Cell image segmentation for P53 immunohistochemistry in bladder inverted papilloma

Zhao, Peng

2007

Zhao, P. (2007). Cell image segmentation for P53 immunohistochemistry in bladder inverted papilloma. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/46968>

<https://doi.org/10.32657/10356/46968>

Nanyang Technological University

Downloaded on 26 Apr 2023 17:58:54 SGT

**Cell Image Segmentation for P53
Immunohistochemistry in Bladder Inverted
Papilloma**

Zhao Peng

School of Electrical & Electronic Engineering

**A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy**

2007

RB
46.6
Z63
2007

*To my dear parents,
for their love and encouragement.*

Acknowledgements

I would like to express my deepest appreciation and gratitude to my supervisor Dr. Mao Kezhi for his guidance, enlightenment, encourage, and help throughout the whole Ph.D study of mine. I will be always indebted to him for everything that I learned from his professional ability and great personality.

I also take this opportunity to express my sincere thanks to Nanyang Technological University (NTU) for the financial support of research scholarship.

Earnest thanks also go to Dr. Koh Tong San from NTU, and Dr. Tan Puay-Hoon from Department of Pathology, Singapore General Hospital, for the cooperation on the medical image acquisition and analysis.

Finally, I wish to acknowledge the staff and students at Biomedical Instrumentation Laboratory of NTU for their technical support and kind friendship.

Table of Contents

Acknowledgements	i
Table of Contents	vi
Summary	vi
List of Figures	x
List of Tables	xiv
1 Introduction	1
1.1 Motivation and Objective	1
1.2 Major Contribution of the Thesis	3
1.3 Organization of the Thesis	6
2 Medical Background and Literature Review	7
2.1 Medical Background	7
2.1.1 Bladder Inverted Papilloma	7

TABLE OF CONTENTS	iii
2.1.2 Diagnosis of Bladder Inverted Papilloma	10
2.1.3 P53 Immunohistochemistry Staining Techniques	13
2.2 Automatic Cell Image Analysis	14
2.3 Review of Generic Color Image Segmentation	18
2.3.1 Segmentation in the Color Space	20
2.3.1.1 Histogram Thresholding	21
2.3.1.2 Clustering	24
2.3.1.3 Support Vector Machines (SVMs)	26
2.3.1.4 Multiresolution based Segmentation	30
2.3.1.5 Statistical model based Segmentation	31
2.3.2 Segmentation in the Image Plane	33
2.3.2.1 Edge Detection	33
2.3.2.2 Region-based Segmentation	39
2.3.2.3 Watershed Transformation	40
2.4 Review of Cell Image Segmentation	44
2.5 Evaluation for Segmentation Results	46
2.5.1 Evaluation with Gold Standard	46
2.5.2 Evaluation without Gold Standard	47
2.6 Concluding Remarks	48
3 Supervised Learning-based Two-Step Procedure for Color Cell Im- age Segmentation	49

TABLE OF CONTENTS		iv
3.1	Introduction	49
3.2	A Brief Review of Color Image Conversion	51
3.3	Supervised Learning-based Method for Color Image Conversion	55
3.4	Comparison of C-G-T procedure with Pixel Classification-based Color Image Segmentation	63
3.5	Experimental Results and Discussion	66
3.5.1	Performance Evaluation of Learning-based Methods for Image Conversion	66
3.5.2	Performance Evaluation of Segmentation	69
3.6	Concluding Remarks	74
4	Marker Detection for Watershed-Based Separation of Overlapped Nuclei	78
4.1	Introduction	78
4.2	Commonly Used Approaches for Marker Detection	80
4.3	The Basic Idea of the Proposed Algorithm	81
4.4	Detecting Markers using Both Grayscale and Complemented Distance Images	84
4.4.1	Input Variables to Classifier	89
4.4.1.1	Step 1: Extract Valley Line between Two Minima	90
4.4.1.2	Step 2: Extract Characteristic Variables from the Valley Line	92
4.4.1.3	Step 3: Select Features based on Class Separability	94

<i>TABLE OF CONTENTS</i>	<i>v</i>
4.4.2 Multiple Minima Merging based on Pattern Classification . . .	97
4.5 Experimental Results and Discussion	100
4.6 Concluding Remarks	105
5 Region Merging for Watershed-Based Image Segmentation	108
5.1 Introduction	108
5.2 Brief Review of Region Merging	110
5.3 Topographic Characteristics of the Two Adjacent Regions	114
5.4 Supervised Learning-based Region Merging	117
5.4.1 Feature Selection on the Characteristic Variables	117
5.4.2 Region Merging based on Supervised Classification	119
5.5 Unsupervised Learning-based Region Merging	120
5.6 Experimental Results and Discussion	124
5.6.1 Performance Evaluation of Supervised Learning-based Region Merging	124
5.6.2 Performance Evaluation of Unsupervised Learning-based Re- gion Merging	127
5.7 Conclusion	127
6 Conclusions and Recommendations	132
6.1 Conclusions	132
6.2 Recommendations for Further Research	134

<i>TABLE OF CONTENTS</i>	vi
Author's Publications	136
Bibliography	137

Summary

Bladder inverted papilloma is a histologically benign tumor located in the bladder, however it has also been reported that some bladder inverted papilloma could recur and develop into cancer. Diagnosis of bladder inverted papilloma primarily depends on biopsy, where p53 immunohistochemistry is the most frequently used marker. In clinical practice, diagnosis of bladder inverted papilloma is determined human-visually in terms of the percentage of positively stained cells in tissue samples, which is subjective or/and inconsistent. To enable quantitative, objective, and reproducible measurements, image analysis techniques are introduced to computerize cell analysis. The success of an automatic cell analysis system largely depends on the quality of segmentation. However, segmentation of cell images is far from easy due to the complex natures of histology images. Development of sophisticated cell image segmentation algorithms is necessary to ensure success of automatic cell analysis. This is the objective of the present study.

Considering the characteristics of the images in our project, i.e., hundreds of cells in one image, we chose color space-based segmentation, which views segmentation as a pixel classification problem. We notice that the pixel classification-based segmentation in color space is equivalent to a procedure that maps color pixel vectors to scalars through a transform first and then thresholds the scalars. Based on this result, a supervised learning-based C-G-T procedure is developed, which converts color image (C) to grayscale (G) by using supervised classification techniques and

then thresholds (T) cells from grayscale image. Note that the threshold T is learned online from the grayscale image currently studied, while the corresponding parameter by using pixel classification directly is set after the training procedure and it can't change according to the different images. Therefore the C-G-T procedure provides better adaptability. The parameters for color image conversion in our method are learned off-line from our data set, which is equivalent to using pixel classification directly. The popular classification techniques have been tried, such as support vector machines (SVMs), linear discriminant analysis (LDA), and principle component analysis (PCA). The experimental results demonstrate that the C-G-T procedure achieved better results, such as the lower boundary disagreement, compared to using pixel classification directly.

Besides large numbers of cells in the images, cells are often overlapping or touching that the C-G-T procedure can't deal with and therefore a further processing is needed. A new method for overlapping cells separation is proposed which is based on the marker-controlled watershed algorithm. The proposed method utilizes the information from both intensity and shape, and combines them naturally in a framework of pattern classification to extract the most crucial features for separation and then makes decision by classifying the selected features to see whether the current region should be separated or not. In the step of feature selection, Mahalanobis distance is employed to measure class separability and the orthogonal sequential forward selection is used for search algorithm. The supervised classifiers LDA and SVM (linear SVM, SVM with a polynomial kernel, and SVM with a Gaussian kernel) have been tried for classification. The experimental results show that the proposed method performs very well to prevent both over-segmentation and under-segmentation.

An alternative method to separate overlapping cells is developed which is based on the watershed-based region merging algorithm. The traditional watershed-based region merging methods make use of region or edge information separately, or com-

bine them in a weighted fashion with user-selected values. The proposed method utilizes both region and edge information, and combines them in a framework of pattern classification. Namely, the inputs to the classifier are characteristics of adjacent regions including both region and edge information, and the output of the classifier is the merging possibility. Therefore the relationship among the different information is learned through classification training, rather than manual selection. In addition, the stop criterion for region merging is also learned through pattern classification. The supervised and unsupervised learning are both tried and the experimental results show that the supervised learning-based region merging outperformed the traditional region merging methods and the unsupervised learning-based region merging.

