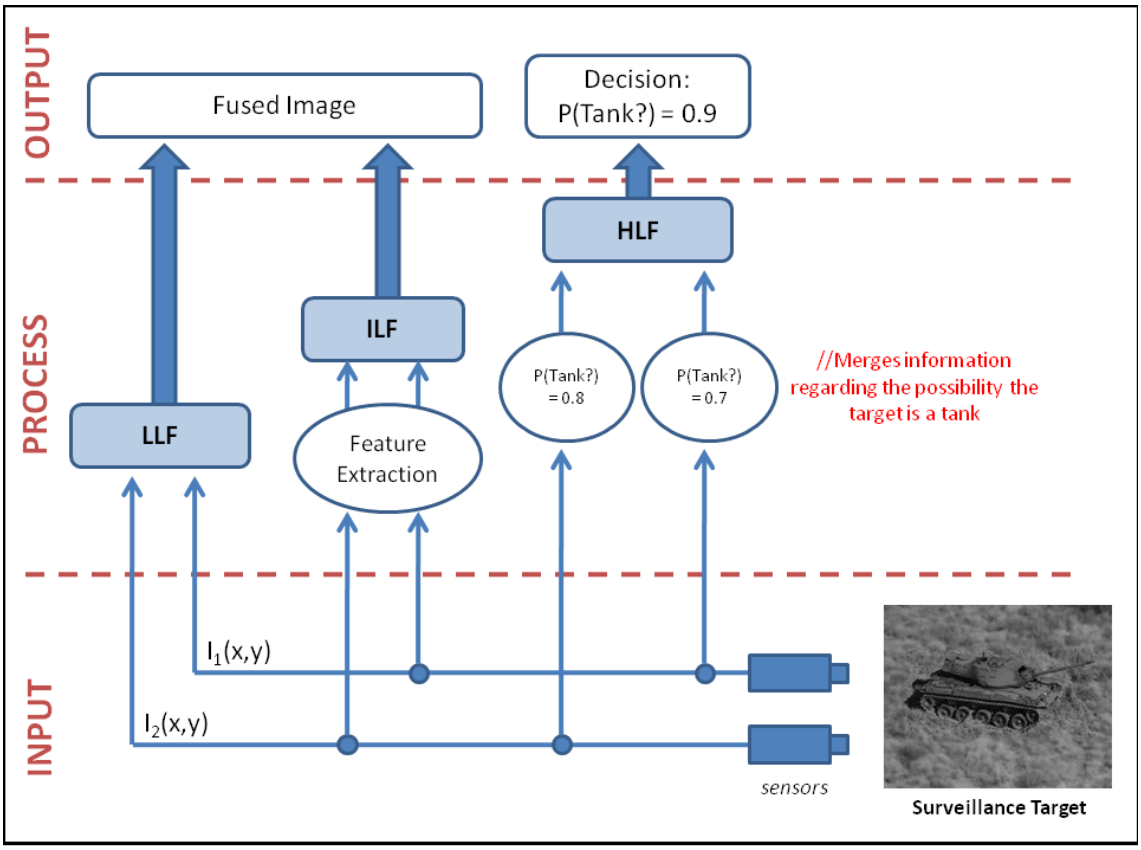


Figure 2.2: Fusion abstraction levels comprising LLF, ILF and HLF

2.3 History of Image Fusion

The advent of multisensory applications in the 1980's particularly in the field of remote sensing, coinciding with extensive research discoveries in pyramid-based transform methods [31, 51], introduced image fusion as a research area for the acquisition of higher quality images for human

visualisation [52].

As a research topic, image fusion can be uniquely perceived as in two ways:

(i) Abstraction-wise, as a sub-branch of data fusion, which also includes fusion of other data types such as audio, video and multi-dimensional and complex numerical data [36, 53]. Several examples have been put in practice, namely an intrusion detection system in cyberspace which fuses network data. In advanced vehicles, location estimates from a global positioning system (GPS) chip are merged with the on-board diagnostics (OBD) system of the car for automated navigation [36]. In medicine, electroencephalography (EEG) signals are combined with electrooculography (EOG) and respiratory signals for fatigue modelling of patients [54, 55].

(ii) Algorithm-wise, as an extension of image analysis methods that also comprise other image processing tools such as compression/coding [51], denoising [6], feature extraction [56], registration [57, 58], recognition [59] and segmentation. As it is, many of the image fusion approaches discussed in the thesis can generally be applied towards other image and signal processing applications. Examples include ICA for blind source separation of EEG and ECG signals in medicine [60, 61] and for facial recognition [59, 62] and the wavelet transform for dimensionality reduction for use in image coding.

2.3.1 Early Fusion Systems

The primary purpose of fusion was initially restricted to human observation and decision making. The earliest and most basic form of fusion is pixel averaging [31], whereby each pixel $I(x, y)$ of all input images is individually summed up and their average pixel value is incorporated into the fused image:

$$I_{\text{fused}}(x, y) = \frac{I_1(x, y) + I_2(x, y) + ... + I_n(x, y)}{n} \tag{2.1}$$

However this method is extremely crude and its results proved unsatisfactory [63]. The averaging method introduces artefacts especially when features present in only one input image is ‘super-

imposed' on the fused output, as can occur in photographic multiple exposure. It also causes pattern cancellation and contrast reduction [64] in the case where two inputs have features of equal salience but opposite contrast.

2.3.2 Pyramid-based Methods

A more tangible approach to image fusion is by pyramid decomposition [51, 65]. An image pyramid, an early form of multi-resolution analysis (MRA), comprises a set of filtered and scaled representations of the image. Fusion is performed through selection of coefficients at every scale from the source image pyramids, followed by the inverse transform of the resulting pyramid [66].

The pyramid method was first proposed by Burt (1984) [51, 67], who introduced the low-pass Laplacian pyramid for binocular fusion. A Laplacian pyramid is the bandpass equivalent of the Gaussian pyramid [45], and is obtained by the subtraction between two successive lowpass Gaussian pyramid levels. In ratio of low-pass (RoLP) pyramid [68] the level is scaled to a ratio of two from its preceding level, whilst contrast pyramids [69] is similar to RoLP but measures the ratio of luminence of a certain region within an image to the local background luminence. Eventually a host of improved pyramid-based schemes, including filter-subtract-decimate (FSD) [70], morphological [71] and gradient pyramids [72, 73] have been proposed and used in fusion literature.

1988 saw the first application of fusion on visible, thermal and infrared images through works by Lillquist [74], Nandhakumar and Aggarwal [75] and Rogers et al. [76], whilst Ajjimarangsee and Huntsberger [77] have suggested utilising neural networks for fusion of these modalities. A weakness of the neural network method is the large overhead entailed from processing whole images. MRA techniques overcome this by decomposing images into details and average channels, where fusion can be performed in the wavelet coefficient space [46].

To that end, pyramid decomposition approaches have been in widespread use within the image fusion community, though it is not without drawbacks. It was often found that fused images tend to contain blocking artifacts particularly in regions where the multi-sensory input data are significantly different in modality [64]. Another problem of pyramids is their lack of flexibility [78], i.e. lack of anisotropy and directional information.

2.3.3 Wavelet-based Methods

In 1993 Hunstberger and Jawerth [46] introduced us to a wavelet-based image fusion approach. Thereafter in 1995 a study by Li et al. [64] also used wavelets as alternative basis functions for multisensor image fusion, which is able to overcome the limitations of pyramid-based schemes by virtue of its directionality in which an image is decomposed into its low frequency approximation and horizontal, vertical and diagonal edges. Essentially, wavelets enable spatial information to be incorporated into the transform process. These efforts, and another work by Chipman et al. [79] kick-started a trend of using wavelet transforms for image fusion, on which a majority of current fusion algorithms in existence is based.

As opposed to the non-local and non-finite sinusoidal representations used by Fourier analysis, wavelets can be suitably used in finite domains and are a good fit for approximating data with sharp discontinuities (or edges). The wavelet transform works on the same premise as the pyramid, where the subband coefficients of the corresponding frequency content are merged and the subsequent inverse transform generates a synthesised fused image. Wavelet methods that are based on critically-sampling image signals, namely the discrete wavelet transform (DWT), suffer from shift variance whereby a discontinuity by a source signal could adversely affect its transform equivalent. This is a direct consequence of the downsampling operation of wavelet transforms. An alternative to this is to use oversampled schemes, though that would increase redundancy and add to the processing time and cost. A solution was proposed by Rockinger [80] in 1997 for a shift-invariant DWT (SI-DWT) method which discards the subsampling step, therefore rendering it overcomplete [81]. Further, Chibani [82] introduced a redundant wavelet transform (RWT), using an undecimated form of the dyadic filter tree which is best implemented via the *à trous* algorithm. Elsewhere the Haar wavelet [83] solves the shift variant DWT by circumventing the down-sampling move in the decomposition process and utilising a set of new filters throughout the process of each decomposition [72].

The development of complex-based wavelet transforms, namely the dual-tree complex wavelet transform (DT-CWT) [84] was able to overcome poor directional and frequency selectivity issues surrounding previous wavelet models, in addition to reducing overcompleteness and easily achieving perfect reconstruction [81, 85]. Its complex property means the phase information de-

rived from the transformation can be utilised for further analysis if necessary. Offshoots derived from the wavelet transform include contourlet [2], ridgelet [86] and curvelet [87] transforms that incorporate anisotropic behaviour and directional sensitivity to better facilitate the analysis of essential image features like edges.

2.3.4 Data-driven Methods

A perceived weakness of the wavelet transform, and similar transforms such as Fourier and Gabor, is the constant dependence of the basis functions on a fixed mathematical property that bears no relation statistically to the input data at hand, which often are non-linear and non-stationary [88]. In this regard, independent component analysis (ICA), empirical mode decomposition (EMD) [88] and other non-parametric and data-driven methods are considered superior as its features are directly derived from the training of data. Instead of a standard bases system using wavelets, a set of bases that are suitable for particular types of images may be trained for ICA [33]. A further explanation of the ICA process is found in Chapter 4.

In contrast, EMD is an entirely adaptive fusion approach that makes no assumptions of the data. EMD works in the spatial domain where it recursively deconstructs an image into intrinsic mode functions (IMF) at different frequencies. The decomposition method utilises envelopes associated with the local maxima and minima respectively. Fusion takes place through a weighted combination of IMF's of input images. Compared to wavelet coefficients, the IMF's are not fixed and can be suited to fit the data at hand [48].

ICA and EMD are both examples of non-parametric regression that requires a larger sample size of data, which in turn supplies the model structure and model estimates. The works on data modelling have since prompted a study into using biologically-inspired models, namely colour vision for fusion in 1995 by Waxman and colleagues [89, 90], in which opponent processing was applied on fusing visible and infrared images. Since then, new directions in image fusion have led to the development of a number of biological models, which are based on the way the human brain processes and combines information obtained by many different senses [49]. ICA, for example, assumes the solution to the 'cocktail party' problem in which auditory signal sources are distinguished by the human cognitive system.

2.3.5 State of the Art in Image Fusion

Another emerging trend in image fusion is the application of region-based methods [78, 91]. This approach is borne from the understanding that regions, or wholesome objects within an image tend to carry information of interest. It therefore makes sense to focus on regions as opposed to just individual pixels, since pixels can be processed more efficiently if they are treated as a collective group within a region rather than separate entities. Region-based fusion may therefore help to overcome some drawbacks of pixel-based fusion, like blurring, susceptibility to noise and misregistration [91].

Other notable approaches for image fusion include those based on statistical and estimation theory [16, 92], as first proposed by Sharma et al. [93] using Bayesian fusion. Particle models and Bayesian-based fusion achieve superior performance with high requirement applications, though this often comes at a cost of higher computational complexity [94].

Today a plethora of image fusion methods are in existence, each with their advantages and suitability designed for certain fusion applications. Significant improvements have taken place to close the gap between computer vision and the human perception of image quality; however the search for the gold standard in fusion remains.

2.4 Image Fusion Applications

It is generally acknowledged that all imaging applications that comprise analysis of multiple image inputs may benefit from image fusion [49]. Indeed, image fusion has been found to be very useful in a variety of critical applications such as remote sensing, medical imaging, industrial defect detection and military surveillance.

Digital Photography

In photography, the dynamic range of digital cameras tend to be restricted and far less than one would experience in real life. As a result, an acquired image may be under or over-exposed to light which detracts from its quality and makes the image undesirable for further use. Exposure

fusion [95] resolves this by merging a sequence of bracketed exposure images, by which the traditional steps involving computation of a high dynamic range and then tone mapping adjustments for display, can be circumvented.

A similar solution may be applied to multi-focal fusion [96–98]. Multiple focus images are an example of a depth-of-field system, as found in camera design or industrial inspection applications, where it is often difficult for optical lenses to simultaneously capture multiple objects in focus due to their limited depth-of-field. A way to overcome this is to capture multiple images, each focusing on localised objects, then using fusion to produce a globally-focused composite [2, 99]. This was summarised in [100], where Zhang and Blum have presented a comprehensive investigation of several multi-resolutionary fusion methods for a digital camera application.

Similarly, solutions to problems such as object occlusion and fog or smoke removal may also be found in image fusion.

Digital Watermarking

The increased availability and pervasive use of digital image data on the internet requires a system to manage and protect them from issues such as counterfeiting, manipulation and illegal duplication. This is achieved by digital watermarking [101], in other words digitally adding an invisible mark onto the host image to denote its ownership. A popular approach to watermarking is via the transform-level fusion of the host image and the watermark [102], due to the intrinsic similarity with common fusion approaches.

Remote Sensing

Remote sensing (RS) applications are concerned with the acquisition of geo-spatial images using aerial photography by satellites and airborne sensors, such as SPOT, QuickBird, IKONOS and IRS. RS aims to deliver high quality geographic images in terms of both spatial and spectral resolutions for purposes such as urban planning, agriculture and geology. Developing a high performance sensor camera to perform such tasks is unfeasible due to factors such as the radiation energy absorbed by the sensor and the limited data transfer rate from satellite platform to ground [52]. Rather, signal processing methods are utilised to achieve similarly high quality results [103, 104].

In [42] a detailed review of fusion techniques for RS was presented. In pan-sharpening, acquired data of a given scene comprise two modalities: a panchromatic (PAN) image depicting the scene in a high spatial resolution but in a single frequency, and a multispectral/hyperspectral (MS/HS) image that captures the landscape in a multitude of spectral resolutions across the wavelength spectrum though at 1:4 the spatial resolutions of PAN [105]. Fusion offers a practical and cost effective method to aid in distinguishing objects with salient information from RS imagery, by means of injecting the detailed spatial resolutions of PAN into a resampled version of multispectral images using methods such as the wavelet transform [106].

Classical fusion techniques in RS applications also include the intensity-hue-saturation (IHS) method [107] in which the red-green-blue (RGB) coloured domain of the original MS imagery is transformed into IHS to obtain a better separation of colour for fusion with PAN images, though it often produces spectral degradation.

Others include the principal component analysis (PCA), in which the MS image is decorrelated into several components. Fusion occurs by replacing the first/principal MS component with the PAN image, coupled with the Brovey transform that multiplies each MS band by the PAN image, and finally by the division of each product by the sum of the MS bands [27, 52, 105]. However these methods tend to ignore the need for high quality outputs of spectral information, which has proven essential in applications such as lithology and soil and vegetation analysis [108, 109]. High pass filtering (HPF) or modulation (HPM) of PAN inputs added to multispectral images are able to overcome this drawback [23, 105]. More recently, given the conciliatory nature of RS fusion between spatial resolution of PAN and spectral resolution of MS images, wavelet-based fusion techniques were found to be better equipped to handle this trade-off [110].

Medical Imaging

In medicine, image fusion and other technological advances are increasingly being relied upon for diagnostics and treatment of patients. An overview of medical image fusion was given by Pattichis et al. [55]. Fusion aids medical imaging by providing a complementary composite of various image formats stemming from multiple modalities, such as ultrasound, magnetic resonance image (MRI), computed tomography (CT), positron emission tomography (PET), and single pho-

ton emission computed tomography (SPECT), which in turn helps to delineate and distinguish targeted objects of interest such as tumours and blood vessels [111]. In radiation oncology, a treatment plan for radiotherapy involves CT data primarily for patients' dose calculation, while the outlines of tumour are better represented in MRI. For medical diagnosis, CT best illustrates denser tissues with low distortion, while MRI offers more comprehensive information on soft tissues with higher distortion and PET provides better information on blood flow with a generally low spatial resolution [54]. Using image fusion helps to distinguish important anatomical objects of interest of both sources.

Security

One of the most important goals of image fusion is to enhance object detections and target recognition [64] and reduce errors for applications such as military surveillance, visual sensor networks [112] and tracking [113, 114], and concealed weapon detection [78, 92] in the area of security. The use of multiple cameras belonging to different sensors has led to greatly increased performance in those areas [115], as well as improving the user's situational awareness [116] to facilitate better decision making for users [117]. A detailed example of surveillance-based fusion can be found in Section 2.1.

Another application is found in inspection systems such as the X-ray machine responsible for security at airports and federal buildings [21]. Detection failures and false alarms can be mitigated by the combined process of fusion, denoising and enhancement of dual-energy X-ray images.

Chapter 3

Chebyshev Polynomials for Image Fusion

This chapter discusses an image fusion method employing Chebyshev Polynomials which incorporates a low-pass filter for approximation. The numerical method assumes the orthogonal polynomials generated from the two-dimensional Chebyshev series expansion as a new set of basis functions, similar to Fourier analysis, where the coefficients are fused. This is especially useful for denoising, in which CP fusion impulsively purges high frequency noise components through a careful selection of the polynomial degree to be utilised.

3.1 Issues in Fusion for Image Denoising

A major concern of fusion is that real-life image acquisition applications are often susceptible to noise, therefore prone to corruption and error. This has been explained in detail in Chapter 1. In order to remove the noisy components, a method is required to transfer only salient features of the image inputs into a composite output image. Conventional fusion processes estimate a denoised image of acceptable quality, with the least amount of transferred noise possible. A standard approach to achieve joint fusion and denoising is to use an energy compact model, such as those in the transfer domain, whereby only prominent basis components are retained. To date there exist few techniques to specifically address this. Methods such as ICA have been successfully

tested for denoising, via sparse code shrinkage (SCS), on numerous occasions but suffer from the limitation of needing to estimate the noise variance beforehand [118]. Furthermore, most methods in existence lack the evaluation in high density noise conditions. This alludes to the notion that their performance tend to decrease dramatically in such conditions, which we will prove later in the chapter.

Our approach of using Chebyshev polynomials as basis functions for image fusion is tailored to address the two issues above. As shall be seen, CP performs a self-determined and broad suppression of image coefficients in order to remove high frequency noise components. Performance evaluation demonstrates the superior performance of CP over other methods in very noisy image scenarios.

3.2 Univariate Chebyshev Polynomials

Tchebichef or Chebyshev polynomials (CP) belong to a family of ultraspherical (Gegenbauer) polynomials, which themselves are a special case of Jacobi polynomials. CP has found its use in a broad range of applications including signal approximation and interpolation, numerical analysis and spectral methods for partial differential equation. This is due to the many similarities it shares with Fourier methods, and also to other advantageous features of Chebyshev as compared to other counterpart polynomials – mainly fast convergence and uniform approximation.

Classical Chebyshev polynomials of the first kind, $T_n(x)$ are defined as the mathematical solution to the Chebyshev differential equation

$$(1 - x^2)y'' - xy' + n^2 y = 0, \tag{3.1}$$

where the degree, n is a non-negative integer.

There is a host of ways as to how CP can be described, though the simplest and most popular is through the recursive generation of the polynomials i.e.

$$T_0(x) = 1, T_1(x) = x, T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{3.2}$$

The first ten CP's are given below,

$T_0(x) = 1$

$T_1(x) = x$

$T_2(x) = 2x^2 - 1$

$T_3(x) = 4x^3 - 3x$

$T_4(x) = 8x^4 - 8x^2 + 1$

$T_5(x) = 16x^5 - 20x^3 + 5x$

$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$

$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x$

$T_8(x) = 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1$

$T_9(x) = 256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x$

Chebyshev polynomials are closely related to sinusoidal functions, where their relations are defined as such [119]:

$$T_n(x) = \cos n\theta \text{ when } x = \cos\theta. \tag{3.3}$$

Chebyshev polynomials oscillate within the interval [1, -1], similar to a trigonometric function except for a few differences. With only two critical values, CP also counts amongst the Shabat polynomials. The polynomials have exactly n zeros and $n + 1$ local extrema in this interval, a distinguishing property that makes them attractive for signal interpolation. This will be touched on later in the chapter. The zeros of $T_n(x)$ are easily solved by

$$x = x_k = \cos\frac{(k - \frac{1}{2}\pi)}{n}, (k = 1, 2, ..., n) \tag{3.4}$$

and the extrema

$$x = \cos\frac{k\pi}{n}, (k = 0, 1, ..., n). \tag{3.5}$$

The extrema were all noted to have equal magnitude (unity) and tend to alternate in signs at the points between -1 and +1.

CP is orthogonal with respect to the weight function $\mu(x) = \frac{1}{\sqrt{1-x^2}}$. It is due to this orthogonality relation that the solution of a mathematical equation can be reduced to a simple inner product, as has been accomplished in [120] and [121]. The local extrema of CP is ± 1, as can be seen in Figure 3.1 [1].

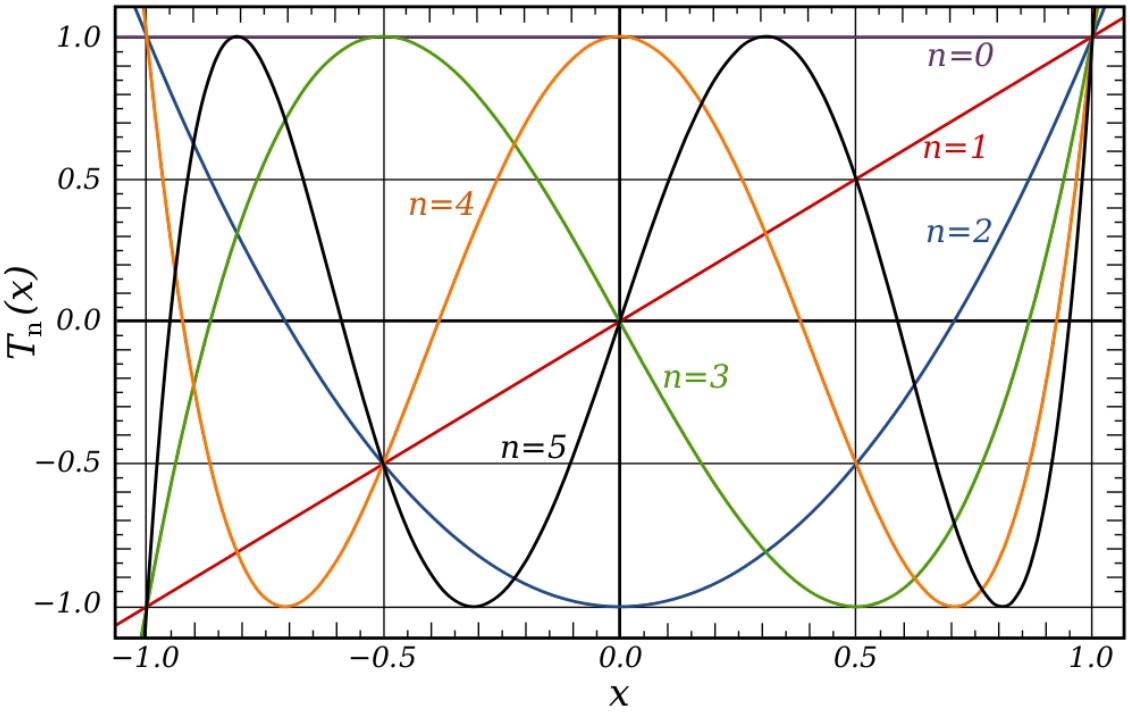


Figure 3.1: Chebyshev polynomials, up to $N = 5$ degrees

Figure 3.2 [2] displays a number of prominent orthogonal polynomials for degrees $n = 1, 2, .., 5$. The advantages of Chebyshev polynomials over Monic-Hermite, Jacobi, Legendre and Laguerre polynomials are evident: they are symmetrical, bounded within the interval $x = [-1, 1]$, have precisely n extrema with critical values $y = [-1, 1]$ and stable, i.e. least deviation from zero. An n-th order CP possesses n roots which can be derived analytically and are suitably used as nodes

[1] image taken from `http://en.wikipedia.org/wiki/Chebyshev_polynomials`
[2] image taken from `http://dlmf.nist.gov/18`

for interpolation.

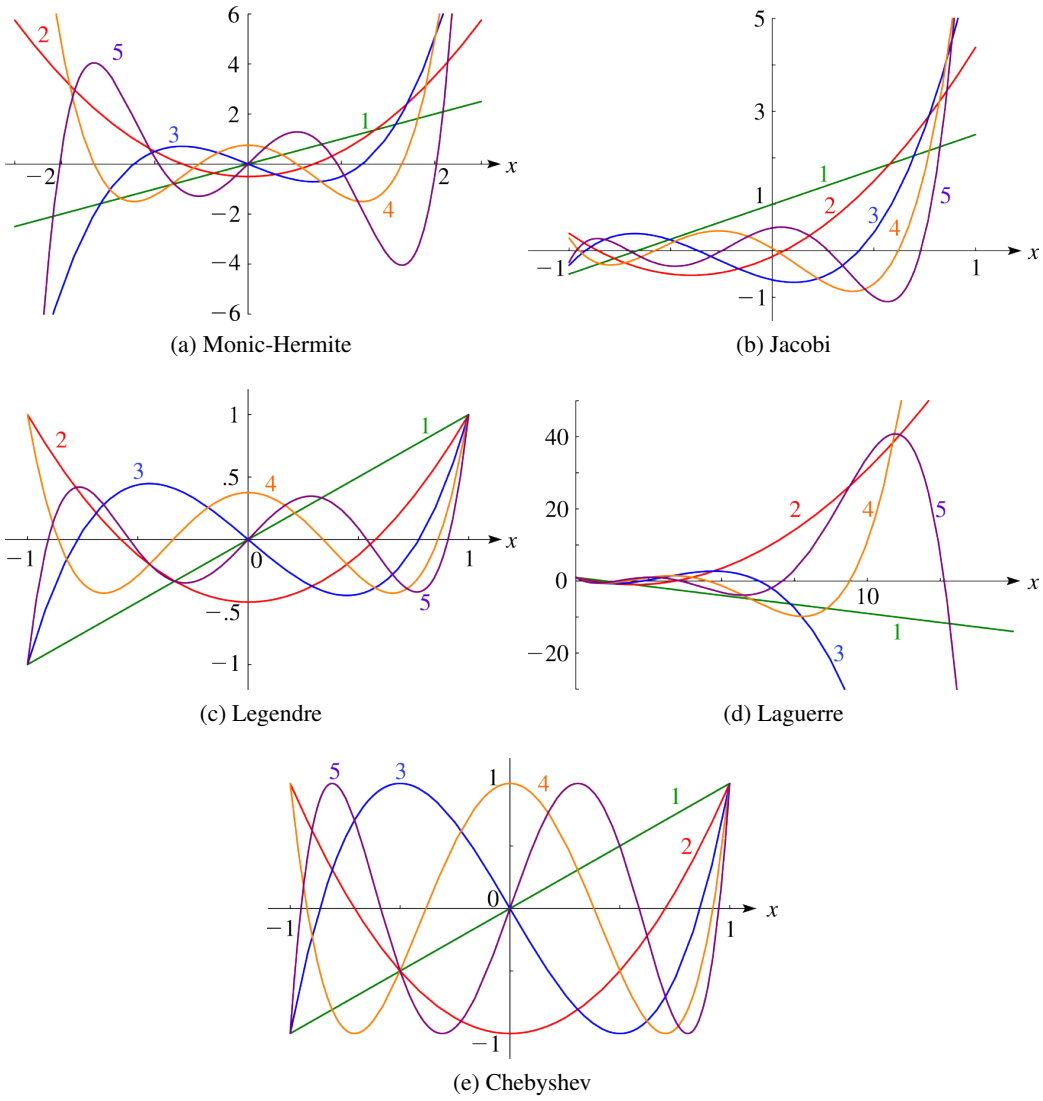


Figure 3.2: Comparison of orthogonal polynomials for degrees $n = 1, 2, ..., 5$

3.3 Chebyshev Polynomials for Signal Approximation

To approximate a function is to replace, replicate or mimic that function with a simpler form of representation, such as a polynomial series. Approximation has proven useful in signal processing theories as not only it is more compact to represent and store, but enables easier manipulation of data.

Minimax approximation of a signal aims to recreate an estimate that is as close as possible to the original signal. Chebyshev polynomials are considered important in the theory of signal approximations because the roots serve as nodes, named Chebyshev nodes, for the polynomial interpolation of virtually any given signal.

Runge's phenomenon is a problem comprising large errors in oscillation near the interval edges caused by high degrees of polynomial interpolation, another similarity shared with Fourier functions that experience ringing from Gibb's phenomenon. Ringing is borne from the fact that good synthesis functions such as wavelet reconstruction filters are normally in oscillation [27], and occurs due to loss of high frequency components. Many polynomials, such as the Taylor series, suffer from Runge's problem because of their use of equidistant interpolation points, in which approximation norms tend to grow larger as the polynomial degree increases. A low set of polynomial degrees, however, tends to limit approximation accuracy and induces low quality outputs.

Chebyshev polynomials mitigate the Runge's effect by introducing nodes that gravitate towards the edges on the interval [-1,1] with the asymptotic density given by $\frac{1}{\sqrt{1-x^2}}$, thereby minimising the oscillation. For CP, errors are evenly spread – it ensures the maximum error in approximation diminishes as the polynomial order increases. The formula is subsequently used as the weight factor for the condition of CP orthogonality.

