# ROBUST DETECTION AND CLASSIFICATION OF BIOMEDICAL CELL SPECIMENS FROM LIGHT MICROSCOPE IMAGES

## SARAVANA KUMAR

(B.Eng. (Hons.), M.Eng., NUS)

### A THESIS SUBMITTED

# FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

### DEPARTMENT OF

#### ELECTRICAL AND COMPUTER ENGINEERING

#### NATIONAL UNIVERSITY OF SINGAPORE

 $\mathbf{2006}$ 

To Anita, With Love

## Acknowledgements

I would like to thank my supervisors Associate Professor Ong Sim Heng and Associate Professor Surendra Ranganath for their many suggestions and constant support.

I am also thankful to my co-supervisor, Dr. Chew Fook Tim from the Faculty of Science for offering his scientific expertise on the identification of air-borne allergens and for providing the resources to effectively undertake this project.

I would like to acknowledge Dr. Kevin Tan from the Faculty of Medicine for his collaboration on our successful development of a program for detecting and classifying malaria infection in humans and rodents.

I am grateful to Dr. Ong Tan Ching for her invaluable help and patience throughout the course of my research.

Francis, the lab officer at the Vision and Image Processing Lab is thanked for being so accommodating and helpful.

I wish to thank my friend and colleague, Subramanian Ramanathan for all

those inspiring talks we had during lunch and tea.

I am grateful to my parents for their patience and *love*. Without them this work would never have come into existence (literally).

Lastly but certainly not the least, I am forever grateful to Anita, my fiancée, for loving me and believing in me.

## Summary

Automated identification of biomedical specimens such as malaria parasites from red blood cells would enable the undertaking of timely preventive measures which could potentially save millions of lives.

However, current automated systems lack robustness as they only work well under fixed operating conditions of the microscope, such as the choice of objective lens, aperture size, z-focus and intensity, but perform poorly when one or more of these settings change. Clumping of cells, when placed on slides, also adversely affects the system accuracy since the entire clump may be erroneously considered as a single specimen.

A robust scheme is developed for automatically identifying biomedical specimens from light microscope images. Contributions are made to the areas of edge detection, segmentation and classification.

A novel edge detection method is proposed which, unlike existing methods,

accurately identifies regions of interest (ROI) in the images under different luminance, contrast and noise levels. This is achieved by developing a new edge similarity measure that incorporates a regularization term. Directional finite impulse response (FIR) hyperbolic tangent (HBT) filters are also proposed as edge detectors and Chapter 2 shows that they achieve better noise tolerance and edge localization compared to Canny's Gaussian first derivative (GFD) filter.

A novel multi-scale edge detection method is proposed which ensures accurate detection of edges under noisy conditions. It is henceforth called the multi-scale *min-product* method (MMPM) as it uses a point-wise operation involving the *min* and *product* operators, in that sequence, to accurately detect step edges while significantly reducing false edges due to noise. Unlike existing multi-scale methods, a wider range of edge filters can be applied in MMPM. The problem of edge drift over successive scales is also avoided by directly applying edge filters of multiple widths on the original image.

The boundary edges enable the identification of the ROIs but each ROI may be a clump comprising two or more specimens. Therefore, a novel binary clump splitting method using is developed using a set of concavity-based rules to accurately split each clump into constituent specimens. The proposed method accurately splits clumps with specimens of diverse sizes and shapes at different degrees of overlap.

A novel texture classification method is presented that is invariant to specimen orientation, scale and contrast. Orientation invariance is achieved by expressing each specimen in an alternate Cartesian space defined by the major and minor axes of the largest ellipse within the specimen. Scale invariance is achieved by mapping the elliptical regions of arbitrary size, to a fixed unit circular region from which a polar map is subsequently constructed.

Edge maps are then extracted from the polar map by applying the edge similarity measure proposed in chapter 2 so that the resultant texture features obtained from these maps are invariant to contrast. The texture features comprise both *local* and *global* norm-1 energy measures since they enable improved classification accuracy.

The techniques proposed in this thesis are validated through experiments and compared against existing methods. They have been successfully applied to light microscope images of airborne spores and cytological specimens. The robustness of the edge detection techniques is also shown by successfully testing them on natural and magnetic resonance (MR) images.

## Contents

A	cknov	vledgements i	ii			
Su	imma	ary	v			
Li	st of	Tables xi	v			
Li	st of	Figures xv	7 <b>i</b>			
Li	st of	Acronyms xxi	ii			
1	Intr	Introduction				
	1.1	Motivations	1			
	1.2	System Overview	2			
	1.3	Limitations of Current Methods	3			
		1.3.1 Staining and fluorescence microscopy	3			
		1.3.2 Contrast and luminance	4			

		1.3.3	Clumping of specimens	5
		1.3.4	Orientation and scale	6
		1.3.5	Noise	8
	1.4	Objec	tives	9
	1.5	Thesis	Contributions	10
		1.5.1	Edge detection: Regularized similarity measure from hyper-	
			bolic tangent filters with finite impulse response $\ldots \ldots$	10
		1.5.2	Edge detection: Multi-scale min-product method $\ldots \ldots$	10
		1.5.3	Robust rule-based approach to clump splitting	11
		1.5.4	Texture classification: Local and global energy measures from	
			non-linear polar map filtering	11
	1.0	Thee	Organization	10
	1.0	1 nesis		12
2	1.0 A L	umina	nce and Contrast-Invariant Edge-Similarity Measure	12 14
2	1.6 <b>A</b> L 2.1	Ratior	nce and Contrast-Invariant Edge-Similarity Measure	12 14 14
2	<ul> <li><b>A</b> L</li> <li>2.1</li> <li>2.2</li> </ul>	Ratior Classic	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16
2	<ul> <li><b>A</b> L</li> <li>2.1</li> <li>2.2</li> <li>2.3</li> </ul>	Thesis Jumina Ratior Classic Edge 1	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16 18
2	<ul> <li><b>A</b> L</li> <li>2.1</li> <li>2.2</li> <li>2.3</li> </ul>	Ratior Classic Edge 1 2.3.1	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16 18 19
2	<ul> <li><b>A</b> L</li> <li>2.1</li> <li>2.2</li> <li>2.3</li> </ul>	Aumina Ratior Classic Edge 1 2.3.1 2.3.2	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16 18 19 21
2	<ul> <li><b>A</b> L</li> <li>2.1</li> <li>2.2</li> <li>2.3</li> </ul>	Aumina Ration Classic Edge 1 2.3.1 2.3.2 2.3.3	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16 18 19 21 22
2	<ul> <li><b>A</b> L</li> <li>2.1</li> <li>2.2</li> <li>2.3</li> </ul>	Thesis Jumina Ration Classic Edge 1 2.3.1 2.3.2 2.3.3 2.3.4	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16 18 19 21 22 23
2	<ul> <li>1.6</li> <li>A L</li> <li>2.1</li> <li>2.2</li> <li>2.3</li> </ul>	Thesis Jumina Ration Classic Edge 1 2.3.1 2.3.2 2.3.3 2.3.4 Edge 1	nce and Contrast-Invariant Edge-Similarity Measure         nale	12 14 14 16 18 19 21 22 23 25

		2.5.1	Uneven illumination	29
		2.5.2	Contrast variation	30
		2.5.3	Noise	31
		2.5.4	Edge localization	33
	2.6	Conclu	ision	36
3	Ster	o Edge	Detection via a Multi-Scale Min-Product Method	37
	3.1	Ration	nale	38
	3.2	Multi-	Scale Min-Product Method	40
		3.2.1	Defining multi-scale edge filters	41
		3.2.2	Implementation of MMPM algorithm	43
	3.3	Multi-	Scale Edge Detection Criteria	47
		3.3.1	Multi-scale SNR, M-SNR	48
		3.3.2	Multi-scale Localization, ML	48
		3.3.3	Multi-scale false edge responses, MFER	49
	3.4	Experi	ments	49
		3.4.1	M-SNR performance	50
		3.4.2	ML performance	51
		3.4.3	MFER performance	54
		3.4.4	Overall performance	54
	3.5	Discus	sion	59
	3.6	Conclu	nsion	61
4	A R	ule-Ba	ased Approach for Robust Clump Splitting	62

4.1	Ratio	nale	62
4.2	Review	w of Concavity Analysis for Clump Splitting	65
	4.2.1	Detection of concavity regions or concavity pixels	65
	4.2.2	Detection of candidate split lines	66
	4.2.3	Selection of best split line	67
4.3	Overv	iew of Methodology	68
4.4	Detect	ting Candidate Split Lines	69
	4.4.1	Concavity depth	70
	4.4.2	Saliency	70
	4.4.3	Alignment	71
	4.4.4	Concavity angle and concavity ratio	73
4.5	Select	ing the Best Split Line	74
4.6	Metho	odology	75
	4.6.1	Training	76
	4.6.2	Implementation of clump splitting	79
4.7	Perfor	rmance on Unseen Data	80
4.8	Perfor	rmance Comparison and Feature Validation	84
	4.8.1	Comparison I	85
	4.8.2	Comparison II	86
	4.8.3	Comparison III	87
4.9	Concl	usion	87

<b>5</b>	Inva	ariant	Texture Classification via Non-Linear Polar Map Filter-	-
	ing			89
	5.1	Ratio	nale	90
	5.2	Stand	ard Polar Map Transform	93
	5.3	Overv	riew of Method	95
	5.4	Identi	fying Elliptical Region	95
	5.5	Orient	tation- and Scale- Invariant Polar Map	97
	5.6	Contr	ast and Luminance Invariant Filter Output	100
	5.7	Local	and Global Energy Measures	103
		5.7.1	Global energy measures, <b>GEM</b>	105
		5.7.2	Normalized global energy measures, <b>NGEM</b>	105
		5.7.3	Local energy measures, $\mathbf{LEM}$	105
	5.8	Exper	imental Results	105
		5.8.1	Texture classification via support vector machines (SVM)	106
		5.8.2	Contrast invariance	112
		5.8.3	Orientation invariance	114
		5.8.4	Scale invariance	114
		5.8.5	Variation of feature extraction area	116
		5.8.6	Validation of energy measures	116
		5.8.7	Variation of regularization parameter	117
	5.9	Discus	ssion	120
	5.10	Concl	usion	121

6	Conclusions		
	6.1	Summary of Contributions	122
	6.2	Future Directions of Research	125
Re	efere	nces	127

# List of Tables

2.1	Optimal $\sigma_W$ , $W = 1$ , 2 and 3 for standard images from USC-SIPI	
	Image Database.	24
2.2	Influence of HBT filter $\sigma_2$ on the noise response	26
2.3	Influence of parameter $c$ on the noise response $\ldots \ldots \ldots \ldots \ldots$	27
2.4	Quantitative performance of edge detection with noise	33
3.1	Comparison of the extent of drift in local maxima (in pixels) be-	
	tween (a) MWPM non-decimated wavelet transform scheme and (b)	
	proposed MMPM	52
4.1	Threshold values assigned to the features that determine validity of	
	split lines	77
4.2	Training results for different penalty factor values	78
4.3	Detailed performance of RBA	83

4.4	Summary of performance comparison and feature validation results.	86
5.1	Overall classification percentage of polynomial SVM for a range of	
	$\lambda, d$ and PF	110
5.2	Overall classification percentage of RBF SVM for a range of $\sigma$ and	
	PF	111
5.3	Confusion matrix of classifying test set using RBF SVM	112
5.4	Overall classification percentage for different combination of energy	
	measures	118

# List of Figures

1.1	Block diagram of automated system.	2
1.2	Overview of image analysis software for robust detection and classi-	
	fication of biomedical cell specimens from light microscope images	12
2.1	Least-squares estimates of PCA eigenvectors using FIR HBT filters.	
	(a) and (c): PCA eigenvectors of second and third largest eigen-	
	values, (b) and (d): Corresponding least-squares estimates using a	
	linear combination of FIR HBT filters	20
2.2	Spatial and frequency properties of HBT filter. (a) 1–D continuous	
	spatial profile of HBT filters $f_2$ ( $\sigma = 0.5$ (dark gray), 1.0 (medium	
	gray) and 2.0 (light gray)). (b) Frequency responses of 1–D discrete	
	FIR filters after normalization by their respective maximum values	
	$(\sigma = 0.5 \ (dark \ gray), \ 1.0 \ (medium \ gray) \ and \ 2.0 \ (light \ gray)) \ and$	
	Gaussian filter ( $s = 1.0$ (dashed lines))	22

. 1

DI

۰.

0

2.3	Influence of $\sigma_W$ on (a) $\varepsilon_{total}$ and (b) $C_W$ for $W = 1$ (light gray), 2	
	(medium gray) and 3 (dark gray). $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$	23
2.4	Comparison between the $P_i$ and $R_i$ measures. (a) Lena image with	
	illumination gradient. (b) Similarity map, $P_i$ measure. (c) Similarity	
	map, $R_i$ measure	29
2.5	Comparison between the $\hat{C}_i$ and $R_i$ measures. (a) Light microscope image	
	of infected red blood cells. (b) Similarity map, measure $\hat{C}_i$ . (c) Similarity	
	map, measure $R_i$ . (d) Edge map, measure $\hat{C}_i$ . (e) Edge map, measure $R_i$ .	30
2.6	Comparison between the FIR HBT and GFD filters on noisy images. (a)	
	Outdoor scene. (b) Similarity map, GFD filter. (c) Similarity map, FIR	
	HBT filter. (d) Edge map, GFD filter. (e) Edge map, HBT filter	32
2.7	Edge localization comparison of the proposed and PC methods. (a) Syn-	
	thetic image. (b) Edge map from proposed method. (c) Edge map from	
	PC method	34
2.8	Edge localization comparison between the proposed and PC methods.	
	(a) Lena image. (b) Similarity map from proposed method. (c) PC map.	
	(d) Edge map from proposed method. (e) Edge map from PC method	35
3.1	Pseudo code for the proposed MMPM scheme	45
3.2	The noisy step signal $I$ with the corresponding similarity signals	
	$G_{1\rightarrow 3}$ , composite similarity signals $CS_{1\rightarrow 3}$ and gradient magnitude	
	signal $\nabla I$ : (a) $I$ . (b) $G_1$ . (c) $G_2$ . (d) $G_3$ . (e) $CS_1$ . (f) $CS_2$ . (g) $CS_3$ .	
	(e) $\nabla I$	46
3.3	The M-SNR performance of MMPM for different $J. \ldots \ldots$	50

3.4	Comparison between the M-SNR performances of MMPM and MPM	
	for different input SNR levels	51
3.5	Localization measure ML of the MMPM for four different filters	
	under different $J$	53
3.6	Localization measure ML of the MMPM for four different filters	
	under different levels of noise.	53
3.7	Multiple false edge response measure MFER of the MMPM for four	
	different filters under different $J$	54
3.8	A noiseless synthetic image of a white rectangular box on a black	
	background.	55
3.9	${\cal F}$ for MMPM edge map, obtained using GFD, as a function of input	
	SNR at different $J$	56
3.10	${\cal F}$ for MMPM edge map as a function of input SNR for the four	
	filters with $J = 10. \dots \dots$	56
3.11	MMPM 2-D edge map as a function of $J$ (a) MR image from an	
	axial head scan (15 dB). MMPM edge map for (b) $J = 1$ . (c) $J = 5$ .	
	(d) $J = 10.$	57
3.12	MMPM edge maps for MR image from Fig. 3.11(a) at $J = 5$ and	
	filter (a) GFD. (b) DOB. (c) HBT. (d) RMP.	58
3.13	Comparison between edge detection results of MR image from Fig.	
	3.11(a) using GFD filter for (a) fixed scale with $W_n = 2, \sigma = 2$ .	
	(b) fixed scale with $W_n = 1$ , $\sigma = 0.3$ . (c) multi-scale with $J = 3$	
	(combining scales from $W_n = 1 \rightarrow 6$ )	59

4.1	Binary clump with convex hull chords $K_1$ , $K_2$ and $K_3$ and corre-	
	sponding boundary arcs, $B_1$ , $B_2$ and $B_3$	66
4.2	Wang's opposite alignment criterion (from Ref. [91]). $\ldots$	67
4.3	Feature space of length of split line vs. concaveness. Dashed line -	
	decision boundary obtained by using two separate thresholds. $Solid$	
	<i>line</i> - correct decision boundary.	69
4.4	Binary clump with concavity pixels, $CV_1$ and $CV_2$ , and correspond-	
	ing concavity depths, $CD_1$ and $CD_2$	70
4.5	Alignment (a) Clump comprising three overlapping specimens. (b)	
	Concavity-concavity alignment, CC and concavity-line alignment, CL.	71
4.6	Concavity angle, CA and concavity ratio, CR	73
4.7	Five species of airborne spore specimens used in the experiments.	76
4.8	Linear decision boundary obtained from the training data set	78
4.9	Sample results of splitting clumps comprising two touching spore	
	specimens (not to scale).	80
4.10	Sample results of splitting clumps comprising two or three touching	
	cytological specimens.	81
4.11	Splitting a clump comprising only one dominant concavity region.	
	(a) Two overlapped <i>Dreschlera</i> specimens. (b) Split line joining the	
	concavity pixel and a boundary pixel	81
4.12	Split results of large clumps comprising several specimens. (a) Nephrole	pis
	clump. (b) $Nephrolepis$ clump after splitting. (c) Two large $Podocar$ -	
	pus clumps. (d) Podocarpus clumps after splitting.	82

4.13	Splitting clumps comprising specimens with different sizes and shapes.	
	(a) Fungal and fern spore. (b) <i>Nephrolepis</i> with attached dirt particle.	82
4.14	False splitting of a clump comprising two <i>Dreschlera</i> specimens	
	crossing each other	84
4.15	Reduction in the sizes of the concavity regions in a <i>Curvularia</i>	
	clump. (a) Overlapping and individual $Curvularia$ specimens. (b)	
	Binary clump of <i>Curvularia</i> specimens after dilation/erosion opera-	
	tion	84
4.16	Shortcomings of concavity measure in ODM. (a) Two overlapping	
	Curvularia specimens; concavity region $S_a$ is not detected. (b)	
	Curvularia specimen with overlapping detritus; concavity region $\mathcal{S}_b$	
	is not detected.	86
4.17	False splitting of a <i>Podocarpus</i> specimen using ODM	87
4.18	Splitting a clump comprising three <i>Curvularia</i> specimens that over-	
	lap along their major axes. (a) False splitting when saliency and	
	alignment conditions are removed. (b) Correct splitting when saliency	
	and alignment conditions are imposed	88
5.1	Transformation of largest circle within textured image $I(x, y)$ to	
	polar map $p(\alpha, r)$	93

5.2	Two step process of finding the largest ellipse within a segmented
	specimen. (a) Determining the ellipse eccentricity from parameters
	$\hat{a}_1, \hat{a}_2$ and $\hat{a}_3$ . (b) Ensuring that the ellipse completely fits within
	the specimen by adjusting its translation and size via parameters
	$\hat{a}_4,  \hat{a}_5 \text{ and } \hat{a}_6,  \ldots  \ldots  \ldots  \ldots  \ldots  \ldots  \ldots  \ldots  \ldots  $
5.3	Ellipse redefined in $x'y'$ Cartesian space and centered at the origin. 98
5.4	Ellipse expressed as a unit circle in the $(a_1u, a_2v)$ Cartesian space. 99
5.5	Influence of affine transformation on polar map. (a)–(c)—Images
	captured under (a) 40×. (b) 60×. (c) 20× objective magnification.
	(d)–(f)—Corresponding polar maps for images (a)–(c). (g)—Image
	(a) rotated by $45^{\circ}$ counter-clockwise. (h)—Corresponding polar map
	of image (g)
5.6	Influence of contrast variation on linear and non-linear filtering out-
	put. (a)-(c)—Images captured under progressively increasing lu-
	minance and contrast. (d)-(f)—Corresponding magnitude of filter
	output, $ C_i $ . (g)-(i)—Corresponding magnitude of filter output, $ R_i $ . 103
5.7	Distribution of local energy features. (a) Elliptical area divided into
	six localized regions. (b) Corresponding six rectangular regions of
	equal area: $A_1$ to $A_6$ in polar map $p$

5.8	Sample images of different species used in the proposed work. (a)	
	Nephrolepis auriculata, NEBI (95 $\mu$ m $\times$ 75 $\mu$ m). (b) Stenochlaena palus-	
	tris, STPA (122 $\mu$ m×85 $\mu$ m). (c) Sorghum halepensis, SOHA (115 $\mu$ m×11	$5\mu m$ ).
	(d) Acacia auriculiformis, ACAU (93 $\mu$ m×84 $\mu$ m). (e) Curvularia	
	brachyspora, CUBR ( $34\mu m \times 50\mu m$ ). (f) Pithomyces maydicus, PIMA	
	$(45\mu \mathrm{m} \times 82\mu \mathrm{m})$	.07
5.9	Comparison chart of the classification percentage for the individual	
	classes	.12
5.10	Overall percentage of individual classes for different contrast stretch-	
	ing factors	.13
5.11	Comparison of overall classification accuracy attributed to non-linear	
	and linear filtering methods for different contrast factors 1	.14
5.12	Overall percentage of individual classes for different orientations 1	.15
5.13	Overall percentage of individual classes for different scales 1	.15
5.14	Overall percentage for different feature extraction areas 1	.17
5.15	Overall classification accuracy of noisy test images for different reg-	
	ularization values	.19
5.16	Overall classification accuracy of test images linearly stretched by	
	contrast factor = 0.2 for different regularization values. $\dots \dots \dots \dots$ 1	.19

# List of Acronyms

1–D	one dimensional
2–D	two dimensional
ACAU	Acacia auriculiformis
AN	angle-based
AS	accurate splitting
CA	concavity angle
CC	concavity-concavity alignment
CCD	charge-coupled device
CD	concavity depth
$\operatorname{CL}$	concavity-line alignment
CR	concavity ratio
CS	composite sub-band
CUBR	Curvularia brachyspora
CV	concavity pixel

DG	concavity degree
DOB	difference of box
DWT	discrete wavelet transform
ED	edge detection
EL	edge localization
FIR	finite impulse response
$\mathbf{FS}$	false splitting
GEM	global energy measures
GFD	Gaussian first derivative
GLCM	gray level co-occurrence matrix
GM	gradient magnitude
HBT	hyperbolic tangent
IIR	infinite impulse response
LEM	local energy measures
LM	light microscope
M-SNR	multi-scale signal-to-noise ratio
MAD	median absolute deviation
MFER	multi-scale false edge responses
ML	multi-scale localization
MMPM	multi-scale min-product based method
MPM	multi-scale product based method
MR	magnetic resonance
MWPM	multi-scale wavelet product based method
NEBI	Nephrolepis auriculata

NGEM	normalized global energy measures
ODM	optimal dissection method
PC	phase congruency
PCA	principal component analysis
$\mathbf{PF}$	penalty factor
PIMA	Pithomyces maydicus
$\mathbf{PS}$	multi-scale product of the first three scales
RBA	rule-based approach
RBF	radial basis function
RGB	red, green and blue
RMP	ramp
ROI	region of interest
SA	saliency
SEM	scanning electron microscope
SF	suppression of false edge responses
SNR	signal-to-noise ratio
SOHA	Sorghum halepensis
STPA	Stenochlaena palustris
SVM	support vector machine
TH	threshold
US	under splitting
USC-SIPI	University of Southern California, Signal and Image
	Processing Institute
WT	normalized concavity weight

## Chapter

## Introduction

### 1.1 Motivations

Fast and accurate identification of biomedical specimens from light microscope (LM) images is an essential step in a wide variety of application domains where the specimens of interest could be asthma-causing allergenic spores [6, 15, 23, 29, 54, 58, 78, 85] or malaria infected red blood cells [82, 87] among others [27]. An early assessment of these specimens enables us to undertake preventive measures which could potentially save millions of lives. The practice of identifying specimens of interest from microscope images even extends to non-biological samples such as the detection of defects in wafers and the analysis of gun shot residues [6].

Manual methods of detecting and characterizing biomedical cell specimens from microscope images can be time consuming due to the large amount of data involved. For example, approximately 5 000 to 50 000 red blood cells need to be inspected for the presence of malaria parasite in order to determine the extent of infection with sufficient accuracy. Similarly, it takes about four to five hours for an experienced technician to determine the total number of allergenic spores on a single microscope slide. The results obtained from manual methods are also inconsistent as they depend on the person's experience and state of mind.

The need for fast and reliable analysis necessitates the development of reliable automated methods for identifying biomedical specimens. It also reduces the need for manpower and enables research personnel to focus on more critical areas of research such as analyzing the output results from the automated system. These results can be generated in large quantities and stored in an image or data file format to be re-examined by different scientists as a form of quality control.

### 1.2 System Overview

The system comprises the image analysis software, 3–axis motorized microscope and an image acquisition module comprising a  $570 \times 760$  3–CCD color video camera and frame grabber as shown in Fig. 1.1. The image analysis software represents the brains of the entire system as it controls the image acquisition and motorized motion of the slides apart from its central role of detecting and characterizing the biomedical specimens.



Figure 1.1: Block diagram of automated system.