List of Figures

1.1	Flowchart of the automatic cell analysis procedure.	4
2.1	The position of the bladder in the urinary system [50,52]	8
2.2	(a) Gross photograph of bladder inverted papilloma [56], (b) micro-graph of bladder inverted papilloma [138].	9
2.3	IVP [58]	11
2.4	CT [58]	11
2.5	MRI [58]	12
2.6	Cystoscopic appearance of an inverted papilloma [139]	12
2.7	Biopsy [58]	13
2.8	Histology image of p53 immunostained bladder inverted papilloma. . .	15
2.9	Automatic cell analysis system	16
2.10	Building dams at the places where the water coming from two differ-ent minima would merge.	41
2.11	Grayscale reconstruction of mask I from marker J	43

<i>LIST OF FIGURES</i>	<i>xi</i>
3.1 Different projection directions for the same dataset (○ - object pixel, ● - background pixel), where projection direction $\mathbf{w}_1$ in (a) provides better class separability than $\mathbf{w}_2$ in (b).	56
3.2 Segmentation results, where (a) to (f) are based on color image conversion followed by thresholding, and (g) is based on minimum distance-based pixel classification. The projections in (a) to (f) for color image conversion are respectively by LDA, maximizing separating margin, PCA, luminance-based transform, intensity-based transform, and value-based transform (continued overleaf).	70
3.3 Figure 3.2 Continued.	71
4.1 h-maxima transform, where grayscale reconstruction of I from $I - h$ is implemented. The regional maxima are marked by the red lines. . .	81
4.2 Different types of images and associated minima. (a) grayscale image, (b) minima of grayscale image, (c) gradient image, (d) minima of gradient image, (e) complemented distance image, (f) minima of complemented distance image.	85
4.3 An example that grayscale image does not work well (severely uneven grey level). (a) grayscale image, (b) markers obtained by h-minima based on grayscale image, (c) segmentation result with markers (b), (d) marker obtained by h-minima based on complemented distance image, (e) segmentation result with marker (d).	86
4.4 An example that grayscale image does not work well (shallow watershed line). (a) grayscale image, (b) topographic surface of the grayscale image.	87

4.5 An example that complemented distance image does not work well (severe overlapping). (a) grayscale image, (b) complemented distance image, (c) minima of complemented distance image, (d) minima of grayscale image, (e) markers obtained by h-minima based on grayscale image. 88

4.6 Flow diagram of the proposed marker detection method 89

4.7 The time required to cover the path $\mathcal{P}$ on the image f equals the integral of the value of f along $\mathcal{P}$, where (a) is the path $\mathcal{P}$ between points p and q on an image f , and (b) is the values of f along $\mathcal{P}$. . . 91

4.8 Extracting valley line between two minima, where valley line between minima 1 and 3, between minima 1 and 4 are shown respectively. . . 92

4.9 (a) Valley line in the grayscale image between minima 1 and 3, (b) valley line in the grayscale image between minima 1 and 4. 93

4.10 Features for describing curve, where 7 features including path length, uphill height, downhill height, uphill area, downhill area, uphill gradient, and downhill gradient are employed. 94

4.11 The relationship between class separability and number of selected features, where class separability becomes stable until the first 6 important features are selected. 98

4.12 User interface of automatic cell analysis system. 101

4.13 Comparison of segmentation results. (a) grayscale cell image, (b) segmentation result by using watershed algorithm directly. (c), (d) and (e) segmentation results by marker-controlled watershed, where marker detection is performed on grayscale image, gradient image, and complemented distance image, respectively, all using h-minima algorithm. In (f), markers are detected by the proposed supervised minima-merging algorithm (continued overleaf). 102

4.14 Figure 4.13 Continued. 103

5.1 The topographic representation of the two adjacent regions. (a) two regions that should be merged, (b) two regions that should not be merged. 115

5.2 Comparison of the segmentation results, (a) the original grayscale image, (b) segmentation result by using watershed algorithm directly, (c) result by marker-controlled watershed with h-minima for marker detection, (d) result by waterfall transform, (e) result by region merging with dissimilarity function based on intensity mean, (f) result by region merging with dissimilarity function based on intensity mean and variance, (g) result by region merging with dissimilarity function based on edge, (h) result by region merging dissimilarity function based on intensity mean and edge, and (i) result by the proposed supervised-based region merging approach (continued overleaf). . . . 129

5.3 The segmentation result of the cell image by the proposed unsupervised-based region merging approach. 131

List of Tables

3.1	Measures of Class Separability	68
3.2	Evaluation of computer-aided segmentation results with manual seg- mentation results	73
3.3	statistical assessment $Q(I)$ of segmentation results on whole image . .	75
4.1	Comparison on accuracy of segmentation results on different types of images	83
4.2	Algorithm of minima-merging	99
4.3	Comparison on accuracy of the proposed method	105
4.4	Statistical assessment $Q(I)$ of segmentation results.	106
4.5	Comparison on Execution Time (in seconds)	107
5.1	Features to describe the characteristics of the two adjacent regions. .	118
5.2	The Algorithm of The Supervised Learning-based Region Merging . .	121
5.3	The Algorithm of Unsupervised Learning-based Region Merging . . .	123
5.4	Comparison on accuracy of the proposed supervised learning-based approach with commonly used methods	125

LIST OF TABLES

xv

5.5	Statistical assessment $Q(I)$ of segmentation results on cell images (the proposed supervised learning-based region merging approach)	126
5.6	Statistical assessment $Q(I)$ of segmentation results on cell images (the proposed unsupervised-based region merging approach)	128

Chapter 1

Introduction

Digital image analysis techniques have been applied to the field of medical research and clinical application to relieve practitioners from the laborious work as well as to provide quantitative, efficient, and reproducible analysis results. This thesis presents a framework for automatic cell image analysis specially for bladder inverted papilloma with staining of p53 immunohistochemistry, and proposes three new methods for cell image segmentation. In this introductory chapter, the motivation and objective of this study are presented and the main contributions are summarized.

1.1 Motivation and Objective

Bladder cancer is a common cancer, and is an important cause of morbidity and mortality. It is estimated that world-wide about 260,000 people are diagnosed with bladder cancer each year [41] and over two million patients worldwide are living with bladder cancer [42]. The American Cancer Society estimates there were 63,210 new cases and 13,180 deaths from bladder cancer in the United States in 2005 [47]. In the UK, bladder cancer is the fifth most common cancer and affects around 12,500 people each year [49]. Bladder cancer can occur at any age, but it is primarily a

disease of advanced age. The median age at diagnosis is 68 years, and the incidence increases with age. In addition, bladder cancer is much more common in men than women, where the male-to-female ratio of occurrence is 3:1, however women generally have a worse prognosis than men [48]. The ‘bladder inverted papilloma’ studied in this project basically is a benign tumor located in the bladder with low grade malignant potential [53]. However, it has also been reported that some bladder inverted papilloma recurred and developed into transitional cell carcinoma. Approximately 6% of inverted papillomas of the urinary tract have been reported to be susceptible to malignant transformation [55]. Therefore, it is very helpful and promising for patients to be diagnosed and treated early in the progression of the bladder inverted papilloma.

The most common symptom of bladder cancer is blood in the urine. The urine may look bright red or rust colored and the degree of bleeding may vary. Frequent urination, painful urination and a constant urge to urinate are also symptoms of bladder cancer. But the best way to monitor the bladder is to examine it with an instrument called a cystoscope by a urologist. Through the cystoscope, the urologist can see inside the bladder. If any suspicious areas are noted, a biopsy procedure is recommended in which a small piece of the bladder wall is scraped away and examined carefully under a microscope [135]. Biopsy may be the only sure way to know whether the problem is in fact cancer [136].

Diagnosis of bladder tumor primarily depends on bladder biopsy [51], where p53 immunohistochemistry is the most frequently used marker. The p53 gene is a tumor suppressor gene whose activity prevents formation of tumors. However, mutations of p53 are found in most tumor types. Mutated p53 gene encodes proteins with a prolonged half-life, and this in turn induces over-expression of p53 proteins in nuclei. Mutations of p53 can be detected by p53 immunohistochemistry, where nuclei with over-expressed p53 proteins are positively stained. The p53 immunohistochemistry

has been used as diagnostic marker for a wide variety of tumor types including bladder inverted papilloma. The diagnosis of bladder inverted papilloma is determined in terms of the percentage of positively stained cells in tissue samples.