The expansion of the Chebyshev series [122], $\tilde{f}(x)$ forms a complete basis set in a Sobolev space and is used to model almost any function $f(x)$ through the formula

$$\tilde{f}(x) = \sum_{n=0}^{N-1} a_n T_n(x) \tag{3.6}$$

where a_n is an approximated coefficient that closely resembles the equation,

$$a_n \approx \frac{2}{\pi} \sum_{x=-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x) T_n(x) \tag{3.7}$$

In the context of digital signal processing, the discretisation process entails the data/pixel values to fit the interval [-1,1] through the scaling transform below,

$$x_a = \frac{2(a-1)}{(L-1)} - 1, a = 1, 2, ..., L. \tag{3.8}$$

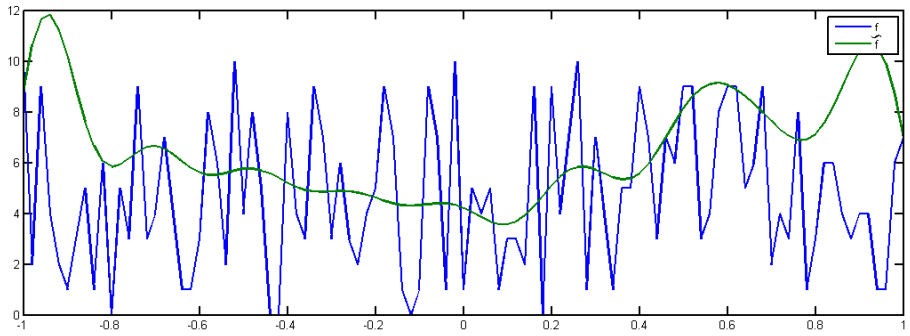
where L is the size of data. For the sake of symmetry, odd-valued choices for L are recommended.

An optimal approximation requires a balanced choice for both L and the order of Chebyshev polynomials, n [121]. Higher order CP's normally equate to a more accurate approximation, whilst the choice of window size L is equally imperative. Short data lengths are naturally unable to capture signal variance comprehensively, whilst an overly large data size may result in a non-stationary signal that is too complex to model. The relationship $R = \frac{n}{L}$ is illustrated in Figure 3.3.

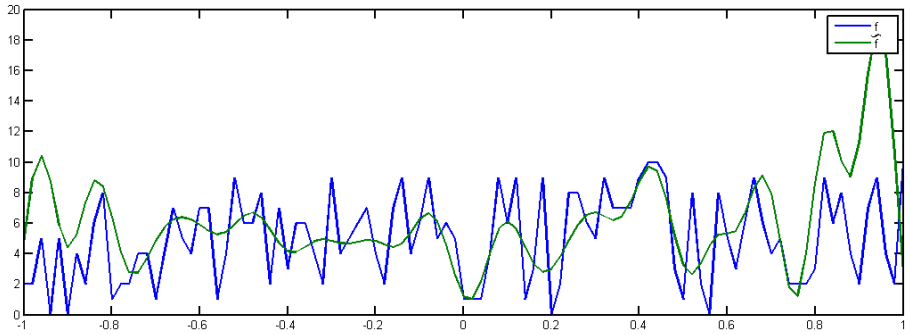
3.3.1 Some Properties of Chebyshev Polynomials

Chebyshev polynomials hold several properties that are of particular benefit in image approximation. An attractive feature of Chebyshev over others, such as the Taylor series and the discrete Chebyshev polynomials (see Appendix B) is its fast rate of convergence. It is this redeeming property that also sees the pervasive use of Chebyshev polynomials for designing relatively inexpensive IIR filters and array beamformers. As with the discrete cosine transform (DCT), CP coefficients can be derived either via analytical expression or by a look-up table.

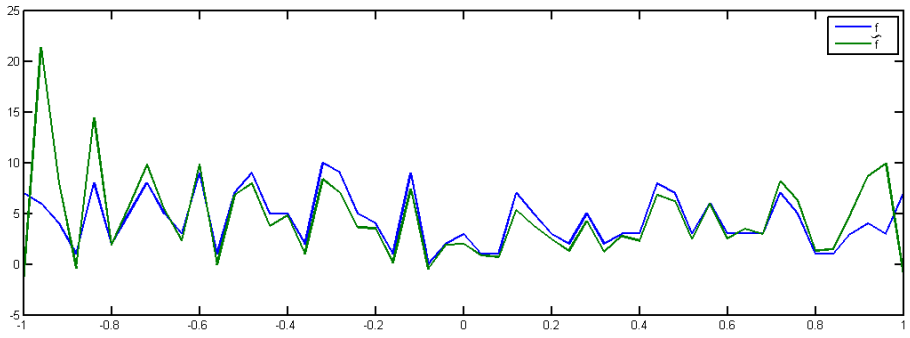
Moreover, the components of Chebyshev, as with all other polynomials, are arranged according to degree order. Thus a finite order n-th degree used in CP expansion enables the most basic signal features to be retained while more complex polynomials are omitted. In this regard CP allows more flexibility in designing a system than wavelets. It also holds an advantage over ICA and EMD methods, among others, in which a signal is decomposed into a set of equally independent basis components that are randomly arranged. For instance, ICA bases perform better with the specific type of images they were trained for. However due to the same reason, CP performs better under medium to heavy noise conditions. This is an important premise in our work, as the difference is evident in performance evaluation of noise corrupted images.



(a) R=0.5



(b) R=1



(c) R=10

Figure 3.3: Univariate Chebyshev approximation ($f(x)$ in blue, $\tilde{f}(x)$ in green). R is the ratio of CP order $n = 10$ over data size L.

3.4 Chebyshev Polynomials as Separable Basis Functions for Image Approximation

3.4.1 Rationale of Approach

The assertion that noise is to be considered as abrupt and random interferences within signals has been established in the first chapter. Theoretically, noise components tend to occupy the higher frequency spectrum. As higher-order polynomials comprise complex and high frequency components, our premise therefore is to truncate the CP coefficients below a certain threshold for a sparse representation. This allows for the removal of noise components at the cost of sacrificing high energy information, including edges and strong texture. Effectively, CP approximation acts as a low-pass filter that eliminates unwanted noise at the expense of lower signal accuracy. This approach has proven useful in image fusion [34, 123].

3.4.2 Derivation of Bivariate Coefficient

Approximation of bivariate (two-dimensional) signals i.e. images [124] is achieved by two separable univariate classical Chebyshev polynomials $T(x)$ and $T(y)$, that exploits its orthogonality in a similar way to DCT and the Haar transform. This allows for a lower computational complexity, as the same principles from univariate CP can now be applied. This definition precedes two other approaches for a two-dimensional Chebyshev polynomial expansion, as clarified in Appendices A and B, where a numerical quadrature formula has to be introduced to properly allow the evaluation of the discrete sum on a suitable points grid [121].

For images, processing is performed row-wise (x) followed by columns (y). We employ the same approximation algorithm and weight factor as those described in Section 3.3. Thus, the following bivariate CP approximation of a signal is introduced:

$$\tilde{f}(x,y) = \sum_{m=0}^{M} \sum_{n=0}^{N} a_{m,n} T_m(x) T_n(y), \tag{3.9}$$

Concerned with finding $a_{m,n}$, the left hand side is multiplied using the orthogonality relation,

$$A = \sum_{x=-1}^{1} \mu(x) f(x,y) T_k(x) =$$
$$\sum_{x=-1}^{1} \mu(x) \sum_{m=0}^{M} \sum_{n=0}^{N} a_{m,n} T_m(x) T_n(y) T_k(x) =$$
$$\sum_{m=0}^{M} \sum_{n=0}^{N} a_{m,n} T_n(y) \sum_{x=-1}^{1} \mu(x) T_m(x) T_k(x) \tag{3.10}$$

From [122], it is known that

$$\sum_{x=-1}^{1} \mu(x) T_m(x) T_k(x) = \frac{\pi}{2} \delta(m-k) \tag{3.11}$$

Given that $\delta(m-k) = 0$ if $m \neq k$ and $\delta(m-k) = 1$ if $m = k$, we now have

$$A = \frac{\pi}{2} \sum_{n=0}^{N} a_{k,n} T_n(y) \tag{3.12}$$

Having completed x, the transformation is applied along y, therefore

$$\sum_{y=-1}^{1} \mu(y) \sum_{x=-1}^{1} \mu(x) f(x,y) T_k(x) T_l(y) =$$
$$\frac{\pi}{2} \sum_{n=0}^{N} a_{k,n} \sum_{y=-1}^{1} \mu(y) T_n(y) T_l(y) = \frac{\pi^2}{4} a_{k,l} \tag{3.13}$$

so finally we have

$$a_{m,n} = \frac{4}{\pi^2} \sum_{x=-1}^{1} \sum_{y=-1}^{1} \mu(x) \mu(y) f(x,y) T_m(x) T_n(y). \tag{3.14}$$

3.5 A Novel Method of Image Fusion with Chebyshev Polynomials

3.5.1 Fusion Rules

The process of selecting important features from the respective input images is referred to as fusion. The outline of the transform domain-based fusion is as follows:

$$I_f(x,y) = T^{-1}\left\{g\left(T\left\{I_1(x,y)\right\},...,T\left\{I_n(x,y)\right\}\right)\right\} \tag{3.15}$$

where $I_f(x,y)$ is the fused image and $I_1(x,y),...,I_n(x,y)$ are the input images, $T(\cdot)$ is the transform operator (in this case Chebyshev approximation) and $g(\cdot)$ is the fusion rule operator.

The fusion rule [125] lays the guideline for the fusion step at its most basic level. Among the regular fusion rules are:

- *Mean*: taking the average value of the input coefficients, ensuring an altogether fair representation of all inputs within the fused image. This also means that edge details tend to be over-smoothed.

$$T\left\{I_f(x,y)\right\} = \frac{1}{n}\sum_{i=1}^{n} T\left\{I_i(x,y)\right\} \tag{3.16}$$

- *Max-abs*: taking the coefficient value with the highest absolute value. This enhances the strong pixels throughout the image region. However the intensity information in constant background areas are prone to slight distortion.

$$T\left\{I_f(x,y)\right\} = \text{sgn}\left(T\left\{I_i(x,y)\right\}\right)\max_i\left|T\left\{I_i(x,y)\right\}\right| \tag{3.17}$$

- *Weighted combination (WC)*: also called weighted average (WA) [31]. An extension of max-abs which addresses the limitations of the previous rules. The fused coefficient $\underline{u}_f(t)$

is given as

$$T\left\{\underline{I}_f(x,y)\right\} = \sum_{k=1}^{T} w_k(x,y) T\left\{\underline{I}_k(x,y)\right\} \tag{3.18}$$

The contribution of each image coefficients is measured via a weight factor $w_k(t)$, represented by an activity indicator. The weights act to emphasise those sources with more intense activity. For WC, a balanced size of image patch is required so as to contain a meaningful degree of information variance within the image.

3.5.2 Fusion Example

An n-th order CP expansion ensures that only the prominent information is retained, while any redundant statistics are discarded [126]. An example of image fusion using Chebyshev polynomials is described below. Fusion is performed by a direct comparison of the coefficients of the two or more input images. As the polynomials are derived from formulae, they can be pre-computed and stored in order to reduce processing complexity.

Let there be two image approximations,

$$f_1(x,y) = \sum_{m=0}^{M}\sum_{n=0}^{N} \alpha_{m,n} T_m(x) T_n(y) \tag{3.19}$$

$$f_2(x,y) = \sum_{m=0}^{M}\sum_{n=0}^{N} \beta_{m,n} T_m(x) T_n(y) \tag{3.20}$$

The fused image approximation $f_{\text{fused}}(x,y)$ is formed by fusing the two coefficients $\alpha_{m,n}$ and $\beta_{m,n}$ via the max-abs fusion rule.

$$f_{\text{fused}}(x,y) = \sum_{m=0}^{M}\sum_{n=0}^{N} \max\{|\alpha_{m,n}|, |\beta_{m,n}|\} T_m(x) T_n(y) \tag{3.21}$$

The coefficient corresponds to global image strength, therefore performing max-abs over the two images enables the preservation and enhancement of strong pixels while suppressing weak ones.

This fusion process essentially results in an improved quality image.

3.6 Performance Evaluation of CP Fusion

Image fusion performance using Chebyshev polynomials is evaluated using a set of two grayscale image inputs. Extension to three or more images and to RGB images can be developed in a similar way. ICA-based fusion and DT-CWT fusion methods using the max-abs rule [127] are used in the experiments as benchmarks. A range of image sets that comprise multimodal and multifocal scenarios were tested. The experiments aim to test our presumption of CP's superior performance under the presence of noise.

Four levels of wavelet decomposition were applied on the DT-CWT method. For ICA, training was carried out on 1000 patches belonging to images of similar content, each measuring 7×7 pixels, and a total of 48 independent bases were derived. After several configurations, we concluded that the size of the patches is large enough without being overly complex, and presents a good balance between information content and computational complexity. As for CP, $m = 11$ and $n = 11$ degree of Chebyshev polynomials were chosen and 7×7 overlapping patches were used – for similar reasons as above. Overlapping is performed by a shift of one pixel per iteration. All fusion methods utilise the max-abs fusion rule.

3.6.1 Experiment 1: Multimodal Images

Images of multiple modalities are tested and the results are shown in Figure 3.4, Figure 3.5 and Figure 3.6. The initial experiments are performed simply to test the plausibility of CP as a fusion algorithm. The results concur with our notion that Chebyshev polynomials present a viable alternative to image analysis techniques, and its application in image fusion is considered valid and meaningful.

3.6.2 Experiment 2: Multifocal Clock

The second experiment involves a multifocal clock example and the results are as illustrated in Figure 3.7. Multiple focus images are an example of a depth-of-field system. As in camera design

(a) Input 1 - visual camera

(b) Input 2 - IR

(c) ICA

(d) DT-CWT

(e) CP

Figure 3.4: Multimodal gun example

(a) Input 1 - front partially obscured

(b) Input 2 - back partially obscured

(c) ICA

(d) DT-CWT

(e) CP

Figure 3.5: Occluded jet and cloud example

(a) Input 1 - visual camera

(b) Input 2 - IR

(c) ICA

(d) DT-CWT

(e) CP

Figure 3.6: Multimodal UN Camp example

or industrial inspection applications, it is often difficult for optical lenses to simultaneously capture multiple objects in focus due to their limited depth-of-field. A way to overcome this is to capture multiple images, each focusing on different objects, then using fusion to produce a fully-focused composite [91].

Preliminary observations suggest that the CP output is of less quality than ICA and DT-CWT. This is attributed to the smooth property of CP approximation that filters out edges and other high frequency components, creating a blurred effect on the image.

The input image Clock 1 is subsequently corrupted with various levels of noise. Additive white Gaussian noise (AWGN) with zero mean and unit variance is used, with various degrees of strength according to the signal-to-noise ratio (SNR):

$$SNR = 10 \log \frac{P_{\text{image}}}{P_{\text{noise}}} \tag{3.22}$$

where P_{image} is the power spectrum of the input image and P_{noise} is the corresponding power spectrum for the Gaussian noise. $\mu = 0$ thus $P_{\text{noise}} = \sigma^2_{\text{noise}}$.

Figures 3.8 to 3.13 show the results. Starting with 21dB SNR, at which level noise is minimal, CP's blurring effect is still evident as opposed to ICA and DT-CWT which displayed very good results. As the noise increases with 17 and 14dB SNR we observe the detriment in quality in ICA and DT-CWT images; whilst the effect on CP has been nominal due to its intrinsic smoothing algorithm. This is more distinct in Figure 3.11, Figure 3.12 and Figure 3.13 where the noise levels are further increased with SNR levels 11, 7 and 4dB. Images for ICA and DT-CWT become heavily degraded in very noisy conditions, despite using sparse code shrinkage for denoising.

To further affirm the visual evaluation, the experiment results for noisy images are subsequently placed under objective evaluation using the Petrovic and Piella fusion metrics respectively [125]. A more thorough description of objective evaluation of image fusion methods is provided in Chapter 5.

Table 3.1 validates our subjective perceptual assessments. CP fusion exhibits a slightly lower score for multifocus images in Figure 3.7, but performs much better against ICA and DT-CWT in tests involving corrupted signals in Figures 3.8 to 3.13. The incongruity is largely due to the finite

(a) Clock 1 - background focused

(b) Clock 2 - foreground focused

(c) ICA

(d) DT-CWT

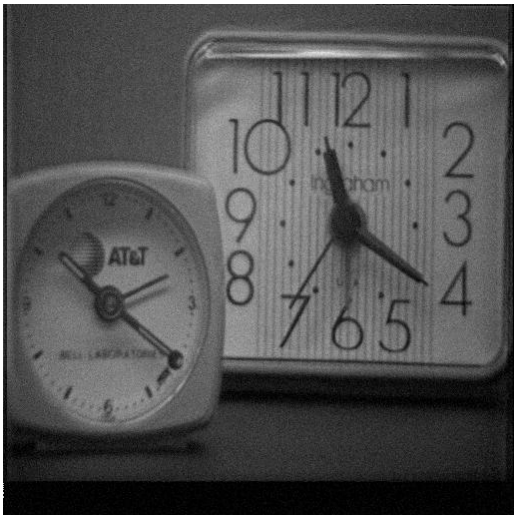
(e) CP

Figure 3.7: Multifocus clock example

(a) Input image, SNR 21dB

(b) CP

(c) ICA



(d) DT-CWT

Figure 3.8: Fusion performance on images with Gaussian noise, SNR = 21dB

(a) Input image, SNR 17dB

(b) CP

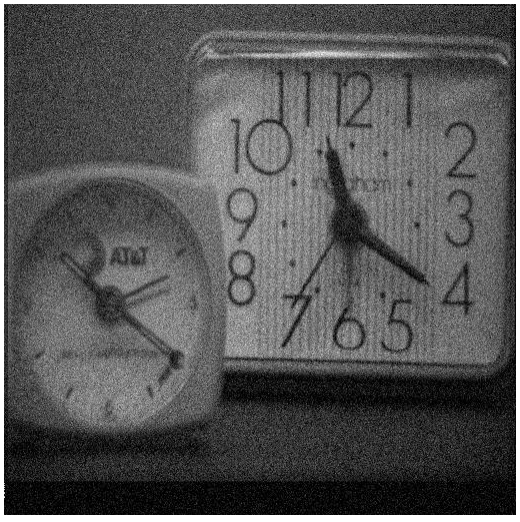
(c) ICA

(d) DT-CWT

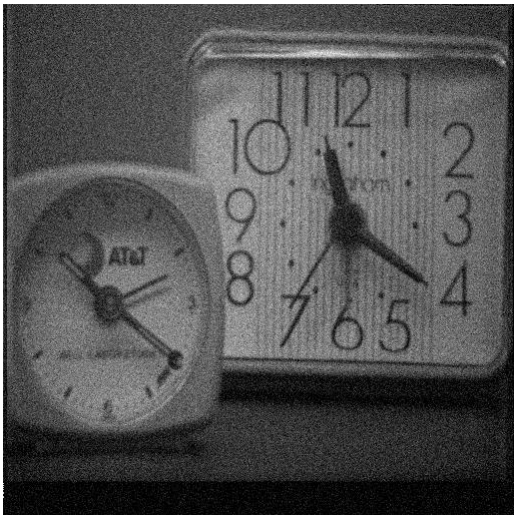
Figure 3.9: Fusion performance on images with Gaussian noise, SNR = 17dB

(a) Input image, SNR 14dB

(b) CP

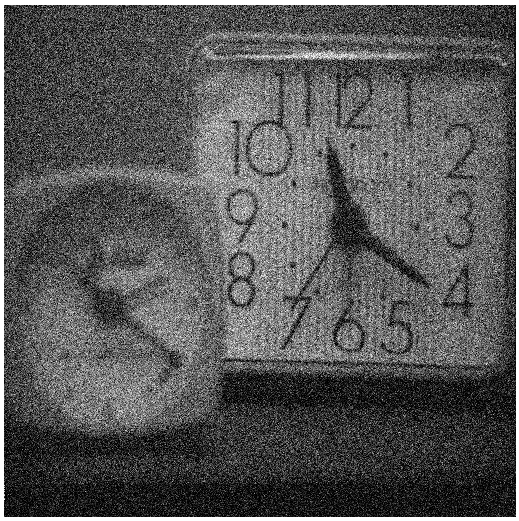


(c) ICA



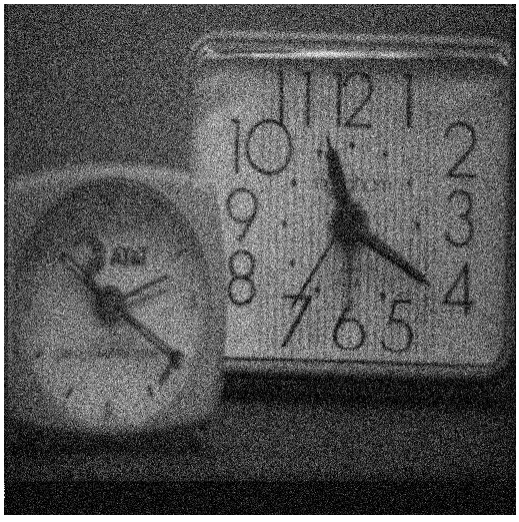
(d) DT-CWT

Figure 3.10: Fusion performance on images with Gaussian noise, SNR = 14dB

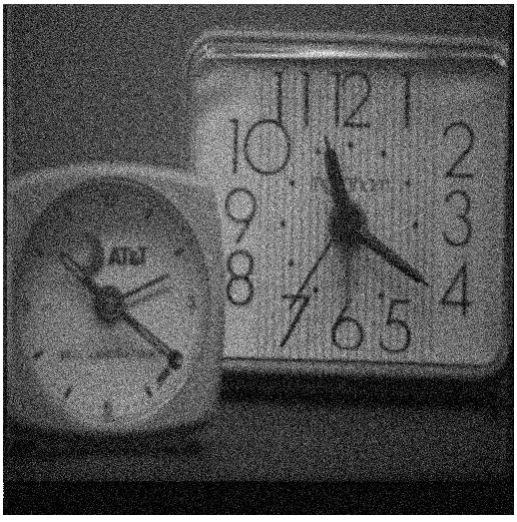


(a) Input image, SNR 11dB

(b) CP

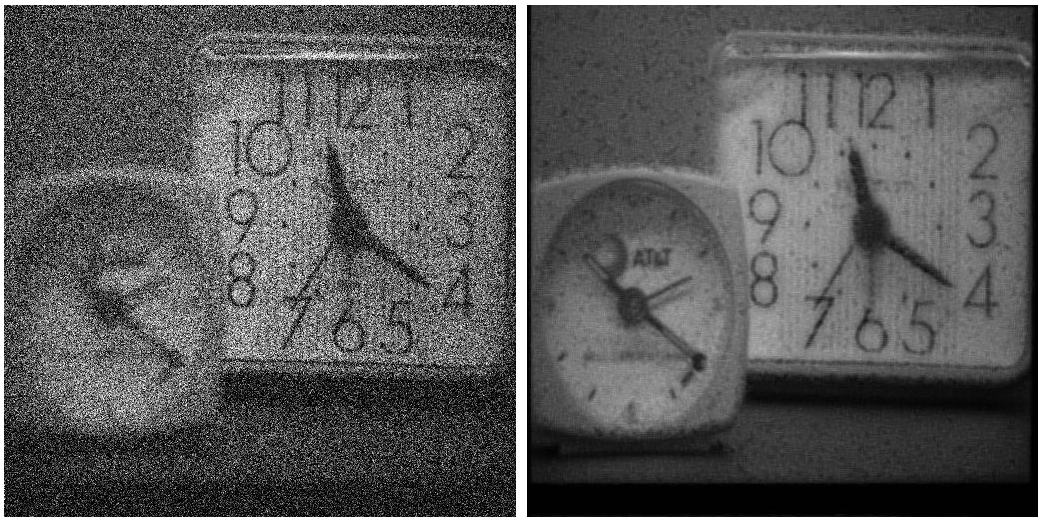


(c) ICA

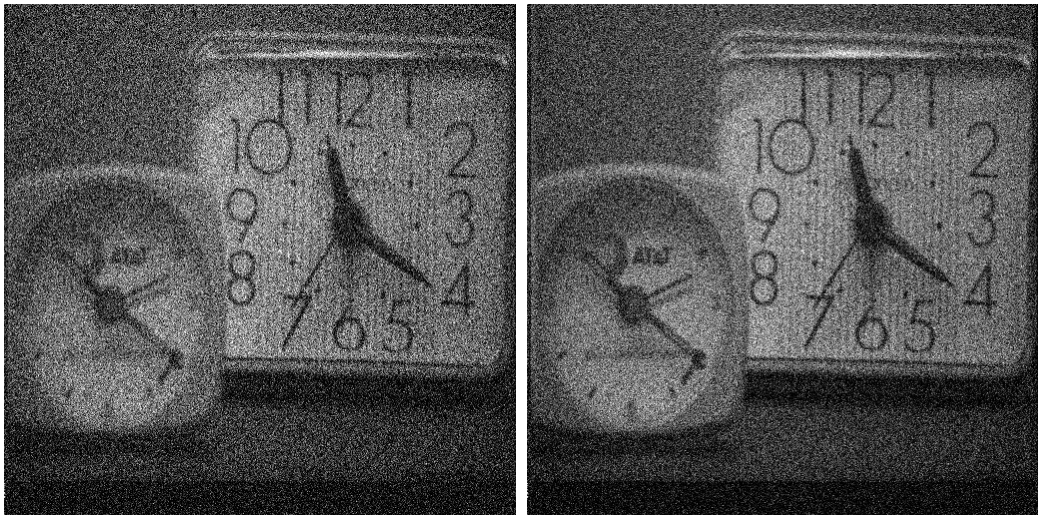


(d) DT-CWT

Figure 3.11: Fusion performance on images with Gaussian noise, SNR = 11dB

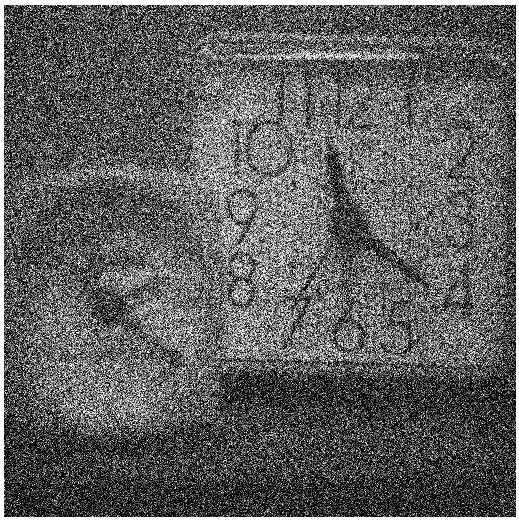


(a) Input image, SNR 7dB (b) CP

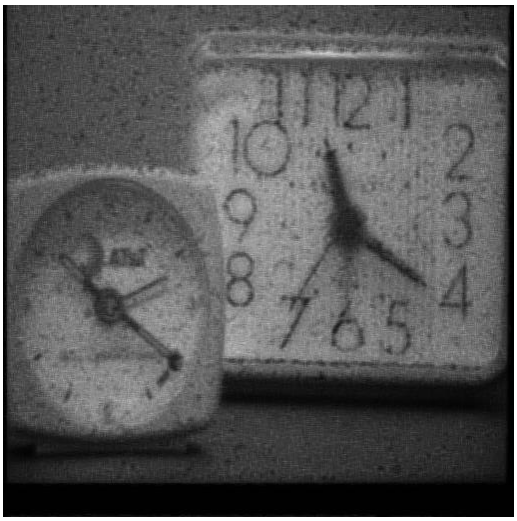


(c) ICA (d) DT-CWT

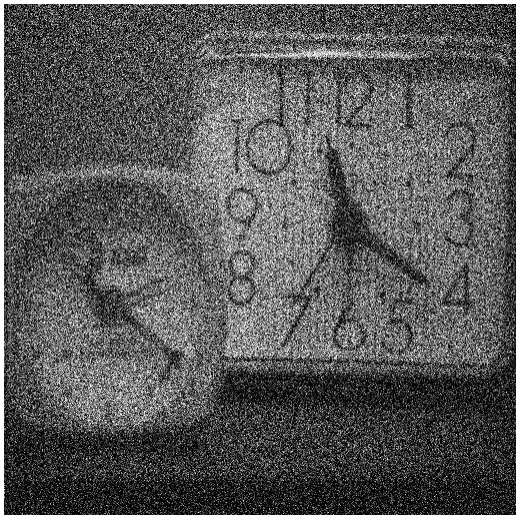
Figure 3.12: Fusion performance on images with Gaussian noise, SNR = 7dB



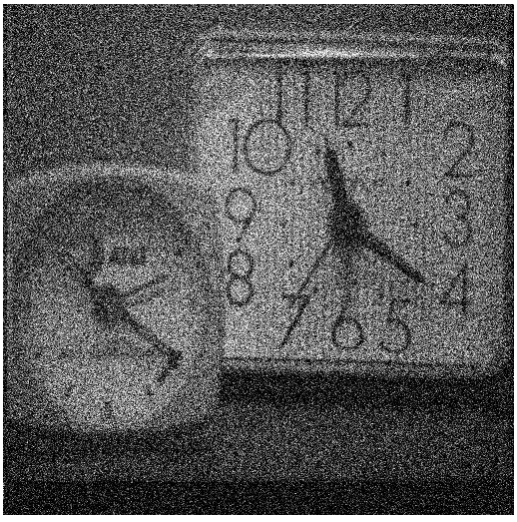
(a) Input image, SNR 4dB



(b) CP



(c) ICA



(d) DT-CWT

Figure 3.13: Fusion performance on images with Gaussian noise, SNR = 4dB

nature of CP approximation, where any large amounts of noise in the higher frequency spectrum are fundamentally filtered out.