The motorized stage is automated and enables the movement of the microscope stage along the x and y axis as well as the vertical z-focus setting. The image analysis software controls the motion of the motorized stage via the stage control unit. The software reads the x, y and z settings of the motorized stage via the stage control and then instructs the stage control to move the stage to a new x, y and zsetting. More importantly, it obtains digitized images from the frame grabber and subsequently processes these images in order to generate the output results.

The processing work basically entails the segmentation of the biomedical specimens from the images followed by the classification of each specimen into its corresponding group based on the specimen features. The definition of the term "group" depends on the problem domain. For example, it denotes the specimen genus/species for spore images or the stage of infection of the specimens for images of malaria infected red blood cells.

### **1.3** Limitations of Current Methods

Efforts to implement automated systems have not been successful since they lack robustness. Existing methods work well under fixed operating conditions of the microscope such as the choice of objective lens, aperture size, z-focus and intensity but perform poorly when one or more of these settings change.

#### **1.3.1** Staining and fluorescence microscopy

Fluorescence microscopy has been used to detect specimens of interest which fluoresce in contrast to the background [27, 78]. However, a shortcoming of this approach is that it does not segment dead specimens since they lack an enzyme required for fluorescing. Staining has also been applied in order to improve the contrast of specimens of interest in the digitized images [1, 23]. However, these methods are only able to discriminate a specific type or family of specimens from the entire range studied.

#### 1.3.2 Contrast and luminance

Automated intensity threshold methods that detect foreground specimens from the background image using fixed threshold values, are sensitive to luminance [6]. The aforementioned methods fail when the image luminance varies and this can be easily caused by a small adjustment to the voltage setting of the lamp since both voltage and luminance share a power law relationship [39]. An increase in voltage results in higher image luminance and vice versa. A reduction in luminance is also observed due to deterioration in the light source over time where the light intensity remains more or less constant over an operation time of 12 hours [6].

It is also observed that a reduction in luminance, due to the aforementioned factors, also causes a decrease in image contrast, which is defined as the difference in luminance between the light and dark areas in an image [34]. This is due to the narrowing of the dynamic range in gray level values of the microscope image. Therefore it is impractical to expect a constant contrast especially when using different microscope systems. However, current cell segmentation methods, based on edge detection, are sensitive to image contrast since the underlying inner products between a predefined edge filter and the local neighborhoods in an image, emphasizes stronger edges and suppresses weaker ones. Frei and Chen [31] have proposed a contrast invariant method for detecting edges. It is termed an angle-based (AN) method since it is based on the computation of the cosine of the projection angles between local neighborhoods and pre-defined edge filters. A problem with this technique is its sensitivity to luminance since it inhibits edges in regions of low luminance or, conversely, enhance them.

Current texture classification methods [89, 98, 57, 58] using filtering methods such as Laws' [55] and wavelet decomposition [16, 62] are sensitive to luminance and contrast since (1) features extracted from the low frequency (approximation) sub-band of these methods contain the luminance information. (2) as in the case of current edge detection methods, the underlying spatial convolution operation, in filtering methods, emphasizes texture patterns of stronger contrast and suppresses those of weaker contrast. Methods [89, 54, 57, 58] based on the gray level cooccurrence matrix (GLCM) [35] are also sensitive to luminance and contrast since the matrix carries this information in the form of co-occurrences between pairs of gray levels a displacement d apart.

#### **1.3.3** Clumping of specimens

Clumping together of specimens in the slide sample also adversely affects the system accuracy since the entire clump may be erroneously segmented as a single specimen. This poses a problem if the aim is to accurately label the constituent specimens in every clump. Various methods such as binary erosion [2, 69, 86, 88], watershed [5], model based [14, 26, 40, 94] and concavity analysis [8, 9, 21, 41, 59, 91, 93] have been applied to split such clumps into the constituent specimens but they all suffer from specific shortcomings.

Erosion-based methods [2, 69, 86, 88] may completely erode a constituent specimen in a clump before a split occurs. Watershed techniques [5] tend to oversplit clumps. Model-based methods [14, 26, 40, 94] are computationally expensive and require initialization of the model parameters. Concavity analysis methods [8, 9, 21, 41, 59, 91, 93] offer an intuitive way of clump splitting and have been applied to the examination of cervical cancer cells [93], plant cells [21], chromosomes [59], and crushed aggregates [91], to name a few. However, tests conducted by Wang [91] and experimental results presented in Section 4.8 of this thesis show that these methods are ad hoc and applicable for objects of specific sizes and shapes.

#### **1.3.4** Orientation and scale

Existing methods are based on explicit or implicit assumption that the microscope images are acquired at the same scale and that the specimens have the same orientation. The scale of microscope images varies depending on the choice of the objective lenses used where each magnification ratio, i.e.,  $10 \times$ ,  $20 \times$ ,  $40 \times$  and  $60 \times$  corresponds to a particular scale. The specimens are also oriented in an arbitrary fashion when viewed under a LM. García-Sevilla [33] has shown that the classification accuracy of features extracted from classical methods such as the gray level co-occurrence matrix (GLCM) [35] and wavelet transform [13] are sensitive to scale.

Various methods have been proposed to reduce the sensitivity of analysis to orientation and scale. Combining the detail sub-bands in wavelet decomposition [73, 98] or using a set of rotated wavelet filters and multi-channel Gabor filters [28] are attempts to reduce orientation sensitivity but the performance of these methods degrades when the number of texture classes/groups increases since they are derived from standard filtering methods which are sensitive to orientation. Muneeswaran et al. [66] exploited the scale invariance property of fractal analysis to characterize textural regions. However, empirical studies show that the fractal dimension is often different at different scales of natural textures, although it may be constant for a range of scales [11]. Circular auto-regressive [44] and the logpolar Gabor filters [56] are computationally intensive especially when the number of classes or size of textural regions increases. More recently, a method combining log-polar transform and shift invariant wavelet packet transform reported by Pun and Lee [76] gave promising results when tested on a set of 25 distinct Brodatz textures with different scale and orientation [10].

The studies mentioned above used rectangular sample regions. Similarly, Langford et al. [54] identified pollen specimens from scanning electron microscope (SEM) images by selecting a rectangular region of approximately 10% of the entire pollen area. Such a small region was representative of the textural pattern since it was manually selected but this is not the case for an automated texture classification scheme where *a priori* information is not available. The use of rectangular sample regions may not be the best choice for biomedical cell specimens such as air-borne spores [52] and red blood cells where most cells can be approximated by a general elliptical form with a suitable choice of eccentricity and size.

#### 1.3.5 Noise

The presence of noise introduced during image acquisition adversely affects the segmentation of cells via edge detection. Image smoothing has been used as a preprocessing step [12, 63] to reduce noise but this can sometimes lead to excessive blurring such that weak edges go undetected [53].

Multi-scale edge detection methods [60, 62, 68, 80, 83, 96] promise accurate detection of edges for a range of scales despite noisy conditions. Rosenfeld et al. pioneered this effort by demonstrating that edges can be enhanced while suppressing noise by taking the direct point-wise products of the image sub-band decompositions [80]. Mallat et al. extended this idea by distinguishing edges from noise and characterizing various edge profiles from the Lipschitz regularity of these edges across scale space [60, 62].

Several other methods have also been developed for detecting edges based on their scale space behavior in the wavelet domain [4, 68, 83, 92, 96, 97]. These methods will henceforth be called the multi-scale wavelet product based method or MWPM since they involve the direct point-wise multiplication of wavelet coefficients at several adjacent scales. Xu et al. [92] applied MWPM to filter noise from images. Subsequently, Sadler and Swami applied this method to step edge detection [83] while Zhang et al. [96, 97] imposed an adaptive threshold on the point-wise products of the wavelet coefficients in order to identify important edge features.

However, MWPM results in the drift of edge maxima from the finer to coarser scales when the low pass filter as used in Mallat's wavelet decomposition method, has an even number of coefficients. This adversely affects the detection of specimen boundaries. Current multi-scale methods resort to the "band-aid" solution of restricting the product operation to the first two or three sub-band decomposition levels. Another drawback of MWPM is that the choice of edge detection filter is restricted to the quadratic spline filter.

#### 1.4 Objectives

The primary objective of this thesis is the development of robust methods for the detection and classification of biomedical specimens from LM images. The methods are to be robust with regards to the following aspects:

- 1. Luminance and Contrast
- 2. Noise
- 3. Clumping
- 4. Orientation
- 5. Scale
- 6.  $Focus^1$

<sup>1</sup>The thesis is based on the premise that the LM images are captured under optimal focus setting prior to the detection and classification of the biomedical cell specimens.

#### **1.5** Thesis Contributions

With the aim of developing robust cell detection and classification methods, key contributions are made in the following areas.

## 1.5.1 Edge detection: Regularized similarity measure from hyperbolic tangent filters with finite impulse response

A novel edge similarity measure is proposed for detecting cell boundaries [47]. It is robust under different luminance and contrast levels and incorporates a regularization term which offers a good compromise between contrast invariance and noise suppression. Hyperbolic tangent (HBT) filters with finite impulse response (FIR) [47, 48] are also proposed as edge detectors as they give better noise tolerance and edge localization for narrow filter widths compared to Canny's Gaussian first derivative (GFD) [12]. The proposed method also shows better edge localization compared to the phase congruency (PC) [46] method.

#### 1.5.2 Edge detection: Multi-scale min-product method

The multi-scale min-product method (MMPM) is proposed as it yields accurate boundary detection in the presence of noise. Unlike existing multi-scale methods, a wider range of edge filters can be used in MMPM. The edge drift problem over successive scales is avoided by directly applying edge filters of multiple widths to the original image. Canny's criteria on edge detection performance are also effectively extended, from its traditional definition in the fixed scale domain to the multi-scale domain. This multi-scale criteria enables us to objective evaluate filter performance in the multi-scale domain. It will be shown that the proposed MMPM method gives a better overall edge detection performance compared to the classical multi-scale product method (MPM). In addition, the superior signal to noise ratio (SNR) performances of the ramp (RMP) and HBT [47, 48] filters over the difference of box (DOB) [75] and GFD [12] filters are also reported in this thesis.

#### **1.5.3** Robust rule-based approach to clump splitting

Detected cells may overlap with one another to form clumps. A robust rule-based approach (RBA) to clump splitting is proposed [50, 51]. The novel concavity-based rule set accurately splits each clump into the constituent cells. The rule set ensures that (1) valid concavities are effectively distinguished from minor boundary irregularities, (2) concavity regions at the ends of split lines are suitable oriented with respect to each other and (3) false splitting of objects with natural concavities is significantly reduced. It is shown that, unlike current concavity analysis methods, RBA accurately splits objects of diverse sizes, shapes and extent of overlap. Experimental results show that the proposed approach is more robust and accurate compared to classical concavity analysis methods [8, 9, 21, 41, 59, 91, 93].

## 1.5.4 Texture classification: Local and global energy measures from non-linear polar map filtering

A novel texture classification routine that is invariant to cell orientation, scale and contrast is proposed. Orientation invariance is achieved by expressing each cell
region in a Cartesian space defined by the major and minor axes of the largest elliptical region within the cell. Scale invariance is achieved by mapping the elliptical region to a unit circle before constructing the polar map. The non-linear filtering method, from Chapter 2, is then applied to the polar map so that the texture features extracted from the filter output are invariant to contrast. The implementation of both local and global energy measures achieves improved accuracy. It is shown that the proposed method consistently achieves an accuracy of over 90% in classifying six species of pollen, fungal and fern spores when orientation, scale or contrast is altered. In contrast, the classification accuracy of methods based on linear filtering can dip below 50% when subjected to the same test.

# 1.6 Thesis Organization

Fig. 1.2 shows an overview of the proposed methodology. In the next four chapters, the thesis develops the rationale and provides a detailed discussion and validation of the various aspects in this methodology



Figure 1.2: Overview of image analysis software for robust detection and classification of biomedical cell specimens from light microscope images.

In Chapter 2, the classical edge detection measures are described followed by

a detailed description of the proposed hyperbolic tangent (HBT) filter and edge similarity measure with the regularization term. The experimental results of this method is presented and compared against current edge detection methods.

In Chapter 3, the multi-scale min-product method (MMPM) is presented and the performance criteria for multi-scale edge detection is also defined. This criteria is then applied to compare the edge detection performance of MMPM and MWPM. The performance of the difference of box (DOB) [38] and HBT filters is also compared against Canny's filter [12].

In Chapter 4, current clump splitting methods are briefly reviewed before the rule-based robust clump splitting method is proposed. The rules are designed for accurate splitting of clumps comprising objects of diverse sizes and shapes. The performance of the method is evaluated on unseen data and also compared against other methods. Each clump splitting rule is also carefully validated.

In Chapter 5, a rotation, scale and contrast invariant method for texture classification is proposed. Experimental results in Section 5.8 establish these invariant properties and validate the choice of texture based features used based on a data-set of air-borne allergens from six species of fungal, fern and pollen spores.

Finally, the conclusions and recommendations for future work in this area of research is presented in Chapter 6.

# Chapter 2

# A Luminance and Contrast-Invariant Edge-Similarity Measure

A novel similarity measure, which is robust to luminance and contrast, is presented for edge-detection. It incorporates a regularization term and employs directional FIR edge filters with hyperbolic tangent profiles to ensure improved noise performance and edge localization compared to classical methods.

# 2.1 Rationale

The accurate detection of edges is often not achieved due to the sensitivity of commonly used methods to image contrast, noise and, to some extent, uneven illumination. Despite the importance of developing edge detection methods that are robust under these conditions, reported research [46, 48, 47, 53, 65, 81] which is suitable for use with light microscope images is limited.

Classical gradient magnitude (GM) methods [12, 63, 67] are usually dependent on edge strength; hence, weaker edges such as those at texture boundaries may not be detected. Frei and Chen [31] have proposed an alternative method of detecting valid edges regardless of their magnitude. Their approach is termed as an anglebased (AN) method in this thesis since it is based on the computation of the cosine of the projection angles between neighborhoods and predefined edge filters. A problem with this technique is its sensitivity to noise and uneven illumination. Methods based on local thresholding of image gradients are also sensitive to uneven illumination since they tend to inhibit edges in regions of low luminance [77] or, conversely, enhance them [43]. In general, edge detection methods that are robust under different contrast levels tend to be more affected by noise.

The spatial profile of the edge filter is another factor that influences the edge detection performance. Canny's Gaussian first derivative (GFD) filter [12] may be regarded as an optimal step-edge detector. However, it is derived for an ideal step edge model [12], when in fact, the images of interest in this thesis have blurred profiles arising from the digital image acquisition process.

Morrone et al. [65] and later Kovesi [46] described a technique in which images are represented in the frequency domain and edges occur at points of maximum phase congruency. Such phase congruency (PC) methods are invariant to changes in illumination and contrast. Although they exhibit better contrast invariance than GM methods, they give poorer edge localization in that false edges are detected in the vicinity of sharp transitions. This is due to the multiple zero crossings in the spatial profile of the log polar Gabor filter.

More recently, Desolneux et al. [18] proposed a contrast-invariant edge detection method based on the Helmholtz principle. It is a parameter-free method that defines edges as geometric structures with large deviations from randomness. The detection of a given edge is sensitive to the size of the windowed region while edge localization is inferior to that of Cannys edge detector.

A robust edge detection algorithm is proposed. It is more robust than the GM and AN methods under different illumination, contrast and noise levels, and results in better localizations of sharp transitions in an image compared to PC methods. It is based on a measure of edge similarity between image neighborhoods and the use of directional finite impulse response (FIR) edge filters with hyperbolic tangent (HBT) profiles. A balance between the conflicting requirements of contrast invariance and noise tolerance is obtained by using a regularization term in the similarity measure.

# 2.2 Classical Edge Detection Scheme

Here, the classical GM and AN methods and their shortcomings is briefly discussed. Both methods are based on the measure of similarity between local image neighborhood and a predefined set of directional edge filters. An image neighborhood that is similar to one of these edge filters has a high similarity value. However, the methods differ in their definition of the similarity measure used. In the GM method, a normalized similarity measure  $\hat{C}_i$  is defined as

$$\hat{C}_{i} = \frac{C_{i}}{\max\{|C_{i}|, 1 \le i \le N\}}$$
(2.1)

where

$$C_i = \langle \mathbf{b}_i, \mathbf{g} \rangle \tag{2.2}$$

is the inner product of  $\mathbf{b}_i$  and  $\mathbf{g}$ , which are the  $(2W + 1)^2 \times 1$  column representations of the  $(2W + 1) \times (2W + 1)$  image neighborhood centered at pixel *i* and the 2–D edge filter, respectively. *N* is the total number of image pixels whereas *W* is the width of the edge filter. The drawback of this method is its sensitivity to edge strength since weak edges tend to be suppressed in the presence of significantly stronger edges in the image. Therefore, it may be difficult to determine suitable thresholds to accurately detect both strong and weak edges in an image.

Frei and Chen [31] introduced an angle-based method [48] that detects edges based on the similarity of the image neighborhoods to a predefined filter irrespective of the edge strength. The measure of similarity at each pixel i is the cosine of the projection angle,  $P_i$ , between  $\mathbf{b}_i$  and  $\mathbf{g}$  [31]:

$$P_i = \frac{\langle \mathbf{b}_i, \mathbf{g} \rangle}{\|\mathbf{b}_i\| \|\mathbf{g}\|} \tag{2.3}$$

A larger magnitude of  $P_i$  means higher similarity between  $\mathbf{b}_i$  and  $\mathbf{g}$ . A consequence of (2.3) is an undesirable dependence on the mean value (average luminance),  $\mu_i$ , of  $\mathbf{b}_i$  since  $P_i(\mathbf{b}_i) \neq P_i(\mathbf{b}_i - \mu_i)$ . Therefore, larger values of  $P_i$  will be obtained for edges in the darker regions of an image (where  $\mu_i$  is low) compared to those in the lighter regions (where  $\mu_i$  is high).

The edge detection performances of both the GM and AN methods are also dependent on the choice of edge filter, which would require a trade-off to be made between noise suppression and edge localization. The AN method in [31] performs poorly in the presence of noise due to the use of the highly localized  $3 \times 3$  Sobel filter. Similarly, the GFD filter in Canny's GM method offers good edge localization but has higher noise sensitivity compared to filters such as the difference-of-boxes (DOB) [75] for the same filter width. The GFD filter therefore requires smoothing as a pre-processing step to minimize noise but this blurring may lead to weak edges being difficult to detect [53].

# 2.3 Edge Detection via HBT Filter

Petrou and Kittler [71] observe that image edges resemble ramp profiles rather than ideal step edges due to the process of digital image acquisition, and so they use the ramp profile with additive noise as the edge model to derive their edge filter. This is obtained by optimizing a performance measure that combines Canny's criteria [12] of accurate edge detection, edge localization and minimization of false edge responses. Here, their observation is confirmed by obtaining an optimal estimate (in the least-squares sense) of image edges by using principal component analysis (PCA). However, unlike their approach, the edge detectors proposed in this thesis are required to resemble the actual profiles of image edges since the emphasis here is on optimizing two of Canny's criteria—accurate edge detection, ED, and localization, EL—without explicitly including the third criterion on suppression of false edge responses, SF, due to noise. Section 2.3.4 will show that such responses to noise can be distinguished from valid edges since the separation between adjacent noise maxima in the filter response exceeds the narrow spatial widths of the proposed edge filters [17]. The criteria ED and EL are defined as

$$ED = \frac{\int_{-W}^{+W} S(-x) f_W(x) dx}{n_0 \sqrt{\int_{-W}^{+W} f_W^2(x) dx}}$$
(2.4)

$$EL = \frac{\int_{-W}^{+W} S'(-x) f'_W(x) dx}{n_0 \sqrt{\int_{-W}^{+W} f'_W(x) dx}}$$
(2.5)

where S(x) is the natural edge model in an image centered at x = 0,  $f_W(x)$  is the finite impulse response of the edge filter bounded by [-W, W],  $f'_W$  its derivative and  $n_0$  is the standard deviation of the Gaussian white noise. From the Schwarz inequality, the upper bounds of both ED and EL are reached when  $f_W(x) = S(-x)$ , i.e., they are maximized if the filter resembles the natural edges in an image [12].

#### 2.3.1 Similarity to natural edges

Here, the method of obtaining 2–D edge filters is described such that the edge filters optimally approximate these natural edges, in the least squares sense [48, 47]. A sliding window first extracts all  $(2W+1) \times (2W+1)$  (W = 2) local neighborhoods in the image. PCA is then applied to this set of neighborhoods to generate  $(2W+1)^2$ eigenvectors  $\{\mathbf{e}_i, 1 \le i \le n^2\}$ , each of size  $(2W+1) \times (2W+1)$ . A neighborhood  $\mathbf{b}_i$  can be expressed as  $\mathbf{b}_i = \mathbf{m} + \sum_{j=1}^{n^2} u_{ij}\mathbf{e}_j$ , where the average over all local neighborhoods is  $\mathbf{m} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{b}_i$  and  $u_{ij}$  is the projection of  $\mathbf{b}_i - \mathbf{m}$  onto the *j*th eigenvector  $\mathbf{e}_j$ .

Eigenvector  $\mathbf{e}_1$  corresponds to the largest eigenvalue  $\lambda_1$ , where  $\lambda_1, \dots, \lambda_{n^2}$  are in decreasing order of magnitude. This eigenvector behaves as a low pass filter where  $u_{i1}\mathbf{e}_1$  approximates the average luminance  $\mu_i$  of  $\mathbf{b}_i - \mathbf{m}$  given by  $\mu_i = \frac{1}{n^2} \sum_{j=1}^{n^2} (b_i(j) - m(j))$ . The local gray level variation in  $\mathbf{b}_i$  is therefore given by

$$\mathbf{s}_i = \mathbf{b}_i - \mathbf{m} - \mu_i = \sum_{j=2}^{n^2} u_{ij} \mathbf{e}_j$$
(2.6)

In the set of eigenvectors  $\{\mathbf{e}_j : 2 \leq j \leq n^2\}$  in (2.6),  $\mathbf{e}_2$  is the basis function that yields the best mean square error approximation for  $\mathbf{s}_i$  as  $u_{i2}\mathbf{e}_2$  since it gives the smallest mean square error,  $J_{min}$ ,

$$J_{min} = \frac{1}{N} \sum_{i=1}^{N} \|\mathbf{s}_i - u_{i2}\mathbf{e}_2\|^2$$
(2.7)

over all local neighborhoods compared to the other eigenvectors.

PCA is a successive approximation scheme that ensures that the approximation error  $\mathbf{s}_i - u_{i2}\mathbf{e}_2$  is best approximated by using  $\mathbf{e}_3$ , the subsequent error  $\mathbf{s}_i - u_{i2}\mathbf{e}_2 - u_{i3}\mathbf{e}_3$  is best approximated by  $\mathbf{e}_4$ , and so on. Figs. 2.1(a) and (c) show the 5 × 5 eigenvectors  $\mathbf{e}_2$  and  $\mathbf{e}_3$  extracted from the *Lena* image. The eigenvector pair have



Figure 2.1: Least-squares estimates of PCA eigenvectors using FIR HBT filters. (a) and (c): PCA eigenvectors of second and third largest eigenvalues, (b) and (d): Corresponding least-squares estimates using a linear combination of FIR HBT filters

similar profiles but are orthogonal. The number of zero crossings in the 5 × 5 eigenvectors increases from eigenvector pair  $\mathbf{e}_2 - \mathbf{e}_3$  to  $\mathbf{e}_6 - \mathbf{e}_7$  and beyond, indicating that the eigenvectors corresponding to smaller values of eigenvalues capture higher frequency information of the local neighborhoods and hence are more susceptible to noise. The two eigenvectors  $\mathbf{e}_2$  and  $\mathbf{e}_3$  are considered for edge detection since they most accurately approximate the gray level variation in local neighborhoods.

From Fig. 2.1, it is noteworthy that both  $\mathbf{e}_2$  and  $\mathbf{e}_3$  have blurred step edge profiles and can be approximated by  $\hat{\mathbf{e}}_2$  and  $\hat{\mathbf{e}}_3$  (Figs. 2.1(b) and 2.1(d)) given by  $\hat{\mathbf{e}}_2 = \alpha_{21}\mathbf{g} + \alpha_{22}\mathbf{g}^T$ ,  $\hat{\mathbf{e}}_3 = \alpha_{31}\mathbf{g} + \alpha_{32}\mathbf{g}^T$ , where  $\{\alpha\}$  are weights,  $\mathbf{g}^T$  is orthogonal to **g**, and **g** has an HBT profile (within its region of support)

$$f_W(x,y) = \begin{cases} \frac{1-e^{-\sigma_W(x+y)}}{1+e^{-\sigma_W(x+y)}} & |x|, |y| \le W, \\ 0 & \text{otherwise} \end{cases}$$
(2.8)

The region of support for  $f_W$  is limited by W to ensure edge localization (W = 2 in Fig. 2.1).