In clinical practice, the percentage of positively stained cells is usually estimated through human visual inspection, and the result is subjective or/and inconsistent. To alleviate this problem, image analysis techniques are introduced to computerize cell analysis, see for example [1, 81, 143] and the references therein. Segmentation is the crucial step in the automatic cell analysis. However, segmentation of cell images is far from easy due to the complex nature of histology images. Major problems include overlapping or touching of nuclei, variation of nuclei size and shape, non-uniformity of staining and background, and variation of contrast between the nuclei and the background are caused by illumination inconsistencies, etc. [81]. Development of sophisticated cell image segmentation algorithms is necessary to ensure the success of automatic cell analysis. This is the objective of the present study.

1.2 Major Contribution of the Thesis

The success of automatic cell analysis largely depends on the quality of segmentation. Accurate segmentation of nuclei from the background is the first issue. Due to the complex nature of histological images, a large proportion of nuclei overlap or touch. Separation of these overlapping or touching nuclei is another important issue. In the present study, the cell image segmentation procedure illustrated in Fig. 1.1 is adopted. The first stage of the procedure is the initial segmentation, which segments nuclei from the background. The second stage is further processing that separates overlapping or touching nuclei. In this thesis, a new learning-based method is proposed to deal with the first issue, and two watershed-based algorithms are developed to address the second issue. These contributions are summarized as

Figure 1.1: Flowchart of the automatic cell analysis procedure.

follows:

- A supervised learning-based 2-step procedure for color cell image segmentation is developed. Here the color image (C) is first mapped to grayscale (G) via a transform learned through supervised learning, thresholding (T) is then performed on the grayscale image to segment objects from the background. We prove that pixel classification-based color image segmentation in color space is equivalent to performing segmentation on grayscale image through thresholding. The transform used for image conversion in the C-G-T procedure is learned off-line from the global pixel distribution of color images while the threshold is learned on-line from local pixel distribution of the grayscale image. The combination of global and local learning makes the C-G-T procedure particularly efficient.
- A new marker detection algorithm for watershed-based separation of overlapping or touching cells is proposed. The merit of the new algorithm is that it employs both photometric and shape information and combines the two naturally in the framework of pattern classification to provide more reliable markers. Extensive experiments show that it achieved 0.4% and 0.2% over-segmentation and under-segmentation respectively, while the commonly used reconstruction-based method produced 4.4% and 1.1% over-segmentation and under-segmentation respectively.
- A new region merging method for watershed-based image segmentation is proposed. The characteristics of the adjacent regions, including both region and edge information, are employed, and the different kinds of information are combined and processed in the framework of pattern classification, rather than the traditional weighted fashion with manually controlled parameters. Thus, the characteristics of the adjacent regions are input variables to the classifier and the output of the classifier is a potential value for merging. The iterative hierar-

chical region merging is terminated automatically through a threshold learned from data distribution. Both supervised-based and unsupervised-based region merging are discussed.

1.3 Organization of the Thesis

The thesis is organized as follows:

In Chapter 2, related medical background on bladder inverted papilloma and its diagnosis is introduced. Literature reviews on approaches of generic image segmentation and cell image segmentation are given. In addition, the evaluations of segmentation results are described.

In Chapter 3, a supervised learning-based 2-steps color cell image segmentation is proposed. We demonstrate that pixel classification-based color image segmentation in color space is equivalent to performing segmentation on grayscale image through thresholding, while the proposed method is more efficient as it combines the global and local learning from the image.

In Chapter 4, a new marker detection algorithm for watershed-based separation of overlapping or touching cells is proposed. The new algorithm employs both photometric and shape information and combines them naturally in the framework of pattern classification.

In Chapter 5, a new region merging algorithm for watershed-based segmentation is proposed, where both supervised classification and unsupervised classification are discussed.

Finally, conclusions and possible future work are given in Chapter 6.

Chapter 2

Medical Background and Literature Review

2.1 Medical Background

2.1.1 Bladder Inverted Papilloma

The cell, the body's basic unit of life, makes up tissue, and tissue makes up the organs of the body. Normally, cells grow and divide to form new cells as the body needs them. When cells grow old and die, new cells take their place. Sometimes this orderly process goes wrong. New cells form when the body does not need them, and old cells do not die when they should. These extra cells can form a mass of tissue called tumor. If the tumor is confined to a few cell layers and it does not invade surrounding tissues or organs, it is considered benign. In contrast, if the tumor spreads to surrounding tissues or organs, it is considered malignant, i.e., cancerous [43,44]. The bladder is an organ located in the pelvic cavity that stores and discharges urine. The position of the bladder in the urinary system is shown in Fig. 2.1, where the urinary system includes two kidneys, two ureters, the bladder,

two sphincter muscles, and the urethra. These organs, tubes, muscles, and nerves work together to create, store, and carry urine [52].

Figure 2.1: The position of the bladder in the urinary system [50,52]

Bladder tumors are caused by abnormal growth and division of cells in the bladder. There are different types of bladder tumors [57]: 1) inverted papilloma, 2) transitional papillary tumors, 3) nonpapillary carcinomas, such as adenocarcinoma, 4) bladder cancer associated with other disease, e.g., bladder stones, diverticula, 5) metastatic cancer to the bladder, 6) cancer in situ, 7) involvement of ureters, urethra, and renal pelvis, and 8) sarcomas. For the purpose of this study, we will confine our discussion to ‘bladder inverted papilloma’. This lesion was first described in four cases as adenomatoid polyp in the German literature by the Viennese urologist Paschkis in 1927 [162]. Potts and Hirst termed it ‘inverted papilloma’ in 1963 [163]. Grossly, bladder inverted papilloma appears as a smooth surfaced nodule as Fig. 2.2(a) shows, and is usually located in the bladder neck, sidewall, or trigone with tumor size ranging from about 1.0 to 3.5 cm in diameter. Microscopically, bladder

Figure 2.2: (a) Gross photograph of bladder inverted papilloma [56], (b) micrograph of bladder inverted papilloma [138].

inverted papilloma is comprised of papillary structures with inward growth within the bladder wall [54], and it is also described as a tumor with smooth outer surface and invagination of epithelial cords into underlying stroma [137] as shown in Fig. 2.2(b). A thorough literature review on bladder inverted papilloma can be found in [139] and the references therein.

Although the exact cause of bladder inverted papilloma is uncertain, there may be some environmental or occupational risk relationships to it. For example, employees who work with dyes, metal, paints, leather, textile, and organic chemicals have been suggested to be at risk [51]. Cigarette smoking is also considered to be associated with bladder inverted papilloma.

The biological potential of bladder inverted papilloma remains unclear [55]. Most urologists and pathologists will recognize it as benign. However, the multiplicity, tendency of recurrence and association with transitional cell carcinoma (TCC) reported in the literature raises suspicion or doubt concerning its benign nature [139].

2.1.2 Diagnosis of Bladder Inverted Papilloma

Diagnostic tools to establish a diagnosis of bladder tumor include radiology, cystoscopy and pathology. The various diagnostic tools are briefly described below [58].

1) *Intravenous Pyelogram (IVP)*

IVP is a conventional x-ray test using dye to examine the pelvis of the kidneys, ureters, and bladder. This x-ray allows visualization of the urinary tract to determine the presence of any abnormality.

2) *Computed Tomography (CT)*

CT scanning is essentially a detailed x-ray of the body. CT shows cross-sections of the body and allows the doctor to see details of the anatomy that would not be seen on regular x-ray.

Figure 2.3: IVP [58]

Figure 2.4: CT [58]

3) *Magnetic Resonance Imaging (MRI)*

MRI is more sensitive than CT scanning. CT and MRI have the added benefit of detecting enlarged lymph nodes near the tumors, which can suggest that a cancer has spread/metastasized to the lymph nodes.