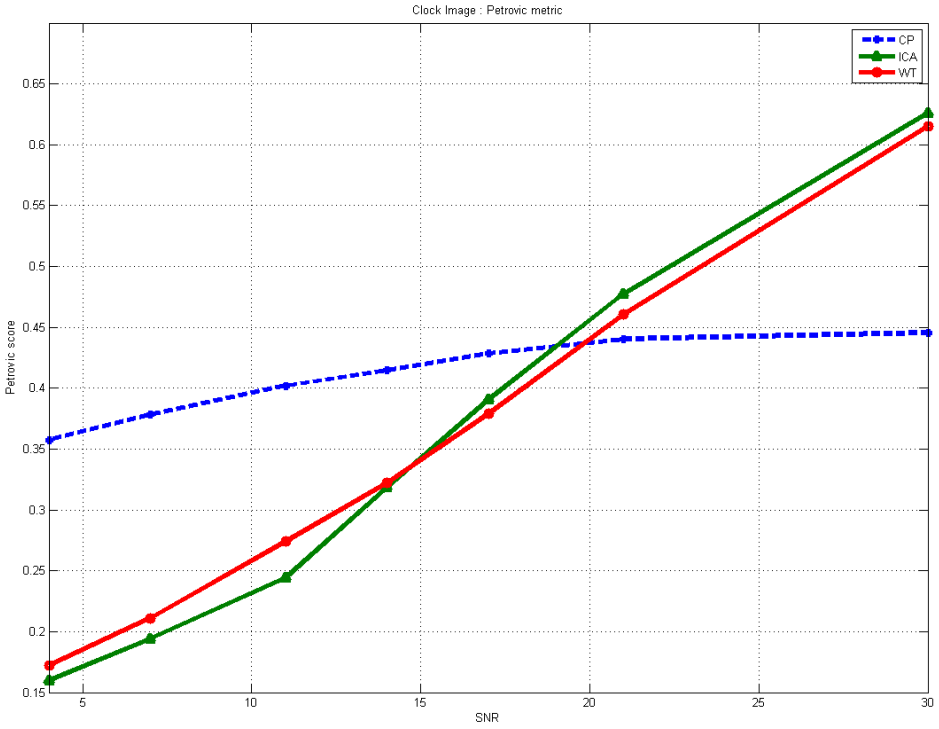
Table 3.1: Fusion metric evaluation for clock images – ICA, CP and DT-CWT

Input	Method	Petrovic	Piella
Multifocal clock	CP	0.4454	0.8258
	ICA	**0.6264**	**0.9165**
	DT-CWT	0.6151	0.9067
SNR 21dB	CP	0.4403	0.8224
	ICA	**0.4775**	**0.8462**
	DT-CWT	0.4606	0.8363
SNR 17dB	CP	**0.4286**	**0.8170**
	ICA	0.3906	0.7353
	DT-CWT	0.3791	0.7252
SNR 14dB	CP	**0.4147**	**0.8078**
	ICA	0.3184	0.5846
	DT-CWT	0.3221	0.6129
SNR 11dB	CP	**0.4020**	**0.7946**
	ICA	0.2441	0.4103
	DT-CWT	0.2739	0.4967
SNR 7dB	CP	**0.3783**	**0.7542**
	ICA	0.1942	0.2785
	DT-CWT	0.2110	0.3418
SNR 4dB	CP	**0.3574**	**0.6947**
	ICA	0.1600	0.1901
	DT-CWT	0.1720	0.2320

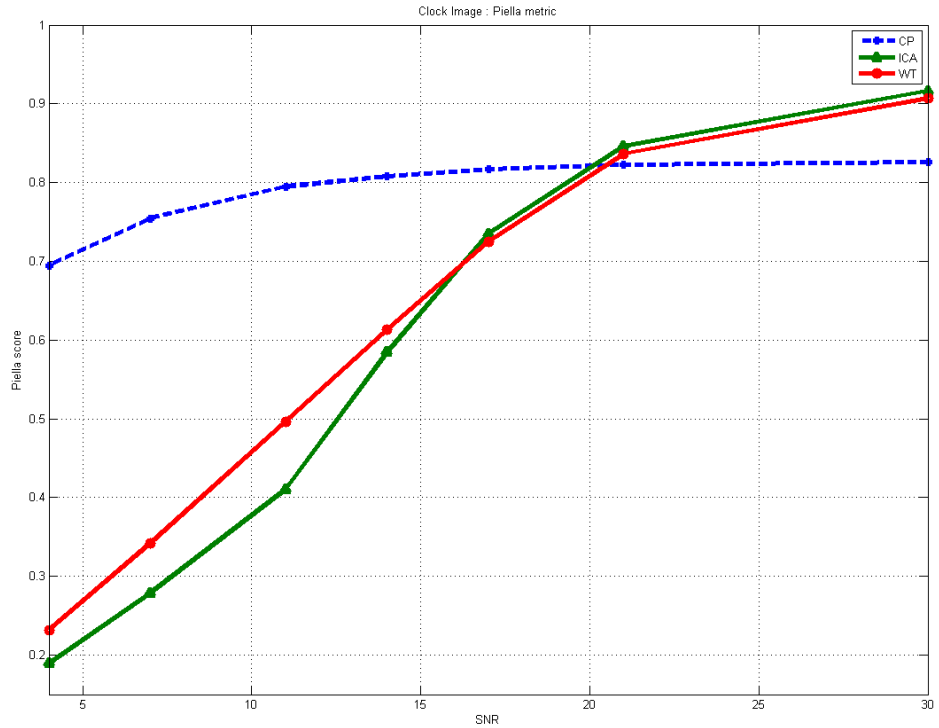
The objective evaluation results are plotted in a graph in Figure 3.14 for Petrovic and Piella metrics respectively. Both graphs display a similar trend in showing that the degradation of fusion quality is significantly more adverse for ICA and DT-CWT than it is for CP, which exhibits an almost linear degradation proportional to the level of noise. On the other hand, both the Petrovic and Piella metric scores for ICA and DT-CWT suffer a sharp decrease as noise is gradually introduced.

3.6.3 Experiment 3: Multifocal Computer

The third experiment shows another multifocal scenario and affirms the superior performance of CP in noisy conditions. Two input images are given, each consisting of a clock in the foreground and a person on the computer in the background. As before, the first input is foreground-focused and the second input is background-focused. Results for non-noisy conditions in Figure 3.15



(a) Petrovic



(b) Piella

Figure 3.14: Comparison of fusion methods for multifocal clock images in noisy conditions using Petrovic and Piella metrics

suggest the plausibility of CP fusion despite its weak performance in comparison to ICA and DT-CWT.

(a) Computer 1 - foreground focused

(b) Computer 2 - background focused

(c) ICA

(d) DT-CWT

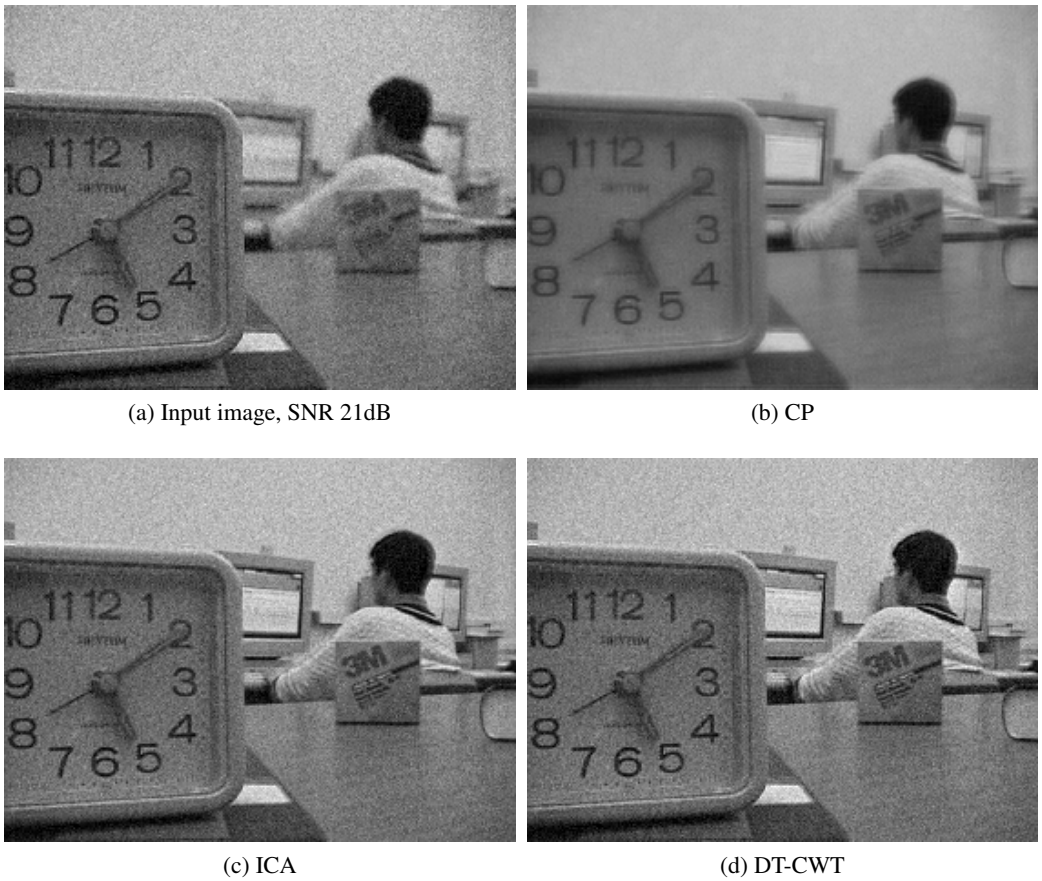
(e) CP

Figure 3.15: Multifocus computer example

The input image Computer 1 is then corrupted with various levels of noise from SNR 21dB to 4dB, which represents extremely high degradation. The fusion results of ICA, DT-CWT and Chebyshev

are displayed in Figures 3.16 to 3.21. These results are subsequently tested via the Petrovic and Piella fusion metric to objectively assess the performance of respective fusion algorithms. The metric scores are recorded in Table 3.2, which is then converted into the graph in Figure 3.22.

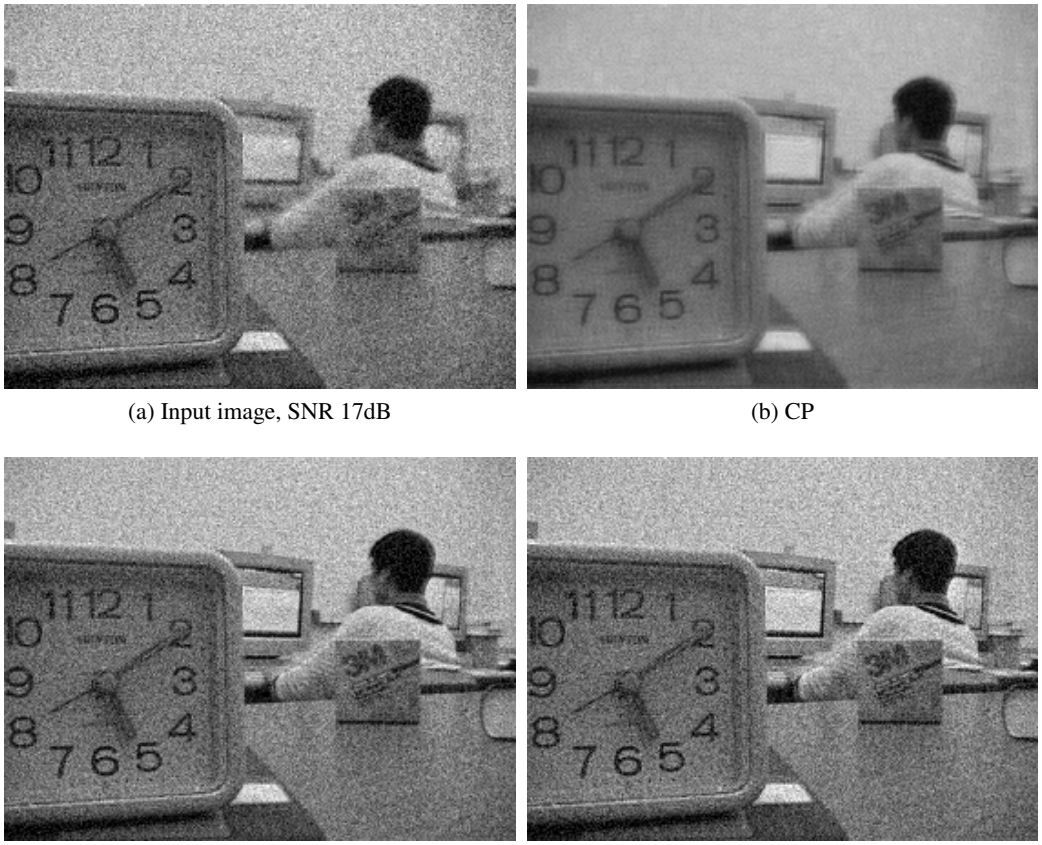
CP's result looks visually inferior to ICA and DT-CWT at 21dB SNR as seen in Figure 3.16, where the noise level is minimal. This is reflected in the metric scores, in which the Petrovic score for CP is lower than Piella. This may be explained by the absence of sharp gradients in a majority of the image area. However as with the clock example, we find that as the noise increases ICA and DT-CWT tend to suffer a significant decline particularly near the 16dB SNR mark. By 4dB, both Petrovic and Piella recorded very low scores for both ICA and DT-CWT whereas CP was able to maintain a moderate score. From the images in Figure 3.21, CP clearly prevails in its robustness to high density noise.



(a) Input image, SNR 21dB (b) CP

(c) ICA (d) DT-CWT

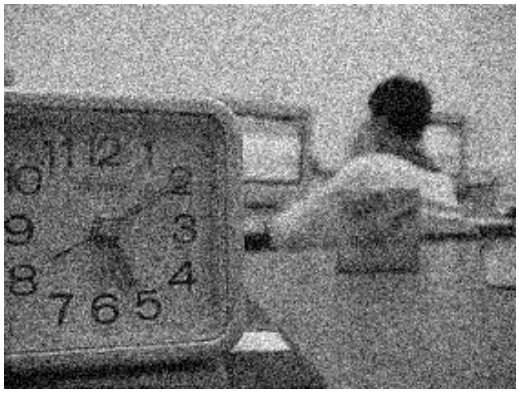
Figure 3.16: Fusion performance on computer images with Gaussian noise, SNR = 21dB



(a) Input image, SNR 17dB (b) CP

(c) ICA (d) DT-CWT

Figure 3.17: Fusion performance on computer images with Gaussian noise, SNR = 17dB



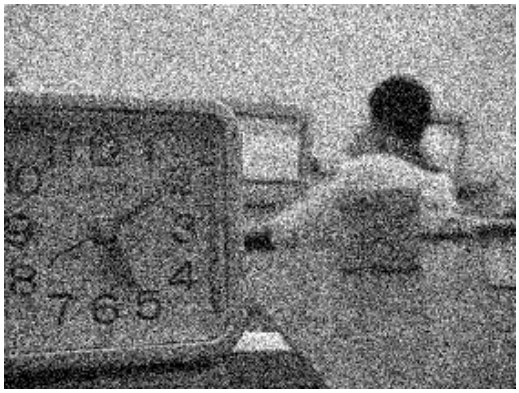
(a) Input image, SNR 14dB

(b) CP

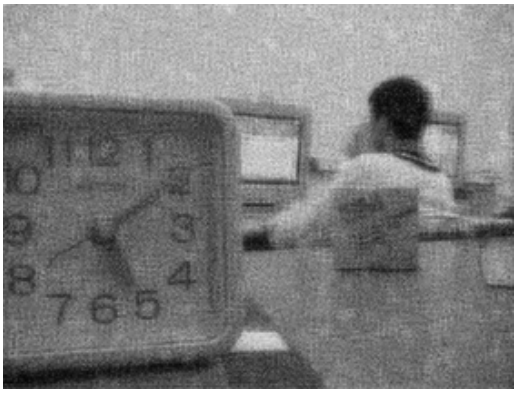
(c) ICA

(d) DT-CWT

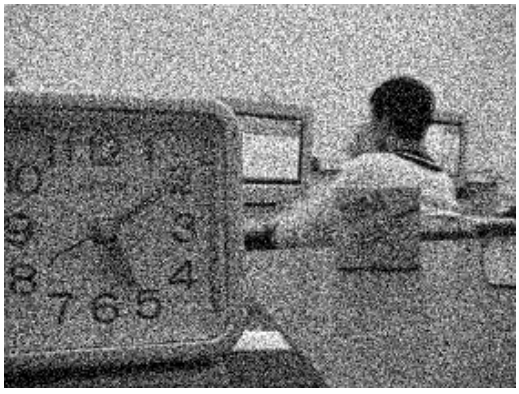
Figure 3.18: Fusion performance on computer images with Gaussian noise, SNR = 14dB



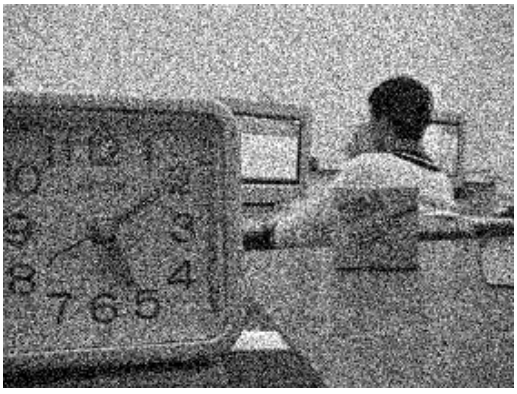
(a) Input image, SNR 11dB



(b) CP

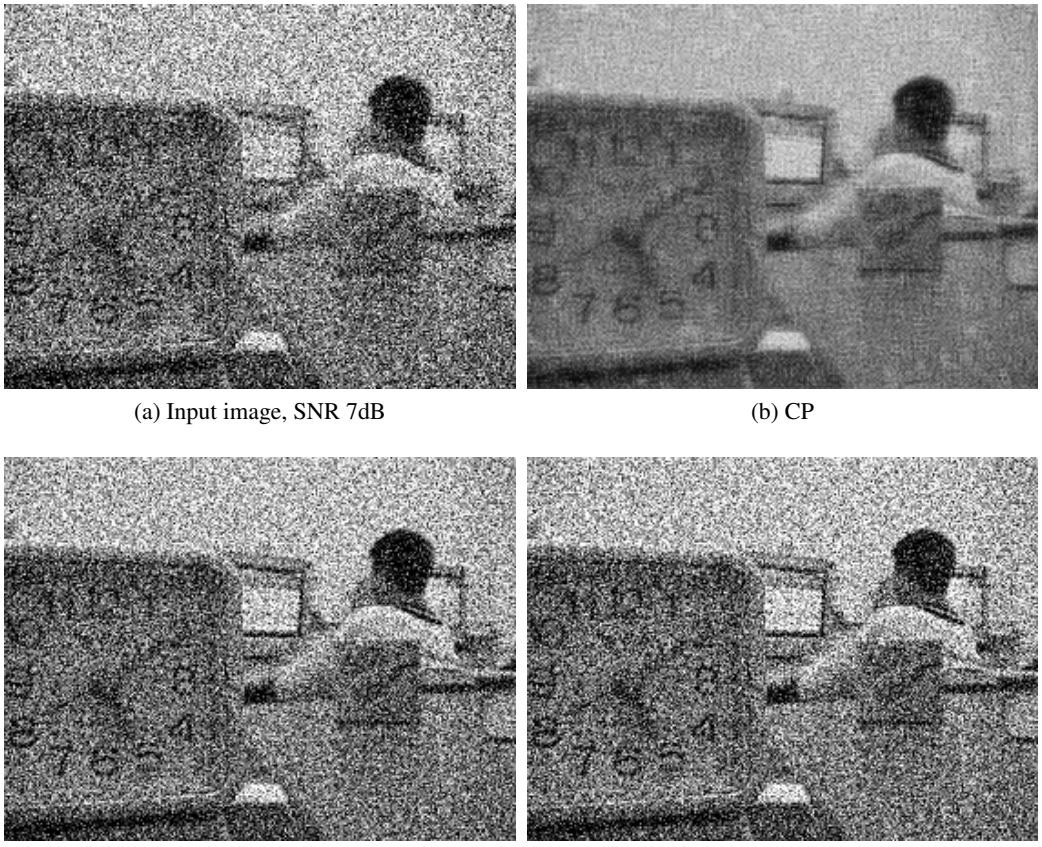


(c) ICA



(d) DT-CWT

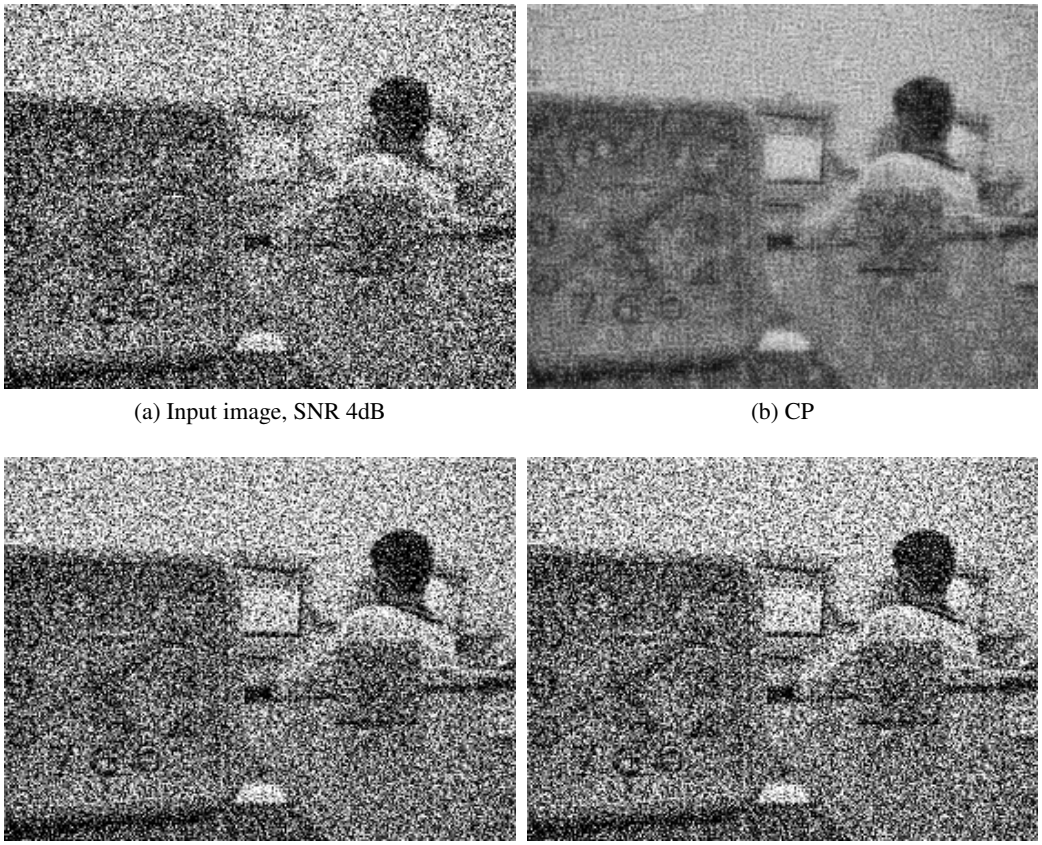
Figure 3.19: Fusion performance on computer images with Gaussian noise, SNR = 11dB



(a) Input image, SNR 7dB (b) CP

(c) ICA (d) DT-CWT

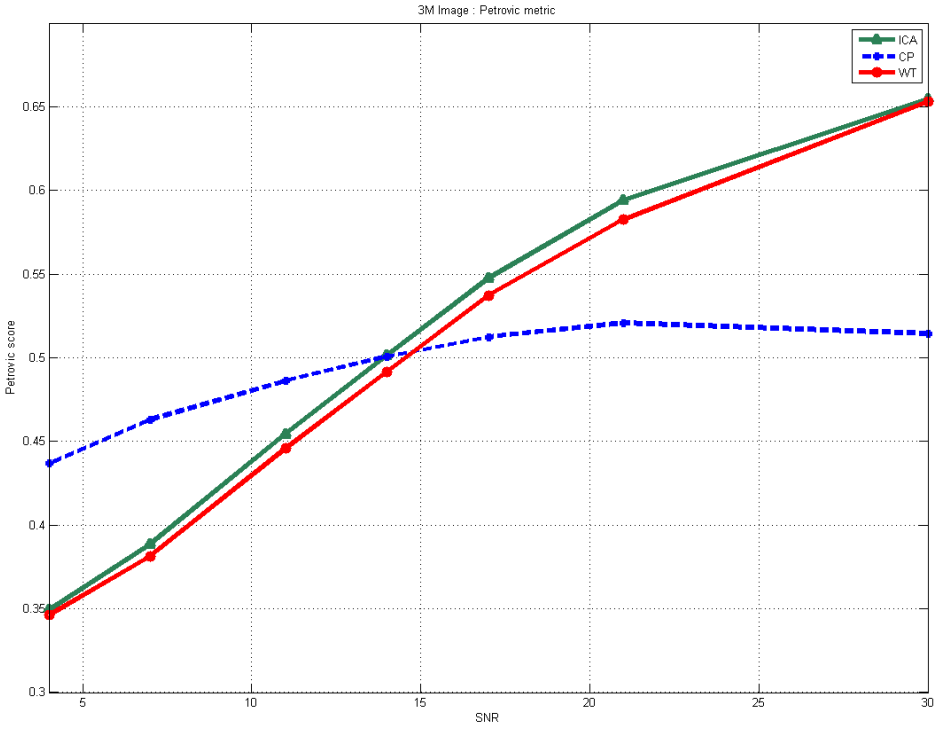
Figure 3.20: Fusion performance on computer images with Gaussian noise, SNR = 7dB



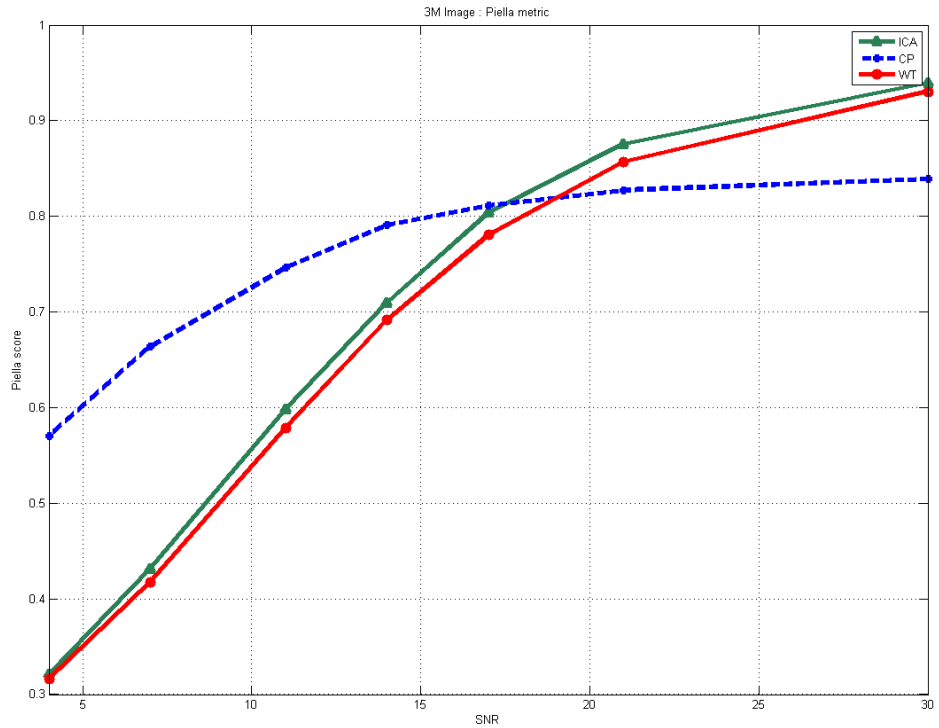
(a) Input image, SNR 4dB (b) CP

(c) ICA (d) DT-CWT

Figure 3.21: Fusion performance on computer images with Gaussian noise, SNR = 4dB



(a) Petrovic



(b) Piella

Figure 3.22: Comparison of fusion methods for multifocal computer images in noisy conditions using Petrovic and Piella metrics

Table 3.2: Fusion metric evaluation for computer images – ICA, CP and DT-CWT

Input	Method	Petrovic	Piella
Multifocal computer	CP	0.5145	0.8388
	ICA	**0.6545**	**0.9394**
	DT-CWT	0.6529	0.9304
SNR 21dB	CP	0.5207	0.8274
	ICA	**0.5943**	**0.8757**
	DT-CWT	0.5828	0.8568
SNR 17dB	CP	0.5124	**0.8110**
	ICA	**0.5475**	0.8037
	DT-CWT	0.5373	0.7806
SNR 14dB	CP	0.5007	**0.7907**
	ICA	**0.5014**	0.7096
	DT-CWT	0.4917	0.6917
SNR 11dB	CP	**0.4862**	**0.7462**
	ICA	0.4545	0.5981
	DT-CWT	0.4459	0.5786
SNR 7dB	CP	**0.4631**	**0.6640**
	ICA	0.3886	0.4316
	DT-CWT	0.3815	0.4174
SNR 4dB	CP	**0.4368**	**0.5707**
	ICA	0.3493	0.3206
	DT-CWT	0.3462	0.3163

3.6.4 Discussion

The efficient denoising capability for CP is a result of the trade-off between low-pass filtering and the degree of CP approximation. A low n-th order Chebyshev polynomials is adequate to estimate an image, hence lower computational complexity and costs. However CP degree is also representative of the accuracy of signal approximation, meaning that using lower degrees will produce a rather basic or smooth approximation. More complex approximations that consist of high frequency polynomial components will require higher orders. Therefore approximating an image with a low CP order essentially means that salient features above certain cutoff frequencies will be filtered out. Whilst this may result in a poorer output, as in the multifocal examples in Figure 3.7 and Figure 3.15, it is also the reason the Chebyshev method is able to successfully suppress the majority of noise components in the subsequent examples. This approach slightly resembles the way ICA performs shrinkage of sparse coefficients for its denoising algorithm – the

approach of CP can thus be considered an autonomous version of sparse code shrinkage, where only a few basis functions are retained so as to suppress the noisy elements. The main difference remains that for CP, only coefficients of a higher complexity/frequency are absolved, hence it results in a lower accuracy approximation.