The orthogonal FIR filter pair,  $\mathbf{g}$  and  $\mathbf{g}^T$ , is obtained by sampling  $f_W$  at integer locations  $(x_d, y_d)$  within [-W, W]. The parameter  $\sigma_W$  defines the steepness of the profile at the zero crossing and its relationship to the filter support W is described in Section 2.3. The weights  $\alpha_{ij}$  are determined by projecting both  $\mathbf{e}_2$  and  $\mathbf{e}_3$  onto  $\mathbf{g}$ and  $\mathbf{g}^T$ , i.e.,  $\alpha_{ij} = \langle \mathbf{e}_i, \mathbf{g}_j \rangle / \langle \mathbf{g}_j, \mathbf{g}_j \rangle$ . In this example, the weights  $\alpha_{21}, \alpha_{22}, \alpha_{31}$  and  $\alpha_{32}$  have values -0.98, -0.18, 0.18 and -0.98, respectively. The error  $\varepsilon_i$  defined by  $\varepsilon_i = ||\mathbf{e}_i - \hat{\mathbf{e}}_i||/||\mathbf{e}_i||$  is only 10.3% for  $\mathbf{e}_2$  and 11.5% for  $\mathbf{e}_3$ . This indicates that the eigenvectors  $\mathbf{e}_2$  and  $\mathbf{e}_3$  obtained from the set of local image neighborhoods can be accurately approximated by the orthogonal pair  $\mathbf{g}$  and  $\mathbf{g}^T$ .

#### 2.3.2 Properties of HBT filters

Here, the influence of  $\sigma_W$  on the spatial and frequency characteristics of the continuous FIR HBT filters is discussed by using the 1–D filter representation

$$f_W(x) = \begin{cases} \frac{1-e^{-\sigma_W x}}{1+e^{-\sigma_W x}} & |x| \le W, \\ 0 & \text{otherwise} \end{cases}$$
(2.9)

since the 1–D case may be extended in a straightforward manner to 2–D. The filter  $f_W$  is odd-symmetric with a single zero-crossing at the origin with the slope at the zero crossing point given by  $\sigma_W/2$ . The spatial profile of filter  $f_W$  is plotted in Fig. 2.2(a) for W = 2. The filter resembles a ramp for  $\sigma_2 \leq 0.5$ , a blurred edge



Figure 2.2: Spatial and frequency properties of HBT filter. (a) 1–D continuous spatial profile of HBT filters  $f_2$  ( $\sigma = 0.5$  (dark gray), 1.0 (medium gray) and 2.0 (light gray)). (b) Frequency responses of 1–D discrete FIR filters after normalization by their respective maximum values ( $\sigma = 0.5$  (dark gray), 1.0 (medium gray) and 2.0 (light gray)) and Gaussian filter (s = 1.0 (dashed lines)).

for larger values of  $\sigma_2$ , and approaches a DOB filter for  $\sigma_W > 5.0$ .

Fig. 2.2(b) compares the frequency responses of 1–D FIR HBT filters with that of Canny's GFD filter,  $f(x) = xe^{-x^2/2s^2}$  (with s = 1.0 to limit the truncation at the tail ends of the Gaussian function). Both filters are discretized by sampling them at  $x_d = \{-2, -1, 0, 1, 2\}$ . It is clearly seen that the family of FIR HBT filters has a narrower bandwidth, indicating better noise reduction compared to Canny's GFD.

#### 2.3.3 Tuning of HBT filter parameters

The parameter value of  $\sigma_W$  is determined for a given filter width W such that the HBT filter pair can best approximate the natural step edges in an image and therefore ensure that the filter pair meets the objectives of good SNR and accurate edge localization [12, 17]. This is done by selecting  $\sigma_W$  for a given W such that its corresponding approximation error  $\varepsilon$  is the smallest. For the eigenvectors extracted from the *Lena* image, the relationship between  $\varepsilon_{total}$  (=  $\varepsilon_2 + \varepsilon_3$ ) and  $\sigma_W$  for W =1,2,3 are shown in Fig. 2.3(a). The optimal values of  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$  are 2.5, 1.0 and 0.7, respectively. Table 2.1 lists the optimal values obtained for five images from the USC-SIPI image database [84].



Figure 2.3: Influence of  $\sigma_W$  on (a)  $\varepsilon_{total}$  and (b)  $C_W$  for W = 1 (*light gray*), 2 (*medium gray*) and 3 (*dark gray*).

# 2.3.4 Average distance between adjacent noise maxima, $C_W$

Canny's third criterion, SF, aims to limit the detection of false edges due to noise in the vicinity of a valid step edge. Ideally, it requires the mean distance between adjacent noise maxima to approximate the width of the filter response to a single step edge, i.e., 2W, although a fraction of this may suffice. Demigny and Kamle [17] define Canny's third criterion in the discrete domain as

$$C_W = \frac{2\pi}{\arccos(-\rho)} \tag{2.10}$$

No.	Images	$\sigma_1$	$\sigma_2$	$\sigma_3$
1	Aerial	2.3	1.1	0.8
2	Airplane(F-16)	3.1	1.2	0.8
3	Baboon	2.0	1.4	1.0
4	Couple	3.1	1.4	1.0
5	House	3.4	1.2	0.8
6	Stream and Bridge	2.6	1.4	0.9
7	Level Step Wedges	2.3	1.4	0.9
8	Man	2.5	1.2	0.8
9	Moon Surface	2.2	1.1	0.8
10	Boat	3.2	1.4	0.4

Table 2.1: Optimal  $\sigma_W$ , W = 1, 2 and 3 for standard images from USC-SIPI Image Database.

where  $C_W$  is the average distance between adjacent noise maxima for filter width W and  $\rho$  is expressed as

$$\rho = \frac{\sum_{k=-\infty}^{k=+\infty} (g_W(k) - g_W(k-1))(g_W(k) - g_W(k+1))}{\sum_{k=-\infty}^{k=+\infty} (g_W(k) - g_W(k-1))^2}$$
(2.11)

Fig. 2.3(b) shows the relationship between  $C_W$  and  $\sigma_W$  for filter widths W = 1, 2, 3. The maximum values of  $C_1 (\approx 3.4)$  and  $C_2 (\approx 4.8)$  are greater than the corresponding values of 2W (2.0 and 4.0, respectively) while  $C_3 (\approx 4.8)$  is slightly smaller than 6.0. This shows that compact filters with widths up to W = 3 effectively prevent the detection of false edges in the vicinity of the true edge.

# 2.4 Edge Detection Scheme Incorporating New Similarity Measure

The undesirable sensitivity to illumination in the AN method [31] may be ameliorated by modifying their edge-similarity measure  $P_i$  in (2.3) to

$$P_i' = \frac{\langle \mathbf{b}_i - \overline{\mathbf{b}}_i, \mathbf{g} \rangle}{\|\mathbf{b}_i - \overline{\mathbf{b}}_i\| \|\mathbf{g}\|}$$
(2.12)

where  $\overline{\mathbf{b}}_i$  is the average of  $\mathbf{b}_i$ . The use of  $P'_i$  ensures that the detection scheme is invariant to the average luminance and contrast of the neighborhood. However, some smooth regions may yield a high  $P'_i$  if they approximate the form of the attenuated spatial profile of the edge filter and, thereby, lead to erroneous edge points. Therefore, a modified AN method is presented in this thesis where it is robust under different illumination, contrast and noise levels. The aim here is to (a) detect meaningful edges regardless of illumination, and (b) significantly reduce the sensitivity to noise while preserving responses to true edges.

The proposed method utilizes the following cosine measure  $R_i$  with a regularization parameter  $\gamma$  and an empirically determined constant c:

$$R_{i} = \frac{\langle \mathbf{b}_{i} - \overline{\mathbf{b}}_{i} + c\gamma, \mathbf{g} \rangle}{\|\mathbf{b}_{i} - \overline{\mathbf{b}}_{i} + c\gamma\|\|\mathbf{g}\|}$$
(2.13)

An estimate,  $\hat{\gamma}$ , of  $\gamma$  may be obtained by using the median absolute deviation (MAD) [19] of the  $(2W + 1) \times (2W + 1)$  image neighborhoods

$$\hat{\gamma} = \frac{\text{median}(Y_i : 1 \le i \le N)}{0.6745} \quad \text{where} \tag{2.14}$$

$$Y_i = \sqrt{\frac{1}{n^2} \sum_{j=1}^{n^2} [\mathbf{b}_i(j) - \bar{\mathbf{b}}_i]^2}, \quad i = 1, 2, \dots, N.$$
(2.15)

Neighborhoods corresponding to smooth regions have smaller  $Y_i$  compared to those containing edges. A small value of  $\hat{\gamma}$  increases robustness to different edge strength but results in heightened sensitivity to noise, and *vice versa*.

Table 2.2 shows the influence of  $\sigma_2$  on the 5 × 5 HBT filter response to random noise. The energy of the response is given by the mean sum squared filter response,  $\sum_i R_i^2/N$ , averaged over ten different random noise records, where each record is a  $512 \times 512$  image with zero mean and unit standard deviation. The results agree with the qualitative evaluation of Fig. 2.2(b), which shows the noise response decreasing as  $\sigma_2$  increases. Similar results are obtained for  $\sigma_2 = 0.5$  and 1.0 since  $\sigma_2 = 0.5$  gives rise to a larger side lobe although its main lobe is smaller than that of  $\sigma_2 = 1.0$ . All four HBT filters also exhibit a smaller noise response compared to the GFD filter value of  $2.41 \times 10^{-3}$  obtained with s = 0.8.

$\sigma_2$	$\sum_i R_i^2/N \ \left(\times 10^{-3}\right)$
0.5	2.24
1.0	2.24
1.5	2.29
2.0	2.35

Table 2.2: Influence of HBT filter  $\sigma_2$  on the noise response

Table 2.3 shows the influence of parameter c in (2.13). As observed, the filter response to noise decreases as the parameter c increases. Typically, c > 1 in order to effectively suppress noise from smooth regions such as those observed when using (2.12). However, large values of c may result in (2.13) being insensitive to the presence of weak edges. It is empirically determined that c = 4 offers a good compromise value.

с	$\sum_i R_i^2/N~\left(\times 10^{-2}\right)$
0	7.88
1	2.47
2	0.82
3	0.39
4	0.22
5	0.14

Table 2.3: Influence of parameter c on the noise response

The details of the method [47] are as follows.

#### Determine HBT parameter, $\sigma_W$

- Step 1. For a given image, apply PCA to find the eigenvectors  $\mathbf{e}_2$  and  $\mathbf{e}_3$ .
- Step 2. From a set of four 2–D HBT filters oriented along 0°, 45°, 90° and 135° and defined by a given W and  $\sigma_W$ , find the orthogonal HBT filter pair  $\mathbf{g}_i$  and  $\mathbf{g}_j$  that best approximates  $\mathbf{e}_2$  and  $\mathbf{e}_3$ , i.e., yield the smallest  $\varepsilon_{total}$  (Section 2.3.3.) The filter width W ranges from 1 to 3. A smaller width achieves better edge localization but poorer noise tolerance and *vice versa*.
- Step 3. Determine the  $\sigma_W$  corresponding to the smallest  $\varepsilon_{total}$  by using a simple 1–D bisection method where step 2 is repeated for specific  $\sigma_W$  values within the interval [0.01, 5.00]. This is possible since the error curves as shown in Fig. 2.3 have only one minima and the corresponding  $\sigma_W$  offers the ideal HBT slope to optimally approximate the natural image edges.

#### Determine $\hat{\gamma}_i$

Step 4. Obtain an estimate of  $\gamma_i$  (2.14) and set c = 4.

#### Compute similarity maps

Step 5. Using  $\gamma_i$  and c from step 4, apply  $R_i$  to the given image to compute four similarity maps where each map corresponds to one of the four HBT filters defined by W and  $\sigma_W$  from step 3. The use of these four filters ensures a more robust detection of edges that are diagonally oriented.

#### Determine equivalent similarity map

Step 6. At each pixel location, compare the corresponding values from the four similarity maps; the largest magnitude is selected as the pixel value at this point.

#### Detect edge pixels

Step 7. Determine the local maximum in the equivalent map and apply a suitable threshold on the local maxima to determine the edge pixels.

## 2.5 Results and Discussion

The robustness of the proposed method is compared against the AN [31] and GM methods [12, 63, 67] under different illumination, contrast and noise levels. The noise reduction property of the FIR HBT filter is also compared with the GFD filter. In the results shown later in this Section, pixel values are scaled to the gray

level range [0, 255] and image intensities are inverted so that pixels with strong edge responses are shown as dark points.

#### 2.5.1 Uneven illumination

Fig. 2.4(a) shows a 512 × 512 *Lena* image with an artificially imposed illumination gradient. Figs. 2.4(b) and 2.4(c) are similarity maps obtained using steps 1-6 of the algorithm in Section 4 with W = 2 and  $\sigma_W = 1.0$ . Fig. 2.4(b) is computed using the similarity measure  $P_i$  [31] and Fig. 2.4(c) using  $R_i$ .



Figure 2.4: Comparison between the  $P_i$  and  $R_i$  measures. (a) Lena image with illumination gradient. (b) Similarity map,  $P_i$  measure. (c) Similarity map,  $R_i$  measure.

As expected, the similarity map of Fig. 2.4(b) is more distinct in the darker rather than in the lighter regions of the image. The similarity map of Fig. 2.4(c)) is equally distinct regardless of image illumination. This example illustrates the fact that  $P_i$ , unlike  $R_i$ , is sensitive to variation in illumination.

#### 2.5.2 Contrast variation

Fig. 2.5(a) shows a 570 × 760 light microscope image of malaria-infected red blood cells. The malaria parasites appear as tiny black spots within the red blood cells while the relatively larger black blob just to the right of the image center is a white blood cell. Figs. 2.5(b) and 2.5(c) are similarity maps obtained using the algorithm in Section 2.4 with W = 2 and  $\sigma_W = 0.5$ ; Fig. 2.5(b) is computed using the similarity measure  $\hat{C}_i$  from (2.1) and Fig. 2.5(c) computed using  $R_i$  from (2.13).



Figure 2.5: Comparison between the  $\hat{C}_i$  and  $R_i$  measures. (a) Light microscope image of infected red blood cells. (b) Similarity map, measure  $\hat{C}_i$ . (c) Similarity map, measure  $R_i$ . (d) Edge map, measure  $\hat{C}_i$ . (e) Edge map, measure  $R_i$ .

It is observed in Fig. 2.5(b) that using  $\hat{C}_i$  gives rise to large values of the similarity measure at the boudaries of the white blood cell and parasite regions at the cost of suppressing the valid edges at the boundaries of the red blood cells. In contrast, both strong and weak edges are accurately detected in the similarity

map of Fig. 2.5(c).  $R_i$  ensures the robustness of the proposed method in handling a wide range of edge strength values. This is a significant advantage in edge detection since a fixed threshold can be easily determined for automatically thresholding the similarity maps of a large set of images, whereas the large difference in magnitude between the strong and weak edges in Fig. 2.5(b) complicates the search for a suitable edge threshold.

The edge maps in Figs. 2.5(d) and (e) correspond to (b) and (c), respectively, and they are obtained via Canny's non-maxima suppression and hysteresis thresholding [12]. As observed in Figs. 2.5(d) and (e), some cell edges are undetected in the classical GM method whereas the proposed method accurately detects all edges.

#### 2.5.3 Noise

Here, the performances of the FIR HBT filter and the GFD filter are compared using steps 1 and 4-6 of the proposed algorithm. Both filters have a spatial dimension of  $5 \times 5$  (W = 2), parameter s of the GFD filter is fixed at 0.8 [12] and  $\sigma_2$  of the HBT filter is found to be 1.11 from steps 2 and 3. Fig. 2.6(a) shows an image of an outdoor scene containing a wide range of edge strengths.

The image is corrupted by Gaussian noise with an image SNR of approximately 10 dB. It is evident that the similarity map using the FIR HBT filter (Fig. 2.6(c)) gives superior edge response and noise performance compared to the GFD filter (Fig. 2.6(b)). This is explained by the careful selection of  $\sigma_2$  to ensure that HBT has a high similarity to the edges in the image as well as a narrower bandwidth ( $\sigma_2 = 1.11$  from Steps 2 and 3 of Section 4) compared to GFD (s = 0.80). Figs. 2.6(d) and (e) are the resultant edge maps corresponding to the GFD and HBT filters, respectively. They were obtained by applying suitable hysteresis thresholds to the local maxima in Figs. 2.6(b) and (c). It is seen that more accurate detection and localization of edges is obtained with the HBT filter.



Figure 2.6: Comparison between the FIR HBT and GFD filters on noisy images. (a)Outdoor scene. (b) Similarity map, GFD filter. (c) Similarity map, FIR HBT filter. (d)Edge map, GFD filter. (e) Edge map, HBT filter.

The noise sensitivity of both GFD and HBT filters are also compared by applying them to a synthetic binary image comprising a white square on a black background and corrupted with white Gaussian noise. The comparison is made over a range of SNR levels for a filter width of W = 2. Using steps 2 and 3 from the algorithm in Section 4,  $\sigma_2 = 1.36$  for SNR  $\geq 5$  dB and  $\sigma_2 = 1.66$  for SNR < 5 dB for the HBT filter. The s values for the GFD filters remain fixed at 0.8.

Since it is known *a priori* that 3.4% of the image comprises edge pixels, which are located at the boundary of the white square, a threshold is imposed such that only pixels with edge strengths in the top 3.4% are admitted as edge pixels. The results of the comparison are shown in Table 2.4 where the total number of true edge pixels is 260. The GFD filter, in general, detects more true edge pixels than the HBT filter. However, HBT detects fewer false edge pixels and tends to perform better than GFD for SNR levels below 5 dB. The performance of GFD degrades more rapidly as the noise level increases, indicating a lower robustness to noise.

		GFD			HBT	
$\mathrm{SNR}(\mathrm{dB})$	Correct	Missed	False	Correct	Missed	False
	Edges	Edges	Edges	Edges	Edges	Edges
No noise	256	4	0	252	8	0
10	255	5	2	232	28	0
5	254	6	28	222	38	3
0	165	95	256	176	84	62

Table 2.4: Quantitative performance of edge detection with noise.

#### 2.5.4 Edge localization

Fig. 2.7(a) shows a  $512 \times 512$  synthetic image from the USC-SIPI database [84]. Figs. 2.7(b) and 2.7(c) represent, respectively, the corresponding edge maps obtained using the proposed method and phase congruency (PC) [46]. The PC method employs filters at four scales with six filter orientations at each scale to generate a map (which is termed the PC map) from the phase congruency information. The edge maps are obtained by applying non-maxima suppression to the PC map and similarity map using  $R_i$ , followed by a threshold of 0.1 on the local maxima.



Figure 2.7: Edge localization comparison of the proposed and PC methods. (a) Synthetic image. (b) Edge map from proposed method. (c) Edge map from PC method.

A quantitative measure of edge detection performance is provided by Pratt's figure of merit [74],

$$F = \frac{1}{\max(E_I, E_D)} \sum_{i=1}^{E_D} \frac{1}{1 + \alpha d^2(i)}$$
(2.16)

where  $E_I$  and  $E_D$  are the number of ideal and detected edge pixels, respectively, d(i) the Euclidean distance between the *i*th edge pixel detected and the ideal edge pixel nearest to it, and  $\alpha$  a scaling constant set to 1. The proposed method gives significantly better edge localization (F = 0.86) than the PC method (F = 0.63), which detects false edges in the vicinity of sharp transitions in an image and gives poor edge localization. The value of F for the proposed method is smaller than 1 since, being a step edge detector, it does not detect the corners in the image (the junctions where the edges meet) The performance of the proposed method is also compared with the PC approach using the *Lena* image of Fig. 2.8(a). Figs. 2.8(b) and (c) show the similarity map using  $R_i$  and the PC map, both normalized to the range [0, 1]. Figs. 2.8(d) and (e) show the corresponding edge maps obtained by first detecting the local edge maxima, followed by applying Canny's hysteresis thresholding. Compared to PC, the proposed method results in a higher gradient magnitude response (Figs. 2.8(d) and (c)) and more accurate detection of both strong and weak edges (Figs. 2.8(d) and (e)).



Figure 2.8: Edge localization comparison between the proposed and PC methods. (a) *Lena* image. (b) Similarity map from proposed method. (c) PC map. (d) Edge map from proposed method. (e) Edge map from PC method.

# 2.6 Conclusion

The proposed edge-similarity measure has been shown to be simultaneously robust to image illumination and contrast unlike traditional GM and AN methods, which are sensitive to contrast and illumination, respectively. The subtraction of mean luminance from the image neighborhoods ensures that the AN method is invariant to contrast changes but results in increased sensitivity to noise. The use of a regularization term  $\hat{\gamma}$ , estimated using the MAD criterion, achieves a good compromise between the opposing objectives of reduced noise sensitivity and contrast invariance. In addition, the edge filter used has a FIR HBT profile, which offers better edge localization and reduced noise sensitivity compared to the classical GFD filter. The proposed method shows better edge localization compared to PC and has been successfully applied to both synthetic and natural images.

# Chapter 3

# Step Edge Detection via a Multi-Scale Min-Product Method

In Chapter 2, a luminance and contrast-invariant edge detection scheme is proposed and validated. An undesirable side effect of having contrast invariance is the heightened sensitivity to noise. This issue is addressed by introducing a regularization parameter which achieves a good compromise between the conflicting goals of contrast invariance and robustness to noise. In this Chapter, a multi-scale method is presented which ensures accurate detection of edges under high levels of noise by the novel use of point-wise *min* and *product* operators on coefficients at successive scales. The method is based on the observation that the best edge localization is achieved at the finest scale and edge coefficients increase in magnitude from the finest to the coarsest scale whereas noise coefficients decreases in magnitude. Unlike existing multi-scale methods, edge filters are directly applied, with multiple spatial widths, on the original image as this avoids the drifting of edge maxima that occurs across successive scales.

# 3.1 Rationale

Edge detection schemes that operate at a fixed scale suffer from a trade-off in achieving the conflicting goals of good signal-to-noise ratio (SNR) and edge localization of the processed image. It is often difficult to find an edge detector with a fixed scale which detects all edges in an image. Smoothing has been used as a pre-processing step to minimize noise [12, 63] but this results in blurring which then makes weak edges harder to detect [53]. Multi-scale edge detection methods offer a solution to this problem since they utilize edge operators at different scales to find edges in the image. Rosenfeld et al. [80] demonstrated that edges can be enhanced while suppressing noise by taking the direct point-wise products of the image sub-band decompositions. Their method was then further developed by Mallat et al. and successfully applied in the wavelet transform domain [61, 62].

The wavelet transform [16] can be used as an effective multi-scale edge detection tool as it is well adapted to finding edges in an image. Mallat et al. distinguish edges from noise and characterize various edge profiles from the Lipschitz regularity of these edges across scale space [61, 62]. The Lipschitz regularity is based on the observation that step edges have large wavelet coefficients over many adjacent scales whereas noise dies out swiftly with increasing scale [61, 62].

Several methods have been developed for detecting edges based on their scale space behaviour in the wavelet domain [92, 4, 68, 96, 97, 83]. These methods will henceforth be called the multi-scale wavelet product based method or MWPM since they involve the direct point-wise multiplication of wavelet coefficients at several adjacent decomposition scales. Xu et al. [92] applied MWPM to filter noise from images. Subsequently, Sadler and Swami [83] applied this method to detect step edges while Zhang [96, 97] imposed an adaptive threshold on the point-wise products of the wavelet coefficients in order to identify important edge features.

A drawback of MWPM is that it results in the drift of edge maxima from the finer to coarser scales when the low pass filter has an even number of coefficients. This adversely affects the performance of MWPM in step edge detection since odd-symmetric high pass filters, used for detecting step edges, have even-numbered coefficients in their corresponding low pass filters. In order to minimize the drift of edge maxima, the point-wise product operation was restricted by Xu et al. to the first two or three sub-band decomposition levels [92]. Another drawback of MWPM is that the choice of edge filter is confined to the quadratic spline filter [4] although the chosen filter may not necessarily perform well under noisy conditions.

A multi-scale min-product method (MMPM) is proposed that addresses the aforementioned shortcomings. It is based on the following two observations: (1) At a given location, the magnitude of edge coefficients increases from the finest to the coarsest scale but conversely, noise coefficients decreases in magnitude and (2) coefficients at the finest scale offer the best edge localization. These observations are reflected in MMPM via the point-wise *min* and *product* operators, applied in that sequence, across successive scales to accurately detects edges under high levels of noise.

Unlike MWPM, a wider choice of filters which include the difference-of-box (DOB) [38] and the hyperbolic tangent filters (HBT) of finite impulse response (FIR) is available for use in MMPM. MMPM also avoids the edge drift problem by directly applying edge filters of multiple spatial widths on the original image instead of following the wavelet decomposition algorithm.

The MMPM algorithm is briefly described as follows: A stack of similarity maps are first computed using the similarity measure as in Chapter 2 where each similarity map is obtained using an edge filter, of a specific width. The stack is ordered such that the filter width progressively increases from the top to the bottom of the stack. The *min* operator reduces the effects of noise in the first three maps of the stack by replacing the coefficients at every location in each map by the smallest magnitude at that location over another subset of the stack. This subset starts from the map which is currently subjected to the *min* operation and extends downwards to include a fixed number of maps. Finally, a composite similarity map is obtained by taking the point-wise product of these three maps.