4) *Cystoscopy*

Cystoscopy is performed by the urologist. It evaluates the bladder by direct visual examination with a specialized instrument called cystoscope, which is placed in the bladder via the urethra during the examination. The purpose of routine outpatient

Figure 2.5: MRI [58]

cystoscopy is to evaluate the lining of the lower urinary tract. If abnormalities such as tumors, stones, or patches of abnormal appearing tissue are discovered during cystoscopy, a biopsy may be taken at that time.

Figure 2.6: Cystoscopic appearance of an inverted papilloma [139]

5) Pathology

Despite the various tools, a definitive diagnosis can only be made by examining bladder tissue which is performed by a pathologist. The diagnosis of bladder cancer is based on examining cells from the bladder, either from a urine specimen or biopsy of the bladder. The correct diagnosis is critical, as appropriate treatment of bladder cancer is dependant upon the type of cancer seen.

The diagnosis can sometimes be made by examining urine cytology. A cytopathologist looks at individual cells from the urine, which are spread into a thin layer onto glass microscopic slides. These procedures have the benefit of not requiring an operation or general anesthesia.

Biopsy of the bladder, performed through the cystoscope, is the more common means of diagnosing these tumors. Biopsy analysis has long been used in medical diagnosis. In the analysis, cells are extracted and are planted on a microscope slide. The slide is then stained to enhance the features in the specimen and is examined by a cytologist. The pathologist will examine a small sample, i.e., a biopsy, of the bladder tissue under a microscope. The pathologist identifies whether the tumor is benign or malignant and the type of tumor. This is essential because tumors of different types behave very differently and require different treatment regimens.

Figure 2.7: Biopsy [58]

2.1.3 P53 Immunohistochemistry Staining Techniques

There are numerous methods to stain cells, and new techniques are developed continuously [160]. P53 immunostaining is probably the most commonly used in biopsy diagnosis. The p53 gene is a tumor suppressor gene whose activity prevents formation of tumors. Mutation of the p53 gene or altered p53 expression are common and

potentially important prognostic events in bladder cancer and other tumors [59]. Mutations prolong the half-life of the p53 protein, and mutant protein accumulates within the nuclei of malignant cells in varying amounts. The over-expression of p53 protein can be detected by immunohistochemistry [60]. An example image captured from p53 immunostained sample of bladder inverted papilloma is shown in Fig. 2.8, where the ellipse-shaped objects in the image are nuclei. Based on the color and intensity of nuclei staining, cells can be classified into four categories: negative for blue stained nuclei, 1+, 2+ and 3+ for light-brown, medium-brown and dark-brown stained nuclei respectively. Based on the percentage of cells in each category, diagnosis or prognosis is determined. In clinical practice, the proportion of cells in each category is usually estimated through human visual inspection. The accuracy of the estimation depends on the experience of pathologists, and the results are subjective and often inconsistent. To alleviate this problem, image analysis techniques have been introduced to computerize cell analysis.

2.2 Automatic Cell Image Analysis

Employing automatic cell analysis seems advisable because of the large number of cells that must be counted and analyzed to achieve reasonable accuracy. The measurement of cell properties with the aid of computer, e.g., measuring the size of the cells or counting the number of stained spots in each cell, is often referred to ‘digital image cytometry’ [142]. The main steps in image cytometry include image acquisition, pre-processing, segmentation, feature extraction, and data analysis, where segmentation is often the most vital and most difficult step. In addition, automatic or semi-automatic analysis of digital microscopy data, so called ‘micrographs’, is also associated with cell analysis, while micrographs often include several imaging techniques on the cellular level [145].

Figure 2.8: Histology image of p53 immunostained bladder inverted papilloma.

The computerized cell analysis system in our project consists of an optical microscope, a charge-coupled device (CCD) camera and a computer as Fig. 2.9 shows. The images are collected from Department of Pathology, Singapore General Hospital. Bladder inverted papilloma samples are acquired from patients by medical staff, and the samples are stained with immunohistochemistry technique, and then the images are captured from each histological slide through microscope and digital camera. Currently, the images are not available in the public domain due to the patients' medical privacy. The magnification of the microscope is set to 40, and 20 images with size of 1392 x 1040 pixels are captured from each slide automatically using an auto-stage.

Figure 2.9: Automatic cell analysis system

Many papers have been published on computer-aided image-based cell analysis [1, 6, 13, 65, 66, 78, 143, 144]. For example, ref. [143] presents the architecture of a system of quantitative microscopy which yields an integrated platform for the morphological image analysis, pattern recognition, and visual content-based indexing of peripheral blood smear wide microscopic fields. Ref. [144] employs morphological operations to vectorial image processing for cell quantification. Ref. [13] describes a system for detecting and classifying malaria parasites in images of Giemsa stained blood slides in order to evaluate the parasitaemia of the blood. Ref. [1] presents

a method for quantification of images of immunohistochemically stained cell nuclei by computing area proportions. Ref. [6] describes the key component of the multimodal cell analysis approach, a novel cytologic evaluation method for early cancer detection. In [78], an automatic cell counting method based on image segmentation technique is introduced to measure the proportion of cells that survive photodynamic therapy to evaluate the effectiveness of treatment conditions. In [65], a general method is proposed for constructing template of cells in differential interference contrast microscopy, where template matching methodology based on the Fourier transform is used. In [66], a fully automatic image analysis method is proposed for cell segmentation, feature extraction, and classification of cells according to their activation.

Most automatic cell analysis systems concentrate on the issue of segmentation. The objective of segmentation is to segment nuclei from non-nuclei regions, and classification aims to categorize each cell based on the color and intensity of the stained nucleus. In the process of automatic cell analysis, segmentation is the most critical because the success of classification or property analysis of cells largely depends on the correct segmentation of the cell image. In theory, an appropriate segmentation algorithm could accurately obtain the number and type of cells by identifying and classifying each cell on microscope images. However, segmentation is a difficult and challenging problem due to the complex nature of cells and the uncertainty in the microscopic images. The main difficulty for individual cell segmentation arises from cell aggregation, which hinders the detection of the cell contours. This is a difficult task for most image-processing algorithms and it is believed that there has been few successful attempt [61].

2.3 Review of Generic Color Image Segmentation

Image segmentation, i.e., identification of homogeneous regions in the image, has been the subject of considerable research activity over the last four decades [70]. Image segmentation is very essential and critical to image processing, and is considered as one of the most difficult tasks in image processing because it determines the quality of the final result of analysis. Image segmentation can be defined as follows [69]: If P is a homogeneity predicate defined on groups of connected pixels, then segmentation is a partition of the set F into connected subsets or regions $(S_1, S_2, \dots, S_n)$ such that

$$\bigcup_{i=1}^n S_i = F \quad \text{with} \quad S_i \cap S_j = \Phi \quad (i \neq j). \quad (2.1)$$

The uniformity predicate $P(S_i) = \text{true}$ for all regions S_i , and $P(S_i \cup S_j) = \text{false}$, when $i \neq j$ and S_i and S_j are neighbors.

A number of comprehensive surveys on image segmentation are available [69, 70, 146, 147] and there are many papers on segmentation techniques. Basically, the segmentation approaches are viewed and categorized in a few ways:

In [3], color image segmentation techniques are categorized into two main classes, whether the distribution of the pixel colors is analyzed either (1) in the image plane or (2) in the color space. Here, image plane-based color segmentation includes two main strategies (a) edge detection and (b) region reconstruction, while color space-based color segmentation includes two main categories (a) histogram analysis and (b) cluster analysis.

Ref. [146] categorizes color segmentation in a similar way as [3]: (1) image-domain based, which consists of split-and-merge techniques, region growing techniques, edge based techniques, and neural-network based classification techniques, (2) feature-

space based, which contains histogram, clustering, and adaptive k-means clustering, (3) physics based, where the underlying physical model accounts for the reflections properties of colored matter.

Ref. [69] points out that most published results of color image segmentation are based on grey level image segmentation approaches applied to different color representations. And it views color image segmentation in the categories of (1) histogram thresholding and color space clustering, (2) region based approaches, (3) edge detection, (4) fuzzy techniques, (5) physics based approaches, and (6) neural network approaches.

Ref. [70] categorizes color segmentation approaches into (1) pixel based segmentation, which includes histogram, clustering, and fuzzy clustering in the color space, (2) area based segmentation, which consists of splitting and/or merging, and region growing, (3) edge based segmentation, which contains local techniques and global techniques, and (4) physics based segmentation.