A strong advantage of CP fusion is its independence from *a priori* noise variance. However it has also been accepted that in certain cases of fusion application, knowledge regarding noise estimates is readily available. A possible extension to this work therefore would be to exploit such information in order to facilitate a more adaptive version of CP approximation. Such information would especially be useful for the estimation of Chebyshev orders, such that the trade-off between noise suppression and computational complexity can be optimised.

We also noted the comparative processing complexity between the three algorithms. ICA requires pre-processing and training of basis components which take up time, while DT-CWT consists of extensive forward and inverse transform steps. CP on the other hand takes some time to process the 7×7 patches individually, but the universal polynomials need only be generated once. CP generally comprises fewer steps and thus requires low computational complexity.

3.7 Chapter Summary

This chapter has introduced us to a novel method for image fusion where Chebyshev polynomials are used as basis components. The theory and algorithms behind Chebyshev polynomials, its properties and advantages over other methods were discussed. Performance tests revealed that while incorporating fewer steps, CP-based fusion has displayed encouraging results in noise corrupted images, due to its intrinsic smoothing property. The use of bivariate Chebyshev polynomials for fusion may also influence other image processing applications – namely denoising, coding and features segmentation.

The limitations attributed to the experiments are acknowledged. Grayscale images were employed on all three methods; though RGB images should produce similar results albeit it being more computationally expensive, due to more coefficients required for CP; and may be considered in the future. Gaussian noise used in the corrupted images presented the best result for CP, though

Poisson noise produced results of near-equal quality. The experiments with Salt & Pepper and speckled noises did not find improved outcomes with any of the methods, and hence they were omitted.

The works on ICA and DT-CWT have long been established in image processing literature and both have extensively addressed the problem of image denoising through approaches like sparse code representation. Our work on Chebyshev polynomials does not undermine that fact; rather, the focus of our advantage is on self-regulated and global denoising of Chebyshev polynomials. This is a highly beneficial property of our approach that other methods do not have. As CP automatically filters high frequency image components, the approximation of images by CP efficiently removes noise within the high frequency spectrum.

Chapter 4

Region-based Chebyshev-ICA Fusion Method

The previous chapter has introduced us to a first application that employs Chebyshev polynomials for image fusion, which performs notably well in heavily corrupted images. This is achieved by truncated approximation using finite orders, which sacrifices representation of salient features and edges for noise removal. To overcome this trade-off, a novel combinatory fusion approach is proposed. CP is utilised alongside ICA for a new fusion method that shall be discussed in this chapter.

4.1 Limitations of CP Fusion

CP is a type of low-pass filter that removes noise in heavily corrupted images at a cost of lower signal accuracy [34]. This particular limitation has been highlighted in the previous chapter. Though its performance has been shown to be significantly robust in noise conditions as high as 4dB SNR, perceptual evaluation suggests that important edge details have been smoothed out from the fusion result. On the other hand, independent component analysis (ICA) has proven to be an excellent image fusion method in non-noisy conditions, particularly due to its ability to preserve edges and gradient saliency within images.

This gives rise to a new type of fusion technique, which can be viewed as an algorithm fusion

process that manages to combine the best aspects of both ICA and CP. In this chapter an image fusion approach is proposed mainly for the denoising of digital images corrupted with noise. We provide a new technique for image fusion that complements the best aspects of both ICA and CP, which we call the Region-based Chebyshev-ICA (RBCI) combinatory fusion method [128].

4.1.1 A Proposed Region-based Approach

The main aspect of RBCI is detecting edges, boundaries and texture. This is achieved by dividing the pixels of a given image into active and inactive regions, in other words distinguishing edges and rich texture as opposed to uniform background or weak texture. We employ the Canny edge detection method for this classification task. The primary purpose of classification is to create a boundary among regions which correspond to different aspects of an image. Fusion is then applied individually on each bounded region.

There are two main attractions of this technique: pixels can be processed more efficiently if they are treated as a collective group within a region, rather than separate entities. Region-based fusion may therefore help to overcome some drawbacks of pixel-based fusion, like blurring, susceptibility to noise and misregistration [91]. The second advantage is that this approach enables us to perform image fusion using a mixture of CP and ICA, hence getting the best of both approaches.

The method takes inspiration from other successful algorithm fusion approaches such as in remote sensing, where IHS results (that possesses good spatial qualities) is merged with wavelet transform fusion output (good spectral qualities). Essentially, RBCI is regarded as a second-level fusion that combines the outcome of incumbent fusion methods in the post processing stage. This is illustrated in Figure 4.1.

4.2 Introduction to Independent Component Analysis

Independent Components Analysis is a popular method for distinguishing underlying components in a random data set [129]. The rationale behind ICA may be explained by the central limit theorem, whereby the linear mixture of two non-Gaussian and independent source signals results in a signal that is more Gaussian-like than any of its sources. ICA searches through the given

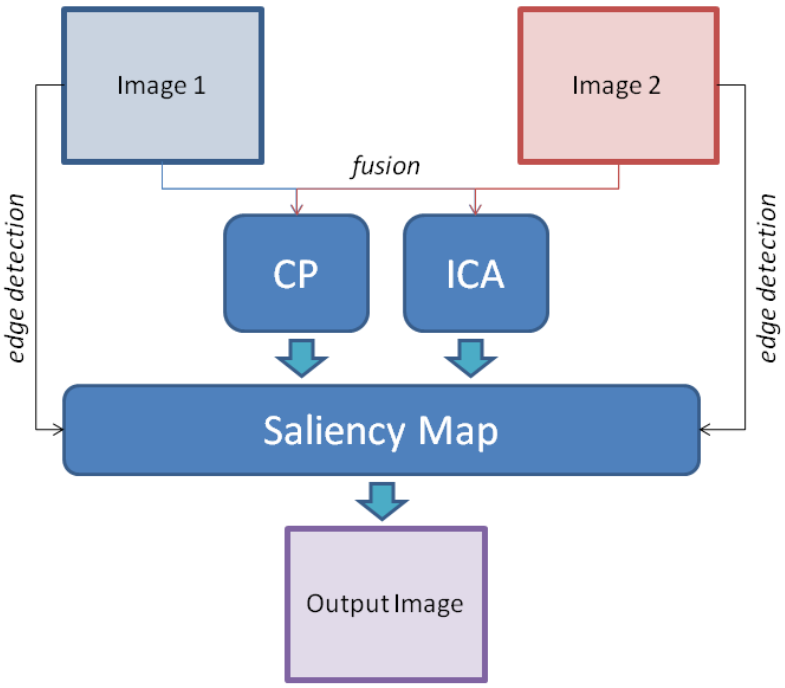


Figure 4.1: Framework of region-based combinatory CP-ICA fusion

mixed signal to precisely identify independent components – which are the source signals. Given a number of observed mixed signals x, ICA attempts to identify the source signals s, by means of deriving a mixing matrix A [118].

$$x = As \tag{4.1}$$

We solve for A by optimising the objective function of the distribution pertaining s, either by maximising its non-Gaussianity or minimising the mutual information between the independent components. For this purpose, kurtosis or the fourth-order cumulant can be used as a suitable measure of non-Gaussianity [118, 130] due to its theoretical and computational simplicity [47]. The kurtosis of a function x is defined as

$$\text{kurt}(x) = E(x^4) - 3(E(x^2))^2 \tag{4.2}$$

If x is zero-mean and has a unit variance then $\text{kurt}(x) = E(x^4) - 3$. Furthermore for a Gaussian x, the fourth moment can be re-written as $3(E(x^2))^2$. Thus kurtosis is zero for a Gaussian random variable, whilst non-zero for a non-Gaussian random variable [47]. By maximising the kurtosis through use of an optimisation/cost function, we aim to draw out the independent (source) components contained in the original signal.

A study by Hoyer and Hyvarinen [131] noted that performing ICA on a given image patch will yield a set of very localised, independent components that tend to resemble the human's visual cortex in analysing scenery and detecting edges. They are also very similar to wavelet bases, and are therefore suitable for image analysis applications [127]. The aim is thus to estimate a finite number of bases that are able to represent most of the image patch's structure.

The new domain is described by a set of 2D basis functions or basis images. The image of interest is then represented as a linear combination of these bases. Bases are trained from a selection of varying image patches, belonging to training images that are similar in content and type to the complex set of images of interest. Estimation and optimisation methods are then employed to solve the mixing matrix and subsequently generate the source signals via the equation

$$y = Wx \tag{4.3}$$

where W is the inverse of A, and y is the derived signal aimed to approximate the source s.

Using the kurtosis method in practice, we begin by defining an arbitrary value for W and calculate the direction in which the kurtosis of $y = Wx$ is increasing the most, based on the available samples of the observed vector x. A gradient method is then used to find a new vector W. The process is reiterated until the minimax criterion is achieved and the source signal's kurtosis is found. Theoretically, kurtosis can thus be used as an optimisation tool to address the ICA problem [47]. For a more detailed description of ICA, [47, 118, 132] are excellent references.

The denoising problem is addressed by ICA through its close relation with the concept of sparse code shrinkage (SCS) [47, 133]. SCS works by removing less significant components in between the independent bases by thresholding a value that corresponds to the noise variance. Each basis component is independent and corresponds to a coefficient value attributed to the mixing matrix. SCS removes those bases with low coefficients, leaving only a 'sparse' representation of basis components in which only very few strong coefficients remain. This method has proven quite effective but still suffers from the limitation of requiring the exact knowledge of noise variance *a priori*.

4.2.1 Comparison of ICA and CP

Both approaches belong to a family of signal decomposition techniques, but with several meaningful differences which we will try to highlight. The data-driven ICA method utilises its independent bases as basis functions, which are derived from the training data and can fit arbitrarily onto image patches. On the other hand, CP uses standard orthogonal polynomials as its basis functions. The mixing matrix W of ICA is subsequently developed analytically by a cost function, whereas CP coefficients are derived by mathematical equations that benefit from the orthogonality relation. Both approaches make use of a linear combination of bases for bivariate signal approximation. However, a notable difference is ICA's bases are all independent and of equal complexity, whereby in SCS stronger bases are sparsely retained. For CP, higher degree polynomials denote

more complex signal components. Therefore the accuracy of image approximation relies heavily on the appropriate polynomial representation. It is due to this then that ICA manages to retain most salient edge features, resulting in an excellent approximation in normal conditions. However it also requires a noise variance estimation step for sparse code shrinkage for denoising. As for CP, its smoothness property ensures noise suppression at a cost of lower signal accuracy. This is alleviated by an autonomous high frequency removal for denoising in signal approximations. In a way, CP utilises an autonomous and indiscriminate shrinkage of sparse coefficients for denoising as opposed to ICA.

4.3 A Novel Region-based CP-ICA Method for Image Fusion

4.3.1 Fusion Rule

The combination of important features from input images is based on a possible set of fusion rules. These include max-abs, mean and weighted combination (WC) [125], as has been described in Section 3.5.1. The development of fusion rules is an ongoing field and is specifically a research topic on its own. In practice, the fusion of images is done either in the spatial or transform domain, based on two approaches – pixel-based and region-based fusion. For pixel-based fusion, only single or arbitrary groups of pixels are considered under the fusion rule. This poses a significant drawback as it ignores the image structure, scenery outlines and relationship between neighbouring pixels, thereby inadequately representing the overall fusion result. In contrast, region-based approaches efficiently adapt the fusion process around object intricacies and make use of the image structure. It is therefore apt to consider region-based fusion as the framework of fusion approach in this chapter.

4.3.2 Region-based Fusion

In region-based schemes, pixels are segmented into larger regions. Each region is formulated to be based on such criteria as objects of interest or degree of activity. We are then able to determine the actual contribution of regions from each input. There have been numerous research involving region-based fusion as can be found in [91, 125, 134].

We define an image or scene as being composed of active and inactive regions. A region is considered to be active if and only if there happens to be sufficient amount of 'interesting' information in it, such as edges or strong texture. On the other hand, inactive regions imply a lack of interesting information or a plain area with weak textural details.

In images corrupted with noise, the presence of noise is prone to spread through the entire image, triggering a notable increase in the overall activity. Our premise here is that noise is more distinguishable, and therefore more easily suppressed, in inactive regions than active ones. Noise components near an edge region, for example, would be mixed with the textural details associated with that edge. It would therefore be difficult to differentiate and subsequently remove noise without affecting the regional texture. It is due to this that blurring tends to occur in some image denoising applications such as Chebyshev polynomials.

Of the previous fusion methods, ICA is extremely adept at modelling edges and textured regions by virtue of its ability to capture essential image features comprehensively. However despite its denoising capability by use of sparse code features [130], the performance of ICA has been shown as being prone to decline especially in very noisy examples. This has been highlighted in [34] and also in tests in the previous chapter. In contrast, the smoothness property of CP makes them better suited to filter noisy signals and as such, would be of convenient use on inactive regions mentioned above.

The ideal solution in dealing with noisy scenarios would therefore be to use ICA on active regions which contain important structural image details, whilst inactive or background regions are processed by CP. We will attempt to implement this idea in the chapter. In a way, we are performing a second-level fusion on the results of the fusion methods that were obtained. The method is flexible enough to replace ICA and CP with a variety of other image fusion methods; though for the purpose of denoising, ICA and CP are appropriate choices.

4.3.3 Determining Active and Inactive Regions

Two input images from the UN Camp sequence database in Figure 2.1 are set as examples of the region defining process. Both images are processed by Canny edge detection [9], which is mostly robust to noise and is essentially a multistage edge detection algorithm that locates critical edge

pixels at various scales. The filtered product, named the edge map, is meant to contain salient information of the image. The respective edge maps for the inputs are then combined via the logical OR operator to form a single, global edge map. This step aims not to accurately determine correct edges or regions from the source images, but rather to ensure all possible edges and regions have been preserved for further processing [127]. This is similar to a scheme proposed by Drajic in [15], where active regions are defined by an activity indicator, $E(t)$, of an $N \times N$ image patch. A region is considered active if $E(t) > \frac{2}{T} \sum_{k=1}^{T} E_k(t)$.

The resulting map is shown in Figure 4.2. The final edge map provides a guide of the areas to be processed by either ICA or CP. Active regions, denoting all hard edges and texture, are represented by white pixels and will be fused by ICA whereas inactive regions i.e. those in black are fused by CP. However this is an overly simple model – blocking artefacts and pixel discontinuities tend to appear in the result which require refining. The enhanced process comprises a morphological opening step, as explained below.

4.3.4 Mathematical Morphology for Edge Map Enhancement

To preserve edge details, an intermediary step comprising an opening morphological operation is added to the edge map that enables us to incorporate more pixels from ICA. Mathematical morphology can be defined as a series of non-linear processes which enhance, alter or simply remove details from a region of interest [135]. The fundamental morphological operations are dilation (enhance) and erosion (remove). Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the so-called structuring element used to process the image. The erosion and dilation of a binary image A in an Euclidean space E, by a structuring element B may be expressed as follows:-

$$\text{Erosion:} A \ominus B = z \in E | B_z \subseteq A \tag{4.4}$$

$$\text{Dilation:} A \oplus B = \bigcup_{b \in B} A_b \tag{4.5}$$

where B_z is the translation of B by a vector z. A simpler way to express morphology is as given in [135]:-

$$\text{Erosion:} I_{\text{new}}(i,j) = min(P) \tag{4.6}$$

$$\text{Dilation:} I_{\text{new}}(i,j) = max(P) \tag{4.7}$$

where P refers to all pixels covered by the structuring element centred at $I(i,j)$.

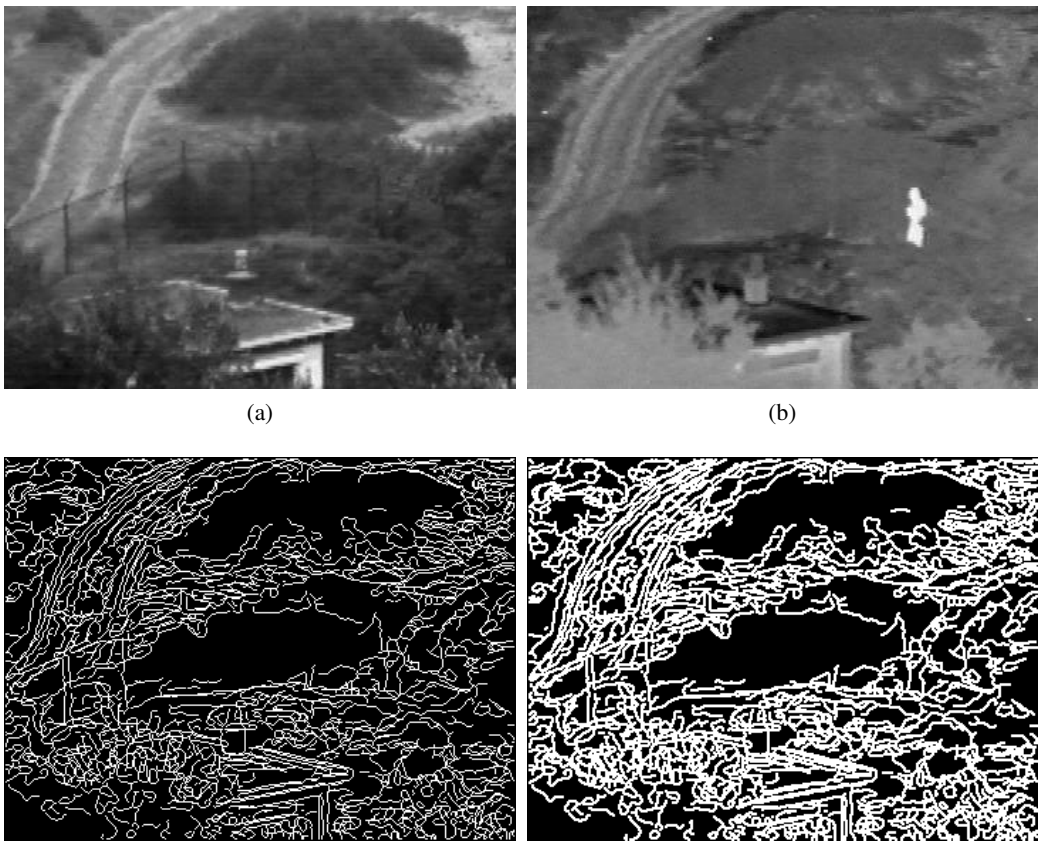
Opening is a combination of such operations that consists of dilating an image after eroding it. The appeal of this step is that it discards small pixel artefacts and discontinuities, whilst smoothing larger pixel structures. In reference to the edge map, the opening step will suppress any single active pixel or small active region while at the same time increase the breadth of larger active regions. This is clearly evident in Figure 4.2d where the edges have been thickened, as compared to the original edge map in Figure 4.2c.

The map also illustrates inactive areas which form the majority of the image space. This manouver allows us to maximise the use of CP for autonomous noise suppression on the majority of the scenery, and at the same time enables ICA to process edge lines and hard textural regions which represent important features in an image. In general the additional step provides a clearer and more balanced usage ratio between pixels belonging to ICA and CP.

In practice, consider two output images of ICA and CP, F_{ICA} and F_{CP} respectively. The edge map is a binary map with binary values, $S \in \{0,1\}$ with white being 1 and black being 0. The expression used to derive RBCI, F_{RBCI} can be written as a Hadamard product,

$$F_{\text{RBCI}} = (F_{\text{ICA}} \circ S) + (F_{\text{CP}} \circ \bar{S}) \tag{4.8}$$

$$F_{\text{RBCI}} = \sum_{x=1}^{X} \sum_{y=1}^{Y} \left[F_{\text{ICA}}(x,y) \circ S(x,y) + F_{\text{CP}}(x,y) \circ \bar{S}(x,y) \right] \tag{4.9}$$



(a) (b)

(c) (d)

Figure 4.2: Edge map derived from selecting all hard and soft edges from both input images (a) UN visual camera, (b) Infrared, (c) Edge map, (d) Enhanced edge map

where $x = 1, 2, ..., X$ and $y = 1, 2, ..., Y$ are the pixel coordinates and $(X \cdot Y)$ is the image size.

The concept brought forth in this method is similar to the way humans generally perceive objects of interest within our line of vision. Uniform areas in the background tend to be plain and carry little or no information of high frequency. The general appeal of our method is that it can enhance strong texture and edges, thus giving priority to meaningful regions whilst de-prioritising the less important regions in the background.

4.4 Performance Evaluation of RBCI

We tested our algorithm using corrupted images, to reflect real world conditions where the transmission of data is prone to noise. Incremental noise was added to a set of input images, ranging

from 16dB to 10dB SNR to represent the various degrees of image corruption. Scenarios were used from the multifocal Clock and multimodal/multispectral UN Camp image datasets respectively. For each experiment, we obtained two grayscale images as source signals, from which the fusion will generate a composite output image.

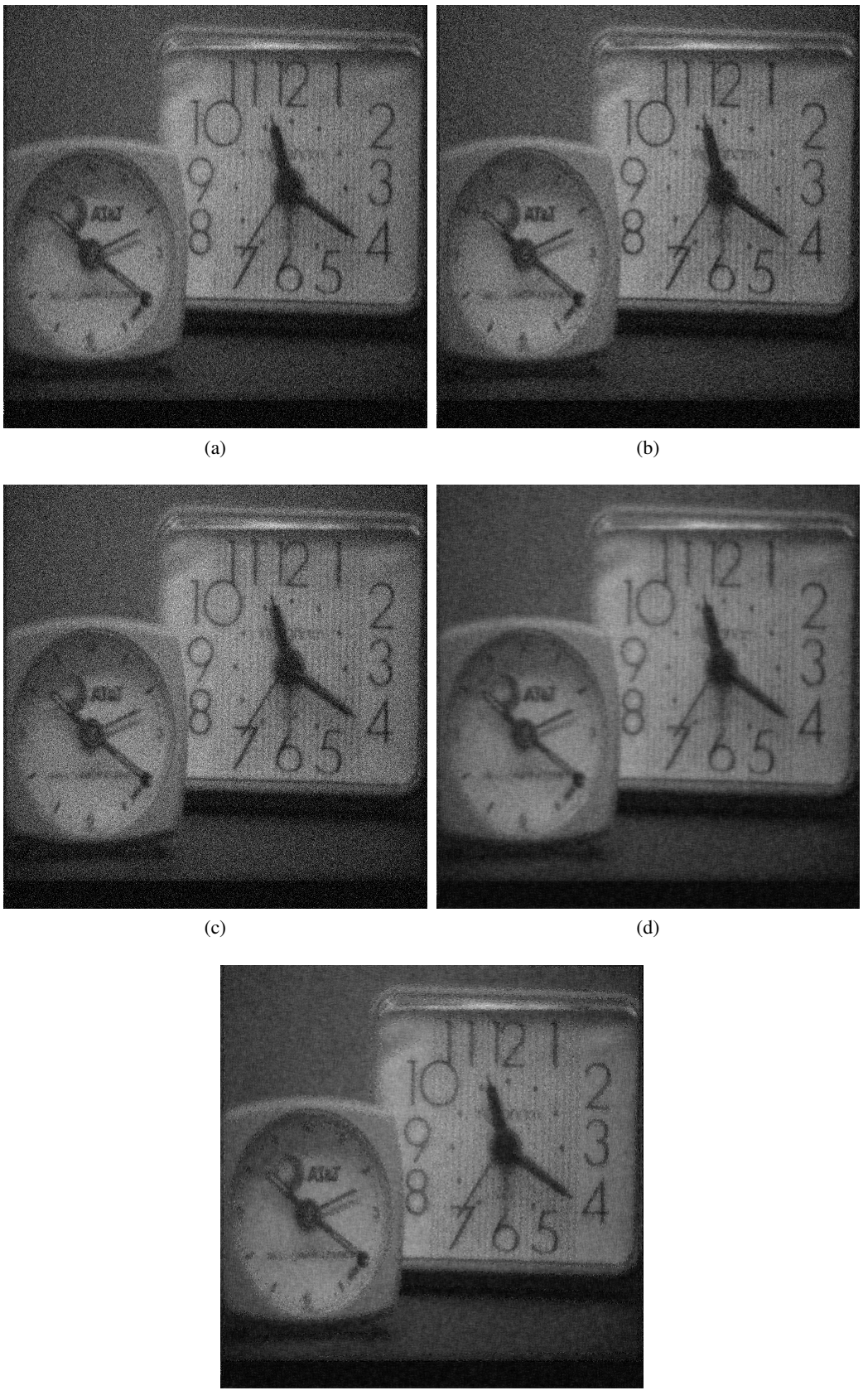
For benchmarking purposes, we tested the performance of RBCI against several common fusion methods namely ICA, CP and the wavelet transform (DT-CWT); and also against very basic fusion techniques namely max-abs, which is simply a pixel-by-pixel application of the maximum operator, and mean, or average operator.

Similar to the tests carried out in Chapter 3, ICA training was carried out on 1000 patches belonging to images of similar content, each patch measuring 7×7 pixels. A total of 48 independent bases were derived. Similarly for CP, $m = 11$ and $n = 11$ degree of Chebyshev polynomials were chosen and 7×7 overlapping patches were used. Preliminary configurations concur that the size of patches and polynomial order used were suitable for the experiments. Four levels of decomposition were applied on the DT-CWT. All fusion methods utilised the max-abs, mean and weighted average fusion rule [125].

It is noted that the RBCI image is a mesh of the regions from the fused results of both ICA and CP, based on the edge map. The method aims to combine the best aspects from both techniques and form a singular output image. The definitive fusion step would still require separate applications of ICA and CP respectively; and their results are subsequently reconstructed under the proposed framework. It is due to this factor that RBCI is computationally more complex than the previous methods. Separate steps are required to carry out individual fusion processes of ICA and CP, whilst the region classification and reconstruction procedures involve further computation. However, we believe that the favourable performance of RBCI in contrast to its predecessors outweighs the aforementioned limitation.

4.4.1 Experiment 1: Multifocal Images

The experiment, as shown in Figure 4.3 deals with multifocal examples, whereby two images of the same scene but with different focal lengths are tested. The aim is to produce a fused image where all its objects are in focus.



(a) (b)

(c) (d)

(e)

Figure 4.3: Multifocal Clock image fusion results for different methods for SNR = 10dB. a) ICA, b) DT-CWT, c) Mean, d) CP, e) RBCI

We tested the multimodal clock image under various noisy conditions i.e. from medium (SNR = 16dB) to high (SNR = 10dB) levels of noise. A range of additive white Gaussian noise (AWGN) were added to both source images to mimic the signal degradation phenomena in real applications. This is similar to the tests performed in Chapter 3. The inferences are thus also alike. In normal circumstances, ICA and DT-CWT tend to produce a superior quality output due to their intrinsic ability to distinguish independent image features. However as the noise level increased we noticed a steady decline in image quality for these methods.

Conversely in non-noisy environments, CP's approximation results in a smooth fusion with less discernible edges. CP filters out other high frequency components, creating a blurred artefact on the image. Yet through the incremental inclusion of noise elements the image quality only degrades slightly. This has been explained by the universal property within CP that automatically removes additive noise and other components of high frequency. The limitation of CP however is that it does not yet solve the problem of blurred edge details.

Therefore a combined output containing the best aspects of the above two methods is desired. An output image should clearly detect and differentiate edges and important texture, whilst at the same time adequately suppressing the noise. The RBCI method is able to accomplish this by combining edge and textural regions from ICA and background regions from CP [128]. The result is an enhanced output with distinctively better subjective visual quality and higher scores in objective evaluation.

Objective Fusion Evaluation

Three fusion rules were used: max-abs, mean and weighted average. Employing the Petrovic metric on our set of images produces the following results in Table 4.2.

The graphs displaying the performance of each method under the different noise levels is given in Figures 4.4 to 4.6. The graphs for CP depict a very low gradient against the rising level of noise. The reason behind this has been discussed earlier in the thesis. In contrast ICA, DT-CWT and pixel-based methods tend to suffer drastically as the noise increases. The exception is RBCI, which manages the highest score for 12dB and higher. The promising results of RBCI is likely due to the combinatory factor of our technique. Strong edge details belonging to ICA in active

Table 4.1: Average Petrovic score for multifocal Clock example for noise levels 16dB to 10dB.

Fusion rule	Method	E(Q)
Max-abs	ICA	0.2893
	CP	0.3144
	DT-CWT	0.2780
	Max-abs	0.2957
	RBCI	**0.3174**
Mean	ICA	0.3218
	CP	0.3381
	DT-CWT	0.3223
	Mean	0.3227
	RBCI	**0.3413**
Weighted average	ICA	0.3293
	CP	0.3342
	DT-CWT	0.3354
	Mean	0.3227
	RBCI	**0.3453**

regions has been retained whilst inactive and plain regions have been fittingly smoothed.

Perceptual Fusion Evaluation

Comparisons by subjective perceptual assessment tend to verify the metric findings. The images at SNR = 10dB in Figure 4.3 are observed: ICA, DT-CWT and pixel-based method's images are beset with speckles of noise, most notably in plain and monotonous regions. Nevertheless the edges, for instance the numerical digits on both clocks, are still very distinguishable. In contrast, CP manages to capture the overall features of the image but generally looks blurry. This may present a problem when it is used in applications that are trying to distinguish small, important objects. As can be seen, RBCI manages to find a compromise between the two methods by combining ICA's sharp edges with CP's smooth background, thus forming a self adaptive fusion method that targets and enhances principal objects of an image, even in noisy conditions.