The rest of this Chapter is organized as follows. Section 3.2 discusses the MMPM method and its application to edge detection. Section 3.3 redefines the Canny criteria for optimal edge detection in the multi-scale edge detection domain. Section 3.4 applies the performance criteria for evaluating the MMPM and MWPM and compares the edge detection performance of the DOB [38] and HBT filters against the Canny filter [12]. Section 3.5 discusses the experimental results. Section 3.6 concludes this Chapter.

### 3.2 Multi-Scale Min-Product Method

As noted above, the MMPM is based on the observation that the magnitude of the similarity measure at valid edge locations, increases from the finest to the coarsest scale. Conversely, their magnitude decreases from the finest to coarsest scale at noise locations. This is attributed by Meyer [64] to the Lipschitz regularity of step edge profiles being greater than or equal to zero whereas that of noise is less than zero.

Here, edge filters of smaller width are used at valid edge pixel locations in order to preserve edge localization and filters of larger width at non-edge pixel locations to suppress noise. This is achieved by applying the *min* and *product* operators on the similarity measure coefficients across a stack of similarity maps. Section 3.2.1 defines the multi-scale edge filters which will be used to compute the similarity measures at the various scales.

#### 3.2.1 Defining multi-scale edge filters

In the wavelet transform domain [60, 62, 61], the low and high pass filters at scale n are denoted by  $\mathbf{h}_n$  and  $\mathbf{g}_n$  where  $\mathbf{h}_n$  functions as a smoothing filter and  $\mathbf{g}_n$  functions as an edge filter. These two filters can be obtained by convolving the low pass filter,  $\mathbf{h}_1$ , at the finest scale, with the low and high pass filters,  $\mathbf{h}_{n-1}$  and  $\mathbf{g}_{n-1}$  from the finer scale, n-1 as:

$$\mathbf{h}_n = \mathbf{h}_1 \otimes \mathbf{h}_{n-1} \tag{3.1}$$

$$\mathbf{g}_n = \mathbf{g}_1 \otimes \mathbf{g}_{n-1} \tag{3.2}$$

where n = 2, 3, ... and  $\otimes$  denotes the convolution operator. The low pass filter performs image smoothing to regularize the subsequent ill-posed differentiating effect of the high pass filter. In the wavelet transform domain, the high pass filter for step edge detection is confined to the quadratic spline filter [61] since the filter is separable in the spatial domain and its low pass pair, the Gaussian function, is the only filter that does not create zero-crossings as the scale increases. This ensures that the edge detection process does not introduce new features as the scale increases thus enabling it to effectively track edge pixels over a range of scales. However, this requirement does not ensure that the chosen filter performs well under noisy conditions.

In order to enable the implementation of a wider choice of filters such as the difference-of-box filter (DOB) [38] in the multi-scale edge detection framework and to avoid the edge drift problem, an alternative method is proposed for generating multi-scale filters:

$$h_{W_n}(x) = [u(x + W_n) - u(x - W_n)] \cdot h_{W_n}^{\infty}(x)$$
(3.3)

$$g_{W_n}(x) = [u(x + W_n) - u(x - W_n)] \cdot g_{W_n}^{\infty}(x)$$
(3.4)

where  $h_{W_n}(x)$  and  $g_{W_n}(x)$  denote the coefficient values at location x of FIR filters  $\mathbf{h}_{W_n}$  and  $\mathbf{g}_{W_n}$ , respectively.  $h_{W_n}^{\infty}$  and  $g_{W_n}^{\infty}$  are the corresponding infinite impulse response (IIR) filters at scale n and u(x) is an ideal step function at x = 0.  $W_n$  is the width of the FIR filters corresponding to scale n. The IIR low pass filter  $h_{W_n}^{\infty}$ is defined as  $h_{W_n}^{\infty}(x) = 1$  for all values of x and scale n.

The IIR high pass filters considered in this thesis are listed as follows:

- First derivative of Gaussian (GFD) filter, [12]  $g_{W_n}^{\infty}(x) = x e^{\frac{-x^2}{\sigma_{W_n}^2}}$  where  $\sigma$  increase with scale n.
- Difference-of-box (DOB) filter, [38

 $g_{W_n}^{\infty}(x) = \begin{cases} \cdots & 1 & 1 & 0 & -1 & -1 & \cdots \\ [\cdots & 1 & 1 & -1 & -1 & \cdots ] & \text{for odd length FIR} \\ [\cdots & 1 & 1 & -1 & -1 & \cdots ] & \text{for even length FIR} \end{cases}$ 

Hyperbolic tangent (HBT) filter from Chapter 2,

$$g_{W_n}^{\infty}(x) = \frac{1 - e^{-\sigma x}}{1 + e^{-\sigma x}}$$

#### 3.2.2 Implementation of MMPM algorithm

The horizontal and vertical sub-bands,  $G_{W_n,H}$  and  $G_{W_n,V}$ , at scale *n* are computed as follows:

$$G_{W_n,H}(x,y) = \sum_{n_1} g_{W_n}(x-n_1) \sum_{n_2} h_{W_n}^T(y-n_2) I(n_1,n_2)$$
(3.5)

$$G_{W_n,V}(x,y) = \sum_{n_1} h_{W_n}(x-n_1) \sum_{n_2} g_{W_n}^T(y-n_2) I(n_1,n_2)$$
(3.6)

where I is the image, the  $\mathbf{g}_{W_n}$  and  $\mathbf{h}_{W_n}$  1–D filter pair is aligned along the x direction and  $\mathbf{g}_{W_n}^T$  and  $\mathbf{h}_{W_n}^T$  is the transposed pair, aligned along the y direction. A corresponding composite pair of horizontal and vertical sub-bands,  $\mathbf{CS}_{W_n,H}$  and  $\mathbf{CS}_{W_n,V}$ , is then obtained by point-wise selection of sub-band coefficients from the optimal scale  $j_o$  corresponding to the optimal filter width  $W_{j_o}$ , at each image location (x, y) as follows:

$$CS_{W_n,H}(x,y) = G_{W_{j_o},H}(x,y), j_o = \arg\min_{\Omega} |G_{W_j,H}(x,y)|$$
(3.7)

$$CS_{W_n,V}(x,y) = G_{W_{j_o},V}(x,y), j_o = \arg\min_{\Omega} |G_{W_j,V}(x,y)|$$
(3.8)

where  $\Omega: W_n \leq W_j \leq J + W_n - 1$  given that J is the number of filter widths considered and is empirically determined. The optimal scale at a particular image location, from the range of filter widths defined in  $\Omega$ , corresponds to the one with the minimum coefficient magnitude. The low and high pass filters at each scale is normalized to unit magnitude so that the magnitude of coefficients at the different scales can be directly compared via the point-wise operation. The composite subbands are computed for the three most localized window widths,  $1 \leq W_n \leq 3$ . These three sub-bands are multiplied together in order to further enhance the edges while reducing noise. The multi-scale product of the first three scales is given by

$$PS_{H}(x,y) = \prod_{n=1}^{3} CS_{W_{n},H}(x,y)$$
(3.9)

$$PS_V(x,y) = \prod_{n=1}^{3} CS_{W_n,V}(x,y)$$
(3.10)

The gradient magnitude map  $\nabla I$  and angle map  $\alpha$  are defined as

$$\nabla I(x,y) = \sqrt{\mathrm{PS}_{H}^{2/3}(x,y) + \mathrm{PS}_{V}^{2/3}(x,y)}$$
(3.11)

$$\alpha(x,y) = \arctan \frac{\mathrm{PS}_{V}^{1/3}(x,y)}{\mathrm{PS}_{H}^{1/3}(x,y)}$$
(3.12)

An edge is therefore detected at location (x, y) if  $\nabla I(x, y)$  has a local maximum in the direction perpendicular to that of the gradient vector, given by  $\alpha(x, y)$ . The overall multi-scale edge detection algorithm is summarized in Fig. 3.1. Fig. 3.2 shows a simulated 1-D data set I of a noisy step edge signal and its corresponding 1-D similarity signals at three successive scales,  $G_1$ ,  $G_2$ ,  $G_3$  and corresponding composite similarity signals,  $CS_1$ ,  $CS_2$ ,  $CS_3$  as well as the final gradient magnitude signal  $\nabla I$ . The input signal is an ideal step edge in additive white Gaussian noise of zero mean and standard deviation of 1 yielding an SNR of 15 dB.

As observed, the similarity coefficients at the finest scale,  $G_1$ , shows a good localization of the step transition at the centre of the waveform but contains noise. The edge response increases from  $G_1$  to  $G_3$  but noise response remains high. The composite outputs represent coefficients with minimum magnitudes over J = 10adjacent decomposition scales. Noise is significantly reduced while edge localization and magnitude is preserved. Finally,  $\nabla I$  represents the geometric mean of the three

```
Initialize the MMPM outputs PS_H(x, y) and PS_V(x, y) to 1's for all x and y
Loop from j = 1 to 3
   {
     Initialize the composite outputs CS_{W_i,H}(x,y) and CS_{W_i,V}(x,y) to very large
     values \rightarrow \infty
     Loop from W_n = j to J + j - 1
        {
           From (3.3) and (3.4) construct the 1-D separable low pass and
           high pass filters with filter width W_n at scale n
           Normalize both filters to unit magnitude
           Compute G_{W_n,H} and G_{W_n,V} by convolving input image I with the 1-D
           separable filters
           Update CS_{W_i,H}(x,y) and CS_{W_i,V}(x,y) for all x and y:
             if |CS_{W_i,H}(x,y)| > |G_{W_n,H}(x,y)|
                {
                     new |CS_{W_i,H}(x,y)| = |G_{W_n,H}(x,y)|
                 }
             if |CS_{W_j,V}(x,y)| > |G_{W_n,V}(x,y)|
                {
                    new |CS_{W_j,V}(x,y)| = |G_{W_n,V}(x,y)|
                }
        }
     Update PS_H(x, y) and PS_V(x, y) for all x and y:
     New PS_H(x, y) = PS_H(x, y) * CS_{W_i, H}(x, y)
     New PS_V(x,y) = PS_V(x,y) * CS_{W_i,V}(x,y)
   }
```

From (3.11) and (3.12), compute  $\nabla I(x, y)$  and  $\alpha(x, y)$ .

Figure 3.1: Pseudo code for the proposed MMPM scheme.



Figure 3.2: The noisy step signal I with the corresponding similarity signals  $G_{1\to3}$ , composite similarity signals  $CS_{1\to3}$  and gradient magnitude signal  $\nabla I$ : (a) I. (b)  $G_1$ . (c)  $G_2$ . (d)  $G_3$ . (e)  $CS_1$ . (f)  $CS_2$ . (g)  $CS_3$ . (e)  $\nabla I$ .

composite signals and achieves a slight improvement in noise suppression compared to the composite signals.

The peak in the composite sub-band output corresponds to the step edge which is prominent and localized due to the use of the fine scale filter at this location. Noise reduction is achieved by applying coarse scale filters at the other locations. A larger number of decomposition scales will give a better noise reduction. The edge detection performance is also influenced by the type of filter used and the noise level.

### 3.3 Multi-Scale Edge Detection Criteria

A set of criteria is provided to objectively evaluate the performance of three multiscale product based methods, MMPM, MWPM and MPM (similar to MMPM except that J = 1). The set of criteria is also used to compare the performances of the DOB, RMP and HBT filters against the GFD filter in the proposed MMPM.

Canny has proposed the use of SNR, edge localization and multiple false edge response criteria to evaluate the performance of step edge filters which are continuous at x = 0 and defined for a fixed scale. Here, the SNR, edge localization and multi-scale false edge responses criteria are redefined in the multi-scale domain for a wider range of filters that are not necessarily continuous at x = 0 e.g. DOB. A performance index is formulated for each criterion based on the MMPM output PS (1–D case). The multi-scale SNR, M-SNR, and localization, ML, indices are computed from the MMPM output of an ideal 1–D step edge signal immersed in additive white Gaussian noise while the multi-scale false edge responses, MFER, index is computed from a pure additive white noise input. Since all three indices are computed in the presence of white noise, a statistical average of each performance index is obtained from over 1000 realizations of noise records.
#### 3.3.1 Multi-scale SNR, M-SNR

The M-SNR value is directly computed from the statistical average of the ratio of the MMPM step edge response to that of the filter response over:

M-SNR = 
$$\frac{1}{N_n} \sum_{i=1}^{N_n} \frac{\sqrt[3]{|PS_i(0)|}}{\frac{1}{2N_s+1} \sum_{x=-N_s}^{N_s} \sqrt[3]{|PS_i(x)|}}$$
 (3.13)

where  $0 \leq \text{SNR} \leq 1$ ,  $2N_s + 1$  represents the number of discrete samples in the noise immersed step edge signal,  $N_n$  is the number of realizations of noise records indexed by i and  $\text{PS}_i(0)$  is the 1–D step edge response corresponding to the  $i^{th}$  noise record.

#### 3.3.2 Multi-scale Localization, ML

The localization index ML is the distance between the expected local maxima corresponding to the step edge at x = 0 and the actual local maxima observed in the MMPM output nearest to x = 0. The local maxima in the MMPM output, PS, are required to have a magnitude exceeding an automatically determined binary threshold to ensure that valid edges are detected. The threshold is obtained using Otsu's method [70]. The index ML is formally defined as

$$\mathrm{ML} = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{x_i} \tag{3.14}$$

where  $\delta_{x_i}$  is the distance between the input step edge and the detected edge for the i<sup>th</sup> realization of noise. The larger ML is, the poorer the localization and vice versa.

#### 3.3.3 Multi-scale false edge responses, MFER

The multi-scale false edge responses index MFER is defined as the average distance between two adjacent noise maxima in the MMPM output. The larger MFER is, the lower the number of false edge responses in the vicinity of the actual edge and vice versa. The formal definition is [17]

$$MFER = \frac{1}{N_n} \sum_{i=1}^{N_n} \frac{2\pi}{\arccos\left(-\rho_i\right)}$$
(3.15)

where MFER  $\geq 2$  since  $-1 \leq \rho_i \leq 1$  and  $0 \leq \arccos(-\rho_i) \leq \pi$  given that

$$\rho_i = \frac{\operatorname{cov}(X_i, Y_i)}{\sqrt{\operatorname{var}(X_i)\operatorname{var}(Y_i)}}$$
(3.16)

 $X_i$  and  $Y_i$ , corresponding to the i<sup>th</sup> instance of noise, have zero mean and similar variance:

$$X_{i}(x) = \sqrt[3]{\text{PS}_{i}(x)} - \sqrt[3]{\text{PS}_{i}(x-1)}$$
(3.17)

$$Y_i(x) = \sqrt[3]{PS_i(x)} - \sqrt[3]{PS_i(x+1)}$$
(3.18)

where the input signal is a Gaussian white noise. The average distance MFER  $\rightarrow \infty$ when  $X_i$  and  $Y_i$  are increasingly negative correlated ( $\rho_i \rightarrow -1$ ). It approaches the lower bound (MFER  $\rightarrow 2$ ) when  $X_i$  and  $Y_i$  are increasingly positive correlated ( $\rho_i \rightarrow 1$ ).

## **3.4** Experiments

The M-SNR, ML and MFER performance of the proposed MMPM method are compared against the classical multi-scale product method (MPM) under different levels of decomposition and noise. The performance of the Canny, DOB, hyperbolic tangent and ramp filters are also observed for both these methods under the aforementioned conditions.

#### 3.4.1 M-SNR performance

MMPM is applied to a step edge signal corrupted by Gaussian noise with an SNR of 10 dB. Fig. 3.3 shows the output M-SNR corresponding to the various filters applied in MMPM for different values of J. As observed, the RMP filter gives the best M-SNR performance followed closely by HBT, DOB and GFD in that order.



Figure 3.3: The M-SNR performance of MMPM for different J.

Setting J = 1 in MMPM is equivalent to applying the classical multi-scale product method, MPM, as it only involves the product operation of MMPM. The proposed MMPM clearly gives better M-SNR performance than MPM for values of J > 1. Fig. 3.4 shows that the MMPM using the RMP filter at a decomposition level of J = 10 gives significantly better output M-SNR for different levels of input SNR.



Figure 3.4: Comparison between the M-SNR performances of MMPM and MPM for different input SNR levels.

#### 3.4.2 ML performance

Both MMPM and MWPM are based on the point-wise correlation of the similarity coefficients across a given range of scales. Therefore, good localization can only be achieved in the point-wise scheme if all local maxima in the finest decomposition scale propagate to the coarser scales within half a pixel from their location. However, it is observed that MWPM is susceptible to edge drift when the low pass filter used in the non-decimated wavelet transform algorithm is of even length. Table 3.1 compares the extent of drift between the proposed MMPM and MWPM when tested on an ideal step edge signal.

The results indicate the displacement at a particular decomposition scale of the local maxima obtained, by either method, from the actual local maxima of the step edge. As observed in Table 3.1, MWPM suffers from edge drift when the low pass filter used has an even length. Conversely, the local maxima in MMPM remains within half a pixel of the actual edge for all possible lengths of low and high pass filters.

MWPM							
Filter	length	Scale $n$					
Low pass filter, $h_n$	High pass filter, $g_n$	1	2	3	4	5	
Even	Odd	-0.5	0.0	0.5	1.0	1.5	
Even	Even	0.0	0.5	1.0	1.5	2.0	
Odd	Odd	-0.5	-0.5	-0.5	-0.5	-0.5	
Odd	Even	0.0	0.0	0.0	0.0	0.0	
MMPM							
Filter	Window $W_n$						
Low pass filter, $h_{W_n}$	High pass filter, $g_{W_n}$	1	2	3	4	5	
Even	Odd	-0.5	0.0	-0.5	0.0	-0.5	
Even	Even	0.0	-0.5	0.0	-0.5	0.0	
Odd	Odd	-0.5	-0.5	-0.5	-0.5	-0.5	
Odd Even		0.0	0.0	0.0	0.0	0.0	

Table 3.1: Comparison of the extent of drift in local maxima (in pixels) between (a) MWPM non-decimated wavelet transform scheme and (b) proposed MMPM.

Fig. 3.5 shows the localization performance of the MMPM for the four different filters under different decomposition levels. The MMPM is applied to a noisy step edge signal with an SNR of 10 dB.

As observed, the localization measure ML defined in Section 3.3.2 is approximately similar for all four cases with GFD performing slightly better, followed by DOB, HBT and RMP. The localization measure of about 0.5 pixel is achieved for all cases, which is obviously small enough to accurately locate the step edge and is largely insensitive to the type of filter used.



Figure 3.5: Localization measure ML of the MMPM for four different filters under different J.

Fig. 3.6 shows the localization measure ML of the MMPM at J = 10 for four different filters when applied to the step edge signal with different levels of additive noise. The relative filter performance is the same as before except that the difference in their localization measure becomes more distinct for SNR values below 5 dB.



Figure 3.6: Localization measure ML of the MMPM for four different filters under different levels of noise.

#### 3.4.3 MFER performance

Fig. 3.7 shows the average MFER distances between adjacent noise maxima for the four different filters under different decomposition levels of MMPM. The larger the distance, the fewer is the number of false edge responses. All four filters give a very similar number of false responses in MMPM, with GFD performing slightly better followed by DOB, HBT and RMP, in that order. MMPM gives better results compared to the classical multi-scale product approach (J = 1) for decomposition levels J = 2, ..., 7 in the case of GFD, and J = 2 for DOB, RMP and HBT. The distance measure MFER becomes approximately constant for all four filters after a decomposition level of MFER = 14.



Figure 3.7: Multiple false edge response measure MFER of the MMPM for four different filters under different J.

#### 3.4.4 Overall performance

The overall performance of MMPM on a synthetic image is evaluated under different levels of noise, types of filters, and decomposition level J. The original synthetic image is a noiseless  $128 \times 128$  binary image as shown in Fig. 3.8 where the white square of size  $64 \times 64$  is located at the center of the image.



MMPM finds the local maxima in the image via (3.11) and (3.12). Since it is known *a priori* that approximately 1.5% of the image comprises edge pixels, a threshold is imposed such that only local maxima with gradient magnitude, from (3.11), that fall within the largest 1.5% of all pixels in the image are admitted as edge pixels. The Pratt's figure of merit (F) 3.19 is then applied to evaluate the accuracy of the edges detected against the ideal edge map,

$$F = \frac{1}{\max(E_I, E_D)} \sum_{i=1}^{E_D} \frac{1}{1 + \alpha d^2(i)}$$
(3.19)

where  $E_I$  and  $E_D$  are the number of ideal and detected edge pixels, respectively, d(i) is the Euclidean distance between the  $i^{th}$  edge pixel detected and the ideal edge pixel nearest to it, and  $\alpha$  is a scaling constant set to 1.

In Fig. 3.9, F of the GFD filter is computed for input SNR levels ranging from -10 dB to 20 dB at different J. In Fig. 3.9, it is observed that the F increases for input SNR values ranging from -10 dB to 5 dB and then saturates at approximately 0.75, for all values of J. For input SNR values below 5 dB, F increases for larger number of decomposition levels J. A similar trend is observed for the other three filters. The proposed MMPM clearly enables a higher F for low input SNR values compared to the classical point-wise product based method corresponding to J = 1.





Figure 3.9: F for MMPM edge map, obtained using GFD, as a function of input SNR at different J.

In Fig. 3.10, F of the GFD, DOB, HBT and RMP filters is computed for input SNR levels ranging from -10 dB to 20 dB at J = 10. It is observed that all four filters show similar F values for input SNR values greater than 5 dB at J = 10. However, F values of RMP and HBT filters and to a lesser extent the DOB filters are markedly better than GFD.



Figure 3.10: F for MMPM edge map as a function of input SNR for the four filters with J = 10.

The 2-D edge detection results in Figs. 11 and 12 corroborate the findings for the 1-D case in Figs. 3.9 and 3.10, respectively. Fig. 3.11(a) is an axial MR

image of a head scan. The image is corrupted by additive white Gaussian noise and has an overall SNR of 15 dB. Figs. 3.11(b) - (d) show the edge detection results of the proposed MMPM, using the GFD filter corresponding to J = 1, 5and 10 respectively. The edges are found by first scaling the gradient magnitude map from (3.11) such that all coefficients lie between 0 and 1 and then detecting all local maxima that exceed an empirically determined threshold of 0.1. It is observed that the number of false edges reduces while the valid edges are retained as the number of decomposition levels increases.





Figure 3.11: MMPM 2-D edge map as a function of J (a) MR image from an axial head scan (15 dB). MMPM edge map for (b) J = 1. (c) J = 5. (d) J = 10.

Fig. 3.12 compares the performance of the four filters on the MR image of Fig. 3.11(a). The RMP filter provides the best edge detection accuracy as it detects the least false edges followed by the HBT, DOB and GFD filters in that order while all four filters give comparable detection of valid edges.





Figure 3.12: MMPM edge maps for MR image from Fig. 3.11(a) at J = 5 and filter (a) GFD. (b) DOB. (c) HBT. (d) RMP.

Fig. 3.13 compares the performance of the proposed MMPM against the fixed scale scheme. It is observed that only strong edges are detected at the coarse scale corresponding to  $W_n = 6$  while noise and fine edges are suppressed whereas both noise and fine edges are retained together with the strong edges at the fine scale corresponding to  $W_n = 1$ . The proposed MMPM effectively combines the fixed scale results to accurately retain both strong and fine edges while suppressing noise.



Figure 3.13: Comparison between edge detection results of MR image from Fig. 3.11(a) using GFD filter for (a) fixed scale with  $W_n = 2$ ,  $\sigma = 2$ . (b) fixed scale with  $W_n = 1$ ,  $\sigma = 0.3$ . (c) multi-scale with J = 3 (combining scales from  $W_n = 1 \rightarrow 6$ ).

# 3.5 Discussion

The SNR performances of the four filters correlate with their expected performances in the fixed scale case, i.e., RMP and HBT filters have significantly better SNR performance than DOB and GFD. As in the fixed scale case, the SNR performance of the filters in the multi-scale domain is also related to the frequency bandwidth of the filter in that a narrower bandwidth results in better noise suppression and vice versa. MMPM performs better than MPM in noise suppression since the output SNR increases as the number of levels (J) is increased. This is attributed to the *min* operator which further suppresses the noise for larger values of J while preserving the valid edges.

Unlike MMPM, MWPM suffers from edge drift when the low pass filter is of even length. The local edge maxima drifts away from the desired location by half a pixel for every decomposition level. This phenomenon is due to the discrete wavelet transform (DWT) algorithm, where a drift of half a pixel occurs everytime an approximate sub-band is computed by convolving the approximate sub-band from the previous decomposition level with a low pass filter of even length. Both MMPM and MPM have comparable localization performances and the number of levels J have little influence on localization. All four filters in the MMPM method also have similar localization performance with the GFD filter performing incrementally better than the other filters for very low input SNR values around 0 dB.