Ref. [151] reviews the existing techniques in six major classes: pixel-based, edge-based, region-based, model-based, and physics-based.

Ref. [147] divides color segmentation methods into three main periods, namely, early (before 1990), middle (1990-2000), and recent stages (since 2000). In the early years, most of the segmentation approaches were mainly devoted to grey-level images, and therefore monochrome segmentation is based on discontinuity and /or homogeneity of grey-level values in a region. In the middle stages, segmentation algorithms using color information are considered in four main classes: pixel-based, area-based, contour-based, and physics-based methods, respectively. In the recent stage, more recent color image segmentation approaches are studied and developed, where the outstanding feature defining a segmentation is that the decomposition of an image into regions should be significant in relation to the application that is using such results. On the other hand, possible applications of color segmentation are stretched

to include multimedia applications, information transmission, image compression, and so on. These works classify segmentation techniques into three main categories: feature-based techniques, image-based techniques, and physics-based techniques.

Next, the commonly used generic color image segmentation techniques are reviewed in two main categories: segmentation in the color space and in the image plane respectively.

2.3.1 Segmentation in the Color Space

The segmentation in the color space can be viewed as pixel classification problem [69], where the image is segmented into regions by assigning individual pixels of the image to specific classes such as class of nuclei or background. Therefore, pattern classification techniques can be used to deal with the problem of pixel classification. The methods based on the classification of the pixel properties in the color space take advantage of the characterization of each pixel by its trichromatic components R, G, and B, or their linear or nonlinear transformations. It is assumed that homogeneous regions in the image plane give rise to clusters in the color space, and each cluster corresponds to a class of pixels which share similar color properties. Pattern classification techniques are powerful and have been widely used to the tasks of image segmentation, object recognition, image compression, etc. Pattern classification techniques can be categorized into supervised and unsupervised learning/classification procedures. In the supervised learning, there is a teacher who provides a category label or cost for each pattern in a training set, and it seeks to reduce the sum of the costs for these patterns. But in the unsupervised learning or clustering, there is no explicit teacher and the system forms clusters or 'natural groupings' of the input patterns. Commonly used pixel classification methods include histogram thresholding and clustering techniques, which are unsupervised-based. Supervised algorithm include maximum likelihood, decision tree, K near-

est neighbor, neural networks, etc., where support vector machines (SVMs) is the promising supervised approach in the recent years, since it can provide a separating hyperplane with maximal margin for the class separation. A brief review on commonly used classification methods for image segmentation is given below.

2.3.1.1 Histogram Thresholding

Histogram thresholding is one of the simplest methods of image segmentation. This attribute lends it great consideration in current segmentation research when a rough segmentation of the image is needed [151]. Histogram thresholding is one of the widely used techniques for monochrome image segmentation. It assumes that images are composed of regions with different grey level ranges, the histogram of an image can be separated into a number of peaks, each corresponding to one region, and there exists a threshold value corresponding to valley between the two adjacent peaks.

Many algorithms for threshold selection have been published, which are based on the exploitation of histogram shape information, measurement space clustering, histogram entropy information, image attribution information, spatial information, and local characteristics [152]. Among these, the Otsu method, which is based on clustering techniques, remains one of the most referenced thresholding methods.

Otsu Thresholding. An image in 2D gray level intensity can be represented by function $f(x, y)$, where the values of $f(x, y)$ range from 0 to $L - 1$ and L is the number of distinct gray levels. Let the number of pixels with gray level i be n_i , and n be the total number of pixels in the given image, the probability of occurrence of gray level i is [199, 200]

$$p_i = n_i/n. \quad (2.2)$$

The mean gray level of the entire or total image is

$$m_{to} = \sum_{i=0}^{L-1} ip_i. \quad (2.3)$$

In the case of single thresholding, the pixels of an image are divided into two classes, i.e., foreground class $C_f = \{0, 1, \dots, T\}$ and background class $C_b = \{T + 1, T + 2, \dots, L - 1\}$, T is the threshold value. The probabilities of the two classes are

$$\omega_f(T) = \sum_{i=0}^T p_i = P(T) \quad (2.4)$$

and

$$\omega_b(T) = \sum_{i=T+1}^{L-1} p_i = 1 - P(T). \quad (2.5)$$

The means of the two classes are

$$m_f(T) = \sum_{i=0}^T ip_i / \omega_f(T) \quad (2.6)$$

and

$$m_b(T) = \sum_{i=T+1}^{L-1} ip_i / \omega_b(T). \quad (2.7)$$

The variances of the two classes are

$$\sigma_f^2(T) = \sum_{i=0}^T (i - m_f(T))^2 p_i / \omega_f(T) \quad (2.8)$$

and

$$\sigma_b^2(T) = \sum_{i=T+1}^{L-1} (i - m_b(T))^2 p_i / \omega_b(T). \quad (2.9)$$

The class separability is measured by

$$\eta = \sigma_{be}^2 / \sigma_{to}^2, \quad (2.10)$$

where σ_{be}^2 is the between-class variance

$$\sigma_{be}^2 = \omega_f(T)(m_f(T) - m_{to}(T))^2 + \omega_b(T)(m_b(T) - m_{to}(T))^2 \quad (2.11)$$

$$= \omega_f(T)\omega_b(T)(m_f(T) - m_b(T))^2 \quad (2.12)$$

$$= P(T)(1 - P(T))(m_f(T) - m_b(T))^2, \quad (2.13)$$

and σ_{to}^2 is the total variance

$$\sigma_{to}^2 = P(T)\sigma_f^2(T) + [1 - P(T)]\sigma_b^2(T). \quad (2.14)$$

Otsu [159] suggests minimizing the weighted sum of within-class variances of the foreground and background pixels to establish an optimum threshold. Recall that minimization of within-class variances is tantamount to the maximization of between-class scatter. The thresholding function for Otsu thresholding algorithm is

$$T_{\text{opt}} = \operatorname{argmax} \left\{ \frac{P(T)[1 - P(T)][m_f(T) - m_b(T)]^2}{P(T)\sigma_f^2(T) + [1 - P(T)]\sigma_b^2(T)} \right\}. \quad (2.15)$$

As for color images, the situation is different from monochrome image because of multi-features. Since the color information is represented by tristimulus R, G, B or their linear/nonlinear transformations, representing the histogram of a color image in a 3-dimensional array and selecting threshold in this histogram is not a trivial job.

One way to solve this problem is to project the 3-D space onto a lower-dimensional space, such as 1-D. The problem can be solved if a line can be found that when the points in the 3-D space are projected onto it and the projected points can be well separated. Ref. [148] utilizes the linear discriminant analysis (LDA) to find such a line for 1-D thresholding. The cluster distribution in the 3-D space is first

estimated using only 1-D histogram. Then, for 1-D thresholding, the clusters are projected onto the line determined by LDA which can minimize the clustering error rate. This permits the simultaneous utilization of all color information, rather than performing thresholding on only one color component at a time where the correlation among the three color components is neglected. In [1], the image is transformed by a principal component transform, and then the resulting first component image is used to segment the objects from the background using thresholding.

2.3.1.2 Clustering

Clustering techniques are regarded as the multidimensional extension of the concept of thresholding. The rationale of the clustering technique is that, typically, the colors in an image tend to form clusters in the histogram, one for each object in the image. In the clustering-based technique, a histogram is first obtained by the color values at all pixels and the shape of each cluster is found. Then, each pixel in the image is assigned to the cluster that is closest to the pixel color. Many different clustering algorithms are in existence today [121]. Among them, the K-means algorithm is the classic approach and has received extensive attention.