4.4.2 Experiment 2: Multispectral Images

In surveillance-based applications such as military surveillance, satellite imagery and remote sensing, a single image acquisition sensor is often inadequate to fully capture salient information

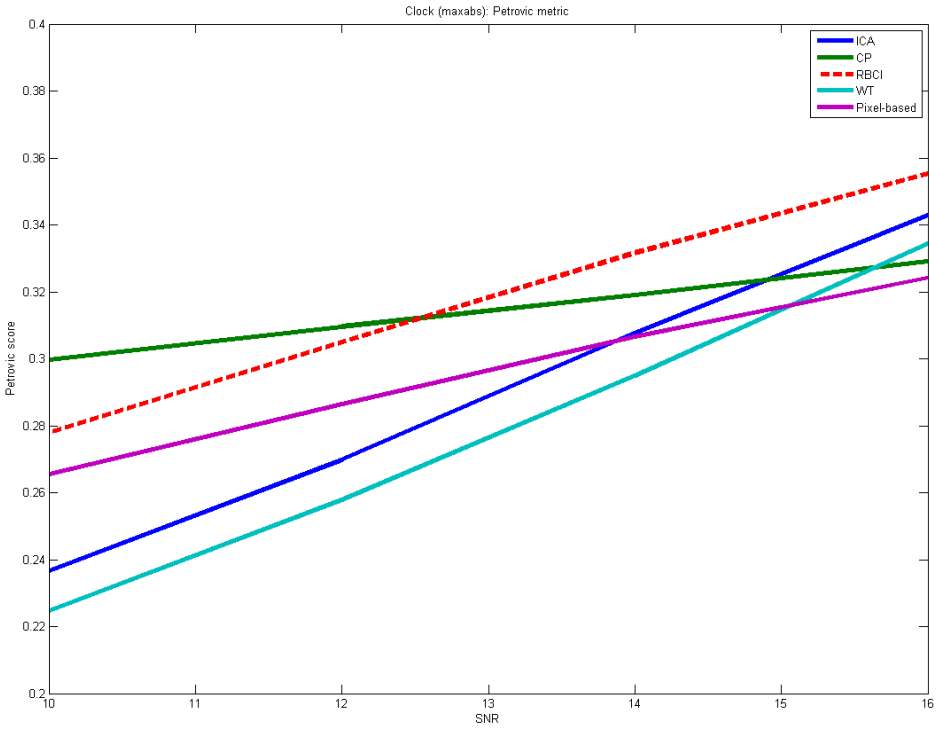


Figure 4.4: Multifocal Clock fusion performance for different fusion methods for SNR = 16dB to 10dB using the Max-abs fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon)

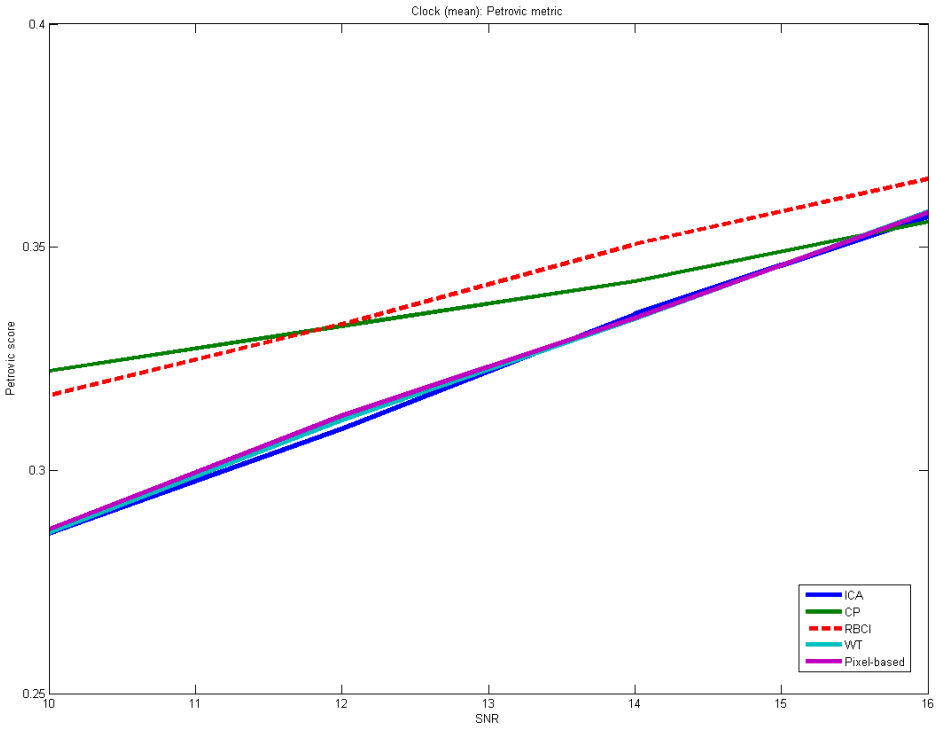


Figure 4.5: Multifocal Clock fusion performance for different fusion methods for SNR = 16dB to 10dB using the Mean fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon)

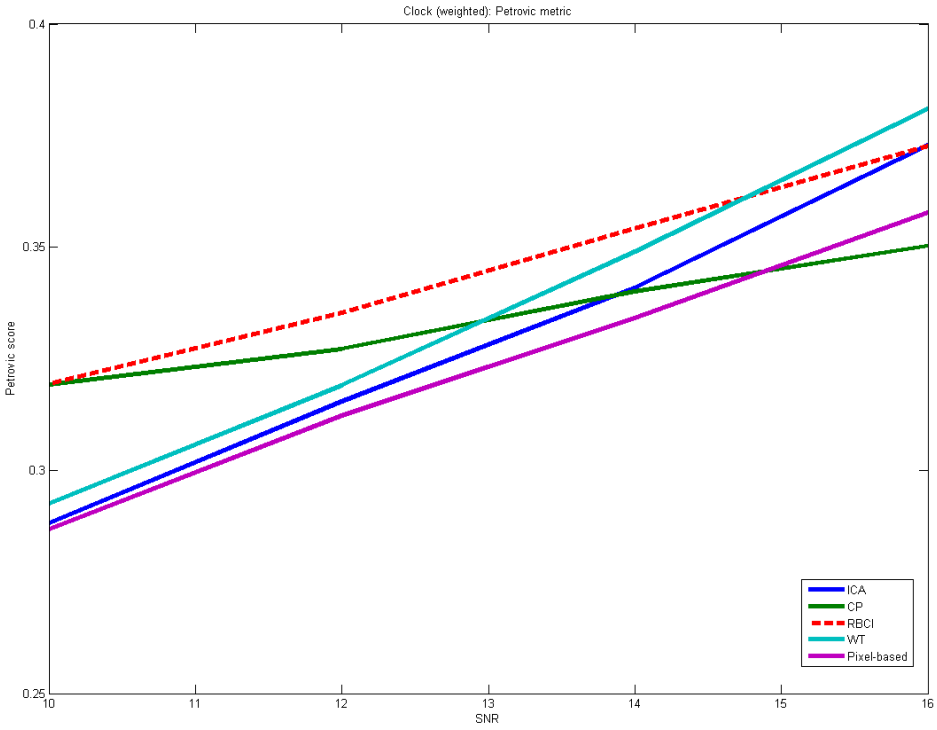


Figure 4.6: Multifocal Clock fusion performance for different fusion methods for SNR = 16dB to 10dB using the Weighted Average fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon)

within a landscape or scenery. The approach commonly used in practice is to acquire an image using different sensors comprised of varying spectra, and to perform fusion of the images in a post-processing step. In remote sensing, fusion between multispectral (MS) images and panchromatic (PAN) images is a notable example. The fusion process complements the properties of MS images, which exhibit high spectral density but low resolution, with PAN images which have very high resolution but lower spectral density [136].

We apply fusion on two input images from the UN Camp dataset in Figure 4.7. The visual and NIR image each possesses a wavelength of $\lambda = 420 - 700nm$ and $\lambda = 750 - 1400nm$, respectively. Similar to Experiment 1, a series of AWGN were incrementally added to the image, ranging from 16dB to 10dB. The objective of fusion is to generate a composite image with the best features from the visual and NIR camera. This ideally means the human presence is fully captured in the scene whilst foliage and other strong textural details are suitably retained.

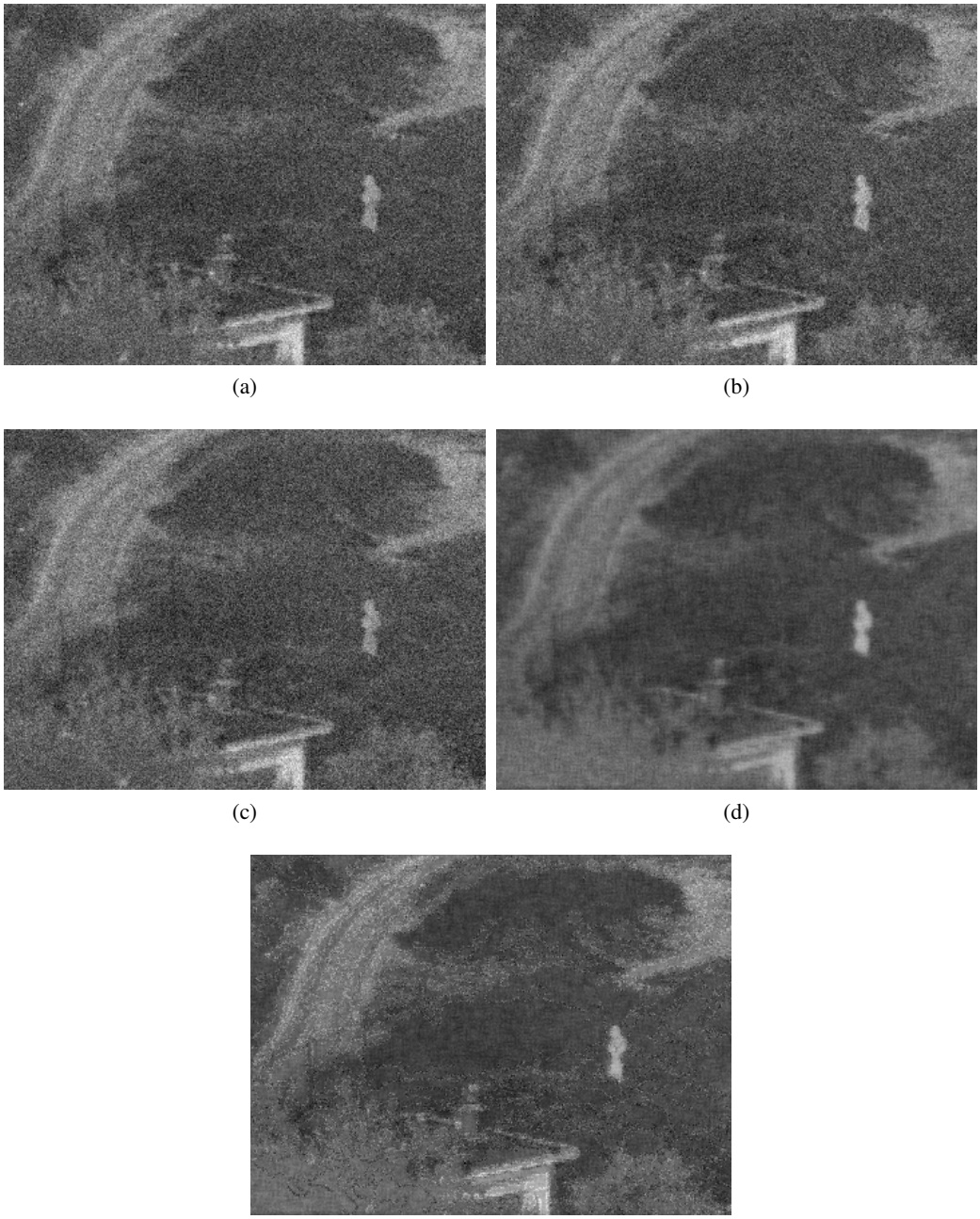
Objective Fusion Evaluation

The findings for multispectral images, measured using the Petrovic metric, are given in Table 4.3 and plotted in a graph in Figures 4.8, 4.9 and 4.10. The performance of ICA, DT-CWT and pixel-based method matches that of RBCI at 16dB but dips notably as the noise level increases. In contrast, CP's scores increases with added noise. Again however, RBCI achieves the best overall performance under mean and weighted average rules, whilst also scoring high under max-abs rule.

Perceptual Fusion Evaluation

Subjective evaluation is performed on UN Camp fused images at SNR level 10dB in Figure 4.7. The human outline has been best captured in DT-CWT, followed by mean and CP. Due to the high degree of noise however, images for ICA and DT-CWT have been inundated with speckles of black and white which spoils the scenery. Getting the correct result for such images is of paramount importance in military surveillance, and the consequence of error could have serious implications in critical decision making systems.

As usual CP features a smooth, noise-free landscape. There is almost no noise speckles evident in the image. However the blurriness remains as do the absence of foliage and terrain texture. RBCI



(a) (b)

(c) (d)

(e)

Figure 4.7: Multispectral UN Camp image fusion results for different methods for SNR = 10dB. a) ICA, b) DT-CWT, c) Mean, d) CP, e) RBCI

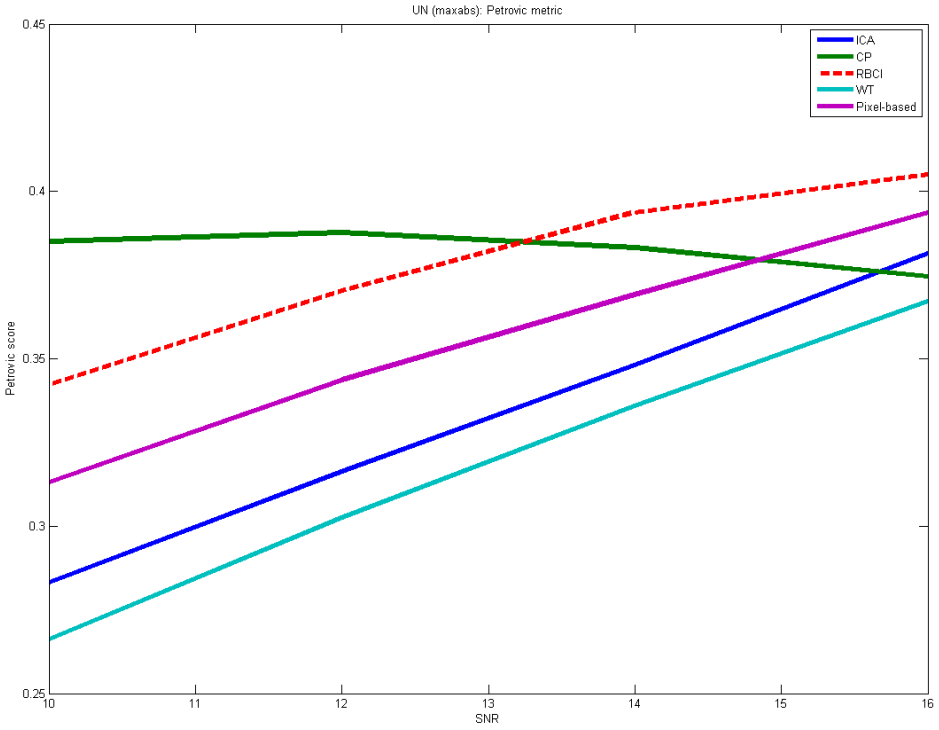


Figure 4.8: Multispectral UN Camp fusion performance for different fusion methods for SNR = 16dB to 10dB using the Max-abs fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon)

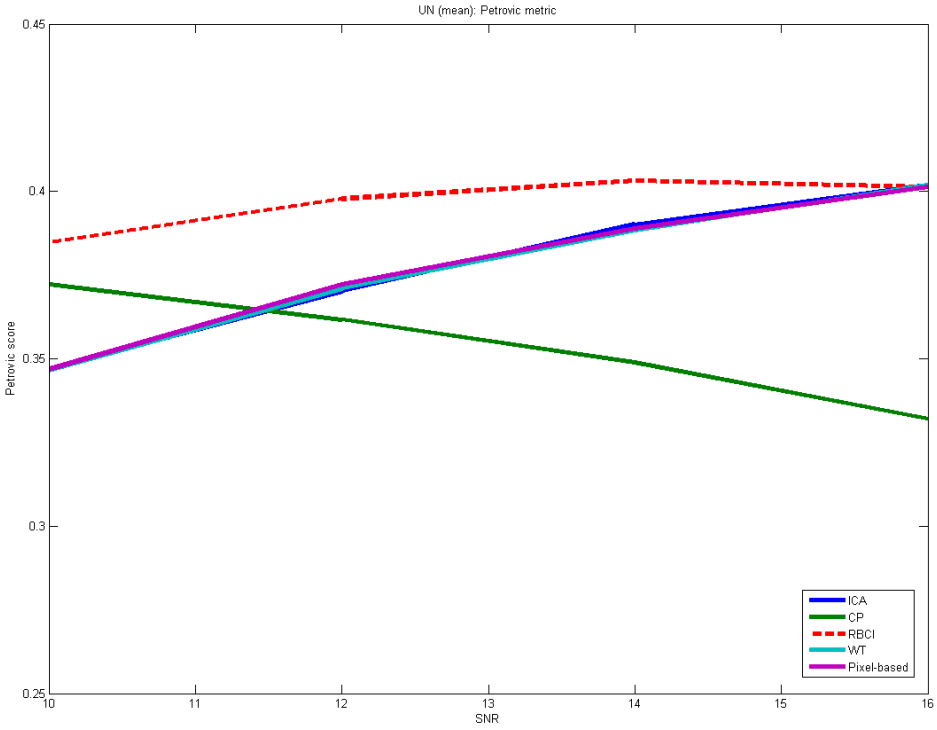


Figure 4.9: Multispectral UN Camp fusion performance for different fusion methods for SNR = 16dB to 10dB using the Mean fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon)

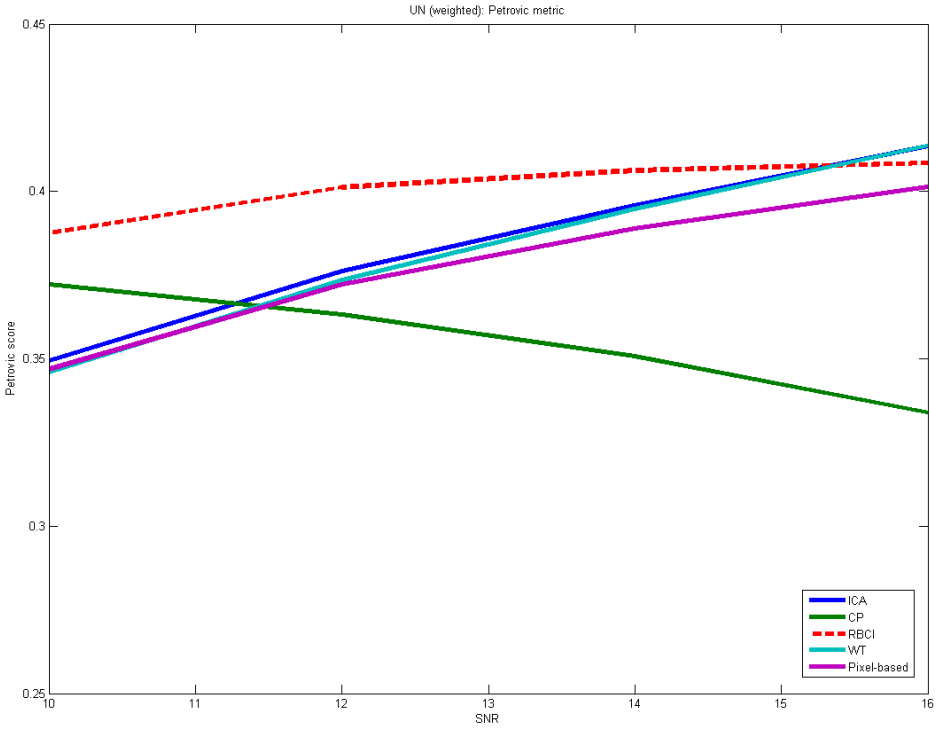


Figure 4.10: Multispectral UN Camp fusion performance for different fusion methods for SNR = 16dB to 10dB using the Weighted Average fusion rule. (ICA=blue, CP=green, RBCI=red, WT=aqua, Pixel-based=maroon)

Table 4.2: Average Petrovic score for multispectral UN Camp example for noise levels 16dB to 10dB.

Fusion rule	Method	E(Q)
Max-abs	ICA	0.3323
	CP	**0.3826**
	DT-CWT	0.3180
	Max-abs	0.3549
	RBCI	0.3779
Mean	ICA	0.3773
	CP	0.3537
	DT-CWT	0.3769
	Mean	0.3773
	RBCI	**0.3968**
Weighted average	ICA	0.3837
	CP	0.3550
	DT-CWT	0.3819
	Mean	0.3773
	RBCI	**0.4009**

can be seen as a compromise of the methods above. It succeeds in fusing the human movement and salient background details, whilst managing to suppress noise elements better than ICA. In addition, the trees in the foreground, the building, the fence and the road are all visibly sharper than in CP.

4.5 Chapter Summary

We have described a novel combinatory method for fusing images and have successfully tested RBCI's ability alongside existing fusion methods. In our tests we have looked into fusing two grayscale images. The positive results are proof that our approach is valid and meaningful. The final output was derived by a piece-wise composition of two prominent fusion techniques: ICA and CP. It is therefore entirely possible to extend the scope of our application to three or more methods and not necessarily the aforementioned two. Our original aim was to generate the best possible image from a second-level fusion step. This is a feasible direction in which to take our research.

Essentially RBCI can be defined as a patch composite of the fusion results of ICA and CP, whereby

selected pixels of ICA and CP are injected into an image framework based on the regions from the edge map. The criteria for selection is defined according to active regions, translated as areas with edge and textural details, and inactive regions. It is due to this therefore that the findings are positive; and RBCI is able to score higher than other methods when tested using the objective fusion measure for two different image scenarios. The work presented here introduces us to new possibilities for improvements especially in the sub-area of edge and saliency-based image fusion.

Chapter 5

Textural-based Fusion Metric

In the previous chapter it was shown that a novel application of combinatory fusion methods has been found to yield an enhanced performance, especially in noise-corrupted multimodal image scenarios. Established fusion performance measures are used to verify the results. Metrics such as Petrovic that heavily emphasise on edge-based calculations tend to concur favourably with these results. An area of fusion assessment that is yet to be explored however is one that focuses on textural features, defined as weak edges. This chapter proposes a new fusion metric that prioritises textural content, which may prove essential in surveillance and remote sensing applications.

5.1 The Need for a New Fusion Metric

Since its advent in the nineties, advancements in image fusion for critical applications such as remote sensing, medical imaging and military surveillance have given rise to the development of several fusion metrics to assess their performance [125]. Nevertheless, this has not always been the case. Early fusion systems rely on visual inspection and perceptual evaluation to assess their performance [137] – in other words, a good image is ultimately determined by the user. Whilst humans still remain the best adjudicator of visual quality, their method of assessment is unfeasible for real-time systems that require fast processing of data. As such, signal processing methods are used to develop a fusion quality metric that processes images in real time and correlates well with human evaluation in general.

In recent times there have been considerable interest in fusion metric schemes, based on various measures. A notable lack of 'ground truth' images in fusion applications means that reference based techniques such as signal-to-noise ratio (SNR), standard deviation (SD) and mean square error (MSE) are rendered unsuitable. These however have been replaced by the emergence of methods that emphasise the relationship and influence of each input image towards the fused output. Qu et al. [26] demonstrated a metric that utilises mutual information and entropy. Petrovic and Xydeas [138] exploited edge information for fusion evaluation purposes. Elsewhere, Piella [25] developed a novel metric that is based on the Structural Similarity Index (SSIM) proposed by Wang and Bovik [139].

There has not yet been a standard approach to formulating a capable metric for image fusion. As it is, a given metric is not necessarily competent for all image scenarios due to the fact that image fusion is a very application-specific field. A multimodal scene for instance may generate vastly different results from a multifocal scene despite using the same fusion algorithm. The task therefore is to accommodate the fusion metric to each fusion scenario. In general, a good fusion metric requires the following:

(a) Has no prerequisites for ground truth. This is highly relevant as most image fusion applications, such as night vision and hyperspectral fusion, are performed without the knowledge of ground truth.

(b) Has good correlation with the human visual system (HVS) and perceptual evaluation. For instance, as HVS is extremely sensitive to edges, a good metric should also be able to exploit this property.

(c) Is quantitative and based on a pre-determined, bounded or normalised score. This helps to standardise the performance of a fusion algorithm when applied on a wide range of image types.

(d) Has the ability to maximise good results and minimise bad results, and has a stronger separation between the two. This requirement concurs with the HVS and has been modelled by a learning curve in [138].

(e) The focus of a fusion metric is a standardised score rather than flexibility or optimisation. As

such, it requires minimal parametric input. Ideally the input values should only consist of source images and the fused image.

(f) Easy to calculate and fast to process.

In this chapter we attempt to introduce a novel method for assessing image fusion quality that fulfills most of the above criteria and works well in outdoor and natural scenery. The new metric is based on textural representations, which we believe possess an advantage over current metrics [140, 141]. It rewards fusion schemes that best preserve textural content which may therefore be crucial in the surveillance of natural terrains.

5.1.1 Our Definition of Texture

It is important here to distinguish texture from edges. Texture is defined as the variation of data at scales smaller than those of interest [135]. A popular portrayal is that texture is typically associated with 'patterns' while edges and boundaries are not, as they are defined using local constraints and seldom by global re-occurrences. Edges and boundaries are typically formed by high frequency contents while texture can be associated with both higher and lower end of the spectrum. This is true in most cases; however, in surveillance and sensor applications which form the focus of our research, the notion of texture differs somewhat. We are less concerned with texture classification based on shape or pattern than determining the presence of texture over edges i.e. texture segmentation, rather than classification or recognition. In effect, we will be ignoring conventional measures such as texels, fractals and shape grammars. Texture is thus regarded as weak edges; that is, regions where the degree of gradient is not as high as edges, but not so low so as to be called a plain region.

This is appropriate especially as edges denote a boundary between two different layered objects whereas texture records shade variation or 'coarseness' within the same object [142], as illustrated with the brick example in Figure 5.1.

The idea draws parallel with works from the previous chapter, whereby hard edges can be defined as active regions. Texture is then considered as part of the inactive region, though not as weak so as to be labelled plain. As shall be discussed, the approach divides an image's gradient content

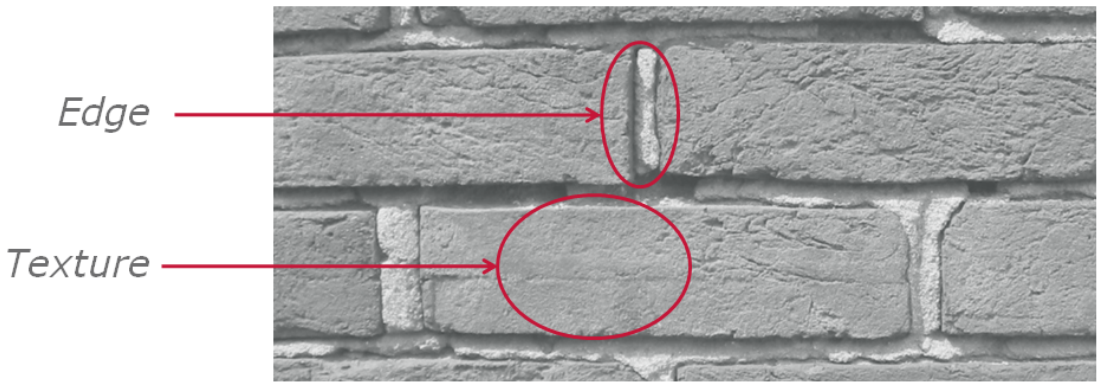


Figure 5.1: Brick example of edges versus texture

into three levels rather than two in order to accommodate the medium-level texture. A distinct difference between both schemes however is that RBCI delineates specific regions that contain active pixels, whereas textural content in this work is strictly based on a statistical histogram and lacks any spatial information.

5.2 A Review of Image Fusion Assessments Metrics

The need for a fusion metric has been widely acknowledged. It enables an efficient and impartial comparison between different fusion methods. Moreover, it allows users to optimise fusion algorithms for the best performance according to their particular needs. Theoretically, evaluating the performance of fusion is akin to measuring the degree of image degradation or enhancement. This is what is called a degradation model [143]. A modified image, considered an output, is directly measured using a specified metric against the original image or 'ground truth'. As such, common image quality assessment (IQA) methods such as [139, 144] can be used. The obvious difference is that fusion involves comparing an output image with multiple input images, rather than the singular comparison as normally used in IQA schemes. An essential issue of fusion therefore is to measure the amount of contribution (or activity level measurement [145]) from each input. In such circumstances, measures like SNR and standard deviation are considered impractical. A good fusion metric should be able to correctly estimate the respective contributions regardless of their scenery, modality or type.

Below are some popular image fusion assessment algorithms that are currently in use:

1. **Piella Metric:** In [25] Piella has extended the SSIM to be used on fused images by including a weighted average of input images to produce a fusion quality index $Q(a, b, f)$. Weights are calculated based on saliency which may be defined as contrast, sharpness or entropy. The SSIM is a measure of quality defined as:

$$Q_0 = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{\bar{x}^2 + \bar{y}^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \tag{5.1}$$

   The first component is the correlation coefficient between two real-valued sequences x and y, with $\bar{x}$ and $\bar{y}$ being the expected value and σ_x^2 and σ_y^2 the variance respectively. The second component corresponds to luminence distortion while the third measures contrast distortion. The SSIM is highly effective at measuring structural distortions of images. Piella has taken upon this concept to define a quality index for image fusion:

$$Q(a, b, f) = \frac{1}{|W|} \sum (\lambda(w) Q_0(a, f) + (1 - \lambda(w)) Q_0(b, f)) \tag{5.2}$$

   where $\lambda(w)$ is the weight given to input images a and b according to their saliency. $|W|$ refers to the cardinality of windows or local image patches used. In another work, Cvejic et al. [146] further defined saliency as the covariance between input and output images. The choice of covariance as a saliency measure is suitable as it concurs with the fourth objective above.