The multiple false edge response performances among all four filters are comparable. This is in contrast to the fixed scale domain where the GFD performs significantly better than the other three filters. The similarity in performance is due to the effective suppression of noise maxima via the *min* operator, thus increasing the average distance between adjacent noise maxima.

MMPM gives a better overall performance compared to MPM while the overall performance of the RMP and HBT filters is better than DOB and GFD for SNR values below 5 dB. However, the performances of MMPM and MPM as well as there of the four filters become comparable for larger SNR values. The MMPM method combined with the RMP or HBT filters gives the best overall performance as its SNR performance is significantly better while its localization and multiple false edge response performance is similar to the other combinations.

# 3.6 Conclusion

The edge response, localization and multiple false edge response criteria have been effectively redefined in the multi-scale domain. They allow the comparison of the multi-scale performance of step edge detection filters which were previously compared at a fixed scale. A robust multi-scale min-product based method, MMPM, has also been presented for step edge detection. In addition to Canny and the quadratic spline filters, a wider range of edge filters can be applied in the proposed method. MMPM gives better noise suppression compared to the classical MWPM. Unlike MWPM, MMPM does not suffer from the problem of edge drift over successive scales and gives better suppression of Gaussian noise compared to MWPM. The four filters considered here give an approximately similar performance for edge localization and number of false edge responses in the multi-scale domain. However, a marked difference is observed in their SNR performance with both RMP and HBT showing significantly better results compared to the other two.

The choice between the fixed-scale method of Chapter 2 and the multi-scale method in this Chapter is largely dependent on the noise levels and difference in magnitude between strong and weak edges. In general, both fixed-scale and multiscale methods work well on images with low to medium levels of noise. However, the proposed multi-scale method is also effective for detecting edges in images with high levels of noise. Conversely, the fixed-scale method shows better robustness compared to the multi-scale method when there is a large difference in magnitude between strong and weak edges.

# Chapter

# A Rule-Based Approach for Robust Clump Splitting

Chapters 2 and 3 of this thesis explore robust techniques for accurately detecting edge pixels of cell boundaries. However, these boundaries may be attributed to clumps comprising two or more cell regions of interest. In this Chapter, a robust rule-based approach is presented for the splitting of binary clumps that are formed by objects of diverse shapes and sizes. First, the deepest boundary pixels, i.e., the concavity pixels in a clump, are detected using a fast and accurate scheme. Next, concavity-based rules are applied to generate the candidate split lines that join pairs of concavity pixels. A figure of merit is used to determine the best split line from the set of candidate lines. Experimental results show that the proposed approach is robust and accurate.

# 4.1 Rationale

The clumping together of objects of interest is a common phenomenon in a wide variety of image data, e.g., cytological [8, 14, 69, 93] and remotely sensed images [45]. Although a human operator may be able to detect the constituent objects of interest based on prior knowledge and perception of texture and structure, it is difficult for a computer-based algorithm to do this automatically. This poses a problem if the aim is to label these objects correctly and perform a population count of each class. The splitting of clumps into constituent objects is thus a vital step that must be performed accurately to ensure the overall success of the vision task.

Clump splitting methods that are available include binary erosion [2, 69, 86, 88] watershed techniques [5], model-based approaches [14, 26, 40, 94] and concavity analysis [8, 9, 21, 41, 59, 91, 51, 93]. A difficulty with erosion-based methods is that they may completely erode a constituent object in a clump before a split occurs. Watershed techniques tend to over-split the clumps. Model-based approaches [14, 40, 94, 26] besides being computationally expensive, require initialization of the model parameters [40].

Concavity analysis methods offer an intuitive way of clump splitting. Such methods have been successfully implemented in a variety of application domains such as cervical cancer cells [93], plant cells [21], chromosomes [59], and crushed aggregates [91], to name a few. However, tests have shown that these methods are only applicable for objects of specific sizes and shapes. Wang [91] reported 90% accuracy in splitting clumps comprising overlapping convex and compactly shaped objects. Fernandez et al. [21] assumed that the grey level variation along the split line was minimal. This may be true for images in a particular application domain but is not generally valid. Liang [59] implemented a scheme for splitting chromosomes that reportedly worked well but required heuristics incorporating shape and grey level information. The method is thus not sufficiently general for splitting other types of clumps.

The clump-splitting method proposed in this thesis [50] addresses the aforementioned drawbacks. It enables the accurate splitting of clumps composed of objects of different sizes, shapes and degrees of overlap. It is a general method that can be applied to a wide variety of application domains. This is achieved via the implementation of a set of features that guide each decision to split the clump. First, the concavity pixels<sup>1</sup> are detected using a fast and accurate scheme. Next, candidate split lines are selected from the set of all possible lines joining any two concavity pixels. A candidate split line is one that connects two concavity pixels that are close together and lie in concavity regions that are appropriately aligned with respect to each other. A candidate split line could also connect a concavity pixel with a non-concavity boundary pixel on the clump contour if the binary clump has only one concavity region, or if no candidate split line can be found. Finally, a figure of merit is introduced to determine the best split line from the set of candidate lines.

A review of recent concavity analysis methods in Section 4.2 provides the background. Sections 4.3 and 4.4 give an overview of the proposed method and define the features used for detecting concavity pixels and candidate split lines. In Section 4.5, the size-invariant feature used for selecting the best split line from the set of candidate split lines is described. Section 4.6 presents training and implementation details. Section 4.7 evaluates the performance of the algorithm on unseen data while Section 4.8 compares its performance against another method and validates

<sup>&</sup>lt;sup>1</sup>The pixel on the boundary arc (Fig. 4.1) that has the largest perpendicular distance from its corresponding convex hull chord.

each of the features used. Section 4.9 concludes this thesis.

# 4.2 Review of Concavity Analysis for Clump Splitting

In methods based on concavity analysis, a clump is split by the line joining two concavity pixels on the clump contour. These methods vary with respect to the technique for locating the concavity pixels and the cost function used to detect a split path. In general, there are three sequential steps: detection of concavity regions, detection of candidate split lines and selection of best split line. The best split line is obtained recursively until a specific stopping criterion is met.

#### 4.2.1 Detection of concavity regions or concavity pixels

This step detects regions or pixels along the boundary where the degree of concavity is high. Such regions or pixels are regarded as valid concavity regions or pixels. Yeo et al. [93] define a concavity region,  $S_i$ , as any region bounded by a boundary arc  $B_i$  and its corresponding convex hull chord (Fig. 4.1). A concavity region is taken to be valid if its concavity degree, DG<sub>i</sub>, and normalized concavity weight, WT<sub>i</sub>, exceed their respective threshold values:

$$DG_i = |B_i|/|K_i|, DG_i > DG_T,$$

$$(4.1)$$

$$WT_i = |B_i|/|B_{max}|, WT_i > WT_T,$$

$$(4.2)$$

where  $|\cdot|$  denotes length and  $|B_{max}|$  is the length of the longest boundary arc in the clump. However, the use of thresholds  $DG_T$  and  $WT_T$  removes valid concavity regions when  $B_{max}$  and  $K_i$  are unusually large. Fernandez et al. [21] and Liang [59] used concaveness measures to identify concavity pixels. These measures place more emphasis on the sharpness of the region surrounding the concavity pixel rather than on its depth (measured by the distance of the concavity pixel from the convex hull). Consequently, their definition often leads to the detection of invalid concavity regions.

Wang applies a polygonal approximation method followed by corner detection to find the concavity regions [91]. The polygonal approximation, however, results in distortion to the clump contour and the natural shape of the constituent objects.



Figure 4.1: Binary clump with convex hull chords  $K_1$ ,  $K_2$  and  $K_3$  and corresponding boundary arcs,  $B_1$ ,  $B_2$  and  $B_3$ .

#### 4.2.2 Detection of candidate split lines

This step detects candidate split lines from all possible lines joining any two concavity regions. Yeo et al. considers a line joining two concavity regions to be a valid split line if its length is less than or equal to those between any two pixels that are immediately adjacent to the pixel pair at the ends of the split line [93]. This approach is computationally expensive and results in some incorrect splitting due to boundary irregularities.

Wang [91] requires the concavity regions at the two ends of the split line to be "oppositely aligned" to each other. Given a concavity region, he defines another concavity region to be in opposite alignment to the first one if the second region lies within the cone (grey triangular region in Fig. 4.2) obtained by extending two vertex lines from the first concavity region. The example in Fig. 4.2 illustrates a situation where his method fails to identify the line connecting concavity pixels  $CV_1$  and  $CV_2$  as a candidate split line.



Figure 4.2: Wang's opposite alignment criterion (from Ref. [91]).

Both Fernandez et al. [21] and Liang [59] require candidate split lines to connect two concavity pixels such that the distance between the two pixels and the intensity variation along the split line are below predefined thresholds. Their use of a distance threshold implicitly assumes that the objects of interest have similar sizes and shapes but this may not be true in many domains. The threshold on intensity variation also fails if the objects of interest are textured.

#### 4.2.3 Selection of best split line

The best split line is selected from the set of candidate split lines. Generally, in all concavity analysis methods, the line that maximizes a predefined cost function is selected as the best split line from a set of candidate split lines. The methods vary in the choice of this cost function.

Yeo et al. [93] and Wang [91] define the best split line to be the shortest of all the candidate lines but do not impose a maximum distance threshold. False splitting may therefore arise if the distance is not small enough to warrant a split.

The refinement proposed by Fernandez et al. [21] and Liang [59] imposed two conditions for a split: (i) the "concaveness" at each end of the split line exceeds a set threshold TH<sub>1</sub>, and (ii) the length of the split line is less than another threshold TH<sub>2</sub>. Each pair of concavity pixels in a clump can be represented by a point in the 2–D space defined by the above two features (Fig. 4.3). The decision boundary obtained by using thresholds TH<sub>1</sub> and TH<sub>2</sub> (dashed lines) will lead to false splitting<sup>2</sup> and under splitting<sup>3</sup> if this partitioning does not conform to the underlying data. In Section 4.6, it is shown that the effective separation of the split and no-split cases for the image data requires a straight line. Hence the use of two thresholds will result in both false splitting (region FS) and under splitting (regions US<sub>1</sub> and US<sub>2</sub>) where the accurately split feature subspace (region AS) is smaller than the ideal case (US<sub>1</sub>  $\cup$  US<sub>2</sub>  $\cup$  AS).

# 4.3 Overview of Methodology

The proposed algorithm [50] splits a binary clump into two smaller clumps and repeats the process on each of them until no more split lines can be detected. The method begins with the detection of concavity pixels in a clump. The boundary arcs  $B_i$  and the convex hull segments  $K_i$  of the clump (Fig. 4.1) are first obtained

<sup>&</sup>lt;sup>2</sup>Instances when an incorrect split line is made.

<sup>&</sup>lt;sup>3</sup>Instances when a correct split line is not made.



Figure 4.3: Feature space of length of split line vs. concaveness. *Dashed line* - decision boundary obtained by using two separate thresholds. *Solid line* - correct decision boundary.

using the methods from [30] and [42]. On each boundary arc,  $B_i$ , the pixel with the largest perpendicular distance from the corresponding convex hull,  $K_i$ , is selected as the concavity pixel,  $CV_i$ .

Next, a set of candidate split lines is selected from the set of split lines obtained by joining all possible pairs of concavity pixels. A pair of concavity pixels  $CV_i$  and  $CV_j$  forms a candidate split line if: (i) the pixels are in close proximity, (ii) they are located in high concavity regions, and (iii) their concavity regions are suitably aligned. A cost function is also introduced for determining the best split line from the set of candidate split lines. The cost function is obtained from a linear classifier and combines the concaveness information at the ends of the split line as well as the length of the line.

# 4.4 Detecting Candidate Split Lines

This Section describes a set of features for detecting concavity pixels and candidate split lines.

#### 4.4.1 Concavity depth

Concavity depth, CD, is a concaveness measure proposed by Rosenfeld [79]. Each pixel on a boundary arc has a concavity depth value equal to its perpendicular distance from the corresponding convex hull segment,  $K_i$ . For each boundary arc,  $B_i$ , the concavity pixel,  $CV_i$ , is defined to be the pixel with the largest concavity depth,  $CD_i$  (Fig. 4.4), provided  $CD_i$  exceeds a threshold  $CD_T$ . In experiments, the threshold  $CD_T$  has a fixed value that is typically small since the aim is only to distinguish concavity pixels from boundary irregularities.



Figure 4.4: Binary clump with concavity pixels,  $CV_1$  and  $CV_2$ , and corresponding concavity depths,  $CD_1$  and  $CD_2$ .

#### 4.4.2 Saliency

The set of possible split lines obtained by joining the concavity pixels is culled with the help of a feature called "saliency", SA. Split lines are more likely to be valid if the concavity regions at both ends of the line have large concaveness measures and the distance between the two regions is small. The saliency of a split line joining a pair of concavity pixels,  $CV_i$  and  $CV_j$ , is defined as

$$SA_{ij} = \frac{\min(CD_i, CD_j)}{\min(CD_i, CD_j) + d(CV_i, CV_j)},$$
(4.3)

where  $\min(\text{CD}_i, \text{CD}_j)$  is the smaller of the two concavity depths  $\text{CD}_i$  and  $\text{CD}_j$  and  $d(\text{CV}_i, \text{CV}_j)$  is the distance between concavity pixels  $\text{CV}_i$  and  $\text{CV}_j$ . The value of  $\text{SA}_{ij}$  lies within the range 0 to 1. Each candidate split line is required to have a saliency measure,  $\text{SA}_{ij}$ , that exceeds a threshold,  $\text{SA}_T$ .

#### 4.4.3 Alignment

Using only saliency is not sufficient for selecting candidate split lines. Fig. 4.5 illustrates a situation where the saliency  $SA_{12}$  is high but line  $CV_1CV_2$  should not be considered a candidate split line since the concavity regions  $S_1$  and  $S_2$  are not "oppositely aligned". The term opposite alignment as defined variously in [93, 59, 91] is domain specific. A generic definition is offered by making use of the features concavity-concavity alignment (CC) and concavity-line alignment (CL). Figure 4.5(b) shows a clump with two concavities  $S_1$  and  $S_j$ . The orientation of



Figure 4.5: Alignment (a) Clump comprising three overlapping specimens. (b) Concavity-concavity alignment, CC and concavity-line alignment, CL.

concavity  $S_i$ , denoted by the unit vector  $\mathbf{v}_i$ , is defined by the line joining the midpoint of its convex hull chord to concavity pixel  $CV_i$ . The concavity-concavity alignment  $CC_{ij}$  is the angle that represents the relative orientation between the pair of concavity regions,  $S_1$  and  $S_j$ .  $CC_{ij}$ , defined by

$$CC_{ij} = \pi - \arccos(\mathbf{v}_i, \mathbf{v}_j), \tag{4.4}$$

has the minimum value of 0 when the concavities are directly facing each other and the maximum value of  $\pi$  when they are oriented in the same direction. A small value of  $CC_{ij}$  indicates a good split line. Concavity-line alignment,  $CL_{ij}$ , is a measure of the difference in directions of the two concavity regions  $S_i$  and  $S_j$ with respect to the split line (whose direction is denoted by the unit vector  $\mathbf{u}_{ij}$ ) connecting them.  $CL_{ij}$  is defined by the larger of the two angles  $\phi_i$  and  $\phi_j$ :

$$CL_{ij} = \max(\phi_i, \phi_j) = \max(\arccos(\mathbf{v}_i, \mathbf{u}_{ij}), \arccos(\mathbf{v}_j, -\mathbf{u}_{ij})), \qquad (4.5)$$

where  $\phi_i$  is the angle between  $\mathbf{v}_i$  and  $\mathbf{u}_{ij}$  and  $\phi_j$  the angle between  $\mathbf{v}_j$  and  $-\mathbf{u}_{ij}$ .  $\mathrm{CL}_{ij}$  is the larger of the angles  $\phi_i$  and  $\phi_j$  since a more conservative estimate for the concavity-line alignment is preferable. Angles  $\phi_i$  and  $\phi_j$  are small if the concavities are well aligned with the split line and large if the directions of the two concavities are distinctly different from their split line. A small value for each angle is indicative of a good split.

For the clump in Fig. 4.5(a), the candidate split lines are most likely to be the lines joining concavity regions  $S_1$  and  $S_3$  as well as  $S_2$  and  $S_4$  since both these concavity pairs have very small values of CC and CL (approximately ). The concavity regions  $S_1$  and  $S_4$  exhibit good opposite alignment (CC<sub>14</sub>  $\approx$  0) but poor concavity-line alignment (CL<sub>14</sub>  $\approx \pi/3$ ) whereas concavity regions  $S_1$  and  $S_4$  are neither oppositely aligned (CC<sub>14</sub>  $\approx \pi$ ) nor laterally aligned (CL<sub>14</sub>  $\approx \pi/2$ ).

In the ideal case, a pair of concavity regions  $S_i$  and  $S_j$  is considered to be perfectly aligned if both  $CC_{ij}$  and  $CL_{ij}$  are 0. However, since candidate split lines connect concavities that are generally misaligned to some extent, mandatory conditions are imposed, namely,  $CC_{ij} < CC_T$ , and  $CL_{ij} < CL_T$ , where  $CC_T$  and  $CL_T$  are preset thresholds.

#### 4.4.4 Concavity angle and concavity ratio

Concavity angle CA and concavity ratio CR are the two features used to decide if a clump is to be split along the line joining a concavity pixel  $CV_i$  to a boundary pixel P. Such a split is considered only if no candidate split lines can be found after applying the above mentioned features. As shown in Fig. 4.6, the split line is the line passing through the midpoint of the convex hull chord  $K_i$  and concavity pixel  $CV_i$ , and intersecting the far side of the boundary at P. The concavity angle,



Figure 4.6: Concavity angle, CA and concavity ratio, CR.

CA and concavity ratio, CR, are defined respectively by

$$CA = \angle C_{i1} C_i C_{i2}, \tag{4.6}$$

$$CR = CD_m/CD_n, (4.7)$$

where  $CD_m$  and  $CD_n$  represent the largest and second largest concavity depths, respectively. CA is a measure of the sharpness of a concavity region and CR a measure of the size of the major concavity (depth  $CD_m$ ) relative to the other concavities in the clump. The second largest concavity depth,  $CD_n$ , assumes the value of the concavity depth threshold ( $CD_T$ ) in Section 4.4.1 if only one valid concavity region is detected. The use of preset thresholds  $CA_T$  and  $CR_T$  ensure that a split is made only if a concavity is sufficiently sharp  $CA < CA_T$  and significantly larger than all the other concavities in the clump  $CR > CR_T$ .

## 4.5 Selecting the Best Split Line

The best split line is selected from the set of candidate split lines. It is a general observation that, apart from satisfying the alignment conditions (Section 4.4.3), valid split lines connect the two concavity pixels  $CV_i$  and  $CV_j$  that are closest to each other and at the same time come from the largest concavities (large  $CD_i$  and  $CD_j$ ). With these considerations, a figure of merit, the "measure of split"  $\chi$ , is proposed:

$$\chi = \frac{c_1 \text{CD}_i + c_1 \text{CD}_j + c_2}{d(\text{CV}_i, \text{CV}_j) + c_1 \text{CD}_i + c_1 \text{CD}_j + c_2},$$
(4.8)

where  $c_1$  and  $c_2$  are appropriate weights.  $CD_i$  and  $CD_j$  are given the same weight,  $c_1$ , since they should have equal influence. It is noted that  $\chi$  lies in the range [0, 1]and

- $\chi \to 1$  when  $d(CV_i, CV_j) \to 0$  or either  $CD_i, CD_j \to 0$
- $\chi \to 0$  when  $d(\mathrm{CV}_i, \mathrm{CV}_j) \to \infty$

A split line is regarded as the best choice if it has the largest  $\chi$  that exceeds the empirically obtained threshold of 0.5, which simplifies (4.8) to

$$d(\mathrm{CV}_i, \mathrm{CV}_j) < c_1(\mathrm{CD}_i + \mathrm{CD}_j) + c_2.$$
(4.9)

It follows from (4.9) that the decision boundary for ascertaining whether a split should be made is a straight line in the 2–D feature space defined by  $d(CV_i, CV_j)$ and  $CD_i + CD_j$ . This is verified experimentally, as described in Section 4.6. The values of weights  $c_1$  and  $c_2$  can be determined using any linear classifier. The linear SVM classifier is used here since it ensures maximum separation between the two subsets (split and no-split classes) and minimizes decision errors [20, 37]. The decision rule of (4.9) can be expressed in the form

$$\mathbf{w}^T \mathbf{z} + b > 0, \tag{4.10}$$

where w is the weight vector, b is the bias and z is the feature vector comprising the Euclidean distance,  $d(CV_i, CV_j)$ , and the total concavity depth,  $CD_i + CD_j$ :

$$\mathbf{z} = [d(\mathrm{CV}_i, \mathrm{CV}_j), \mathrm{CD}_i + \mathrm{CD}_j]^T.$$
(4.11)

# 4.6 Methodology

The appropriate values for the parameters used in the proposed algorithm are first determined and then validated. The training and test data sets comprise binary clumps extracted from microscope images of the five different species of pollen (*Acacia and Podocarpus*), fungal (*Dreschlera and Curvularia*) and fern (*Nephrolepis*)



spores shown in the composite image of Fig. 4.7.

Figure 4.7: Five species of airborne spore specimens used in the experiments.

#### 4.6.1 Training

To ensure robust clump splitting, the training set must contain spores of different sizes and shapes that may overlap with debris or with one another. The process comprises (i) the selection of threshold values of the features defined in Section 4.4, and (ii) the determination of the weight constants  $c_1$  and  $c_2$  of the measure of split,  $\chi$ , defined in Section 4.5.

The threshold values (Table 4.1) are obtained via inspection of the training set and should be effective in detecting the candidate split lines:

- Thresholds  $CD_T$  and  $SA_T$  are chosen to be greater than the CD and SA values arising from minor boundary irregularities.
- Thresholds CC<sub>T</sub> and CL<sub>T</sub> are selected from known cases of correct split lines in the training set where the concavity regions at the ends of each correct split line are well aligned in opposing directions.

- A split line joining a concavity pixel and a boundary pixel will be considered if the clump has only one significant major concavity; this is ensured by having a sufficiently large threshold CR<sub>T</sub>.
- Threshold  $CA_T$  is selected to be greater than the CA values due to the natural concavities of objects such as *Dreschlera* and *Podocarpus*.

 Table 4.1: Threshold values assigned to the features that determine validity of split

 lines.

Rules	Threshold	Value	
$\mathrm{CD} > \mathrm{CD}_T$	$\mathrm{CD}_T$	3	
$SA > SA_T$	$SA_T$	0.12	
$\mathrm{CL} < \mathrm{CL}_T$	$\operatorname{CL}_T$	$70^{\circ}$	
$\mathrm{CC} < \mathrm{CC}_T$	$\mathrm{CC}_T$	$105^{\circ}$	
$CA < CA_T$	$CA_T$	90°	
$CR > CR_T$	$CR_T$	6	

A training set of 1,100 2–D feature vectors, where the features are total concavity depth and distance between a pair of concavity pixels on a clump, is used as the inputs to a linear SVM classifier [37] to determine suitable values for  $c_1$  and  $c_2$ . The data samples were extracted over the five spore species and comprise 100 samples taken from pairs of concavity pixels that form valid split lines and 1,000 samples from pairs that form invalid split lines. (The disparity in sample size is due to a much lower occurrence of the former.) Classification accuracy is computed using threefold cross validation of the training set for different penalty factor values. The results are shown in Table 4.2, where the weight constants  $c_1 = 1.72$  and  $c_2 = -4.70$  give the best generalization performance with an average classification accuracy of 99.5%. Fig. 4.8 shows the decision boundary that optimally separates the training distribution set into the split and no-split classes.

Penalty factor	Accuracy (%)	$c_1$	$c_2$
0.1	99.4	1.73	-5.32
1	99.5	1.73	-4.72
10	99.5	1.73	-4.72
100	99.5	1.72	-4.70
1000	99.5	1.70	-4.54

Table 4.2: Training results for different penalty factor values.



Figure 4.8: Linear decision boundary obtained from the training data set.

#### 4.6.2 Implementation of clump splitting

The clump splitting algorithm [50] is applied to binary clump images. After each round of splitting, the resulting objects are subjected to the same algorithm. This iterative splitting of a clump terminates when the resulting constituent objects do not have any candidate split lines. A split line joining a concavity pixel and a boundary pixel is then attempted if there are no more candidate split lines. The two cases of splitting overlapping regions are, therefore: (i) split line joining two concavity pixels, and (ii) split line joining a concavity pixel and a boundary pixel.

#### Split line joining two concavity pixels

A line is considered a candidate split line if the concavity depths at its two ends,  $CD_i$  and  $CD_j$ , are greater than  $CD_T$ . It must also satisfy the requirements of saliency and alignment, i.e.,  $SA > SA_T$ ,  $CL < CL_T$  and  $CC < CC_T$ . From the set of candidate split lines, the one with the largest measure of split,  $\chi$  ( $\chi > 0.5$ ), is used to split the clump. An exception to the above rule is that the alignment conditions of Section 4.4.3 can be ignored if  $\chi > 0.8$ .