K-means algorithm. The K-means [161] algorithm for cluster-seeking is based on the minimization of a performance index which is defined as the sum of the squared distances from all points in a cluster domain to the cluster center. The basic theory of K-means is as follows:

1. Determine or choose K initial cluster centers $\mathbf{c}_1(1), \mathbf{c}_2(1), \dots, \mathbf{c}_K(1)$.
2. At the k th iteration each pixel $\mathbf{a}$ is assigned to one of the K clusters $C_1(k), \dots, C_K(k)$, where $C_j(k)$ denotes the set of pixels whose cluster center is $\mathbf{c}_j(1)$. $\mathbf{a}$ is assigned

to cluster $C_j(k)$ if

$$\| \mathbf{a} - \mathbf{c}_j(\mathbf{k}) \| \leq \| \mathbf{a} - \mathbf{c}_i(\mathbf{k}) \| \quad \forall i, j = 1, 2, \dots, K, \quad i \neq j. \quad (2.16)$$

3. Computer the new cluster centers $\mathbf{c}_j(\mathbf{k} + 1)$, $j=1,2,\dots,K$, such that the sum of the squared distances from all points in $C_j(k)$ to the new cluster center is minimized

$$J_j = \sum_{\mathbf{a} \in C_j(k)} \| \mathbf{a} - \mathbf{c}_j(\mathbf{k} + 1) \|^2, \quad j = 1, 2, \dots, K. \quad (2.17)$$

The new cluster center is

$$\mathbf{c}_j(\mathbf{k} + 1) = \frac{1}{N_j} \sum_{\mathbf{a} \in C_j(k)} \mathbf{a}, \quad j = 1, 2, \dots, K, \quad (2.18)$$

where N_j is the number of pixels of cluster $C_j(k)$.

4. If $\mathbf{c}_j(\mathbf{k} + 1) = \mathbf{c}_j(\mathbf{k})$ for $j = 1, 2, \dots, K$, the algorithm has converged and the procedure is terminated. Otherwise, go to step 2.

The determination of the initial cluster centers plays a crucial part because the better the initial partition is, the faster the algorithm will converge.

The K-means algorithm can be extended to include fuzzy inference rules. The so-called fuzzy K-means algorithm [121], which is also referred to as the fuzzy c-means algorithm, uses an iterative optimization of an objective function based on a weighted similarity measure between the pixels in the image and each of the K cluster centers. A local extremum of this objective function indicates an ‘optimal’ clustering of the input data.

2.3.1.3 Support Vector Machines (SVMs)

The goal of image segmentation is to split image into regions, which can easily be described as a pixel classification problem. As a powerful classification tool, SVMs can provide maximal separating margin between different classes and provide good generalization performance because of their solid foundations stem from the statistical theory. In the field of medical image processing, SVMs have been applied to the classification of brain volumes [196], brain segmentation [197], prostate image segmentation [198] and so on. For example, ref. [197] and [198] both utilize SVM into the shape deformation segmentation approaches to help with curve evolution.

SVMs, introduced by Vapnik and co-workers, are state-of-the-art pattern classification techniques compared with the traditional classification techniques such as neural network. SVMs have found a wide variety of applications, such as in image classification, optical character recognition, object detection, face verification, text categorization and so on. In recent years, image segmentation with SVMs has been actively studied [153, 154]. SVMs are general algorithms based on guaranteed risk bounds of statistical learning theory, i.e., the structural risk minimization principle. SVMs are considered as good candidates because of their outstanding generalization performance without the need of priori knowledge. They are learning machines capable of implementing a set of functions that approximate best the supervisor's response with an expected risk bounded by the sum of the empirical risk and the Vapnik-Chervonenkis (VC) confidence. The latter is a bound on the generalization ability of the learning machine, which depends on the VC dimension of the set of functions implemented by the machine.

Let's briefly review the basis of the theory of SVMs in the classification problem. There is a two-class pattern classification problem with a set of training examples

$$\{(\mathbf{x}_i, y_i), \quad i = 1, 2, \dots, l\}, \quad (2.19)$$

where $\mathbf{x}_i \in R^n$ is the input vector for the i th example, $y_i \in \{+1, -1\}$ is the corresponding desired response, l is the total number of training examples and n is the dimension of the input space. The problem is how to construct a classifier, i.e., a decision function $f(\mathbf{x})$, that can correctly classify an input vector $\mathbf{x}$. The nonlinear SVM first performs a nonlinear mapping $\phi(\cdot) : R^n \rightarrow R^m$. The nonlinear SVM classifier is defined as

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b, \quad (2.20)$$

which is linear in terms of the transformed data $\phi(\mathbf{x})$, but nonlinear in terms of the original data $\mathbf{x}$. The formulation of this specific problem may not be useful in practice because the training data may not be completely separable by hyperplane. In this case, slack variables, denoted by ξ_i , can be introduced to relax the separability constraints as follows:

$$y_i[\mathbf{w}^T \phi(\mathbf{x}_i) + b] \geq 1 - \xi_i, \quad (2.21)$$

where $\xi_i \geq 0$, $i = 1, 2, \dots, l$. The parameters of the decision function $f(\mathbf{x})$ are determined by the following minimization:

$$\min \quad \Phi(\mathbf{w}, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \xi_i \quad (2.22)$$

where C is a user-specified, positive, regularization parameter. Using the technique of Lagrange multipliers, one can show that a necessary condition for minimizing $\Phi(\mathbf{w}, \xi)$ in Eq. 2.22 is that the vector $\mathbf{w}$ is formed by a linear combination of the mapped vectors $\phi(\mathbf{x}_i)$, i.e.,

$$\mathbf{w} = \sum_{i=1}^l \lambda_i y_i \phi(\mathbf{x}_i), \quad (2.23)$$

where $\lambda_i \geq 0$, $i = 1, 2, \dots, l$, are the Lagrange multipliers associated with the con-

straints in Eq. 2.21. Substituting Eq. 2.23 into Eq. 2.20 yields

$$f(\mathbf{x}) = \sum_{i=1}^l \lambda_i y_i \phi^T(\mathbf{x}_i) \phi(\mathbf{x}) + b = \sum_{i=1}^l \lambda_i y_i K(\mathbf{x}_i, \mathbf{x}) + b, \quad (2.24)$$

where the function $K(\cdot, \cdot)$ is defined as

$$K(\mathbf{x}_i, \mathbf{x}) = \phi^T(\mathbf{x}_i) \phi(\mathbf{x}). \quad (2.25)$$

$K(\cdot, \cdot)$ is a kernel function and it must satisfy the condition of Mercer's theorem so that it corresponds to some type of inner product in the high-dimensional feature space. The Lagrange multipliers λ_i , $i = 1, 2, \dots, l$, are solved from the dual form of Eq. 2.22, which is expressed as:

$$\max_{\Lambda} F(\Lambda) = \sum_{i=1}^l \lambda_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \lambda_i \lambda_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j), \quad (2.26)$$

$$\text{subject to } 0 \leq \lambda_i \leq C, \quad i = 1, 2, \dots, l, \quad (2.27)$$

$$\sum_{i=1}^l \lambda_i y_i = 0, \quad (2.28)$$

where $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_l)$. The cost function $F(\Lambda)$ is convex and quadratic in terms of the unknown parameters λ_i . In practice, this problem is solved numerically through quadratic programming.

Kernel Function. The kernel function in SVMs classifier plays an important role of implicitly mapping the input vector into a high-dimensional feature space. Common choices of kernel function are the linear kernel, polynomial kernel and Gaussian RBF kernel:

1) Linear kernel

$$K(\mathbf{x}_i, \mathbf{x}) = \mathbf{x}_i^T \mathbf{x}. \quad (2.29)$$

2)Polynomial kernel

$$K(\mathbf{x}_i, \mathbf{x}) = (\mathbf{x}_i^T \mathbf{x} + 1)^d, \quad (2.30)$$

where $d > 0$ is constant that defines the kernel order.

3)Gaussian RBF kernel

$$K(\mathbf{x}_i, \mathbf{x}) = \exp \left[-\frac{\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2} \right], \quad (2.31)$$

where $\sigma > 0$ is a constant that defines the kernel width.

Model Selection. To optimize these parameters, usually the k-fold cross validation is applied to the training set by minimizing the estimated generalization error. In k-fold cross validation, the training data is randomly split into k mutually exclusive subsets of approximately equal size. The SVMs decision rule is obtained using k-1 of the subsets and then tested on the subset left out. The procedure is repeated k times and in this fashion each subset is used for testing once. Averaging the testing errors over the k trials gives an estimate of the expected generalization error. In the end, the model with the smallest generalization error will be adopted.

The SVMs have been demonstrated to outperform other traditional classification methods such as neural network or fuzzy c-means not only in the classification error but also the execution time [153,155]. SVMs are considered as a good candidate for image segmentation [195].