2. **Qu Metric:** A measure for image fusion was proposed in [26] in which mutual information (MI) is utilised as the basis. Ranjith and Ramesh [147] expanded on the concept of MI by introducing Fusion Factor (FF) and Fusion Symmetry (FS) to quantify image quality. MI measures the statistical dependence of two random variables, defined by the Kullback-Leibler distance. For two variables A and F, it calculates the amount of information F contains about A and vice-versa:

$$I_{FA}(f, a) = \sum_{f,a} P_{FA}(f, a) \log \frac{P_{FA}(f, a)}{P_F(f) P_A(a)} \tag{5.3}$$

$$I_{FB}(f,b) = \sum_{f,b} P_{FB}(f,b) \log \frac{P_{FB}(f,b)}{P_F(f)P_B(b)} \tag{5.4}$$

therefore the measure of image fusion performance is defined as

$$M_F^{AB} = I_{FA}(f,a) + I_{FB}(f,b) \tag{5.5}$$

However the measure makes no assumptions regarding the relation between both input modalities [26]. Further, MI treats an image as a global entity (histogram) and attributes to it a single score, without taking into account the spatial information of individual pixel intensities and regional structures. Another drawback of MI is its unbounded score, which may contribute to MI's inconsistent performance in recent tests [148].

3. **Petrovic Metric:** The objective image fusion performance measure [24], also called the Petrovic metric, assesses the quality of visual information obtained from the fusion of input images. The metric works by extracting all perceptually important information in the inputs, and subsequently measuring quantitatively the ability of the fusion method to transfer this information to the output image. The metric attributes salient information with the edge strength and orientation of each pixel, which are derived by the Sobel operator. A normalised weighted performance metric $Q^{AB/F} \in [0,1]$ is used to measure fusion performance.

$$Q^{AB/F} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M} Q^{AF}(n,m)w^A(n,m) + Q^{BF}(n,m)w^B(n,m)}{\sum_{i=1}^{N}\sum_{j=1}^{M}(w^A(i,j) + w^B(i,j))} \tag{5.6}$$

The performance metric is written as a weighted combination of edge preservation values of the input images A and B, respectively. Q^{AF} and Q^{BF} are the edge preservation values, while $(N \cdot M)$ is the image size. w^A and w^B are the weight factors, by which the pixels with a higher edge strength should influence the metric more than those with a lower strength.

Compared to MI, Petrovic metric has a normalised score and processes each pixel individually. The exact weights and constants used in this method can be adaptively modified so as to increase its robustness. In [149] the metric was further developed to incorporate

segmentation to identify important regions for fusion.

4. Other metrics that are found in fusion literature include [150] based on the Quantitative Correlation Analysis (QCA) which can be applied to scenarios involving a large number of source images, such as hyperspectral fusion. An improved version was proposed in [151] that analyses the general (including non-linear) relationship between source and output images.

5.2.1 Limitation of Current Metrics

From the above schemes, we noted a lack of depth in the study of texture features as the basis for a fusion metric. The closest is possibly the Petrovic metric which calculates the edge information transferred between input and output images. This is insufficient however as a specific measurement of 'texturedness', rather than simply edges, is desired.

Hence in this chapter we aim to maximise the textural content within images to develop a novel metric for image fusion assessment. The metric intends to combine textural quality, thus rewarding source images with rich texture, with the information degradation model, which measures the perceptual loss of information from input images into the fused output. The approach is based on GLCM, whereby grayscale transitions of neighbouring pixels and edge details are stored. Overall we believe the metric would be feasible for use in surveillance operations that involves a multitude of sensors which require the fusion process.

5.3 Introduction to Gray-level Co-occurrence Matrix

Gray-level co-occurrence matrix (GLCM) has proven to be a very effective tool for texture analysis [152, 153]. It defines the frequency of a pixel intensity appearing in a specified spatial linear relationship with another pixel intensity within a region. It is generally based on two parameters – distance d and orientation ϕ. d refers to the spatial distance between a reference pixel and its neighbour, whereas ϕ is quantised in four directions namely 0° (horizontal), 45° (diagonal), 90° (vertical) and 135° (inverted diagonal), with their respective symmetrical counterparts. In practice the average of all four orientations are normally used.

For an image I of size $[X, Y]$, let m represent the intensity value of pixel (x, y) and n represent the intensity of pixel $(x \pm d\phi_1, y \pm d\phi_2)$. L is the total number of gray levels for I, $0 \leq m, n \leq L-1$ whereas $0 \leq x \leq X-1$ and $0 \leq y \leq Y-1$. The GLCM $C_{m,n}$ is defined as

$$C_{m,n} = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} P\left\{I(x, y) = m \& I(x \pm d\phi_1, y \pm d\phi_2) = n\right\} \tag{5.7}$$

where $P\{.\} = 1$ if the argument remains true and 0 if false. In other words, Equation 5.7 counts the total number of occurrence for the pixel pair (m, n) over the relative distance d and in a specified direction ϕ. For instance, consider the pixel block and its GLCM equivalent in Figure 5.2, with $d = 1$, $\phi = 0°$, image size $[4, 4]$ and $L = 4$. $C_{m,n}(1, 2) = 1$ which means pixel pair $0 - 1$ occurs once in I. Due to the symmetrical property, $C_{m,n}(2, 1) = 1$ also.

I

1	2	3	3
0	3	1	0
3	2	1	2
0	2	1	3

$C_{m,n}$

0	1	1	1
1	0	4	2
1	4	0	2
1	2	2	2

Figure 5.2: GLCM

The concept of GLCM was first proposed by Haralick et. al in [154]. The paper describes fourteen features related to texture and from those, six are the most relevant: energy, entropy, contrast, variance, correlation and inverse difference moment. Each feature analyses the textural representations in the image; for example, contrast measures the intensity contrast and rewards high variations (strong edges) between neighbouring pixels. A detailed description of the main features can be found in [155]. The advantage of GLCM lies not only in the multitude of available textural features but also its flexibility in allowing new features to be derived, as was done in [156].

5.4 A Novel Texture Enhancement Feature for GLCM

As GLCM is symmetrical along its diagonal we will only consider the upper pyramid. GLCM defines the middle diagonal as the pixel-pairs of the same value (i.e. no variance). The adjacent diagonal consists of pixel-pairs of one intensity variance (e.g. 127 and 128). The following diagonal is those with two intensity variance, and so on until the last diagonal where the variance is maximum (e.g. 0 and 255). The first diagonal is attributed to regions containing uniform texture; the edge magnitude increments marginally throughout each diagonal until the top-right element, which consists of very strong edges.

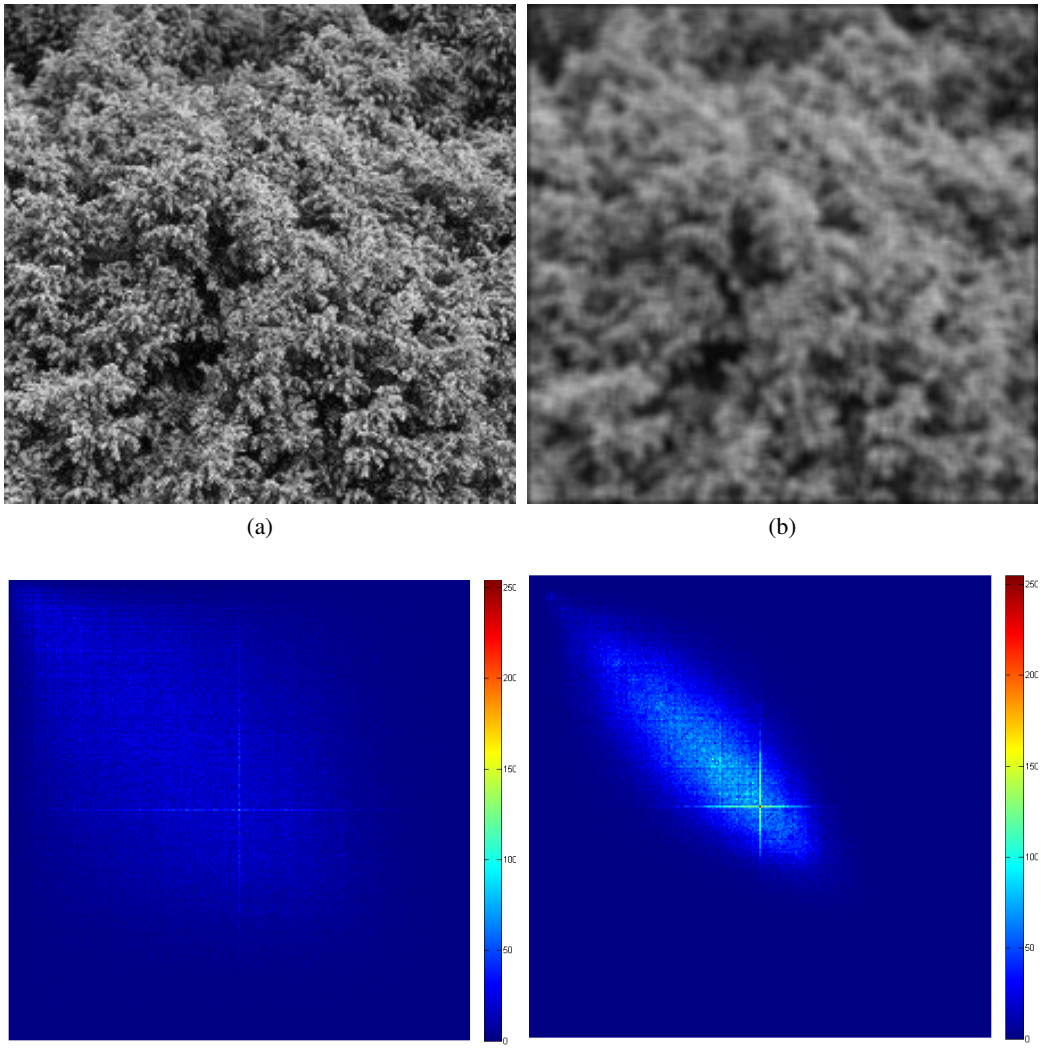
Our premise is founded on studies on IQA based on a three-component image model [157] which segmentise the different levels of edge into strong edges, texture and weak edges, and plain region. The thresholds for each category are given as $T_1 = 0.06G_{max}$ and $T_2 = 0.12G_{max}$ where G_{max} is the maximum gradient or edge strength [144]. As such,

1. Plain region: $0 \leq G \leq 0.06G_{max}$

2. Textural region: $0.06G_{max} \leq G \leq 0.12G_{max}$

3. Edge region: $0.12G_{max} \leq G$

Referring back to the texture definition in Section 5.1, GLCM is able to reflect the disparity between texture and edges quite well. An example is given in Figure 5.3. Images depicting foliage and its blurred counterpart are shown along with their respective GLCM maps. Edges denoted by high frequency pixel changes are more evident in the original foliage, therefore its GLCM values are more 'spread out' throughout the map. Conversely, lower frequency pixel intensities in the blurred foliage means that the GLCM is more concentrated towards the middle diagonals which represent lower gradients. From this it is clear that the method can be useful to elicit meaningful information from textures, rather than just high frequency edges within the image. A comparison of GLCM content in the textural region for foliage (21.6%) versus blurred foliage (35.1%) affirms this.



(a) (b)

(c) (d)

Figure 5.3: GLCM texture example (a) Bush original image, (b) Bush filtered image, (c) GLCM of original image and (d) GLCM of filtered image

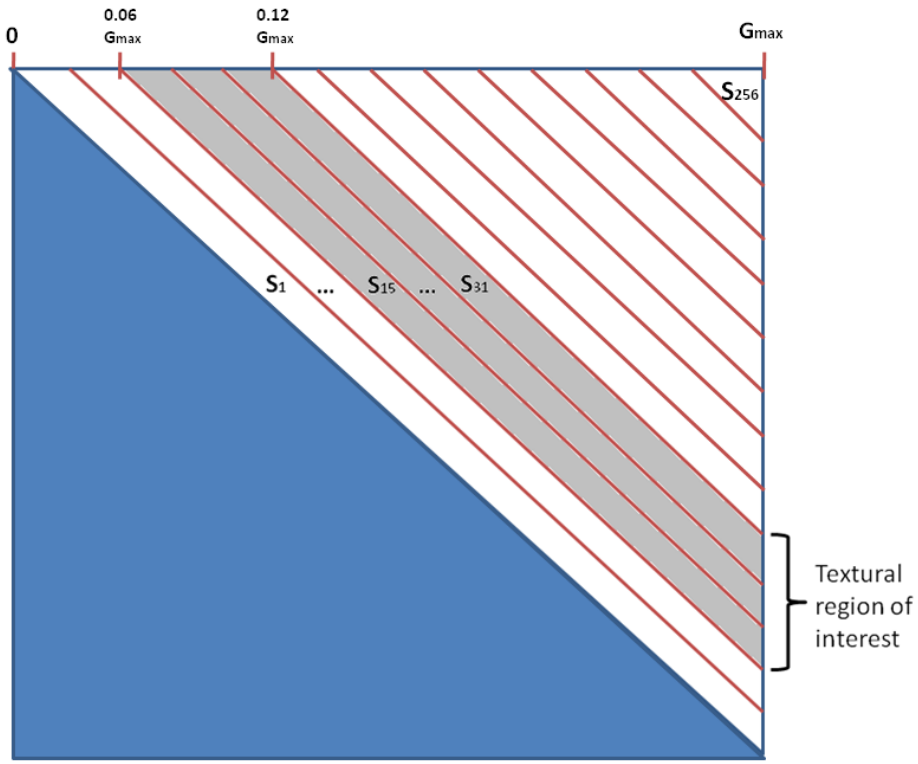


Figure 5.4: Texture definition in GLCM

The idea is incorporated into GLCM as shown in Figure 5.4. We are only interested in information contained within the scope of the textural region, denoted in gray. With this in mind, a textural-based image quality measure is derived. Vector **S** contains the sum of the elements of each diagonal. Assuming a pixel range of $p = [0, 255]$, our texture measure, T is defined as

$$T = \frac{\sum_{m=0}^{255-15} \sum_{n=m+15}^{255} C(m,n) - \sum_{m=0}^{255-31} \sum_{n=m+31}^{255} C(m,n)}{\sum_{m=0}^{255} \sum_{n=m}^{255} C(m,n)} \tag{5.8}$$

T ascribes a ratio of an image's textural content over its general saliency. That is, the remaining area when the triangular bordered by $p = 15$ is subtracted by the triangular bordered by $p = 31$.

5.5 Textural-based Fusion Metric

We aim to develop a performance assessment of image fusion methods that is based on the textural measure. The objective evaluation by Petrovic and Xydeas [138] is used as the template, with gradient strength, g and orientation, σ replaced with T. The metric was chosen as it comprehensively measures edge information, which closely relates to the textural features that form the focus of this work through the notion that texture is regarded as weak edges – though the salient differences of the two have been discussed in the previous section. The new metric is expected to perform well in natural imagery, in which textural content is abundant, for applications such as remote sensing and military surveillance.

The relative textural strength of an input image A with respect to the fused image F is obtained,

$$T_{AF} = \begin{cases} \frac{T_A}{T_F} \text{ if } T_A \leq T_F \\ \frac{T_F}{T_A} \text{ otherwise} \end{cases} \tag{5.9}$$

The relative measure is applied to generate the image's texture preservation value that relates explicitly to perceptual information in humans [138],

$$Q_T^{AF} = \frac{\Gamma}{1 + e^{-\kappa(T_{AF} - \sigma)}} \tag{5.10}$$

where Γ, κ and σ are constants that determine the exact shape of the sigmoid function representing the learning curve used to form Q_T^{AF}. It models the human visual response to the loss of edge quality in images. The texture preservation value models the information transfer from A to F, and the perceptual loss sustained. It represents how well the textural details are preserved in the fused image. Hence the weighted fusion metric, $Q_T^{AB/F}$ can be written as

$$Q_T^{AB/F} = \frac{Q^{AF} w_A + Q^{BF} w_B}{w_A + w_B} \tag{5.11}$$

The weights w_A and w_B enable us to prioritise and reward the input image with higher textural features. Theoretically the preservation values with high textural content should influence the metric more than those with relatively low textural content. We define $w = [T]^\lambda$ where λ is a constant.

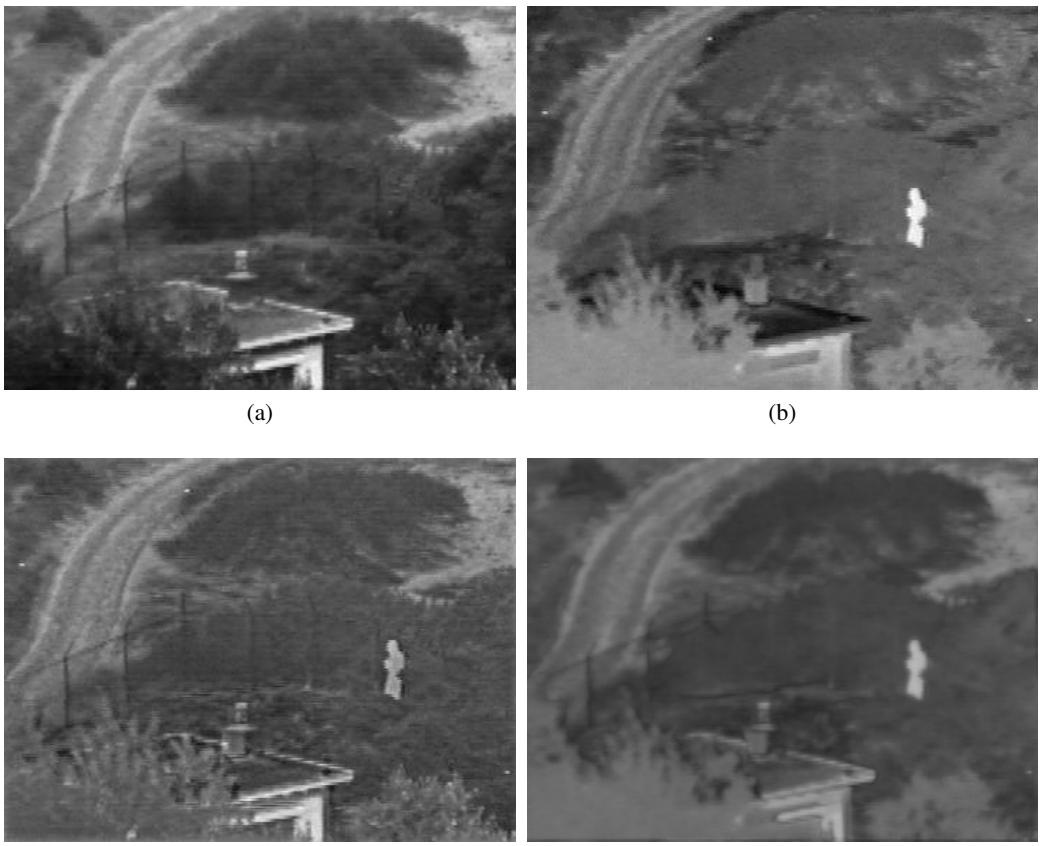
5.6 Performance Evaluation of Textural Fusion Metric

The algorithm was tested along with other state of the art fusion metrics on a number of image fusion examples, all of which were designed to mimic real world applications of surveillance, remote sensing and photography. As with the UN example, we aim to preserve textural details from the fusion process. Therefore the metric yields a high score for fusion methods that contain minimal loss of textural information.

Two fusion methods that mainly featured in the previous chapters were compared for their performance – independent component analysis (ICA) [127] and Chebyshev polynomials (CP) [34]. Figures 5.5 to 5.8 display the results. ICA is a popular signal decomposition technique that maximises the independence or non-Gaussianity of multiple mixed variables. It performs well in image fusion due to its intrinsic ability to differentiate signal components and correctly fuse them. CP on the other hand is developed mainly to address denoising in fusion applications. Its smoothing property and low-pass filtering enables enhanced performance under noisy conditions, though at the cost of lower signal accuracy. This drawback tends to be more evident in non-noisy scenarios where edges are indiscriminately smoothed out.

The GLCM representations of the input images and output were derived using $d = 3$, average ϕ

and $L = 255$. The distance selected is important to properly describe textural content. A short distance ($d = 1$) would detect mainly edges whereas a long one ($d > 5$) would overlook the texture altogether. The value chosen is deemed a reasonable description for textures for our experiments, though it can be optimised according to any particular user requirements. The sigmoid constants are valued as follows [138]: $\Gamma = 0.9994$, $\kappa = 15$, $\sigma = 0.5$ and $\lambda = 1.5$ for the proposed and Petrovic metrics respectively. The Piella and Qu metric do not require external parameters. The results are put in Table 5.1 and subsequently displayed in a chart in Figure 5.9.



(a) (b)

(c) (d)

Figure 5.5: Source images and fusion results a) UN Visual, b) UN NIR, c) UN Camp ICA and d) UN Camp CP.

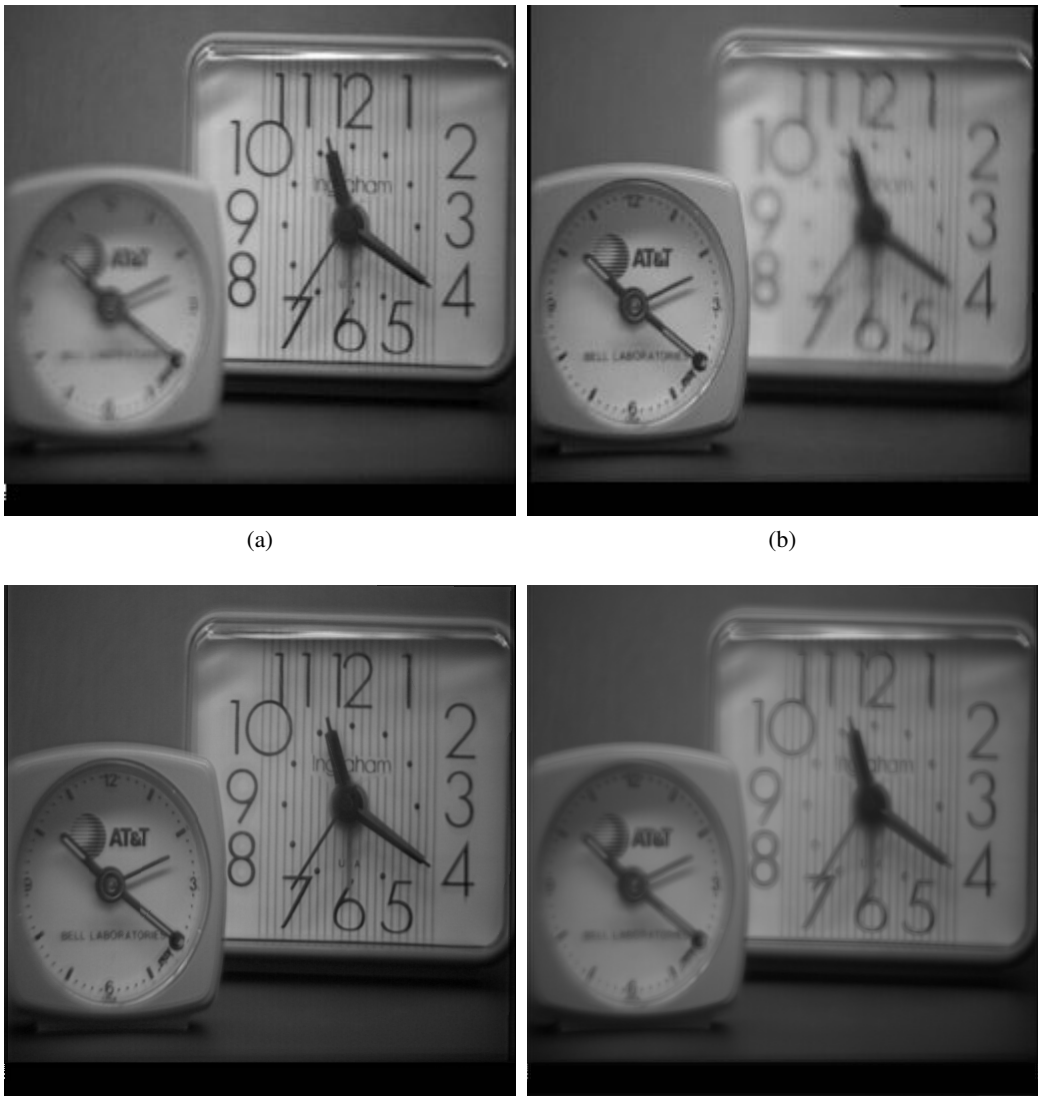
Multimodal Examples

The table and chart display scores for ICA and CP images according to the different metrics and the disparity in scores between both methods. Overall, ICA clearly displays a superior performance

(a) (b)

(c) (d)

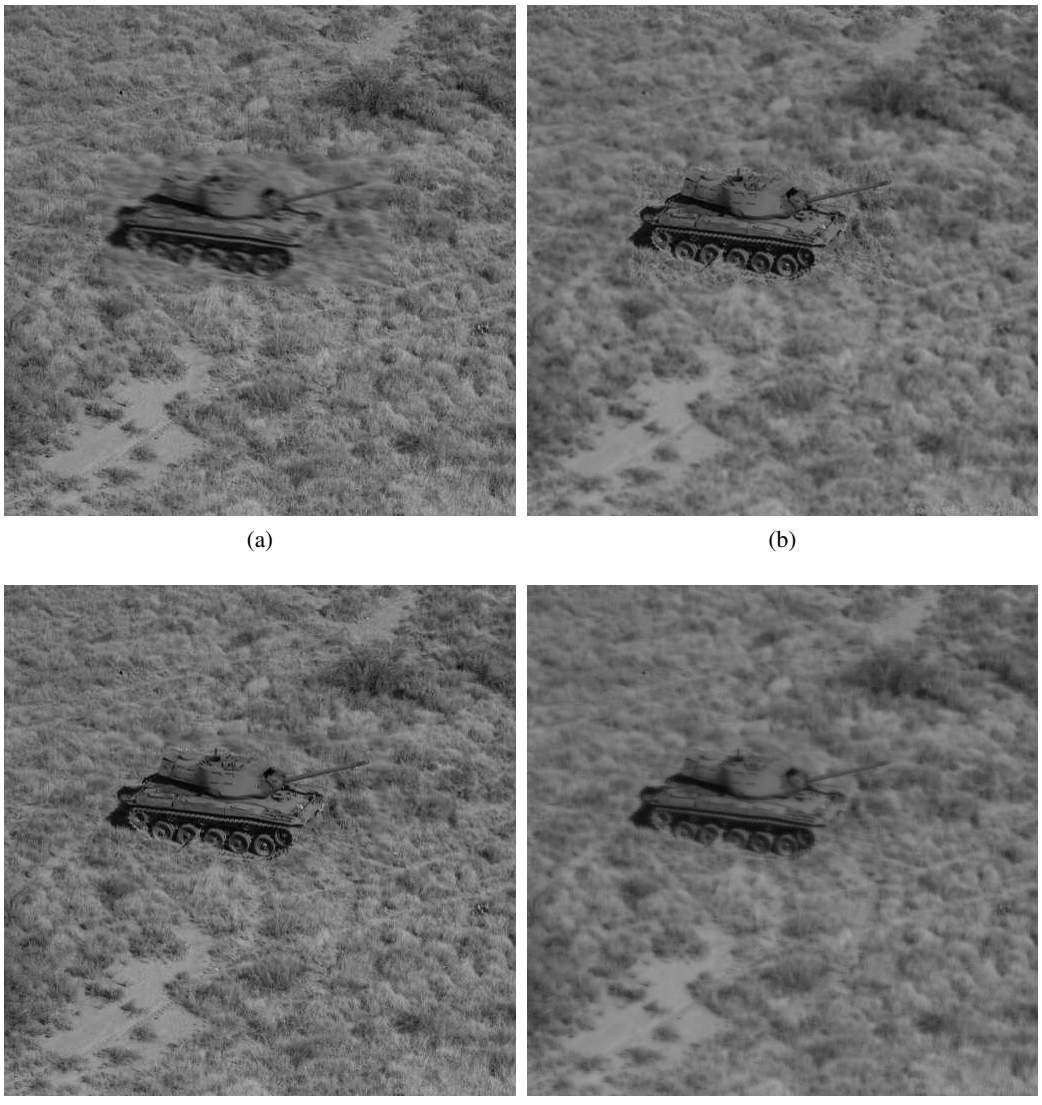
Figure 5.6: Source images and fusion results a) Gun Visual, b) Gun MMW, c) Gun ICA and d) Gun CP.



(a) (b)

(c) (d)

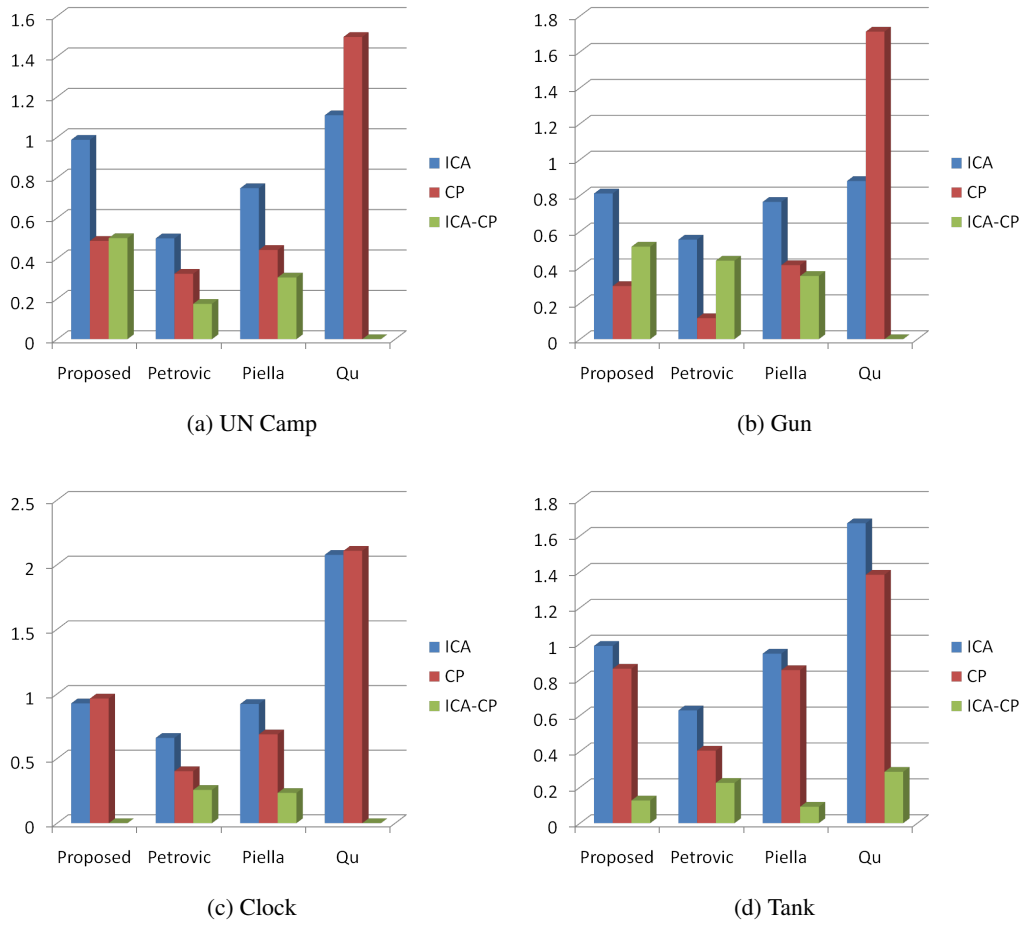
Figure 5.7: Source images and fusion results a) Clock background focus, b) Clock foreground focus, c) Clock ICA and d) Clock CP.