#### Split line joining one concavity pixel and one boundary pixel

A split line joining a concavity pixel  $CV_i$  and a non-concavity boundary pixel P is considered when (i) there is only one concavity pixel, or (ii) there is more than one concavity pixel but no candidate split line can be found. A split is made in the first case if  $CA < CA_T$ , and in the second case if  $CA < CA_T$  and  $CR > CR_T$ .

## 4.7 Performance on Unseen Data

A total of 112 overlapping and 140 non-overlapping binary regions across all five species were involved in the validation of the rule-based approach (henceforth referred to as RBA). These test images were extracted from 8-bit airborne-spore images via the application of edge detection, thresholding and morphological routines.

Some of the splitting results for clumps involving two spore specimens are shown in Fig. 4.9. RBA was also applied to a set of cytological images to demonstrate its versatility. Fig. 4.10 shows the results for some overlapping specimens. The



Figure 4.9: Sample results of splitting clumps comprising two touching spore specimens (not to scale).

splitting of overlapping clumps with only one major concavity region is illustrated in Fig. 4.11. The two overlapping *Dreschlera* specimens have only one major concavity region and are accurately split due to the sufficiently small CA and large CR of the clump. RBA also performs well on large clumps comprising multiple



Figure 4.10: Sample results of splitting clumps comprising two or three touching cytological specimens.



Figure 4.11: Splitting a clump comprising only one dominant concavity region. (a) Two overlapped *Dreschlera* specimens. (b) Split line joining the concavity pixel and a boundary pixel.

objects as shown in Fig. 4.12. Fig. 4.13 shows the accurate splitting of clumps involving specimens of different shapes and sizes. It is also observed that the natural boundaries of the spore specimens sometimes lead to the formation of small concavity regions that are adjacent to one another. The generation of false split lines joining adjacent concavity regions is avoided by the alignment criteria of Section 4.4.3. The overall clump splitting performance is evaluated from the percentages of correct, false, and under splitting (Table 4.3). Considering the diverse sizes and shapes of the specimens, a creditable overall splitting accuracy of 79.5% is obtained. The extremely low percentage of clumps that are falsely





Figure 4.12: Split results of large clumps comprising several specimens. (a) Nephrolepis clump. (b) Nephrolepis clump after splitting. (c) Two large Podocarpus clumps. (d) Podocarpus clumps after splitting.



Figure 4.13: Splitting clumps comprising specimens with different sizes and shapes.(a) Fungal and fern spore. (b) *Nephrolepis* with attached dirt particle.

split (5%) is due to the conditions imposed by SA, CC and CL. The proposed

method works well on both convex (*Acacia* - 100%, *Nephrolepis* - 78.5%) and nonconvex spores (*Podocarpus* - 91%). However, the split accuracies for *Dreschlera* and *Curvularia* are relatively poorer (69.5% and 60%, respectively). The high rate

Species		Clumps			Individual			
	Total	$\operatorname{Correct}(\%)$	False(%)	Under(%)	Total	$\operatorname{Correct}(\%)$	False(%)	
Acacia	6	100.0	0.0	0.0	15	100.0	0.0	
Dreschlera	11	69.5	23.0	7.5	20	100.0	0.0	
Ne phrole pis	29	78.5	0.0	21.5	31	100.0	0.0	
Podocarpus	42	91.0	4.5	4.5	37	100.0	0.0	
Curvularia	24	60.0	6.0	34.5	37	100.0	0.0	
Overall	112	79.5	5.0	15.5	140	100.0	0.0	

Table 4.3: Detailed performance of RBA.

of false splitting for *Dreschlera* (23.0%) is explained by the frequent occurrence of specimens crossing each other (Fig. 4.14). This would result in one of the specimens in a clump to be correctly split at the expense of the other or both specimens to be incorrectly split as observed in Fig. 4.14. Under splitting for *Curvularia* (34.5%) is pronounced because of the removal of some concavities in its clumps by the dilation/erosion operations at the pre-processing stage. This is illustrated in Fig. 4.15, where the concavity regions in 4.15(b) appear smaller than their actual sizes in 4.15(a).


Figure 4.14: False splitting of a clump comprising two *Dreschlera* specimens crossing each other.



Figure 4.15: Reduction in the sizes of the concavity regions in a *Curvularia* clump.(a) Overlapping and individual *Curvularia* specimens. (b) Binary clump of *Curvularia* specimens after dilation/erosion operation.

# 4.8 Performance Comparison and Feature Validation

This Section compares the performances of RBA and the optimal dissection method (ODM) of Yeo et al. [93]. It also validates the importance of the features in RBA by studying the effects on splitting performance when a feature is removed or replaced by another feature from ODM. The following experiments were performed:

• Comparison I - Concavity depth (CD) in RBA vs concavity degree (DG) and normalized concavity weight (WT) in ODM,

- Comparison II Measure of split (χ) in RBA vs Optimal dissection requirement in ODM, and
- Comparison III Effect of removing saliency (SA) and alignment features (CC and CL) in RBA.

The aforementioned comparisons are valid since the features, whose performances are being compared, share a similar role in the concavity analysis scheme. The comparisons are also fair since the other parameters in the scheme remain fixed.

#### 4.8.1 Comparison I

In this experiment, the effects on split accuracy are determined if the concavity pixels, identified using CD in RBA, are detected only from concavity regions which satisfy concavity degree DG and normalized concavity weight WT criteria of ODM with thresholds  $DG_T = 1.15$  and  $WT_T = 0.25$ , respectively. As seen in Table 4.4, the selection of concavity pixels only from these concavity regions results in lower split accuracy of 58.5% compared to 79.5% for RBA. The reason for this is the ineffectiveness of DG and WT in detecting all valid concavity regions, as illustrated in Fig. 4.16 (where the desired split lines are depicted in white). Fig. 4.16(a) shows a binary region of two overlapping *Curvularia* specimens; the concavity pixel in region  $S_a$  is undetected since its corresponding boundary arc is significantly smaller than the longest boundary arc,  $|B_{max}|$ , of the clump. In Fig. 4.16(b), where the clump consists of a *Curvularia* specimen and a long detritus, the concavity pixel in region is undetected since its concavity region lacks sharpness and has a very long convex hull chord,  $|K_i|$ .

		Clumps	Individual		
Experiment	$\operatorname{Correct}(\%)$	False(%)	Under(%)	$\operatorname{Correct}(\%)$	False(%)
RBA	79.5	5.0	15.5	100.0	0.0
Comparison I	58.5	15.0	26.5	99.5	0.5
Comparison II	56.0	28.0	16.0	79.0	21.0
Comparison III	74.0	12.0	14.0	99.5	0.5

Table 4.4: Summary of performance comparison and feature validation results.



Figure 4.16: Shortcomings of concavity measure in ODM. (a) Two overlapping *Curvularia* specimens; concavity region  $S_a$  is not detected. (b) *Curvularia* specimen with overlapping detritus; concavity region  $S_b$  is not detected.

## 4.8.2 Comparison II

In this experiment, the effects on splitting accuracy are determined if the method of selecting the best split line in RBA is replaced by ODM. In the latter, the best split line is the shortest line that satisfies the optimal selection criterion and joins two concavity regions that meet the DG and WT requirements [93]. From Table 4.4, ODM falsely splits the clumps and individual objects 28% and 21% of the time, respectively, compared to the experiment in Comparison I, where they are falsely split 15% and 0.5% of the time, respectively. The higher percentage of false splitting in Comparison II is attributed to the shortcoming of optimal selection criterion in ODM for selecting the best split line. It is sensitive to boundary irregularity and performs poorly on individual objects with large concavities (e.g., *Podocarpus*, Fig. 4.17).



Figure 4.17: False splitting of a *Podocarpus* specimen using ODM.

## 4.8.3 Comparison III

The influence of saliency and alignment is investigated by determining the amount of false splitting when these features are not used. The result is a significantly higher percentage of false splitting (12.0% compared to 5.0%). The lower percentage of false splitting by RBA is due to the validity checks imposed by these two conditions. Fig. 4.18(a) shows an invalid split of overlapping *Curvularia* specimens when these conditions are relaxed as opposed to the accurate split results shown in Fig. 4.18(b) when they are imposed.

## 4.9 Conclusion

A novel approach for splitting binary clumps has been presented using a set of concavity-based rules. The concavity depth CD provides a fast and simple way of detecting concavity pixels. Both CD and saliency SA effectively distinguish concavity pixels that form valid split lines from those that arise due to minor boundary irregularities. In addition, the alignment features, CC and CL, ensure



Figure 4.18: Splitting a clump comprising three *Curvularia* specimens that overlap along their major axes. (a) False splitting when saliency and alignment conditions are removed. (b) Correct splitting when saliency and alignment conditions are imposed.

that the concavity regions at the ends of candidate split lines are suitably oriented with respect to each other. Clumps with only one major concavity region are correctly split with the use of concavity angle CA and concavity ratio CR. By using the measure of split, the best split line can be determined while avoiding the false splitting that often occurs in objects with natural concavities in their boundaries. The proposed method has been shown to be robust by accurately splitting objects of diverse sizes, shapes and degrees of overlap. It has been successfully applied to images comprising objects such as airborne spores and cytological specimens.

# Chapter

# Invariant Texture Classification via Non-Linear Polar Map Filtering

Textural classification of specimens from light microscope images is a daunting problem given the arbitrary orientation, scale and contrast of specimens. A novel texture based classification, which is robust under the aforementioned conditions, is presented for elliptical biomedical specimens in an image. First, the largest ellipse is defined for each segmented specimen from which the textural features are to be extracted. The elliptical region lies completely within the specimen and provides a sufficiently large feature extraction area unlike conventional methods, which define rectangular or circular areas within the specimen. A polar map, which is invariant to the effects of the specimens' orientation and scale, is then constructed from the elliptical region. Non-linear filtering is performed on the polar map to obtain a contrast-invariant similarity map using a  $5 \times 5$  Gaussian first derivative (GFD) filter. Local and global energy measures are extracted from this similarity map to be used in the training and testing of an SVM classifier. Experimental results show that the proposed method achieves an accuracy of over 90% in classifying six species of pollen, fungal and fern spores. The robustness of the method and choice of features are also validated under different orientation, scale and contrast levels.

## 5.1 Rationale

The classification of biomedical specimens is a vital step in studies where the specimens in question could be asthma-causing allergenic spores or malaria-infected red blood cells, to name a few examples [29, 54, 57, 22, 82, 52, 49]. However, manual classification of these specimens via visual inspection of light microscope (LM) images can be a time-consuming exercise. Conversely, automated methods potentially provide fast and reliable classification of these specimens. The texture of the specimen offers vital information to this end provided each class of specimens have their own unique textural properties.

Current methods of texture based classification of specimens in LM images work well under fixed setting of image luminance, scale and contrast but often fail when these are changed. It is impractical to assume that the settings remain fixed since, for instance, image contrast and luminance can change depending on the intensity from the light source, which may deteriorate over time [39].

Current classification methods using linear filtering schemes such as Laws' [55] and wavelet sub-band decomposition [62] are sensitive to luminance and contrast since (1) features extracted from the low frequency (approximation) sub-band are sensitive to luminance changes and (2) the underlying spatial convolution operation emphasizes textured patterns of stronger contrast and suppresses those of weaker contrast although their profiles may be the same.

García-Sevilla [33] has shown that the classification accuracy of features extracted from classical methods such as the gray level co-occurrence matrix (GLCM) [35] and wavelet transform [13] are sensitive to scale. The scale of the specimens in microscope images varies depending on the choice of the objective lense where each magnification ratio, i.e.,  $10 \times$ ,  $20 \times$ ,  $40 \times$  and  $60 \times$  corresponds to a particular scale.

Muneeswaran et al. [66] have used fractal analysis, via wavelet decomposition, to characterize textural regions in their work due to its scale invariance property. However, empirical studies have shown that the fractal dimension is often different at different scales of natural textures, although it may be constant for a range of scales [11].

García-Sevilla [33] also showed that classical methods perform poorly when the spatial orientation of the textured regions is changed. This is a serious limitation since the biomedical specimens, studied in this thesis, are also oriented arbitrarily when viewed under the LM.

Various methods have been proposed to address the effects of both scale and orientation variation of textural regions [95]. Standard wavelet-based algorithms have been modified by combining detail sub-bands [73, 98] or using a set of rotated wavelet filters and multi-channel Gabor filters [28]. However, the performance of these methods degrades when the number of texture classes increases since they are derived from standard filtering methods which are sensitive to orientation.

The circular auto-regressive [44] and the log polar Gabor filters [56] have also been proposed but these methods are computationally intensive especially when the number of classes or size of textural regions increases. More recently, a method combining the log-polar transform and shift invariant wavelet packet transform [76] gave promising results when tested on a set of 25 distinct Brodatz textures [10] with different scale and orientation. The above methods assume that the texture regions are rectangular in shape. Langford et al. [54] identified pollen specimens from SEM scanned images by selecting a rectangular region of approximately 10% of the entire pollen area. Such a small region was representative of the textural pattern since it was manually selected but this will not be the case for an automated texture classification scheme where *a priori* information is not available. The assumption of rectangular regions are often not valid in this case since specimens such as air-borne spores and red blood cells have a general elliptical form with different eccentricity and size.

A texture classification method that is robust under the aforementioned conditions is presented and implemented for the identification of biomedical specimens from LM images. Orientation invariance is achieved by expressing each specimen region in a Cartesian space defined by the major and minor axes of the largest elliptical region within the specimen. Scale invariance is achieved by mapping the elliptical region to a unit circle before constructing the polar map. The non-linear filtering method, from Chapter 2, is applied on the polar map so that the energy measures extracted from the filter output are invariant to contrast. Both *local* and *global* energy measures are extracted to ensure improved accuracy.

The Chapter is organized as follows: Section 5.2 briefly describes the standard polar map transform followed by an overview of the proposed method in Section 5.3. Section 5.4 describes the process of identifying the largest ellipse within each specimen and Section 5.5 describes the use of this ellipse for obtaining orientation and scale invariance. Section 5.6 describes the contrast- and luminance-invariant properties whereas Section 5.7 defines the energy measures used in the classification of the specimens. Section 5.9 discusses results from Section 5.8 on experiments undertaken to validate the invariant properties of the proposed method, choice of texture based features used and overall classification accuracy on a dataset of airborne allergens from six species of fungal, fern and pollen spores. Finally, Section 5.10 concludes this Chapter.

## 5.2 Standard Polar Map Transform



Figure 5.1: Transformation of largest circle within textured image I(x, y) to polar map  $p(\alpha, r)$ .

The polar map transform [3, 76] is used to eliminate the effects of orientation and scale on the accuracy of specimen classification. As illustrated in Fig. 5.1, the largest circle with radius  $\Re$  is first located within the original textured image I of size  $N \times N$ . Next, each pixel (x, y) in the image is mapped from its Cartesian space to the corresponding polar coordinate space where the position of each pixel  $(r, \alpha)$  in the polar map is expressed in terms of its angle  $\alpha$  and distance r from the centre of the circle. The polar map can be formally defined as follows,

$$p(\alpha, r) = I\left(\left[\frac{N}{2} + r\cos\left(\frac{2\pi\alpha}{360^{\circ}}\right)\right], \left[\frac{N}{2} - r\sin\left(\frac{2\pi\alpha}{360^{\circ}}\right)\right]\right)$$
(5.1)

where  $0^{\circ} \leq \alpha \leq 360^{\circ}, 0 \leq r \leq \Re$  and  $[\cdot]$  rounds to the nearest integer. There are

minor variations to the above polar map such as the log polar map,  $\log(r)$  vs  $\alpha$  [76].

There are several drawbacks associated with the above polar map transformation. First, it is observed that any variation in orientation or scale of the specimen region results in a horizontal or vertical shift of the polar map, thus compromising the invariance of the polar map to such changes [95, 76]. Pun et al. [76] proposed the decomposition of the polar map via an adaptive row shift-invariant wavelet packet transform to eliminate the effects of the shift. In Section 5.3 of this thesis, a more straightforward and computationally efficient way is provided for achieving orientation and scale invariance.

Secondly, the polar map is constructed from the largest circular region from within the arbitrarily shaped region of the segmented specimen. The circular region can be considerably smaller than the segmented area if the specimen is highly elongated, which is the case for some specimens. Polar maps constructed from such a small area may not accurately represent the textural characteristics of the entire specimen and this would adversely affect the classification accuracy. This limitation is overcome in Section 5.3 by identifying the largest ellipse within the specimen prior to constructing the polar map.

Lastly, polar maps are not inherently invariant to changes in specimen contrast. However, this shortcoming is addressed by subjecting the polar map to non-linear filtering.

## 5.3 Overview of Method

The proposed method involves a five step process: (1) Identify largest ellipse which is completely within specimen region, (2) generate an orientation- and scale-invariant polar map (3) non-linearly filter the polar map to obtain contrastinvariant filter response (4) extract local and global energy measures from the filter response and (5) classify features using SVM.

First, it is assumed that the specimens in an image have already been accurately segmented. Next, the largest elliptical area within each segmented specimen region is determined. Orientation invariance is ensured by expressing the elliptical region in a Cartesian space defined by the major and minor axes of the ellipse and then scale invariance is achieved by transforming the elliptical area to a circular area of unit radius from which the polar map is finally constructed. Next, the polar map is subjected to non-linear filtering by a  $5 \times 5$  Gaussian first derivative (GFD) filter, to give a contrast-invariant result. This filter is selected it is a band-pass filter that extracts most of the textural information. Finally, local and global energy measures are then extracted from the filter output and used in the training and testing of a SVM classifier.

## 5.4 Identifying Elliptical Region

An ellipse boundary can be represented by a general conic equation as follows

$$\zeta(\mathbf{a}, \mathbf{x}) = \mathbf{a} \cdot \mathbf{x} = a_1 x^2 + a_2 x y + a_3 y^2 + a_4 x + a_5 y + a_6 = 0$$
(5.2)

where  $\mathbf{a} = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6]^T$  and  $\mathbf{x} = [x^2 \ xy \ y^2 \ x \ y \ 1]^T$ .  $\zeta(\mathbf{a}, \mathbf{x}_i)$  is the algebraic distance of a point  $(x_i, y_i)$  to the boundary defined by  $\zeta(\mathbf{a}, \mathbf{x}) = 0$ , where

 $\zeta(\mathbf{a}, \mathbf{x}_i) < 0$  if  $\mathbf{x}_i$  lies within the boundary and  $\zeta(\mathbf{a}, \mathbf{x}_i) > 0$  if it lies outside. Finding an ellipse within a segmented specimen can be posed as a problem of finding a suitable **a** that minimizes the following sum of the squared algebraic distances [36],  $\varepsilon$ ,

$$\varepsilon(\mathbf{a}) = \sum_{i=1}^{N} \zeta^2(\mathbf{a}, \tilde{\mathbf{x}}_i) = \mathbf{a}^T \mathbf{D}^T \mathbf{D} \mathbf{a}$$
(5.3)

The N data points of  $\tilde{\mathbf{x}}_i$ :  $1 \leq i \leq N$  belong to the segmented specimen boundary and the design matrix  $\mathbf{D}$  is defined as  $\mathbf{D} = [\tilde{\mathbf{x}}_1 \ \tilde{\mathbf{x}}_2 \ \cdots \ \tilde{\mathbf{x}}_N]^T$ . In addition, the entire ellipse is also required to be completely within the segmented specimen,

$$\mathbf{Da} \ge 0 \tag{5.4}$$

The objective function in (5.3) needs to be further constrained to avoid the trivial solution of  $\hat{\mathbf{a}} = 0$ . This is achieved by adopting the well known constraint  $a_2^2 - 4a_1a_3 \leq 0$  [24] which can be expressed as

$$\mathbf{a}^T \mathbf{C} \mathbf{a} > 0 \tag{5.5}$$

where **C** is a  $6 \times 6$  constraint matrix:

This ensures that the solution  $\hat{\mathbf{a}}$  which satisfies (5.3), (5.4) and (5.5) defines an ellipse function (5.2). The above constrained problem is difficult to solve in general

as the Kuhn-Tucker conditions [76] do not guarantee a solution. Therefore, a twostep procedure is implemented to find the largest ellipse within the segmented specimen. First, the eccentricity of the ellipse, i.e., parameters  $\hat{a}_1$ ,  $\hat{a}_2$  and  $\hat{a}_3$ , is determined by minimizing (5.3) subject to the constraint  $\mathbf{a}^T \mathbf{C} \mathbf{a} > 0$  [7]. Using the method of Lagrange multipliers and differentiating gives us the equation for stationarity,  $2\mathbf{D}^T \mathbf{D} \mathbf{a} - 2\lambda \mathbf{C} \mathbf{a} = 0$  which reduces to

$$\mathbf{Sa} = \lambda \mathbf{Ca} \tag{5.7}$$

subject to the constraint  $\mathbf{a}^T \mathbf{C} \mathbf{a} = 1$  where the scatter matrix  $\mathbf{S} = \mathbf{D}^T \mathbf{D}$  and  $\lambda$  is the Lagrange multiplier. The system in (5.7) [25] gives a unique solution  $\hat{\mathbf{a}}$  corresponding to eigenvalue  $\lambda = \mathbf{a}^T \mathbf{S} \mathbf{a} > 0$  from (5.7).

The constraint  $\mathbf{a}^T \mathbf{C} \mathbf{a} = 1$  gives a unique solution for the ellipse and ensures that it is affine invariant with low eccentricity bias. Low eccentricity bias is a desirable property as it draws the solution away from outlier points [24]. The parameters  $\hat{a}_4$ ,  $\hat{a}_5$  and  $\hat{a}_6$  in  $\hat{\mathbf{a}}$  are ignored since they will be determined next.

The parameters  $\hat{a}_1$ ,  $\hat{a}_2$  and  $\hat{a}_3$  are set to the values in  $\hat{\mathbf{a}}$  and then the translation and magnification of the ellipse, i.e., parameters  $\hat{a}_4$ ,  $\hat{a}_5$  and  $\hat{a}_6$ , are determined by minimizing (5.3) subject to the constraints (5.4) and (5.8):

$$a_1 = \hat{a_1}, \quad a_2 = \hat{a_2}, \quad a_3 = \hat{a_3}$$
 (5.8)

Fig. 5.2 illustrates the two step process for determining the largest ellipse.

## 5.5 Orientation- and Scale- Invariant Polar Map

The aim of this Section is to construct an orientation- and scale-invariant polar map p from the elliptical region defined in Section 5.4. The orientation invariance



Figure 5.2: Two step process of finding the largest ellipse within a segmented specimen. (a) Determining the ellipse eccentricity from parameters  $\hat{a_1}$ ,  $\hat{a_2}$  and  $\hat{a_3}$ . (b) Ensuring that the ellipse completely fits within the specimen by adjusting its translation and size via parameters  $\hat{a_4}$ ,  $\hat{a_5}$  and  $\hat{a_6}$ .

of the polar map is achieved by redefining the ellipse in an alternative Cartesian space so that the ellipse is centered at the origin and its major and minor axes are aligned along the coordinate axes, x' and y', as shown in Fig. 5.3 where the parameters  $a_1$  and  $a_2$  from (5.2) correspond to the semi-major and semi-minor axes of the ellipse and

$$\mathbf{x}' = \begin{bmatrix} x_1 & y_1 \end{bmatrix}^T, \qquad \mathbf{y}' = \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T \tag{5.9}$$

in xy Cartesian space. The parameters  $a_1, a_2, \mathbf{x}'$  and  $\mathbf{y}'$  are determined by solving



Figure 5.3: Ellipse redefined in x'y' Cartesian space and centered at the origin.

for  ${\bf Q}$  and  ${\bf \Lambda},$  for a known  ${\bf K},$  from the following eigenvalue system

$$\mathbf{KQ} = \mathbf{Q}\mathbf{\Lambda} \tag{5.10}$$

where from (5.2), **K** is expressed as

$$\mathbf{K} = \begin{bmatrix} -a_1/a_6 & -a_2/2a_6\\ -a_2/2a_6 & -a_3/a_6 \end{bmatrix}$$
(5.11)

where the ellipse centroid is assumed to be at the origin  $(a_4 = 0, a_5 = 0)$  and  $\mathbf{Q}$  is a square transformation matrix which contains the eigenvectors  $\mathbf{x}'$  and  $\mathbf{y}'$ , i.e.,  $\mathbf{Q} = [\mathbf{x}', \mathbf{y}']$  and  $\mathbf{\Lambda}$  is a diagonal matrix with eigenvalues of  $\mathbf{K}$  related to the parameters  $a_1$  and  $a_2$  as follows

$$\mathbf{\Lambda} = \begin{bmatrix} 1/a_1^2 & 0\\ & \\ 0 & 1/a_2^2 \end{bmatrix}$$
(5.12)

The directions of the major and minor axes,  $\mathbf{x}'$  and  $\mathbf{y}'$ , vary with the orientation of the specimen, thus ensuring the orientation invariance of the polar map.