Actually, there may be some overlapping when image segmentation approaches are categorized into color space-based or image plane-based. Basically, the color space-based segmentation considers segmentation as a pixel classification problem and regards the information from image as features in some kind of feature space. On the other hand, the image plane-based segmentation utilizes the spatial information, namely, the relationship between the pixels and their neighborhoods, such as the

edge or boundary, or the information from adjacent regions. In practice, there have been a number of hybrid segmentation approaches, which use the information from both feature space and spatial space. As a commonly used classifier, SVM has been used for segmentation in image plane. For example, ref. [197] and [198] both employ SVM into the shape deformation segmentation approaches, which are image plane-based segmentation.

In summary, segmentation in the color space is a good choice and it has been studied further as the development of the pattern classification techniques. Of course, like other segmentation methods, classification-based segmentation has its disadvantages also. Usually, classification-based segmentations do not consider spatial constraints which make them susceptible to noise in the images. The resulting segmentation often contains isolated, small regions that are not presented in noise-free images. In the past decades, many researchers have included spatial constraints in their pixel-based segmentation techniques using statistical models.

2.3.1.4 Multiresolution based Segmentation

Multiresolution based segmentation contains two parts: 1) multiresolution representation of the image and 2) the segmentation algorithm. The image to be segmented is represented as a data structure that produces a successive condensed representation of the information in the image. The image is first segmented at coarse resolution, and then the information at this resolution is used to segment the finer resolution of the image. Multiresolution based segmentation is believed that can reduce the computational burden [166]. As for the part of image representation, Gaussian pyramids, Laplacian pyramids and wavelet are commonly used for image decomposition, where Gaussian and Laplacian representation generally entail some loss of information while wavelet transform provides a complete image representation [165]. As for the part of image segmentation, any popular segmentation

algorithms can be used. For example, ref. [164] uses wavelets to construct the image pyramid and then uses the k-means segmentation algorithm to segment images at each resolution. Ref. [165] also uses a wavelet transform to create multiresolution images, and then watershed algorithm is used to segment the low-resolution image and the region merging algorithm is followed to merge the oversegmented regions, and finally a inverse wavelet transform is used to project the regions label from the low-resolution image to the full-resolution image. Statistical models such as Markov Random Fields (MRF's) have been integrated in the framework of multiresolution, and this topic has gained increasing attention. Let's look at it next.

2.3.1.5 Statistical model based Segmentation

As reviewed in ref. [210], statistical techniques used to solve image segmentation can be categorized into three classes: 1) pixel classification, 2) exploitation of the probability distribution map as a way to model image structure, 3) combination of 1) and 2), i.e., the statistical knowledge about the spatial properties of images are integrated into the pixel classification. Among the statistical approaches to image modeling, recently, MRF has gained significant attention esp. for the texture segmentation [171].

A natural way to incorporating spatial correlations into a segmentation process is to use MRF as a priori model. The MRF itself is a conditional probability model, where the probability of a pixel depends on its neighbourhood. It is equivalent to a Gibbs joint probability distribution determined by an energy function [168, 170]. Let us look at it in detail. The spatial property can be modelled through different aspects, among which, the *contextual constraint* is a general and powerful one. This is achieved by characterizing mutual influences among such entities using conditional MRF distributions. Let $S = \{s_1, s_2, \dots, s_N\}$ be a set of sites. The sites in S are related to one another via a *neighbourhood system*, which is defined as $N = \{N_i, i \in$

$S\}$, where N_i is the set of sites neighbouring i , $i \notin N_i$ and $i \in N_j \Leftrightarrow j \in N_i$. A random field X is an MRF on S with respect to a neighbourhood system N if and only if [167]

$$P(x) > 0, \forall x \in X, \quad (2.32)$$

$$P(x_i | x_{S-\{i\}}) = P(x_i | x_{N_i}). \quad (2.33)$$

According to the Hammersley-Clifford theorem [211], an MRF can equivalently be characterized by a Gibbs distribution [170]. Thus,

$$P(x) = Z^{-1} \exp(-U(x)), \quad (2.34)$$

where

$$Z = \sum_{x \in X} \exp(-U(x)) \quad (2.35)$$

is a normalizing constant called the *partition function*, and $U(x)$ is an *energy function* of the form

$$U(x) = \sum_{c \in C} V_c(x), \quad (2.36)$$

which is a sum of *clique potentials* $V_c(x)$ over all possible cliques C . A clique c is defined as a subset of sites in S in which every pair of distinct sites are neighbours, except for single-site cliques. The value of $V_c(x)$ depends on the local configuration on clique c [175]. The most important feature of the MRF is that the probability of a particular site to assume a certain value depends only on its neighbors, and not on the whole image. The Hammersley-Clifford theorem establishes the equivalence between the conditional probabilities of the local characteristics in the MRF and the local energy potentials in the Gibbs distribution. Therefore the probability that the MRF is in a particular state can be calculated using the local energies over the entire image.

There are many image segmentation methods which use MRF to model images [169,

212]. MRF based segmentation is robust to noise, but is computationally involved [212]. In addition, it is generally difficult to incorporate specific shape information in standard MRF models [172]. The main disadvantage of MRF based methods is that the objective function associated with most nontrivial MRF problems is extremely nonconvex and as such the minimization problems is computationally very taxing. Ref. [174] addressed that the MRF model has a limited ability to describe large scale behaviors. To reduce the computational burden or deal with the large scale data set, multiresolution techniques have been used with MRF image modelling [166,174,175]. For cell image segmentation, ref. [173] utilizes MRF clustering, which is based on a Potts model that is an isotropic field without a *priori* knowledge. To deal with the aggregated cells, the algorithm they used is followed by watershed algorithm and filtering techniques for post-processing. Considering the computational burden, and MRF based segmentation can't solve the overlapping problem unless combining with other methods, MRF based segmentation is not used in this project.

2.3.2 Segmentation in the Image Plane

The methods based on the analysis of the pixel properties in the image plane are developed under two different assumptions. It assumes that adjacent regions representing distinguishable objects have detectable discontinuities in their local properties or those regions are more or less uniform in their local properties. These two assumptions result in three strategies: edge detection, region-based approaches, and their combination.

2.3.2.1 Edge Detection

Edge detection for monochrome images is based on the detection of discontinuity in grey level and locates the points with abrupt changes in grey level. Edge detec-

tion techniques are usually classified into two categories: sequential and parallel. Sequential edge detection means that the result at a point is dependent on the result of the previously examined points, while parallel edge detection means that the decision of whether or not a set of points are on an edge is not dependent on whether other sets of points lie on an edge or not. The commonly used parallel edge detection approaches are simple and straightforward. Let us look at the parallel edge detection first.

Parallel edge detection There are many types of parallel edge detection operators such as Roberts, Sobel, and Prewitt operators [151], which are the first-difference operator, Laplacian operator, and second-difference operator, respectively. The main differences among these operators are the weights assigned to each element of the mask. These operators require that there is a distinct change in grey level between two adjacent points, and only very abrupt edges between two regions could be detected. They can not detect ill-defined edges that are formed by a gradual change in grey level across the edge. Since the computation is based on a small window, the result is quite susceptible to noise.

In color images, an edge should be defined by a discontinuity in a three-dimensional color space. Ref. [156] gives three alternatives for the definition of a color edge:

- A metric distance in some color space is defined and discontinuities in the distance are used to determine edges. This allows color edge detection to still be performed in 1-D space.
- A color image is regarded as composed of three monochrome images formed by the three color components, respectively, and grey level edge detections on these three images are performed separately. Then the edges detected in the three images might be merged by some specified procedures. This is still essentially a grey level edge detection technique and may be unsatisfactory in

some cases.

- Some uniformity constraints are imposed on the edges in the three color components to utilize all of the three color components simultaneously, but the edges in the three color components are allowed to be largely independent. Actually, these constraints directly affect the computation of the three color components.

Because the parallel edge detection methods are generically sensitive to image noise and artifacts, the obtained edges through the parallel edge detection methods are often discontinuous or over-detected. Therefore, the parallel edge detection algorithms usually are not used solely, but followed by some postprocessing, such as edge tracking, gap filling, smoothing, and thinning. Or it can be combined with other approaches, e.g. region-based approaches, to complete the segmentation tasks.