(a) (b)

(c) (d)

Figure 5.8: Source images and fusion results a) Field focus, b) Tank focus, c) Tank ICA and d) Tank CP.



(a) UN Camp (b) Gun

(c) Clock (d) Tank

Figure 5.9: Chart depicting relative fusion metric scores a) UN Camp, b) Gun, c) Clock and d) Tank.

Table 5.1: Analysis of objective fusion metric scores for ICA and CP results of various image sets

Image	Method	ICA	CP	ICA - CP
UN Camp	Proposed	**0.9874**	0.4866	**0.5008**
	Petrovic	**0.4991**	0.3240	0.1751
	Piella	**0.7477**	0.4418	0.3059
	Qu	1.1087	**1.4965**	-0.3878
Gun	Proposed	**0.8117**	0.2962	**0.5155**
	Petrovic	**0.5548**	0.1172	0.4376
	Piella	**0.7650**	0.4129	0.3521
	Qu	0.8814	**1.7132**	-0.8318
Clock	Proposed	0.9273	**0.9647**	-0.0374
	Petrovic	**0.6599**	0.4024	**0.2575**
	Piella	**0.9225**	0.6879	0.2346
	Qu	2.0762	**2.1071**	-0.0309
Tank	Proposed	**0.9878**	0.8609	0.1269
	Petrovic	**0.6284**	0.4040	0.2244
	Piella	**0.9447**	0.8530	0.0917
	Qu	**1.6703**	1.3840	**0.2863**

over CP in terms of texture quality retention. A close observation of the resulting images confirms this. For the UN Camp in Figure 5.5, ICA was able to retain almost all texture and scene densities whereas CP loses most of its details and is visually poor. This gap in quality is well represented by the scores in our proposed scheme.

The Qu metric was poor as not only it gives higher score to CP in the UN Camp, Clock and Gun scenarios, but its unbounded scheme renders it difficult to base the scores on a mutual standard. In this regard the Petrovic, Piella and proposed metrics consistently fare much better than Qu.

A similar observation can also be made for the Gun images in Figure 5.6. These two scenarios are examples of surveillance systems that combine features from multiple modalities or sensor cameras. As such, the aim of localised detection of the target object whilst preserving the natural features and texture is desired in a fusion scheme. Our method has shown to be favourable towards that end.

Furthermore, a closer inspection of the fusion scores reveals that our method was able to record the strongest separation between good and bad fusion. Results for $(\text{ICA} - \text{CP})$, which denotes the simple difference between ICA and CP scores, are highest for the textural metric. Figure 5.9 denotes the textural metric scores for ICA are almost double that of ICA. This is essentially in

compliance with the fourth requirement previously mentioned.

Multifocal Examples

The multifocal examples of the Clock and Tank images in Figures 5.7 and 5.8 are also observed. These consist of near similar input images which use fusion to generate a more widely focused image and overcome the blurring effects. ICA was shown to be very effective in that regard, whereas the CP results are still blurry due to its filtering property. The ICA score was appropriately reflected in our proposed metric, which scored higher than other schemes.

However the proposed metric also conveyed relatively high scores for CP, which goes against our initial hypothesis. In the case of Clock, the CP score exceeds that of ICA altogether. This may be due to the lack of natural scenery in the Clock image, and plain or monotonous regions which form the majority of the image space, unlike the surveillance examples. Because of the low texture requirement, the advantage of the textural metric is made slightly redundant.

5.7 Chapter Summary

The formulation of a novel image fusion assessment via textural preservation has been addressed in this paper. GLCM, a second order statistical method was used for the extraction of inter-pixel relationships and to measure the degree of edges and texture within an image. From this, a measure of 'texturedness' was developed which in turn facilitated the derivation of a new fusion metric based on the Petrovic measure. The proposed metric was assessed on images in comparison to other state of the art fusion metrics. Results confirm that our proposed method is viable and meaningful, though is prone to incorrect results in multifocal scenarios.

It is worth noting that a majority of image fusion algorithms are primarily aimed at transferring edge details from input images into the fused output. This is in accordance to what has already been mentioned in Section 2.2: *"The central idea of transform-based fusion methods is to modify the magnitude of the source images' coefficients, so that edges and gradients are maximised."*. This is well matched by the numerous fusion assessment metrics that are heavily biased on measuring edge features. A possible direction of our research may therefore be to develop a textural-based

image fusion algorithm to match the proposed metric.

Chapter 6

Future Work and Conclusions

This section discusses the merits and limitations related to the research, as well as possible future directions for the work laid out in this thesis. The works presented in this thesis revolve around issues and solutions for image fusion with particular emphasis on its performance in noisy conditions, and more appropriate measures for fusion assessment. Overall, we feel we have provided a comprehensive study in the general area of image fusion that may benefit future endeavours towards this research, though the thesis is by no means without imperfections.

6.1 Discussion and Future Work

Several points have been offered that can provide the basis for future work:

1. Throughout Chapters 3, 4 and 5 the tests conducted involves fusion of two source images. In a standard fusion process, higher level features from both inputs are compared and the best features are retained to generate a composite fused output image. In the case of Chebyshev polynomials and independent component analysis (ICA), or indeed many other fusion methods, this can be extended to three or more inputs through sequential fusion, or simply selecting the best input coefficients x:

$$\text{argmax}_x f(x) \tag{6.1}$$

2. In Chapter 4, the proposed approach of combining algorithms within a region-based framework is strictly limited to two methods. In this thesis, ICA and CP were chosen due to their appealing properties – ICA performs better in non-noisy conditions, while CP does well in medium to heavy noise situations. The selected algorithm is flexible enough to accommodate other fusion methods such as wavelets and pyramid-based fusion, which may be tested and used in their place so as to optimise the framework for best performance. However the current demarcation of active and non-active regions only allows a combination of two fusion methods in place. An expansion of this for future consideration may be to allocate multiple levels of region according to activity level measurement. By doing so, three or more fusion algorithms may be considered. An ideal structure for our approach would be to have the flexibility of interchanging between many fusion algorithms based on their optimal regions. This ensures that the best results from whichever fusion algorithm is suited to any particular image region, which effectively enhances image fusion performance.

3. The scope of fusion for this research encompasses 8-bit grayscale image data sets. Expansion to 32-bit RGB oriented images is certainly possible, though such an operation would be more suited for processing in the IHS, YUV or L$\alpha\beta$ colour spaces. This allows the important chromacity (intensity) channel to be modelled separately from colour information, which reduces our problem to one of grayscale fusion. Works on RGB and intensity channel fusion algorithms and assessment, for which [158–160] are examples, has emerged as a popular trend and is seen as the step forward in applications such as surveillance and remote sensing.

4. Further, GLCM was designed with the 8-bit grayscale in mind. An interesting extension to our work in Chapter 5 would therefore be to expand the GLCM theory to various colour space models, where each channel possesses its own statistical histogram and the relationships between channels can be explored and optimised.

5. A note on the processing performance of Chebyshev polynomials in Chapters 3 and 4 – implementations of $m = 11$ by $n = 11$ degrees CP on the Matlab platform is typically slow, though this owes largely to the nested operations that form the bulk of the algorithm designed for bivariate signals. The computational performance is expected to increase significantly with more sound programming in place.

6. The limitation of the Chebyshev fusion method in Chapter 3 is readily acknowledged. Algorithms in general take inspiration from natural phenomena and try to exploit them with the limited constraints they are given. With regards to CP, the algorithm works well in fusion scenarios where one or more source images is corrupted by medium to high noise levels. This is attributed to the inherent property of the polynomials which tend to approximate smoothly over a signal, albeit with a compromise on signal accuracy. Such qualities are advantageous in high frequency signals that form most of the AWGN. The method however loses its advantage in normal, non-noisy scenarios, where the smooth filtering properties of CP is made redundant. Applications where CP fusion may therefore be suited to include outdoor surveillance that involves high frequency range, low quality and noise prone cameras such as NIR.

7. Further, the properties deemed useful for image approximation is not unique to Chebyshev polynomials alone. It is therefore entirely feasible that a broader class of orthogonal polynomials, such as the Legendre, Laguerre and Zernike may be tested in place of CP to optimise fusion performance. To this end, previous studies have shown Laguerre [121] and Krawtchouk [161] polynomials to have favourable performance on image analysis methods.

8. The textural quality metric described in Chapter 5 bears a closer resemblance to human perceptual evaluation than other metrics. This is reflected by the high disparity in scores between good and bad fusion results, as well as high scores attributed for texture preservation. To extend the concept, a measure of 'resemblance' is desired. The work in [125] provides us with two such measures: correct ranking (CR) and relevance measure (r). Optimisation techniques may then be used to tailor the parameters of the objective metric to concur with HVS. All these provide the groundwork for future exploration in research regarding image fusion assessment.

9. The gold standard in image fusion, where multimodal colour video sequences and even audio data are merged in real time, is still some distance away but not altogether inconceivable. This forms the future direction of image fusion in general. The practicality of such systems is highly dependent not only on the processing tools available, but also the data size at hand. Conventional pyramid-based fusion algorithms are able to process images of size 512×512 pixels in sequence without suffering delay, thus giving the impression of 'real-time' fusion. In contrast, images 2000×2000 pixels in size is much harder to process. Further, the standard

requires a live feed of video input and is expected to provide instantaneous output, which is crucial to detect and track movements and report anomalies in a surveillance application. Currently there have been concerted efforts on studies of video fusion, primarily for nighttime vision as examples [162, 163] have shown.

6.2 Conclusion

One of the major concerns of fusion in real world applications is the presence of natural noise. This is certainly the case with security, surveillance, medical imaging, millitary and remote sensing where cost/quality trade-offs often result in image acquisition tools of lower quality and higher susceptibility to noise. The fusion process is therefore used to mitigate this effect and generate a higher quality output image. This thesis has dealt with a number of issues within the area of image fusion. Most notably, the performance of fusion under high noise conditions was primarily addressed by the application of a new set of robust basis functions for image decomposition. The subject of image fusion performance measure has also been given considerable thought, where novel image features are selected so as to give a more wholesome description of the fusion result's quality.

Chapter 1 discusses the idea of image quality and provides the problem statement, namely corrupted images borne out of the presence of noise elements in low-cost image acquisition systems. Further, it has been reasoned that a combination of multiple modalities within a single image, where all salient features are transferred, should yield a higher quality than any single modality. These two notions give rise to the use of signal processing methods to develop image fusion techniques to enhance image quality.

An introduction and background study of the field of image fusion is given in Chapter 2. The section aims to brief the general reader as to the basics of fusion and its motivation. The history of image fusion, the advent of which coincides with the development of pyramid-based decomposition methods in the 1980's, has also been discussed. The chapter concludes with a brief discourse on the contribution of image fusion in diverse fields, ranging from digital photography to remote sensing.

Chapter 3 provides the first substantial contribution of this research – addressing fusion and denoising issues through the first application of classical Chebyshev polynomials that form the basis functions for univariate signal approximation. The Chebyshev polynomials were chosen for their distinct properties, which have proven effective for denoising especially on signals inundated with high frequency noise. CP is extended to bivariate image signals via the separable application of one-dimensional polynomials along the x and y axis. Initial tests suggest that the proposed method is valid and meaningful, whilst further evaluation on a number of image datasets corrupted by medium to high levels of noise has affirmed the performances of CP to be much more encouraging in these conditions than other methods.

The approach is then further developed in Chapter 4, comprising an algorithm fusion that involves CP and independent component analysis (ICA) to improve upon the original method. The strong edge and boundary representation by ICA is able to mitigate the blurry drawbacks of CP, whilst CP holds a superior performance to ICA in noise-affected regions. The modified scheme selects the best result from either algorithms by virtue of defined active and inactive regions, that conclusively attains an increase in the fusion performance in similar based noisy conditions.

Chapter 5 of this thesis is concerned with the evaluative aspect of image fusion performance. We essentially propose the development of a measure for image fusion assessment that emphasises textural retention, which correlates well with several generic requirements of a good fusion metric. It is rationalised that image textural features are synonymous with natural outdoor imagery, the preservation of which is crucial for fusion applications such as remote sensing, surveillance and the military. Gray level co-occurrence matrix (GLCM), a second-order statistical histogram is utilised to derive new textural features and in turn the new metric. Tests have demonstrated that the proposed metric tends to reward fusion algorithms that best retain texture information, as well as maximise the separation between good and bad fusion results in multimodal scenery.

Appendix A

Two-dimensional Chebyshev polynomials

A two-dimensional description of the classical Chebyshev polynomial of the second kind $P_{k,l}$ was introduced by Koornwinder [164] and its properties were elaborated in [124]. This concept was pursued in [165], where the definitions of the bivariate Chebyshev polynomials were proposed.

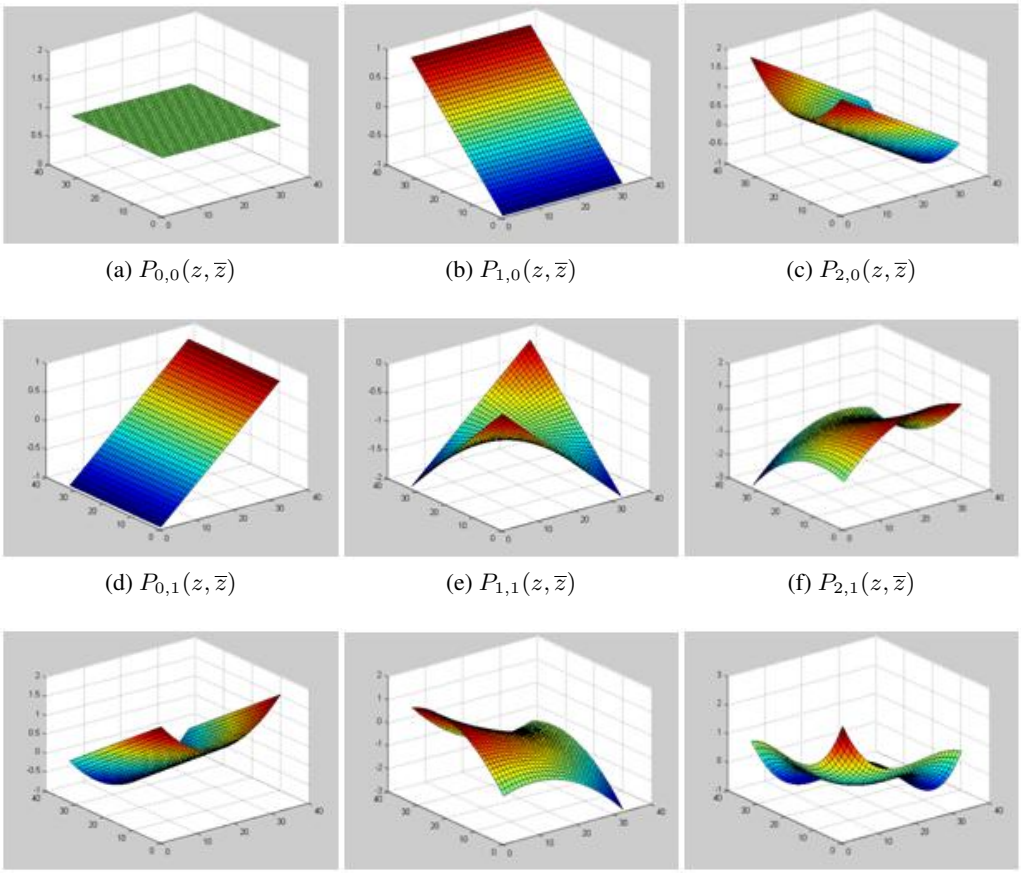
$$P_{-1,l}(z,\overline{z}) = 0,\ P_{k,-1}(z,\overline{z}) = 0,$$
$$P_{0,0}(z,\overline{z}) = 1,$$
$$P_{1,0}(z,\overline{z}) = z,\ P_{0,1}(z,\overline{z}) = \overline{z}$$

A more general form is the following:

$$zP_{k,l}(z,\overline{z}) = P_{k+1,l}(z,\overline{z}) + P_{k-1,l+1}(z,\overline{z}) + P_{k,l-1}(z,\overline{z}) \tag{A.1}$$

$$\overline{z}P_{k,l}(z,\overline{z}) = P_{k,l+1}(z,\overline{z}) + P_{k+1,l-1}(z,\overline{z}) + P_{k-1,l}(z,\overline{z}) \tag{A.2}$$

where $z = x + jy$ and $\overline{z} = x - jy$ are complex correspondents of x and y. The graphs for polynomial order $P_{0,0}(z,\overline{z})$ to $P_{2,2}(z,\overline{z})$ are displayed in Figure A.1. We can observe that they are merely bi-dimensional extensions of the one dimensional CP.



(a) $P_{0,0}(z,\overline{z})$ (b) $P_{1,0}(z,\overline{z})$ (c) $P_{2,0}(z,\overline{z})$

(d) $P_{0,1}(z,\overline{z})$ (e) $P_{1,1}(z,\overline{z})$ (f) $P_{2,1}(z,\overline{z})$

(g) $P_{0,2}(z,\overline{z})$ (h) $P_{1,2}(z,\overline{z})$ (i) $P_{2,2}(z,\overline{z})$

Figure A.1: Graphs of 2D Chebyshev polynomials

The following definition establishes two-variable Chebyshev polynomials as basis functions for 2D signal approximation

$$\tilde{f}(z,\overline{z}) = \sum_{k=0}^{K}\sum_{l=0}^{L} a_{k,l} P_{k,l}(z,\overline{z}). \tag{A.3}$$

Derivation of the coefficient $a_{k,l}$ is given by the following equation:

$$a_{k,l} = \sum_{z=-1}^{1}\sum_{\overline{z}=-1}^{1} \mu(z,\overline{z}) f(z,\overline{z}) P_{k,l}(z,\overline{z}). \tag{A.4}$$

The approximation forms an orthonormal system subject to the weight function:

$$\mu(z,\overline{z}) = \frac{1}{2\pi^2}\sqrt{-z^2\overline{z}^2 + 4z^3 + 4\overline{z}^3 - 18z\overline{z} + 27}. \tag{A.5}$$

However, this entails that the approximated data must reside over the region S inside Steiner's hypocycloid, ∂S.

$$\partial S(\theta) = 2e^{j\theta} + e^{-2j\theta}, 0 \leq \theta \leq 2\pi. \tag{A.6}$$

Solving this requires an integration of the system over $0 \leq \theta \leq 2\pi$ within the hypocycloid. Given that orthogonal polynomials do not correspond to the discrete image coordinate space, a transformation [166] of the Cartesian coordinate such as that from (x, y) to (θ, r) is required, all within the boundaries of a hypocycloid shape. This would not only inevitably create numerical errors, but also geometrical errors due to the severe disparity in the coordinate shape. The complexity in solving this problem has led us to contemplate alternative definitions for two-dimensional Chebyshev polynomials.

Appendix B

Discrete Chebyshev polynomials as separable basis functions

A description of discrete Chebyshev polynomials (DCP) was introduced by Erdelyi et al. [167] and used later by Nikiforov et al. [168], from which orthogonal moments are used in image analysis [126,166,169]. It is reasoned that since orthogonal polynomials such as Zernike, Legendre and the classical Chebyshev polynomials are defined in the continuous domain, they are thus deemed unsuitable for discretised images. A transformation of the coordinate space would be required, which increases computational complexity and is prone to numerical and geometric error.

DCP operates within the interval $[0, N-1]$ and has a unit weight function, $w(x) = 1$, thus making it the simplest of all discrete orthogonal polynomials. Due to exponential growth with n, the original set of definitions provided for DCP is less suited for approximation. This has led to a description of a scaled version of discrete Chebyshev polynomials $\tilde{t}(x)$ [166]

$$\tilde{t}_n(x) = \frac{t_n(x)}{\beta(n, N)} \tag{B.1}$$

in which $\beta(n, N)$, a constant independent of x, is a scale factor that grows as n grows. If we choose $\beta(n, N) = N^n$ then the recursive formula below can be applied

$$\tilde{t}_n(x) = \frac{(2n-1)\tilde{t}_{n-1}(x) - (n-1)(1 - \frac{(n-1)^2}{N^2})\tilde{t}_{n-2}(x)}{n}, n = 2, 3, ...N-1 \tag{B.2}$$

where $\tilde{t}_0(x) = 1$ and $\tilde{t}_1(x) = \frac{(2x+1-N)}{N}$.

A DCP plot for $N = 33$ is shown in Figure B.1.

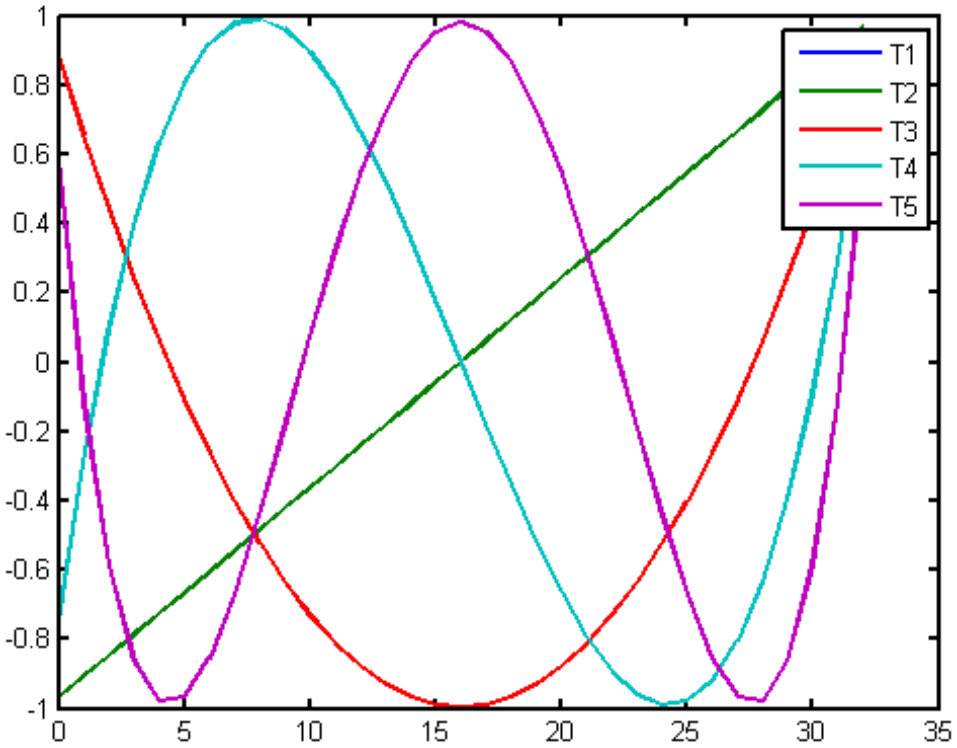


Figure B.1: Discrete Chebyshev polynomials, with $N = 33$ points

The approximation formula for a two-dimensional signal is given below.

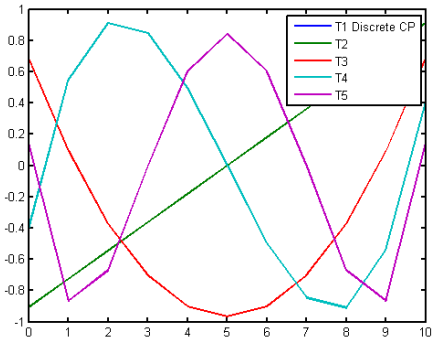
$$\tilde{f}(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{m,n} \tilde{t}_m(x) \tilde{t}_n(y), \tag{B.3}$$

where $a_{m,n}$ is a coefficient defined by the formula below, and X and Y are the size of data x and y respectively.

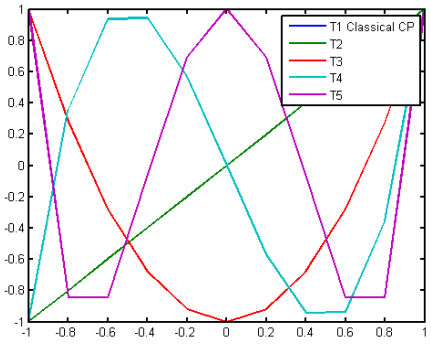
$$a_{m,n} = \sum_{x=1}^{X} \sum_{y=1}^{Y} f(x,y) \tilde{t}_m(x) \tilde{t}_n(y) \tag{B.4}$$

A significant disadvantage of DCP is non-local shift invariance, hence their reliance on data points for judicious approximation. A comparative analysis of DCP versus classical CP (CCP) reveals

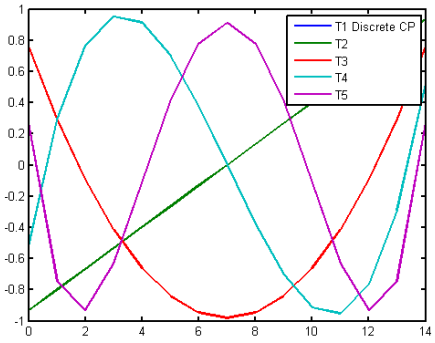
that DCP tend to converge slower with respect to the increase in data size N. This is evident in Figure B.2 below.



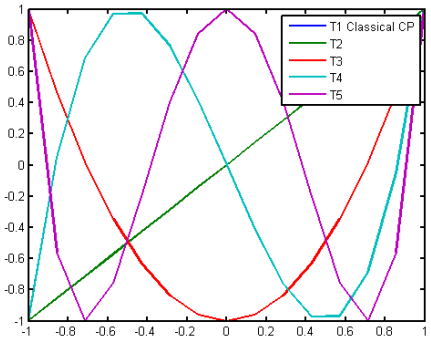
(a) Discrete L=11



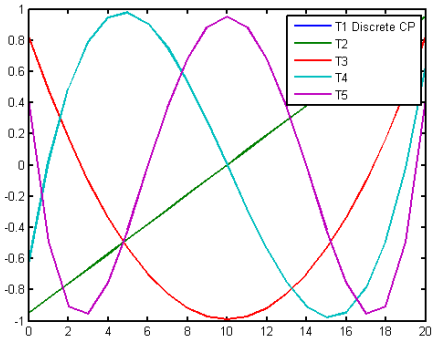
(b) Classical L=11



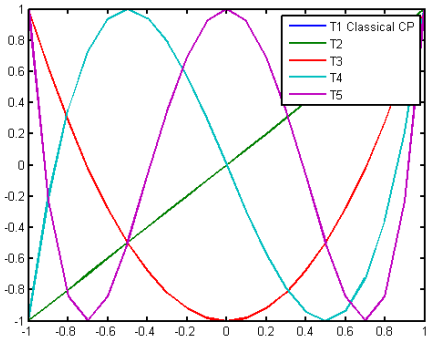
(c) Discrete L=15



(d) Classical L=15



(e) Discrete L=21



(f) Classical L=21

Figure B.2: Comparison of the first five discrete versus classical Chebyshev polynomials for data length $L = 11, 15, 21$

As N increases throughout the figures, CCP tends to converge faster with almost perfect reconstruction at $N = 15$ in Figure B.2d, as opposed to DCP (Figure B.2c). Owing to the disparity in

performance, classical CP is the natural choice for basis functions working with relatively small image patches. This is presented in Chapter 3.

Bibliography

[1] V.S. Petrovic and C.S. Xydeas, "Sensor noise effects on signal-level image fusion performance", *Information Fusion 4*, 2003, pp. 167-183.

[2] J.Z. Cao, Z.F. Zhou, H. Wang and W.H. Liu, "Multifocus noisy image fusion algorithm using the contourlet transform", *International Conference on Multimedia Technology*, 2010, pp. 1-4.

[3] D. Drajic and N. Cvejic, "Adaptive fusion of multimodal surveillance image sequences in visual sensor networks", *IEEE Transactions on Consumer Electronics*, Vol. 53, No. 4, 2007, pp. 1456-1462.

[4] J.H. Park and M.G. Kang, "Noise reduction for gradient-based multispectral image fusion", *Proc. of Int. Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2004, pp. 185-188.

[5] R. Boie and I. Cox, "An analysis of camera noise", *IEEE Transactions of PAMI*, 14(6), 1992, pp. 671-674.

[6] A. Buades, B. Coll and J.M. Morel, "A review of image denoising algorithms, with a new one", *Multiscale Model. Simul.*, Vol. 4, No. 2, 2005, 490-530.

[7] J.S. Hae, P. Chatterjee, H. Takeda, and P. Milanfar,"A comparison of some state of the art image denoising methods", *Conf. Record of 41st Asilomar Conf. on Signals, Systems and Computers*, 2007, pp. 518-522.

[8] J.C. Church, Y. Chen and S.V. Rice, "A spatial median filter for noise removal in digital images", *IEEE Southeastcon*, 2008, pp. 618-623.

[9] J. Canny, "A computational approach to edge detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(6), 1986, pp. 679698.