Next, scale invariance of the polar map is achieved by expressing the elliptical area of Fig. 5.3 as a circular area of unit radius in the  $(a_1x', a_2y')$  Cartesian space defined in Fig. 5.4.



Figure 5.4: Ellipse expressed as a unit circle in the  $(a_1u, a_2v)$  Cartesian space.

Each point  $p(\alpha, r)$  in the polar map p corresponds to the point  $(rcos\alpha - rsin\alpha)^T$ in the unit circle of Fig. 5.4, which in turn corresponds to the point in the ellipse of Fig. 5.3. From (5.9), the same point can be expressed in the original xy Cartesian space as

$$\begin{bmatrix} x_{\alpha,r} & y_{\alpha,r} \end{bmatrix}^T = \begin{bmatrix} a_1 x_1 r \cos\alpha + a_2 x_2 r \sin\alpha & a_1 y_1 r \cos\alpha + a_2 y_2 r \sin\alpha \end{bmatrix}^T (5.13)$$

where  $r \leq \Re = 1$ . The polar map is finally constructed by determining each value  $p(\alpha, r)$  from the pixel value of I at the corresponding point  $(x_{\alpha,r}, y_{\alpha,r})$ . Given that  $(x_{\alpha,r}, y_{\alpha,r})$  is a real-numbered vector and the pixel indices are integer valued, the weighted pixel value of I in the neighborhood of the point  $(x_{\alpha,r}, y_{\alpha,r})$  is computed via cubic interpolation. Both  $\alpha$  and r are quantized into 360 bins each to obtain sufficiently accurate polar maps of size  $360 \times 360$ .

Fig. 5.5 shows the polar maps obtained for the images of a fern spore under different scale (Fig. 5.5(a)-(c)) and orientation (Fig. 5.5(g)) levels. As observed, the polar maps are all similar to each other thus indicating that they are invariant to the affine transformations of the same image. Minor discrepancies between the polar maps are attributed to inaccuracies in the approximation by cubic interpolation.

## 5.6 Contrast and Luminance Invariant Filter Output

In Chapter 2, a similarity measure  $R_i$  was proposed in a contrast- and luminanceinvariant edge detection scheme. In this Section, the same similarity measure is applied on the polar maps to extract textural features that are robust under



Figure 5.5: Influence of affine transformation on polar map. (a)–(c)—Images captured under (a)  $40\times$ . (b)  $60\times$ . (c)  $20\times$  objective magnification. (d)–(f)—Corresponding polar maps for images (a)–(c). (g)—Image (a) rotated by  $45^{\circ}$  counter-clockwise. (h)—Corresponding polar map of image (g).

different contrast and luminance levels. Unlike the classical linear spatial filtering measure  $C_i$  (from Chapter 2),  $R_i$ , is insensitive to image contrast since the filtering operation emphasizes texture patterns of both strong and weak contrast. The similarity measure  $(R_i)$  also incorporates a regularization parameters  $\gamma$  and c for suppressing values of filter outputs corresponding to noise.

Laws [55] generates 2-D filters from a bank of five 1-D separable filters,  $\mathbf{t}_i$ ,  $i = 1, \ldots, 5$ , where  $\mathbf{t}_1 = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix}$ ,  $\mathbf{t}_2 = \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \end{bmatrix}$ ,  $\mathbf{t}_3 = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \end{bmatrix}$ ,  $\mathbf{t}_4 = \begin{bmatrix} -1 & 2 & 0 & -2 & 1 \end{bmatrix}$  and  $\mathbf{t}_5 = \begin{bmatrix} -1 & -4 & 6 & -4 & -1 \end{bmatrix}$ such that each 2-D filter is a template of a specific texture pattern.

From  $\mathbf{t}_1$  and  $\mathbf{t}_2$ , the following 2-D filter  $\mathbf{g}$  is derived

$$\mathbf{g} = \mathbf{t}_1^T \times \mathbf{t}_2 \tag{5.14}$$

which resembles a  $5 \times 5$  Gaussian first derivative (GFD). This filter is chosen since GFD extracts mid-band frequency information that contains most of the textural energy. These frequencies are also lower than some of the other 2-D filters derived from Laws' 1-D  $\mathbf{t}_i$  filters, thus giving GFD more robustness under noisy conditions.

Fig. 5.6 shows the magnitude of GFD filter outputs,  $|C_i|$  and  $|R_i|$  for the same image captured under three different luminance levels. It is also observed, from the corresponding histogram plots, that an increase in luminance from Fig. 5.6(a)–(c) also results in an increase in image contrast.

The regularization constant c for the  $R_i$  case is set to 0 whereas the GFD filter is oriented at  $\theta = 0^\circ$  for both  $R_i$  and  $C_i$ . The brighter pixels in the filter outputs from 5.6(d) to (i) correspond to larger coefficient magnitude. The filter coefficients  $C_i$  is sensitive to contrast since the magnitude of the filter coefficients  $|C_i|$  in the filter outputs increase from 5.6(d) to (f) as seen by the increasing proportion of brighter pixels. Conversely, all three filter outputs from 5.6(g) to (i) have similar coefficient values, indicating a relatively greater robustness of  $R_i$  to contrast variation.



Figure 5.6: Influence of contrast variation on linear and non-linear filtering output. (a)-(c)—Images captured under progressively increasing luminance and contrast. (d)-(f)—Corresponding magnitude of filter output,  $|C_i|$ . (g)-(i)—Corresponding magnitude of filter output,  $|R_i|$ .

## 5.7 Local and Global Energy Measures

At each pixel location *i*, a pair of filter outputs  $R_i^H$  and  $R_i^V$  is obtained corresponding to GFD filters **g** and **g**<sup>T</sup> (5.14) oriented at  $\theta = 0^\circ$  and setting  $\gamma = 0$ . A set of energy measures is extracted from  $R_i^H$ ,  $R_i^V$  and from the corresponding magnitude  $M_i = \sqrt{R_i^{H^2} + R_i^{V^2}}$  resulting in a total of ten features. The feature set is broadly classified into the following three categories: (1) global energy measures, GEM (2) normalized global energy measures, NGEM and (3) local energy measures, LEM.

Global energy measures, GEM, are commonly used in texture classification and effectively represent the total energy measure computed from regions with uniform texture properties such as those found in the Brodatz database [72, 90] or its polar map equivalent [76].

The GEM features capture the overall energy measure of the entire specimen region. However, it is likely that two specimens from different classes may have entirely different polar maps but give similar global feature values. NGEM features enable effective classification of specimens for such cases since they also carry the specimens' area information.

However, textures may not have uniform (stationary) statistics within an entire specimen region. The largest ellipse with each specimen region is therefore divided into several annular sub-regions as shown in Fig. 5.7(a), where the local texture properties within each annulus is assumed to be uniform. Both local features from each annulus and global features are then extracted from the entire ellipse. Both types of features, i.e., local and global, complement each other in accurately classifying the various specimens to their respective classes.

## 5.7.1 Global energy measures, GEM

The GM features,  $T_1$  and  $T_2$  are obtained from the norm-1 energy measures of  $R_i^H$ and  $R_i^V$  as:

$$T_1 = \sum_{i=1}^{N} |R_i^H|, \quad T_2 = \sum_{i=1}^{N} |R_i^V|$$
(5.15)

where N is the number of pixels in the polar map.

## 5.7.2 Normalized global energy measures, NGEM

 $T_1$  and  $T_2$  are both normalized with respect to the specimen area, A, to get the NGEM features,  $T_3$  and  $T_4$ :

$$T_3 = \frac{1}{A} \sum_{i=1}^{N} |R_i^H|, \quad T_4 = \frac{1}{A} \sum_{i=1}^{N} |R_i^V|$$
(5.16)

## 5.7.3 Local energy measures, LEM

Local energy measures LEM obtained from sub-regions within the specimen may facilitate better classification. Fig. 5.7 shows the division of the polar map into six rectangular regions  $A_1$ - $A_6$  of equal area in 5.7(b) where each area corresponds to an annulus of the largest ellipse within a specimen as shown in 5.7(a).

A local feature is computed from each annulus  $A_j$  giving the set LEM comprising six local features  $T_5$  to  $T_{10}$ :

$$T_{j+4} = \frac{\sum_{i \in A_j} M(i)}{\sum_{k \in A_T} M(k)} \text{ where } A_T = \sum_{j=1}^6 A_j$$
(5.17)

## 5.8 Experimental Results

Three experiments were carried out with the following objectives: (1) classify the various specimens to their respective classes using SVM, based on the ten energy



Figure 5.7: Distribution of local energy features. (a) Elliptical area divided into six localized regions. (b) Corresponding six rectangular regions of equal area:  $A_1$ to  $A_6$  in polar map p.

measures  $T_1$  to  $T_{10}$ , (2) test the accuracy of the SVM classifier under different contrast, scale and orientation of the specimens, (3) study the influence of the size of feature extraction area, within the specimen, on the accuracy, (4) validate the relevance of the energy measures and (5) study the influence of  $\gamma$  on contrast invariance and noise robustness.

## 5.8.1 Texture classification via support vector machines (SVM)

The experimental data set comprises airborne spore allergens from six different species. The allergens were captured using the Burkard seven day volumetric recording spore traps. The air-borne allergens were trapped on a silicone grease coated tape that was mounted on a drum which rotates at a rate of 2mm per hour. At the end of the week, the tape was cut into segments that represent each day of the week and mounted on glass slides for examination under an Olympus LM at  $400 \times$  magnification. The microscope is attached to a color video camera which in turn is linked to a frame grabber with RGB channels.

The specimens are classified from a total of six species. A sample specimen for each species is shown in Fig. 5.8. All specimens used in this experiment are gray-scale images with 256 gray levels. For ease of reference, the species names in Figs. 5.8(a)-(f) are henceforth denoted as NEBI, STPA, SOHA, ACAU, CUBR and PIMA respectively.



Figure 5.8: Sample images of different species used in the proposed work. (a) Nephrolepis auriculata, NEBI ( $95\mu m \times 75\mu m$ ). (b) Stenochlaena palustris, STPA ( $122\mu m \times 85\mu m$ ). (c) Sorghum halepensis, SOHA ( $115\mu m \times 115\mu m$ ). (d) Acacia auriculiformis, ACAU ( $93\mu m \times 84\mu m$ ). (e) Curvularia brachyspora, CUBR ( $34\mu m \times 50\mu m$ ). (f) Pithomyces maydicus, PIMA ( $45\mu m \times 82\mu m$ )

The classification study uses a large test set of approximately 1250 and a training set of approximately 2,500 segmented specimens. The segmented specimens have arbitrary orientation but belong to images acquired under fixed luminance, contrast and scale (40× objective magnification) settings. The largest elliptical region was identified within each specimen and then subjected to a polar map transformation. The similarity maps,  $R^H$  and  $R^V$  were computed from the polar map while setting the regularization constant c = 0. The feature vector **T** comprising the ten energy measures  $T_1, \ldots, T_{10}$  was then extracted from the similarity maps.

The classification methodology is as follows. The mean and standard deviation values are extracted from the training set. These values are then used to normalize feature vectors from both the training and test sets where the normalized set  $\hat{\mathbf{T}}$  has zero mean and unit standard deviation.

SVM is then used for supervised classification of the normalized feature vectors  $\hat{\mathbf{T}}$  as it achieves good generalization performance on unseen data. The three types of kernel functions  $K\left(\hat{\mathbf{T}}, \hat{\mathbf{T}}_k\right)$  considered were: (1) linear,  $\langle \hat{\mathbf{T}}, \hat{\mathbf{T}}_k \rangle$ , (2) radial basis function (RBF),  $\exp(-\|\hat{\mathbf{T}} - \hat{\mathbf{T}}_k\|^2/2\sigma^2)$  and (3) polynomial,  $(\lambda \langle \hat{\mathbf{T}}, \hat{\mathbf{T}}_k \rangle + \kappa)^d$ , where  $\hat{\mathbf{T}}_k$  denotes the  $k^{th}$  feature vector from the training phase and  $\sigma$  controls the width of the RBF kernel. The  $\sigma$  values considered in this thesis range from 0.2 to 6.0. The parameters  $\lambda$ ,  $\kappa$  and d belong to the polynomial kernel where  $\kappa = 1$ ,  $\lambda$  ranges from 0.1 to 1.5 and  $d = \{1, 2, 3\}$ .

A pair-wise classification scheme is employed where a dedicated SVM classifier is implemented for each pair of classes. During the testing phase, the SVM outputs a decision value

$$s_{pq} = \sum_{k=1}^{M} \alpha_k^{pq} y_k^{pq} K \langle \hat{\mathbf{T}}, \hat{\mathbf{T}}_k \rangle + b^{pq}$$
(5.18)

where the Lagrange multiplier  $\alpha_k^{pq} > 0$  if it corresponds to a support vector and  $\alpha_k^{pq} = 0$ , otherwise. The *a priori* output label  $y_k^{pq} \in \{\pm 1\}$ , and  $b^{pq}$  is a scalar for

the pq class pair. The values of  $\{\alpha_k^{pq} : 1 \le k \le M\}$  and  $b^{pq}$  are determined during the training phase by minimizing

$$J_{pq} = \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i^{pq} \alpha_j^{pq} y_i^{pq} y_j^{pq} K \langle \hat{\mathbf{T}}, \hat{\mathbf{T}}_k \rangle + \Pr \sum_{i=1}^{M} \xi_i^{pq}$$
(5.19)

subject to the constraint

$$y_i^{pq} \left( \sum_{j=1}^M \alpha_j^{pq} y_j^{pq} K \langle \hat{\mathbf{T}}_i, \hat{\mathbf{T}}_j \rangle + b^{pq} \right) \ge 1 - \xi_i^{pq}$$
(5.20)

where  $\xi_i^{pq} \ge 0$  and PF is a penalty factor which controls the trade-off between model complexity and training error in order to ensure good generalization performance. The PF values considered in this thesis range from 0.1 to 100.

The sign and magnitude of  $s_{pq}$  can be used to predict the winning class which the feature vector  $\hat{\mathbf{T}}$  is more likely to be assigned to and the confidence level of that prediction. A majority voting method by Friedman [32] is used to select the class label with the most number of winning two-class decisions.

Table 5.1 shows the overall classification accuracy of the polynomial SVM for  $\lambda$  ranging from 0.1 to 1.5,  $d = \{1, 2, 3\}$ , PF= $\{0.1, 1, 10, 100\}$  and  $\kappa = 1$ . As observed, the highest overall classification accuracy of 96.4% is achieved for two different cases denoted by the results in bold print and italics, but the parameters,  $\lambda = 0.3$ , d = 2 and PF=1, corresponding to the results in bold print, are selected since the classification accuracy is the highest for both overall and individual classes in the data set.

Table 5.2 shows the overall classification accuracy of the RBF SVM for a subset of the  $\sigma$  values ranging from 3.2 to 6.0 and PF={0.1, 1, 10, 100}. The highest classification accuracy of 96.9% is denoted in bold print and corresponds to parameters  $\sigma = 5.4$  and PF=100. The highest classification accuracy of the linear SVM

	Degree $(d)$												
	1				2			3					
	Penalty factor(PF)				Penalty factor(PF)			Penalty factor(PF)					
$\operatorname{Gamma}(\lambda)$	0.1	1.0	10	100		0.1	1.0	10	100	0.1	1.0	10	100
0.1	94.6	96.0	96.2	96.2		95.5	96.3	95.9	95.8	95.9	96.2	95.4	94.6
0.3	95.1	95.8	96.2	96.2		96.0	96.4	95.8	95.4	95.7	95.5	94.6	92.6
0.5	95.6	96.0	96.0	96.2		95.9	95.6	95.1	95.0	95.3	95.1	93.0	91.2
0.7	95.7	96.1	96.1	96.3		96.1	95.5	95.3	95.3	95.3	94.6	92.6	90.9
0.9	95.9	96.2	96.2	96.2		96.4	95.5	95.4	95.3	95.1	93.7	92.2	91.0
1.1	96.1	96.2	96.2	96.2		96.1	95.5	95.2	95.2	95.0	93.0	91.4	90.7
1.3	96.3	96.2	96.1	96.2		96.1	95.7	95.3	95.0	94.7	92.9	91.3	90.7
1.5	96.2	96.2	96.2	96.2		95.8	95.8	94.9	95.0	94.8	92.9	90.8	90.9

Table 5.1: Overall classification percentage of polynomial SVM for a range of  $\lambda$ , d and PF.

is 96.2% for PF=1. The RBF SVM gives the best overall classification accuracy (96.9%) compared to the other two (96.4% and 96.2%). Although the aforementioned parameter values are by no means the result of an exhaustive search for the optimal classification accuracy, they do provide the best results for the substantial range of parameter values considered.

Fig. 5.9 shows the classification accuracy for individual classes of airborne spores. The RBF SVM gives the highest classification accuracy for three out of the six classes considered. Its performance is consistent over all classes with a mean accuracy of 97.0% compared to 96.2% and 95.7% of the polynomial and linear SVM respectively. Table 5.3 shows the results, in the form of a confusion matrix, for classifying the test set comprising 1250 feature vectors using SVM with

	Penalty factor(PF)					
$\mathrm{Width}(\sigma)$	0.1	1	10	100		
3.2	94.4	96.1	96.7	95.8		
3.4	94.4	96.0	96.8	95.8		
3.6	94.2	95.8	96.8	95.9		
3.8	94.1	95.8	96.6	95.8		
4.0	94.0	95.7	96.5	95.7		
4.2	93.8	95.6	96.4	95.9		
4.4	93.8	95.6	96.5	96.0		
4.6	93.5	95.4	96.6	96.2		
4.8	93.6	95.4	96.6	96.5		
5.0	93.5	95.4	96.6	96.6		
5.2	93.4	95.4	96.6	96.6		
5.4	93.3	95.3	96.6	96.9		
5.6	93.0	95.4	96.6	96.8		
5.8	93.0	95.3	96.6	96.6		
6.0	92.6	95.2	96.6	96.6		

Table 5.2: Overall classification percentage of RBF SVM for a range of  $\sigma$  and PF.

a RBF kernel. There is usually a trade-off between robustness and accuracy. This trade-off is reflected in the classification errors observed in the confusion matrix of Table 5.3. However, these errors are small since the classification accuracies of all six species are above 94%.



Figure 5.9: Comparison chart of the classification percentage for the individual classes.

		Predicted Class						
		NEBI	STPA	SOHA	ACAU	CUBR	PIMA	
	NEBI	97.2	0.5	0.0	0.0	2.3	0.0	
SS	STPA	1.8	94.1	0.0	4.1	0.0	0.0	
al Cla	SOHA	0.0	0.0	98.5	1.5	0.0	0.0	
Actué	ACAU	0.0	1.7	0.7	97.6	0.0	0.0	
ţ	CUBR	1.9	0.0	0.0	0.0	97.8	0.3	
	PIMA	0.0	0.0	0.0	0.0	3.3	96.7	

Table 5.3: Confusion matrix of classifying test set using RBF SVM.

## 5.8.2 Contrast invariance

Here, the RBF SVM is tested on a new test set obtained from images of different contrast from the ones used in training the SVM. A representative set of 50 segmented specimens was collected from each class and subjected to linear contrast stretching [34] by factors of 0.2 to 1.0 in increments of 0.2. Incidentally, the RBF SVM was trained for a contrast factor of 1.0. The classification results are shown in Fig. 5.10. As observed, the classification performance of the proposed method is approximately the same for different contrast stretching factors. Interestingly, the STPA samples show slightly higher sensitivity to contrast variation and this is attributed to minor quantization errors involved in obtaining test images of different contrast factors.



Figure 5.10: Overall percentage of individual classes for different contrast stretching factors.

Another RBF SVM is trained using features generated by applying the linear filtering measure,  $C_i$  from Chapter 2, and the parameters  $\sigma = 4.6$  and PF=100 yielded the maximum accuracy on the same test images. Each test image was subjected to linear contrast stretching [34] by factors of 0.2 to 1.0 in increments of 0.2. Fig. 5.11 compares the overall classification results from the linear and non-linear filtering schemes. As observed, the classification performance with nonlinear filtering is relatively unaffected by contrast stretching compared to that of linear filtering.



Figure 5.11: Comparison of overall classification accuracy attributed to non-linear and linear filtering methods for different contrast factors.

## 5.8.3 Orientation invariance

The RBF SVM is tested on an entirely new feature set obtained from specimens with different orientations. A representative set of 50 segmented specimens was collected from each class where 24 samples were extracted from each image with different orientations (0° to 345° with 15° intervals). In this way, a dataset of 1200 test images (50 × 24) of segmented specimens was created for each class in this experiment. Fig. 5.12 shows the classification results where the twelve bars for each class correspond to the various orientations from 0° at the leftmost bar to 345° at the rightmost. It is found that the test results are approximately similar for different orientation.

## 5.8.4 Scale invariance

The RBF SVM is tested on another feature set obtained by sampling specimens at different scale factors. A representative set of 50 segmented specimens was



Figure 5.12: Overall percentage of individual classes for different orientations.

collected from each class where 10 samples with different scaling factor (0.25 to 2.5 at increments of 0.25) were extracted from each specimen. A dataset of 500 test samples  $(50 \times 10)$  of segmented specimens is therefore created for each class. Fig. 5.13 shows the classification results where the ten bars for each class correspond to the various scaling factors from 0.25 at the leftmost bar to 2.5 at the rightmost.



Figure 5.13: Overall percentage of individual classes for different scales.

Scale-invariance is observed provided that changes in scale do not cause any changes in the visual contents of the specimens. The sharp decline in accuracy for some of the images captured under lower scaling factors such as 0.25 and 0.50 is attributed to the loss in visual content since the RBF SVM classifier was trained for images captured under a higher magnification ratio of 1.0. Conversely, the classification accuracy of all six species remains high for magnification ratios exceeding 1.0 since all visual contents of the images are preserved and no additional information is introduced.

## 5.8.5 Variation of feature extraction area

It is believed that the specimen area, from which features are extracted, may influence the overall classification accuracy. Here, this influence is studied by considering five different cases corresponding to 20%, 40%, 60%, 80% and 100% of the original specimen sizes considered in the previous sections. A set of feature vectors are computed for each case and then used to train its respective RBF SVM classifier. As in Section 5.8.1, a large test set of approximately 1250 and a training set of approximately 2,500 segmented specimens are available for each case.

Fig. 5.14 shows the maximum overall classification results of the test set for different feature extraction areas (expressed as a ratio of the original size in Section 5.8.1) over all values of  $\sigma$  and PF considered. In general, an increase in the feature extraction area has a positive influence on the overall classification accuracy. This is attributed to the larger amount of information available in aiding the classification process.

#### 5.8.6 Validation of energy measures

Here, the effects of including different sets of features on the classification process is studied. As in Section 5.8.1, this study is performed for  $\sigma$  values ranging from 0.2 to 5.0 and PF= {0.1, 1, 10, 100} on a test set of approximately 1250 and a training set of approximately 2,500 segmented specimens.



Figure 5.14: Overall percentage for different feature extraction areas.

Table 5.4 shows the overall classification results for each permutation. From Table 5.4, the combination comprising all three categories gives the highest classification accuracy, indicating that all feature categories contribute towards the overall classification accuracy. Although the feature set comprising exclusively of GEM yields slight higher accuracy than LEM, the combination of NGEM and LEM gives a higher accuracy of 96.2%, compared to 95.9% for the combination of NGEM and GEM. This shows that LEM introduces additional information which is useful in enhancing the classification accuracy. However, NGEM is the most significant of all three energy measures since it encodes the area of the specimens in addition to texture information.

## 5.8.7 Variation of regularization parameter

The influence of the regularization parameter  $\gamma$  on the noise robustness of the proposed method is studied by testing the classifier on images added with Gaussian noise at an SNR of 20 dB. It is reasonable to consider an SNR of 20 dB as it represents the base-line case (worst case scenario) where typical SNR values of

GEM	NGEM	LEM	Accuracy
			(%)
$\checkmark$			85.6
	$\checkmark$		92.5
		$\checkmark$	84.1
$\checkmark$	$\checkmark$		95.9
$\checkmark$		$\checkmark$	89.7
	$\checkmark$	$\checkmark$	96.2
$\checkmark$	$\checkmark$	$\checkmark$	96.9

Table 5.4: Overall classification percentage for different combination of energy measures.

images in the database are significantly higher. The  $\gamma$  values in this study ranges from 0-7 whereas the regularization constant c (from Chapter 2) remains fixed at 1. For each  $\gamma$  value, an RBF SVM classifier is found which gives the highest overall classification accuracy for  $\sigma$  values ranging from 0.2 to 6.0 and PF= {0.1, 1, 10, 100} using the data set from Section 5.8.1.

Fig. 5.15 shows the classification performance of the proposed method on noisy test images for different  $\gamma$  values. It is observed that the classification accuracy increases gradually from 40.0% at  $\gamma = 0$  to its maximum value of 68.3% at  $\gamma = 7$ . The increase in classification accuracy is attributed to the suppression of noise by the regularization parameter.

The regularization parameter also influences the sensitivity of the proposed method to contrast variation. For every  $\gamma$ , the corresponding classifier was tested on a set of test images subjected to contrast stretching by a contrast factor of



Figure 5.15: Overall classification accuracy of noisy test images for different regularization values.

0.2. Fig. 5.16 shows the classification performance of the method on the contrast stretched test images for different  $\gamma$  values. The accuracy drops sharply for  $\gamma > 0$ . The narrowing of the dynamic range in gray levels, causes many local image neighborhoods to have standard deviation values below  $\gamma$  and therefore the  $R_i$  values of these neighborhoods are significantly suppressed from those before contrast stretching.



Figure 5.16: Overall classification accuracy of test images linearly stretched by contrast factor = 0.2 for different regularization values.
### 5.9 Discussion

The non linear filtering method in this thesis is contrast-invariant provided that the contrast variation entails a linear transformation of the pixels. This is because a linear transformation preserves the spatial profile of the local neighborhoods and as such their corresponding  $R_i$  values remain the same.

Section 5.8.3 shows the orientation invariance of the method where minor variations in classification accuracy are attributed to the approximation and quantization errors following the cubic interpolation operation.

The proposed method is scale-invariant provided the changes in scale do not cause a change in the visual contents of the specimens. The visual content is in turn dependent on the image resolution ( $\mu$ m/pixel). Poor resolution may reduce the visual contents and even introduce image artifacts if the texture patterns of the specimens cannot be adequately represented. Therefore a sharp decline is observed in some images captured under lower scaling factors such as 0.25 and 0.50 given that the RBF SVM classifier was trained for images captured under a magnification ratio of 1.0. Conversely, the classification accuracy of all six species remains high for magnification ratios exceeding 1.0 since all visual contents of the images are preserved and no additional information is introduced.

The specimen area from which the features is extracted influences the classification accuracy since a larger area allows more texture information of the specimens to be extracted. This also ensures a more robust feature set against the presence of impurities in the specimen region.

The three energy measures – NGEM, LEM and GEM are required for accurately

classifying the various spore classes. NGEM in particular, facilitates the distinction between CUBR and PIMA, which have similar texture information, as they encode the sizes of the spores.

An increase in the  $\gamma$  values of the non-linear filtering measure  $R_i$  results in a decrease in contrast invariance but results in an increase in noise robustness.

## 5.10 Conclusion

A texture classification method, which is invariant to orientation, scale and contrast, has been successfully implemented. The key points of the proposed classification method are: (1) orientation invariance via a simple yet effective strategy of identifying the major and minor axes of the elliptical region within the specimen, (2) determination of the largest ellipse within a specimen so as to further improve the accuracy, (3) scale invariance by mapping the elliptical area within a unit circle before constructing the polar map, (4) contrast invariance via non-linear filtering of the polar map and (5) improved accuracy by combining local and global energy measures. The various aspects of the proposed method, its invariant properties and classification accuracy have been validated through a series of experiments. This method is an improvement over classical linear filtering schemes employed in Laws' and wavelet decomposition methods as it achieves contrast invariance although it lacks in noise robustness. Its efficacy is demonstrated on a set of air-borne spore allergens.

# Chapter 6

# Conclusions

The thesis focuses on the development of image processing and pattern recognition techniques to ensure the robust detection and classification of biomedical specimens from LM images. In so doing, contributions are made in the following three areas: (1) detection of specimen boundary edges, (2) segmentation of specimens from the background and (3) feature extraction and classification of specimens.

## 6.1 Summary of Contributions

In Chapter 2, the boundary edges of specimens are found via a novel edge detection method [47]. It presents an edge similarity measure that is simultaneously robust to changes in image illumination and contrast unlike traditional GM and AN methods, which are sensitive to contrast and illumination, respectively. It incorporates a regularization term that achieves a good compromise between the opposing objectives of noise reduction and contrast invariance. It is shown that the proposed edge detection filter has a FIR HBT profile that offers better noise reduction compared to the classical GFD filter. Although GFD gives better edge localization (i.e., detection of true edge pixels), the edge localization FIR HBT filters is better for high noise levels with SNR below 5 dB. The reduction of false edge responses by FIR HBT filters is comparable to that of GFD filters for narrow filter widths ranging from  $3 \times 3$  to  $7 \times 7$ . There is also better edge localization of sharp intensity transitions in images compared to the phase congruency (PC) method.

In Chapter 3, an alternative edge detection method, MMPM, is presented in the multi-scale domain for detecting edges under high levels of image noise. It achieves this by using a set of edge filters with multiple spatial widths instead of a filter with fixed width. MMPM uses the *min* and *product* operators, in that sequence, to accurately detect step edges and significantly reduces the number of false edges detected due to noise. Unlike traditional multi-scale edge detection schemes, which are confined to the GFD in Canny's method [12] and the Mallat-Zhong filter, MMPM extends this list to include other filters such as the DOB, ramp filter and the FIR HBT filter proposed in Chapter 2. The MMPM scheme also removes the problem of edge drift, across successive decomposition scales, which afflicts the MWPM scheme.

The edge detection criteria by Canny [12] is also redefined in the multi-scale domain and compares the performance of the various filters based on these criteria. It was observed that the GFD, ramp, HBT and DOB filters considered in this thesis give an approximately similar performance for edge localization and number of false edge responses. However, a marked difference is observed in their SNR performance with both RMP and HBT showing significantly better results compared to the other two. Although it achieves better noise tolerance and reduction of false edge responses than the method in Chapter 2, it is sensitive to contrast variation as it adopts the similarity measure  $C_i$  from 2.2.

In Chapter 4, a robust method for splitting binary clumps is presented via a set of concavity-based rules [50]. The binary clumps represent the specimen regions detected after linking up the edge pixels via binary morphology. Each binary clump comprises two or more overlapping specimens. The concavity-based rules introduced include concavity depth, saliency, concavity-concavity alignment, concavity-line alignment, concavity angle, concavity ratio and measure of split. They ensure the accurate splitting of clumps comprising specimens of diverse sizes and shapes with different extent of overlap. It has been successfully applied to images comprising airborne spores and cytological specimens.

Chapter 5 proposes a texture classification method which is invariant to orientation, scale and contrast. It generates a polar map from each specimen region and extracts textural features from these polar map which are then assigned to their respective classes. Orientation invariance is achieved by redefining the coordinate axes to be aligned along the major and minor axes of specimens. Mapping of the pixels from the largest elliptical area within a specimen to a fixed unit circular area, regardless of the specimen size, ensures that the resultant polar map is invariant to scale. The non-linear filtering of the polar map using the similarity measure  $R_i$ gives the texture classification features its contrast invariance property. Lastly, the use of both local and global(normalized and non-normalized) texture features and their extraction from the largest elliptical area from within the specimen, ensure a high classification accuracy.

#### 6.2 Future Directions of Research

The aim of the thesis is to expand the knowledge base on robust techniques for the automatic identification and classification of biomedical cell specimens from LM images. However, there are several issues that need to be addressed before the proposed techniques can be successfully incorporated into a fully automated system.

#### 1. Tuning of regularization parameter c.

The current selection of the parameter, c from Section 2.13, is an empirical process. Quantifying the relationship of c to noise tolerance and contrast invariance may provide a more principled basis for the selection of its value to be used on a given image.

2. Combining edge detection methods proposed in Chapters 2 and 3. The edge similarity measure  $R_i$  in Chapter 2 is implemented for a 2-D filter of fixed width whereas Chapter 3 proposes a multi-scale edge detection scheme based on the classical similarity measure  $C_i$  and a pair of separable 1-D low pass and high pass filters. Although the multi-scale method has higher noise tolerance compared to the fixed scale method, it lacks contrast invariance due to the use of the similarity measure  $C_i$ . Therefore, replacing  $C_i$  and the 1-D filter pair of the multi-scale method with the measure  $R_i$  and 2-D FIR HBT filter from Chapter 2, respectively may further improve its edge detection performance.

#### 3. Incorporating concavity pixels at the interior boundaries.

The proposed clump splitting method is based on concavity analysis and

therefore works well provided the sizes of 'holes' within the clump is negligible if any. The proposed method only detects concavity pixels at the exterior boundary of the clump whereas concavities at the interior boundaries between the 'holes' and the clump, are ignored. The accuracy of the clump splitting method can be further increased if both types of concavity pixels are taken into consideration. Although the proposed method can be adapted for application on overlapping cells of diverse sizes and shapes, its performance on elongated cells such as the *Dreschlera* spores is markedly poorer than with the more circular cells such as the *Acacia* spores.

#### 4. Increasing robustness of texture classification method.

The proposed method of extracting texture features from the largest elliptical region within a cell, works well provided the cells closely resemble an ellipse. Conversely, the texture extraction area may not be sufficiently large for cells which are highly non-elliptical and this may adversely affect the classification accuracy.

# References

- B. Alain and T. Monique. Object recognition from 3–D blurred images. In *7<sup>th</sup> Internat. Congress on Aerobiology*, page 112, 2002.
- [2] C. Arcelli and G. Sanniti di Baja. Quenching points in distance labeled pictures. In *ICPR84: Proc. 7<sup>th</sup> Internat. Conf. on Pattern Recognition*, pages 344–346, 1984.
- [3] H. Arof and F. Deravi. Circular neighbourhood and 1-D DFT feature for texture classification and segmentation. *IEE Proc.- Vision Image Signal Process*ing, 145:167–172, 1998.
- [4] J.R. Beltran, J. Garcia-Lucia, and J. Navarro. Edge detection and classification using Mallat's wavelet. In *ICIP '94: IEEE Internat. Conf. on Image Processing*, pages 293–297, 1994.
- [5] S. Beucher and C. Lantuejoul. Use of watersheds in contour detection. In CCETT/IRISA: Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation, pages 17–21, 1979.

- [6] J. Bollmann, P.S. Quinn, M. Vela, B. Brabec, S. Brechner, M.Y. Cortés,
  H. Hilbrecht, D.N. Schmidt, R. Schiebel, and H.R. Thierstein. *Image Analysis,* Sediments and Paleoenvironments, chapter 12. Springer, 2004.
- [7] F.L. Bookstein. Fitting conic sections to scattered data. Computer Graphics and Image Processing, 9:56–71, 1979.
- [8] J.E. Bowie and I.T. Young. An analysis technique for biological shape II.
   Acta Cytologica, 21:455–464, 1977.
- [9] J.F. Brenner, T.F. Necheles, and I.A. Bonacossa et al. Scene segmentation for the analysis of routine bone marrow smears from acute lymphoblastic leukaemia patients. J. Histochem. Cytochem., 25:601–613, 1977.
- [10] P. Brodatz. Textures a photographic album for artists and designers, 1965. New York.
- [11] P. Burrough. Fractal dimensions of landscapes and other environmental data. Nature, 294:240–243, 1981.
- [12] J. Canny. A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679–698, 1986.
- [13] T. Chang and C.C.J. Kuo. Texture analysis and classification with treestructured wavelet transform. *IEEE Trans. Image Process.*, 2:429–441, 1993.
- [14] G. Cong and B. Parvin. Model-based segmentation of nuclei. Pattern Recognition, 33:1383–1393, 2000.

- [15] I. Cseke, Z. Fazekas, and T. Holka. Honey qualification an application of the ARGUS image processing system. *Microprocessors and Microsystems*, 17:219–222, 1993.
- [16] I. Daubechies. Orthonormal bases of compactly supported wavelets. Commun. Pure Appl. Math., 41:909–996, 1988.
- [17] D. Demigny and T. Kamlé. A discrete expression of canny's criteria for step edge detector performances evaluation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19:1199–1211, 1997.
- [18] D. Desolneux, L. Moisan, and J-M. Morel. Edge detection by Helmholtz principle. *Mathematical Imaging and Vision*, 14:271–284, 2001.
- [19] D.L. Donoho and I.M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. J. Amer. Statist. Assoc., 90:1200–1224, 1995.
- [20] R.O. Duda, P.E. Hart, and D.G. Stork. Pattern Classification. John Wiley & Sons, New York, 2001.
- [21] G. Fernàndez, M. Kunt, and J.-P. Zrÿd. A new plant cell image segmentation algorithm. In *ICIAP'95: Proc. 8<sup>th</sup> Internat. Conf. on Image Analysis and Processing*, pages 229–234, 1995.
- [22] M. Fernández-Delgado, P. Carrión, E. Cernadas, J.F. Gálvez, and P. Sá-Otero. Improved classification of pollen texture images using SVM and MLP. In VIIP2003: 3<sup>rd</sup> IASTED Int. Conf. on Visualization, Imaging and Image Processing, pages 686–691, 2003.

- [23] H.L.B. Fiona, S. J. Allan, R.T. Euan, and S. Glenn. Differentiation of allergenic fungal spores by image analysis, with application to aerobiological counts. *Aerobiologica*, 15:211–223, 1999.
- [24] A.W. Fitzgibbon and R.B. Fisher. A buyer's guide to conic fitting. In Proc. British Machine Vision Conf., pages 265–271, 1995.
- [25] A.W. Fitzgibbon, M. Pilu, and R.B. Fisher. Direct least square fitting of ellipses. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21:476–480, 1999.
- [26] Y. Fok, J. Chan, and R.T. Chin. Automated analysis of nerve-cell images using active contour models. *IEEE Trans. on Medical Imaging*, 15:353–368, 1996.
- [27] M.G. Forero, F. Sroubek, and G. Cristóbal. Identification of tuberculosis bacteria based on shape and color. *Real-Time Imaging*, 10:251–262, 2004.
- [28] S.R. Fountain and T.N. Tan. Extraction of noise robust invariant texture features via multichannel filtering. In *ICIP '97: IEEE Internat. Conf. on Image Processing*, pages 197–200, 1997.
- [29] I. France, A.W.G. Duller, G.A.T. Duller, and H.F. Lamb. A new approach to automated pollen analysis. *Quaternary Science Review*, 19:537–546, 2000.
- [30] H. Freeman. On the encoding of arbitrary geometric configurations. IRE Trans. EC, 10:260–268, 1961.
- [31] W. Frei and C.C. Chen. Fast boundary detection: A generalization and a new algorithm. *IEEE Trans. Computers*, 26:988–998, 1977.

- [32] F. Friedman. Another approach to polychotomous classification. Technical report, Stanford University, 1996.
- [33] P. García-Sevilla. Analysis of irregularly shaped texture regions. Computer Vision and Image Understanding, 84:62–76, 2001.
- [34] R.C. Gonzales and R.E. Woods. *Digital Image Processing*. Prentice-Hall, New Jersey, 2002.
- [35] R.M. Haralick, K. Shanmugam, and I. Dinstein. Texture features for image classification. *IEEE Trans. Systems Man Cybernet.*, 6:609–622, 1973.
- [36] R.M. Haralick and L.G. Shapiro. Computer and Robot Vision. Addison-Wesley, Boston, 1992.
- [37] S. Haykin. Neural Networks A Comprehensive Foundation. Prentice Hall, New Jersey, 1999.
- [38] A. Herskovitz and T. Binford. On boundary detection. Ai memo, Massachusetts Institute of Technology, 1970.
- [39] S. Inoué and K.R. Spring. Video microscopy: The fundamentals. Plenum Press, New York, 1997.
- [40] H.H.S. Ip and R.P.K. Yu. Recursive splitting of active contours in multiple clump segmentation. *Electronics Letters*, 32:1564–1566, 1996.
- [41] A.K. Jain, S.P. Smith, and E. Backer. Segmentation of muscle cell pictures: a preliminary study. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2:232–242, 1980.

- [42] R.A. Jarvis. On the identification of the convex hull of a finite set of points in the plane. *Information Processing Letters*, 2:18–21, 1973.
- [43] R.P. Johnson. Contrast based edge detection. *Pattern Recognition*, 23:311– 318, 1990.
- [44] R.L. Kashyap and A. Khotanzed. A model-based method for rotation invariant texture classification. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:472–481, 1986.
- [45] T. Kirubarajan and Y. Bar-Shalom. Combined segmentation and tracking of overlapping objects with feedback. In Proc. IEEE Workshop on Multi-Object Tracking, pages 77–84, 2001.
- [46] P. Kovesi. Phase congruency: A low level image invariant. Psychological Research, 64:136–148, 2000.
- [47] S. Kumar, S.H. Ong, S. Ranganath, and F.T. Chew. A luminance- and contrast-invariant edge-similarity measure. *IEEE Trans. Pattern Analysis and Machine Intelligence*, In press, 2006.
- [48] S. Kumar, S.H. Ong, S. Ranganath, F.T. Chew, and T.C. Ong. Segmentation of microscope cell images via adaptive eigenfilters. In *ICIP: Proc. IEEE Internat. Conf. Image Processing*, pages 135–138, 2004.
- [49] S. Kumar, S.H. Ong, S. Ranganath, T.C. Ong, and F.T. Chew. Classification of airspora using support vector machines (SVM). In AAAI: American Academy of Allergy, Asthma and Immunology, Part2, page 111(2), 2003.

- [50] S. Kumar, S.H. Ong, S. Ranganath, T.C. Ong, and F.T. Chew. A rule-based approach for robust clump splitting. *Pattern Recognition*, 39:1088–1098, 2006.
- [51] S. Kumar, S.H. Ong, S.Ranganath, T.C. Ong, and F.T. Chew. Automated clump splitting in digital spore images. In 7<sup>th</sup> ICA: Proc. 7<sup>th</sup> Internat. Congress on Aerobiology, page 112, 2002.
- [52] S. Kumar, T.C. Ong, S.H. Ong, S. Ranganath, and F.T. Chew. A novel approach to texture differentiation and classification of airborne airspora. In *TESA: Third European Symposium in Aerobiology*, page 9, 2003.
- [53] A. Kundu. Robust edge detection. In CVPR: Proc. IEEE Conf. Comp. Vision and Pattern Recognition, pages 11–18, 1989.
- [54] M. Langford, G.E. Taylor, and J.R. Flenley. Computerized identification of pollen grains by texture analysis. *Review of Palaeobotany and Palynology*, 64:197–203, 1990.
- [55] K.I. Laws. Rapid texture identification. In Proc. SPIE Conf. Image Processing for Missile Guidance, pages 376–380, 1980.
- [56] M.M. Leung and A.M. Peterson. Scale and rotation invariant texture classification. In *IEEE. Proc. 26<sup>th</sup> Asilomar Conf. Signals, Systems and Computers*, pages 461–465, 1992.
- [57] P. Li and J.R. Flenley. Pollen texture identification using neural networks. Grana, 38:59–64, 1999.

- [58] P. Li, J.R. Flenley, and L.K. Empson. Differentiation of allergenic fungal spores by image analysis, with application to aerobiological counts. *Journal* of Quatenary Science, 19:755–762, 2004.
- [59] J. Liang. Intelligent splitting in the chromosome domain. *Pattern Recognition*, 22:519–532, 1989.
- [60] S. Mallat and W.L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 32:617–643, 1992.
- [61] S. Mallat and S. Zhong. Characterization of signal from multiscale edges. IEEE Trans. Pattern Analysis and Machine Intelligence, 14:710–732, 1992.
- [62] S.G. Mallat. A theory for multi-resolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.
- [63] D. Marr and E. Hildreth. Theory of edge detection. Proc. Roy. Soc. London B, 207:187–217, 1980.
- [64] Y. Meyer. Ondelettes et Opérateurs. Hermann, Paris, 1990.
- [65] M.C. Morrone, J. Ross, D.C. Burr, and R. Owens. Mach bands are phase dependent. *Nature*, 324:250–253, 1986.
- [66] K. Muneeswaran, L. Ganesan, S. Arumugam, and K.R. Soundar. Texture classification with combined rotation and scale invariant wavelet features. *IEE Proc.- Vision Image Signal Processing*, 38:1495–1506, 2005.
- [67] V. Nalwa and T.O. Binford. On detecting edges. IEEE Trans. Pattern Analysis and Machine Intelligence, 8:699–714, 1986.

- [68] J.C. Olivo-Marin. Extraction of spots in biological images using multiscale products. *Pattern Recognition*, 35:1989–1996, 2002.
- [69] S.H. Ong, Jayasooriah, H.H. Yeow, and R. Sinniah. Decomposition of digital clumps into convex parts by contour tracing and labeling. *Pattern Recognition Letters*, 13:789–795, 1992.
- [70] N. Otsu. A thresholding selection method from gray level histogram. IEEE Trans. System, Man, Cybernetics, 9:62–66, 1979.
- [71] M. Petrou and J. Kittler. Optimal edge detectors for ramp edges. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13:483 491, 1991.
- [72] M. Porat and Y.Y. Zeevi. Localized texture processing in vision: Analysis and synthesis in the Gaborian space. *IEEE Trans. On Biomedical Engineering*, 36:115–129, 1989.
- [73] R. Porter and N. Canagarajah. Robust rotation-invariant texture classification: Wavelet, Gabor filter and GMRF based schemes. *IEE Proc.- Vision Image Signal Processing*, 144:180–188, 1997.
- [74] W.K. Pratt. Digital Image Processing. John Wiley & Sons, New York, 1991.
- [75] J.M.S. Prewitt. Object enhancement and extraction. In *Picture Processing and Psychopictorics*, pages 75–149. 1970.
- [76] CM. Pun and MC. Lee. Log-polar wavelet energy signatures for rotation and scale invariant texture classification. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25:590–603, 2003.

- [77] R.J. Qian and T.S. Huang. Optimal edge detection in two-dimensional images. *IEEE Trans. Image Processing*, 5:1215–1220, 1996.
- [78] O. Ronneberger, U. Heimann, E. Schultz, V. Dietze, H. Burkhardt, and R. Gehrig. Automated pollen recognition using gray scale invariants on 3–D volume image data. In 2<sup>nd</sup> European Symposium on Aerobiology, page 3, 2000.
- [79] A. Rosenfeld. Measuring the sizes of concavities. *Pattern Recognition Letters*, 3:71–75, 1985.
- [80] A. Rosenfeld and M. Thurston. Edge and curve detection for visual scene analysis. *IEEE Trans. Comput*, 20:562–569, 1971.
- [81] P.L. Rosin. Edges: Saliency measures and automatic thresholding. Machine Vision and Applications, 9:139–159, 1997.
- [82] C.D. Ruberto, A. Dempster, S. Khan, and B. Jarra. Analysis of infected blood cell images using morphological operators. *Image Vision Computing*, 20:133–146, 2002.
- [83] B.M. Sadler and A. Swami. Analysis of multiscale products for step detection and estimation. *IEEE Trans. Info. Theory*, 45:1043–1051, 1999.
- [84] Signal and Image Processing Institute University of Southern California. The USC-SIPI image database. http://sipi.usc.edu/database/.
- [85] E.C. Stillman and J.R. Flenley. The needs and prospects for automation in palynology. *Aerobiologica*, 15:1–5, 1996.

- [86] S. Suzuki and K. Abe. New fusion operation for digitized binary images and their applications. In *ICPR82: Proc. 6<sup>th</sup> Internat. Conf. on Pattern Recognition*, pages 732–738, 1982.
- [87] N. Tatsumi and R.V. Pierre. Automated image processing-past, present and future of blood cell morphology identification. *Clinics in Laboratory Medicine*, 22:299–315, 2002.
- [88] D. Thompson, H.G. Bartels, J.W. Haddad, and P.H. Bartels. Scene segmentation in a machine vision system in histopathology. SPIE Proceedings, 1206:40–47, 1990.
- [89] W.J. Treloar. Digital image processing and their application to the automation of palynology. PhD thesis, Hull University, 1992.
- [90] M. Unser. Texture classification and segmentation using wavelet frames. IEEE Trans. On Image Processing, 4:1549–1560, 1995.
- [91] W.X. Wang. Binary image segmentation of aggregates based on polygonal approximation and classification of concavities. *Pattern Recognition*, 31:1503– 1524, 1998.
- [92] Y. Xu, J.B. Weaver, D.M. Healey, and J. Lu. Wavelet transform domain filters: a spatially selective noise filtration technique. *IEEE Trans. Image Processing*, 3:747–758, 1994.
- [93] T.T.E. Yeo, X.C. Jin, S.H. Ong, Jayasooriah, and R. Sinniah. Clump splitting through concavity analysis. *Pattern Recognition Letters*, 15:1013–1018, 1993.

- [94] N. Yokoya and S. Araki. A new splitting active contour model based on crossing detection. In ACCV'95: Proc. Second Asian Conference on Computer Vision, pages 346–350, 1995.
- [95] J. Zhang and T. Tan. Brief review of invariant texture analysis methods. Pattern Recognition, 35:735–747, 2002.
- [96] L. Zhang and P. Bao. A wavelet-based edge detection method by scale multiplication. In *ICPR02: 16<sup>th</sup> International Conference on Pattern Recognition*, pages 501–504, 2002.
- [97] L. Zhang and P. Bao. Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. *IEEE Trans. Medical Imaging*, 22:1089–1099, 2003.
- [98] Y. Zhang, R. Wang, and P. Hunter. Airborne pollen texture discrimination using wavelet transforms in combination with co-occurrence matrices. Int. J. Intelligent Systems Technologies and Applications, 1:143–156, 2005.