[10] J. Harikiran, B. Saichandana, K. Keerthi and D. Anish, "A new method of image fusion technique for impulse noise removal in digital images", *Global Journal of Computer Science and Technology*, Vol. 11, Issue 7, 2011, pp. 27-30.

[11] S. Indu and C. Ramesh, "Image fusion algorithm for impulse noise reduction", *International Conference on Advances in Recent Technologies in Communications and Computing*, 2009, pp. 309-313.

[12] M. Tico and K. Pulli, "Image enhancement method via blur and noisy image fusion," *16th IEEE Int. Conf. on Image Processing (ICIP)*, 2009, pp.1521-1524.

[13] T. Chen, K.K. Ma and L.H. Chen, "Tri-state median filter for image denoising", *IEEE Transactions on Image Processing*, Vol. 8, No. 12, 1999, pp. 1834-1838.

[14] W. Xin, L. Gaolue, Y. Weibo, Y. Limei and X. Zhongyu, "An effective infrared and visible image fusion algorithm restraining noise", *Int. Conf. on Transportation, Mechanical and Electrical Engineering*, 2011, pp. 2323-2326.

[15] D. Drajic and N. Cvejic, "Multimodal image fusion in presence of noise using sparse coding of ICA coefficients", *IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 343-346.

[16] Y. Zhang, S. De Backer and P. Scheunders, "Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, No. 11, 2009, pp. 3834-3843.

[17] B. Liao and L.X. Liu, "Noise resilient image fusion based on orthogonal matching pursuit", *Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics*, 2009, pp. 3112-3115.

[18] S.M.M Rahman, M.O. Ahmad and M.N.S. Swamy, "Contrast-based fusion of noisy images using discrete wavelet transform", *IET Image Processing*, Vol. 4, Issue 5, 2010, pp. 374-384.

[19] S.C. Douglas, A. Cichocki and S. Amari, "A bias removal technique for blind source separation with noisy measurements", *Electronics Letters*, 34, 1998, pp. 1379-1380.

[20] A. Hyvarinen, "Fast independent component analysis with noisy data using Gaussian moments", *Proc. Int. Symp. On Circuits and Systems*, 1999, pp. 57-61.

[21] Z. Chen, Y. Zheng, B.R. Abidi, D.L. Page and M.A. Abidi, "A combinational approach to the fusion, de-noising and enhancement of dual-energy X-ray luggage images," *IEEE Conf. on Computer Vision and Pattern Recognition - Workshops*, 2005, pp. 2.

[22] M. Kumar and R.L. Miller, "An image fusion approach for denoising signal-dependent noise", *Int. Conf. on Acoustics Speech and Signal Processing (ICASSP)*, 2010, pp.1438-1441.

[23] Z. Wang, D. Ziou, C. Armenakis, D. Li and Q. Li, "A comparative analysis of image fusion methods", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 6, 2005, pp. 1391-1402.

[24] V. Petrovic and T. Cootes, "Information representation for image fusion evaluation", *9th International Conference on Information Fusion*, 2006, pp. 1-7.

[25] G. Piella, H. Heijmans, "A new quality metric for image fusion", *Proc. of Intl. Conf. on Computational Vision (ICCV 03)*,Vol. 3, 2005, pp. 173-176.

[26] G. Qu, D. Zhang and P. Yan, "Information measure for performance of image fusion", *Electronics Letters 38* (7), 2002, pp. 313 315.

[27] A. Garzelli, "Possibilities and limitations of the use of wavelets in image fusion", *IEEE International Geoscience and Remote Sensing Symposium*, Vol. 1, 2002, pp. 66-68.

[28] H.R. Sheikh and A.C. Bovik, "Information Theoretic Approaces to Image Quality Assessment", in: *A.C. Bovik (Ed.), Handbook of Image and Video Processing*, Elsevier, 2005.

[29] Y. Chen and R.S. Blum, "Experimental Tests of Image Fusion for Night Vision", *Proc. 8th International Conference on Information Fusion*, 2005.

[30] R.S. Blum, Z. Xue and Z. Zhang, "An Overview of Image Fusion", in R.S. Blum (Ed.), *Multi-Sensor Image Fusion and Its Applications*, Marcel Dekker, 2005.

[31] P.J. Burt and R.J. Kolczynski, "Enhanced image capture through fusion", *Proc. 4th Intl. Conference on Computer Vision*, 1993, pp. 173-182.

[32] A. Loza, "Statistical model-based fusion of noisy images", *Computer Science Departmental Seminar*, University of Bristol, 22 June 2007.

[33] N. Cvejic, D. Bull and N. Canagarajah, "Improving fusion of surveillance images in sensor networks using independent component analysis", *IEEE Transactions on Consumer Electronics*, Vol. 53, No. 3, 2007, 1029-1035.

[34] Z. Omar, N. Mitianoudis and T. Stathaki, "Two-dimensional Chebyshev polynomials for image fusion", *28th Picture Coding Symposium*, Japan, 2010, pp. 426 -429.

[35] L. Wald, "Some terms of reference in data fusion", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 3, 1999, pp. 1190-1193.

[36] D.P. Mandic, D. Obradovic, A. Kuh, T. Adali, U. Trutschel, M. Golz, P. De Wilde, J. Barria, A. Constantinides and J. Chambers, "Data fusion for modern engineering applications: An overview", *International Conference on Artificial Neural Networks 2005*, LNCS 3697, 2005, pp. 715-721.

[37] D.B. Henson, "Visual Fields". *Oxford University Press*, 1993.

[38] R. Blake and R. Fox, "The psychophysical inquiry into binocular summation", *Perception & Psychophysics* 14(1), 1973, pp. 16185.

[39] M. Mangolini, "Apport de la fusion dimages satellitaires multicapteurs au niveau pixel en tel ed etection et photo-interpretation", *These de 'doctorat*, Univ. NiceSophia Antipolis, France, 1994, pp. 174.

[40] G. Brandstatter and A. Sharov, "Environmental monitoring in the high Arctic using different types of high-resolution satellite imagery", *Int. Arch. Photogramm. Remote Sensing*, Pt. 7, Vol. 32, 1998, pp. 201216.

[41] B. Csatho and T. Schenk, "Multisensor data fusion for automatic scene interpretation", *Int. Arch. Photogramm. Remote Sensing*, Pt. 7, Vol. 32, 1998, pp. 336341.

[42] C. Pohl and J.L. van Genderen, "Multisensor image fusion in remote sensing: Concepts, methods and applications", *Int. J. Remote Sensing*, Vol. 19, No. 5, 1998, pp. 823854.

[43] N. Mitianoudis amd T. Stathaki, "Adaptive image fusion using ICA bases", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2006.

[44] S. Nikolov, P.R. Hill, D.R. Bull, C.N. Canagarajah, "Wavelets for image fusion". In A. Petrosian and F. Meyer, editors, *Wavelets in Signal and Image Analysis, from Theory to Practice*. Kluwer Academic Publishers, 2001.

[45]] E.H. Adelson, C.H. Anderson, J.R. Bergen, P.J. Burt and J.M. Ogden, "Pyramid methods in image processing", *RCA Engineer*, 29-6, 1984, pp. 33-41.

[46] T. Huntsberger and B. Jawerth, "Wavelet-based sensor fusion", in *Proc. SPIE*, Vol. 2059, 1993, pp. 488-498.

[47] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications", *Neural Networks 13*, 2000, pp. 411-430.

[48] H. Hariharan, A. Gribok, B. Abidi and M. Abidi, "Multi-modal face image fusion using empirical mode decomposition", *The Biometric Consortium Conference*, USA, 2005.

[49] A.A. Goshtasby and S. Nikolov, "Image fusion: advances in the state of the art", *Guest Editorial in Information Fusion 8*, 2007, pp. 114-118.

[50] Q. Tao and R. Veldhuis, "Threshold-optimised decision-level fusion and its application to biometrics", *Pattern Recognition*, 42, 2009, pp. 823-826.

[51] P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code", *IEEE Transaction on Communications*, 31(4), 1983, pp. 532-540.

[52] Y. Zhang, "Understanding image fusion", *Photgrammetric Engineering and Remote Sensing*, 2004, pp. 657-661.

[53] W. Wang et al., "Video assisted speech source separation", *Proc. Intl. Conf. Acoustics, Speech and Signal Processing*, 2005, pp. 425-427.

[54] G. Qu, D. Zhang and P. Yan, "Medical image fusion by wavelet transform modulus maxima", *Optics Express*, Vol. 9, No. 4, 2001, pp. 184-190.

[55] C.S. Pattichis, M.S. Pattichis and E. Micheli-Tzanakou, "Medical imaging fusion application: an overview", *Asilomar Conference on Signals, Systems and Computers*, 2001, pp. 1263-1267.

[56] Z.Q. Hong, "Algebraic feature extraction of image for recognition", *Pattern Recognition*, Vol. 24, Issue 3, 1991, pp. 211-219.

[57] J.P. Djamdji, A. Bijaoui and R. Maniere, "Geometrical registration of images: the multiresolution approach", *Photogrammetric Eng. and Remote Sensing*, Vol. 59, No. 5, 1993, pp. 645-653.

[58] J.P. Heather and M.I. Smith, "Multimodal image registration with applications to image fusion", *Proc. 8th Intl. Conf. Information Fusion*, USA, 2005.

[59] I. Dagher and R. Nachar, "Face recognition using IPCA-ICA algorithm", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 28, No. 6, 2006, pp. 996-1000.

[60] B. Van Dun, J. Wouters and M. Moonen, "Improving auditory steady-state response detection using independent component analysis on multichannel EEG data", *IEEE Trans. On Biomedical Engineering* , Vol. 54, Issue 7, 2007, pp. 1220-1230.

[61] S. Waldert, "Real-time fetal heart monitoring in biomagnetic measurements using adaptive real-time ICA", *IEEE Trans. On Biomedical Engineering*, Vol. 54, Issue 107, 2007, pp. 1864-1874.

[62] K.C. Kwak and W. Pedrycz, "Face recognition using enhanced independent component analysis approach", *IEEE Trans. on Neural Networks*, Vol. 18, No. 4, 2007, pp. 530-541.

[63] E.J. Bender, C.E. Reese and G.S. van der Wal, "Comparison of additive image fusion versus feature-level image fusion techniques for enhanced night driving", *Proc. of SPIE, Vacuum and Solid State Photoelectronic Imagers, Detectors, and Systems*, 4796, 2002.

[64] H. Li, B.S. Manjunath and S.K. Mitra, "Multi-sensor image fusion using the wavelet transform", *Graphical Models and Image Processing 57 (3)*, 1995, pp. 235-245.

[65] J. M. Ogden , E. H. Adelson , J R. Bergen , P.J. Burt, "Pyramid-based computer graphics", 1985.

[66] F. Sadjadi, "Comparative image fusion analysis", *IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum*, 2005.

[67] P.J. Burt, "The pyramid as a structure for efficient computation", in A. Rosenfeld (Ed.), *Multiresolution Image Processing and Analysis*, Springer-Verlag, 1984.

[68] A. Toet, "Image fusion by a ratio of low-pass pyramid", *Pattern Recognition Letters 9*, 1996, pp. 245-253.

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

[70] C.H. Anderson, "A filter-subtract-decimate hierarchical pyramid signal analysing and synthesising technique", *United States Patent 4,718,104*, 1987.

[71] A. Toet, "A morphological pyramidal image decomposition", *Pattern Recognition Letters 9*, 1989, pp. 255-261.

[72] J. Zeng, A. Sayedelahl, T. Gilmore and M. Chouikha, "Review of image fusion algorithms for unconstrained outdoor scenes", *Proc. Intl. Conf. on Signal Processing*, Vol. 2, 2006.

[73] P.J. Burt, "A gradient pyramid basis for pattern-selective image fusion", *Society for Information Display, Digest of Technical Papers*, 1992, pp. 467-470.

[74] R.D. Lillquist, "Composite visible/thermal-infrared imaging apparatus", *United States Patent 4,751,571*, 1988.

[75] N. Nandhakumar and J.K. Aggarwal, "Integrated analysis of thermal and visual images for scene interpretation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4), 1988, pp. 469-481.

[76] S.K. Rogers, C.W. Tong, M. Kabrisky and J.P. Mills, "Multisensor fusion of ladar and passive infrared imagery for target segmentation", *Optical Engineering*, 28(8), 1989, pp. 881-886.

[77] P. Ajjimarangsee and T.L. Huntsberger, "Neural network model for fusion of visible and infrared sensor outputs", in P.S. Schenker (Ed.), *Sensor Fusion, Spatial Reasoning and Scene Interpretation*, The International Society for Optical Engineering, 1003, SPIE, USA, 1988, pp. 153-160.

[78] Z. Zhang and R.S. Blum, "Region-based image fusion scheme for concealed weapon detection", *Proceedings of 31st Annual Conference on Information Sciences and Systems*, pp. 168-173.

[79] L.J. Chipman, Y.M. Orr and L.N. Graham, "Wavelets and image fusion", *Proc. Intl. Conf. Image Processing*, USA, 1995, pp. 248-251.

[80] O. Rockinger, "Image sequence fusion using a shift invariant wavelet transform", *Proc. Intl. Conf. on Image Processing*, 1997, pp. 288-291.

[81] P. Hill, N. Canagarajah and D. Bull, "Image Fusion using Complex Wavelets", *Proc. 13th British Machine Vision Conference*, United Kingdom, 2002, pp. 487-496.

[82] Y. Chibani and A. Houacine, "On the use of the redundant wavelet transform for multisensor image fusion", *Proc. IEEE Int. Conf. Electronics, Circuit and Systems*, 2000, pp. 442-445.

[83] A. Haar, "Zur theorie der orthogonalen funktionensysteme", *Mathematische Annalen*, 69, 1910, pp. 331371.

[84] N. Kingsbury, "A dual-tree complex wavelet transform with improved orthogonality and symmetry properties", *Proc. IEEE Conf. Image Processing*, 2000, pp. 375-378.

[85] F. Jin, P. Fieguth and L. Winger, "Image denoising using complex wavelets and Markov prior models", In M. Kamel and A. Campilho (Eds.), *Image Analysis and Recognition: Second International Conferece*, 2005.

[86] M.N. Do and M. Vetterli, "The finite ridgelet transform for image representation", *IEEE Transactions on Image Processing*, Vol. 12, No. 1, 2003, pp. 16-28.

[87] E. Candes and D. Donoho, "Curvelets a surprisingly effective nonadaptive representation for objects with edges." In: A. Cohen, C. Rabut and L. Schumaker, Eds., *Curves and Surface Fitting*, Vanderbilt University Press, 2000, pp. 105120.

[88] X. Xiaohong and N. Jiping, "Review of EMD-based image fusion", *Intl. Conf. on Intelligence Science and Information Engineering*, 2011, pp. 282-285.

[89] A.M. Waxman, D.A. Fay, A.N. Gove, M. Siebert, J.P. Racamoto, J.E. Carrick and E.D. Savoye, "Color night vision: fusion of intensified visible and thermal IR imagery", *Proc. SPIE Conference on Synthetic Vision for Vehicle Guidance and Control*, 2463, 1995, pp. 58-68.

[90] A.M. Waxman, A.N. Gove, D.A. fay, J.P. Racamoto, J.E. Carrick, M.C. Seibert and E.D. Savoye, "Color night vision: opponent processing in the fusion of visible and IR imagery", *Neural Networks 10*, 1997, pp. 1-6.

[91] G. Piella, "A region-based multiresolution image fusion algorithm", *ISIF Fusion Conference 2002*, Annapolis, 2002.

[92] J. Yang and R.S. Blum, "A statistical signal processing approach to image fusion for concealed weapon detection", *Proc. Intl. Conf. on Image Processing*, Vol. 1, 2002, pp. I-513 - I-516.

[93] R.K. Sharma, T.K. Leen and M. Pavel, "Probabilistic image sensor fusion", *Advances in Neural Information Processing Systems 11*, The MIT Press, 1999.

[94] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford University Press, 2006.

[95] T. Mertens, J. Kautz and F. Van Reeth, "Exposure Fusion", *Pacific Graphics*, 2007.

[96] W.W. Wang, P.L. Shui and X.C. Feng, "Variational models for fusion and denoising of multifocus images", *IEEE Sig. Proc. Letters*, 25, 2008, pp. 65-68.

[97] Z.G. Jiang, D.B. Han, J. Chen and X.K. Zhou, "A wavelet based algorithm for multi-focus micro-image fusion", *Proc. 3rd Intl. Conf. on Image and Graphics*, 2004.

[98] R. Li and Y.J. Zhang, "Level selection for multiscale fusion of out-of-focus image", *IEEE Signal Processing Letters*, Vol. 12, No. 9, 2005, pp. 617-620.

[99] G. Piella and H. Heijmans, "Multiresolution image fusion guided by a multimodal segmentation", *Proc. Advanced Concepts for Intelligent Vision Systems*, Belgium, 2002, pp. 1-8.

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

[101] D. Kundur, K. Su and D. Hatzinakos, "Digital video watermarking: techniques, technology and trends," in *Intelligent Watermarking Techniques*, Chapter 10, P. J.S. Pan, H.C. Huang and L. Jain, eds., World Scientific Publishing Company, 2004, pp. 265-314.

[102] D. Kundur and D. Hatzinakos, "Toward robust logo watermarking using multiresolution image fusion principles," *IEEE Transactions on Multimedia*, Vol. 6, No. 2, 2004, pp. 185-198.

[103] G. Simone, A. Farina, F.C. Morabito, S.B. Serpico and L. Bruzzone, "Image fusion techniques for remote sensing applications", *Information Fusion 3*, 2002, pp. 3-15.

[104] L. Alparone, L. Facheris, S. Baronti, A. Garzelli and F. Nencini, "Fusion of multispectral and SAR images by intensity modulation", *Proc. 7th Intl. Conf. on Information Fusion*, Sweden, 2004, pp. 637-643.

[105] U. Kumar, C. Mukhopadhyay and T.V. Ramachandra, "Fusion of multisensor data: review and comparative analysis", *Global Congress on Intelligent Systems*, 2009, pp. 418-422.

[106] T. Ranchin and L. Wald, "Fusion of high spatial images: the ARSIS concept and its implementation", *Photogrammetric Engineering & Remote Sensing*, Vol. 66, No. 1, 2000, pp. 49-61.

[107] W.J. Carper, T.M. Lillesand and R.W. Kieffer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and and multispectral image data",*Photogrammetric Engineering and Remote Sensing*, 56, 1990, pp. 459-467.

[108] J.G. Liu, "Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details", *Intl. J. Remote Sensing*, Vol. 21, No. 18, 2000, pp. 3461-3472.

[109] B. Garguet-Duport, J. Girel, J.M. Chassery and G. Pautou, "The use of multi-resolution analysis and wavelets transform for merging SPOT panchromatic and and multi-spectral image data", *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 9, 1996, pp. 1057-1066.

[110] J. Zhou, D.L. Civco and J.A. Silander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data", *Intl. J. Remote Sensing*, Vol. 19, No. 4, 1998, pp. 743-757.

[111] M.J. Gooding, K. Rajpoot, S. Mitchell, P. Chamberlain, S.H. Kennedy and J.A. Noble, "Investigation into the fusion of multiple 4-D fetal echocardiography images to improve image quality", *Ultrasound Med Biol.*, 36(6), 2010, pp. 957-966.

[112] O. Perez, M.A. Patricio, J. Garcia, J. Carbo and J.M. Molina, "Fusion of surveillance information for visual sensor networks", *Proc. Intnt. Conf. on Information Fusion*, Italy, pp. 1-8.

[113] L. Snidaro, G. Luca Foresti, R. Niu and P.K. Varshney, "Sensor fusion for video surveillance", *Proc. 7th Intl. Conf. on Information Fusion*, Sweden, 2004, pp. 739-746.

[114] D.D. Sworder, J.E. Boyd and G.A. Clapp, "Image fusion for tracking manouvering targets", *International Journal of System Science 28(1)*, 1997, pp. 1-14.

[115] Y. Zhou and R. Hecht-Nielsen, "Target recognition using multiple sensors", *Proc. IEEE Workshop: Neural Networks for Signal Processing*, 1993, pp. 411-420.

[116] A. Toet, "Fusion if visible and thermal imagery improves situation awareness", *Displays 18*, 1997, pp. 85-95.

[117] B.V. Dasarathy, "Fusion strategies for enhancing decision reliability in multi-sensor environments", *Optical Engineering 35*, 1996, pp. 603-616.

[118] D.P. Acharya and G. Panda, "A review of independent component analysis techniques and their applications", *IETE Technical Review*, Vol. 25, Issue 6, 2008.

[119] A.T. Benjamin, L. Ericksen, P. Jayawant and M. Shattuck, "Combinational Trigonometry with Chebyshev Polynomials", *Journal of Statistical Planning and Inference*, 2008.

[120] M. Alnasser and H. Foroosh, "Rendering synthetic objects in natural scenes", *IEEE International Conference on Image Processing*, 2006, pp. 493-496.

[121] H. Korizis, "Signal processing methods for the modelling and prediction of financial data", *PhD thesis*, Department of Electrical and Electronic Engineering, Imperial College London, 2008.

[122] J.C. Mason and D.C. Handscomb, "Chebyshev Polynomials", *Chapman & Hall* / CRC, Florida, 2003, pp. 105-141.

[123] N. Amthul, "Image fusion using two dimensional Chebyshev polynomials", *MSc Dissertation*, Imperial College London, 2009.

[124] K.B. Dunn and R. Lidl, "Generalizations of the classical Chebyshev polynomials to polynomials in two variables", *Czechoslovak Mathematical Journal*, Institute of Mathematics, Vol. 32, No. 4, 1982, pp. 516-528.

[125] T. Stathaki (Ed.), "Image Fusion: Algorithms and Applications", (*textbook*), *Academic Press*, 2008, pages 500.

[126] M. Cheong and K.S. Loke, "Textile recognition using Tchebichef moments of co-occurrence matrices", *Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues*, Springer Berlin / Heidelberg, Volume 5226/2008, 2008, pp. 1017-1024.

[127] N. Mitianoudis and T. Stathaki, "Pixel-based and region-based image fusion schemes using ICA bases", *Information Fusion*, Vol. 8, No. 2, 2007, pp. 131142.

[128] Z. Omar, N. Mitianoudis and T. Stathaki, "Region-based image fusion using a combinatory Chebyshev-ICA method", *Proc. Intl. Conf. on Acoustics, Speech and Signal Processing*, Prague, 2011, pp. 1213-1216.

[129] J.V. Stone, "Independent component analysis: an introduction", *TRENDS in Cognitive Sciences*, Vol. 6, No. 2, 2002, pp. 59-64.

[130] E. Oja, A. Hyvarinen and P. Hoyer, "Image feature extraction and denoising by sparse coding", *Pattern Analysis & Applications*, Vol. 2, 1999, pp. 104-110.

[131] A. Hyvärinen and P.O. Hoyer, "Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces", *Neural Computation*, Vol. 12, No. 7, 2000, pp. 1705-1720.

[132] A. Hyvarinen, "Fast and robust fixed-point algorithm for independent component analysis", *IEEE Transactions on Neural Networks*, Vol. 10, No. 3, 1999, pp. 626-634.

[133] A. Hyvrinen, P. Hoyer and E. Oja, "Image denoising by sparse code shrinkage". In S. Haykin and B. Kosko (eds), *Intelligent Signal Processing*, IEEE Press, 2001.

[134] N. Cvejic, J. Lewis, D. Bull and N. Canagarajah, "Region-based multimodal image fusion using ICA bases", *International Conference on Image Processing*, 2006, pp. 1801-1804.

[135] M. Petrou and P.D. Sevilla, "Image Processing: Dealing with Texture", Book, *Wiley*, 2006, pages 618.

[136] F. Nencini, A. Garzelli, S. Baronti and L. Alparone, "Remote sensing image fusion using the curvelet transform", *Information Fusion 8*, 2007, pp. 143-156.

[137] A. Toet, N. Schoumans and J.K. Uspeert, "Perceptual evaluation of different nighttime imaging modalities', *Proceedings of the 3rd International Conference on Information Fusion 3*, 2000, pp. TuD3.17-TuD3.23.

[138] C.S. Xydeas and V. Petrovic, "Objective image fusion performance measure", *Electronics Letters*, Vol. 36, No. 4, 2000, pp. 308-309.

[139] Z. Wang and A.C. Bovik,"A universal image quality index", *IEEE Signal Processing Letters*, Vol. 9, No. 3, 2002, pp. 81-84.

[140] T. Scheermesser and O. Bryngdahl, "Texture metric of halftone images", *J. Opt. Soc. Am.*, Vol. 13, No. 1, 1996, pp. 18-24.

[141] Yossi Rubner and Carlo Tomasi, "Texture Metrics", *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, 1998, pp. 4601-4607.

[142] M. Tuceryan and A.K. Jain, "Texture analysis", *The Handbook of Pattern Recognition and Computer Vision (2nd Edition)*, World Scientific Publishing Co., 1998, pp. 207-248.

[143] N. Damera-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans and A.C. Bovik, "Image quality assessment based on a degradation model", *IEEE Transactions on Image Processing*, Vol. 9, No. 4, 2000, pp. 636-650.

[144] C. Li and A.C. Bovik, "Content-partitioned structural similarity index for image quality assessment", *Signal Processing: Image Communication 25*, 2010, pp. 517-526.

[145] R.S. Blum, Z. Xue and Z. Zhang, "An overview of image fusion", In R.S. Blum and Z. Liu (Eds.), *Multi Sensor Image Fusion And Its Applications*, CRC Press, 2006, pp. 1-35.

[146] N. Cvejic, A. Loza, D. Bull and N. Canagarajah, "A novel metric for performance evaluation of image fusion algorithms", *Proceedings of World Academy of Science, Engineering and Technology*, Vol. 7, 2005, pp. 80-85.

[147] C. Ramesh and T. Ranjith, "Fusion performance measures and a lifting wavelet transform based algorithm for image fusion", *International Conference on Information Fusion*, 2002, pp. 317-320.

[148] Y. Chen and R.S. Blum, "A new automated quality assessment algorithm for night vision image fusion", *Conference on Information Sciences and Systems*, 2007, pp. 518-523.

[149] N. Cvejic, D.R. Bull and N. Canagarajah, "Metric for multimodal image sensor fusion", *Electronics Letters*, Vol. 43, No. 2, 2007.

[150] Q. Wang, Y. Shen, Y. Zhang and J.Q. Zhang, "A quantitative method for evaluating the performances of hyperspectral image fusion", *IEEE Transactions on Instrumentation and Measurement*, Vol. 52, No. 4, 2003, pp. 1041-1047.

[151] Q. Wang and Y. Shen, "Performances evaluation of image fusion techniques based on non-linear correlation measurement", *Proc. of IEEE Instrumentation and Measurement Technology Conference*, Vol. 1, 2004, pp. 472-475.

[152] D. Gadkari, "Image quality analysis using GLCM", *MSc Thesis, University of Central Florida*, 2004.

[153] Y. hu, C.X. Zhao and H.N. Wang, "Directional analysis of texture images using gray level co-occurrence matrix", *IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, 2008, pp. 277-281.

[154] R.M. Haralick, K. Shanmugam and I. Dinstein, "Textural features for image classification", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3, No. 6, 1973, pp. 610-621.

[155] M. Hall-Beyer, `http://www.fp.ucalgary.ca/mhallbey/the_glcm.htm`, *Accessed on 9/12/11*.

[156] M.M. Mokji and S.A.R. Abu Bakar, "Adaptive thresholding based on co-occurrence matrix edge information", *Proceedings of the First Asia International Conference on Modelling & Simulation*, 2007.

[157] X. Ran and N. Farvardin, "A perceptually motivated three-component image model-Part I: description of the model", *IEEE Transactions on Image Processing*, Vol. 4, No. 4, 1995, pp. 401-414.

[158] R. Pavlovic, V. Petrovic and B. Bondzulic, "Fusion of colour and monochromatic images with chromacity preservation", to appear in *Proc. Int. Conf. on Information Fusion*, 2012.

[159] V. Petrovic, B. Bondzulic and R. Pavlovic, "Study of objective evaluation of natural colour image fusion", to appear in *Proc. Int. Conf. on Information Fusion*, 2012.

[160] K.A. Kalpoma and J.-i. Kudoh, "Image fusion processing for IKONOS 1-m color imagery", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, No. 10, 2007, pp. 3075-3086.

[161] P.T. Yap and R. Paramesran, "Image analysis by Krawtchouk moments", *IEEE Transactions on Image Processing*, Vol. 12, No. 11, 2003, pp. 1367-1377.

[162] A. Haq, I. Gondal and M. Murshed, "Automated multi-sensor color video fusion for night-time video surveillance," *IEEE Symposium on Computers and Communications (ISCC)*, 2010, pp. 529-534.

[163] Y. Chen and C. Han, "Night-time pedestrian detection by visual-infrared video fusion," *7th World Congress on Intelligent Control and Automation*, 2008, pp. 5079-5084.

[164] T.H. Koornwinder, "Orthogonal polynomials in two variables which are eigenfunctions of two algebraically independent partial differential operators". I, Nederl. Akad. Wetensch. Proc. Ser. A 77 = Indag. Math. 36, 1974, pp. 48-58.

[165] L.J. Wojakowski, "Moments of measure orthogonalizing the 2-dimensional Chebyshev polynomials", *Banach Center Publications 73*, 2006, pp. 429-433.

[166] R. Mukundan, S.H. Ong and P.A. Lee, "Image analysis by Tchebichef moments", *IEEE Transactions on Image Processing*, Vol. 10 No. 9, 2001, pp. 1357-1364.

[167] A. Erdelyi et al., "Higher Transcendental Functions", New York: *McGraw-Hill*, Vol. 2, 1953.

[168] A.V. Nikiforov, S.K. Suslov and V.B. Uvarov, "Classical Orthogonal Polynomials of a Discrete Variable", Berlin: *Springer-Verlag*, 1991.

[169] R. Mukundan, "Some computational aspects of discrete orthonormal moments", *IEEE Transactions on Image Processing*, Vol. 13 No. 8, 2004, pp. 1055-1059.