Next, let us look at the sequential edge detection techniques. Most approaches in this type are *active contour model* or *deformable model*, which is based on the idea of curve or surface evolution. The algorithms start with an initial curve or point, and usually utilize heuristic search and dynamic programming to develop the curve, and finally achieve a closed and accurate boundary for the interested object. Active contour-based approaches have been widely used in the recent years and they demonstrate great ability including in the application of medical images. There are a number of active contour-based approaches such as snake, level sets, geodesic active contour, and globally optimal geodesic active contour. Let us review them in detail.

Snake The classical active contours were initially introduced in the form of snakes [176]. The snake approach simulates the motion of a series of point masses connected by springs and thin plates, with passive forces drawing them toward image features such as edges [177]. Let $v(s) : [0, 1] \rightarrow R^2$ be a parameterized planar curve.

The energy function for image segmentation using snake [193] is

$$E_{snake}^* = \int_0^1 E_{snake}(v(s))ds \quad (2.37)$$

$$= \int_0^1 E_{int}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))ds, \quad (2.38)$$

where E_{int} is the internal energy of the snake including contour spacing and smoothing, E_{image} represents the information derived from the image, and E_{con} is the external constraint energy imposed during user interaction. The internal energy is given by

$$E_{int} = \frac{1}{2}(\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2), \quad (2.39)$$

where $|v_s(s)|^2$ represents the spring force as the contour is stretched or compressed, and $|v_{ss}(s)|^2$ represents the spline force causing the contour to become smooth. α and β are weights which control the effects of these terms on the regularity of the contour at each point. The image energy is given by

$$E_{image} = -|\nabla G_\sigma \star I(x, y)|^2, \quad (2.40)$$

where $G_\sigma \star I(x, y)$ represents the convolution of a Gaussian of scale σ with the image I . Ignore the term of user interaction and assume that α and β are constants, the snake in practice can be represented as

$$E_{snake} = \int_0^1 \frac{1}{2}\alpha|v_s|^2 + \frac{1}{2}\beta|v_{ss}|^2 - |\nabla G_\sigma \star I(x, y)|^2 ds. \quad (2.41)$$

Taking the first variation of E_{snake} and introducing an artificial time parameter t for the evolution, we obtain the following Euler-Lagrange equation for the minimization of E_{snake}

$$\frac{\partial v}{\partial t} = \alpha v_{ss} - \beta v_{sss} + \nabla|\nabla G_\sigma \star I(x, y)|^2. \quad (2.42)$$

The snake can be viewed as Lagrangian geometric formulation wherein the boundary of the model is represented in a parametric form. These parameterized boundary representations will encounter difficulties when the dynamic model embedded in a noisy data set is expanding/shrinking along its normal field [178].

Level Sets In 1988, Osher and Sethian proposed to represent image boundaries or interfaces implicitly and model their propagation using appropriate partial differential equations to overcome the disadvantages of the explicit techniques [179, 181]. The Lagrangian view of curves as splines in the snake approaches was replaced by an Eulerian view of curves as level sets which demonstrate topological independence and better stability [178–180, 189]. Level set methods are numerical techniques for computing the position of propagating fronts. They rely on an initial value partial differential equation for a propagating level set function and use techniques borrowed from hyperbolic conservation laws. Topological changes, corner and cusp development, and accurate determination of geometric properties such as curvature and normal direction are naturally obtained in this setting [182]. Given an initial position for an interface Γ , where Γ is a closed curve in R^2 , and a speed function F , which gives the speed of Γ in its normal direction, the level set method takes the perspective of viewing Γ as the zero level set of a function $\phi(x, t = 0)$ from R^2 to R . That is, let $\phi(x, t = 0) = \pm d$, where d is the distance from x to Γ , and the plus/minus sign is chosen if the point x is outside/inside the initial hypersurface Γ . Then, by the chain rule, an evolution equation for the interface may be produced

$$\phi_t + F|\nabla\phi| = 0, \quad (2.43)$$

$$\phi(x, t = 0) = \text{given}. \quad (2.44)$$

This is an initial value partial differential equation in one higher dimension than the original problem. The key in level set methods is to approximate the gradient in the level set equation in a way that satisfies the correct entropy condition. One of

the simplest scheme is namely

$$\phi_{ij}^{n+1} = \phi_{ij}^n - \Delta t F_{ij} (\max(D_{ij}^{-x} \phi, 0)^2 + \min(D_{ij}^{+x} \phi, 0)^2 + \max(D_{ij}^{-y} \phi, 0)^2 + \min(D_{ij}^{+y} \phi, 0)^2)^{1/2}, \quad (2.45)$$

where D is the difference operator, for example, $D_{ij}^{+x} \phi = (\phi_{i+1,j} - \phi_{i,j})/(\Delta x)$.

Fast Marching Sethian introduced fast marching algorithm [182] in 1996 that is a special case of the level set methods. Fast marching has been widely used and has demonstrated great performance. Consider the case of a front moving with speed $F = F(x, y)$ where F is assumed to be always either positive or negative. Let $T(x, y)$ be the time at which the curve crosses the point (x, y) . Then the following equation is satisfied [183]

$$|\nabla T| F = 1. \quad (2.46)$$

Accordingly, an upwind, entropy-satisfying scheme to approximate the gradient is

$$[\max(D_{ij}^{-x} T, 0)^2 + \min(D_{ij}^{+x} T, 0)^2 + \max(D_{ij}^{-y} T, 0)^2 + \min(D_{ij}^{+y} T, 0)^2]^{1/2} F_{ij} = 1. \quad (2.47)$$

Geodesic Active Contour (GAC) GAC was proposed as a simplification of the snake energy model with fewer parameters and less sensitivity to the initial contour [184, 188]. These methods typically use a variational framework to obtain locally minimal contours by gradient descent of an energy function. As a result the final segmentations are dependent on their initialization. Also, they may become caught in irrelevant local minima.

Globally Optimal GAC To seek the geodesic active contour with globally minimal energy, as well as reduce the dependence on the initial curve, globally optimal GAC approaches have been proposed [185] which based on global searching algorithms.

Generally, active contour-based algorithms need user-initialized contour to start curve evolution. To improve the curve initialization [191], dual active contour methods were introduced [187], which propagate the user-initialized two contours, in opposite directions, the inner outside and the outer inside, toward locations with high spatial contrast that correspond to the boundaries to be segmented. Ref. [190] introduced watershed algorithm to help with the contour initialization. Ref. [177, 192] can produce the simple closed curve with globally minimal energy which is required to contain a specified internal point p_{int} , rather than a initial curve. This internal point selects the object of interest and may form the only input parameter to the algorithm.

Active contour-based methods have been widely used for the segmentation of organs in medical images, such as the segmentation for heart [192], brain [183], stomach [178], arterial tree structure [178], left ventricle [193] and so on, which have high geometric variation both between subjects and within a single subject [192]. For the application of cell image segmentation, ref. [186] proposed a Viterbi search-based dual active contour algorithm. However they need user-initialized curve for each cell to start the curve evolution. In addition, they didn't mention the case of overlapping cells which is the commonly encountered problem in the cell image segmentation. User-initialized curves are impractical for our images since there are hundreds of cells to be segmented in one image, therefore we didn't use the curve evolution-based approach in this project.

2.3.2.2 Region-based Segmentation

Region-based segmentation approaches, including region growing, region splitting, region merging, and their combinations, attempt to group pixels into homogeneous regions. Region-based approaches tend to be less sensitive to noise because homogeneity is typically determined statistically. In the region growing approach, a seed

region is first selected, then is expanded to include all homogeneous neighbors, and this process is repeated until all pixels in the image are classified. One problem with region growing is its inherent dependence on the selection of seed region and the order in which pixels and regions are examined. In the region splitting approach, the initial seed region is simply the whole image. If the seed region is not homogeneous, it is usually divided into four squared subregions, which become new seed regions. This process is repeated until all subregions are homogeneous. The major disadvantage of region splitting is that the resulting image tends to mimic the data structure used to represent the image and comes out too square. The region merging approach is often combined with region growing or region splitting to merge the similar regions for making a homogeneous region as large as possible.

The morphological approach to image segmentation combines region growing and edge detection techniques: it groups the image pixels around the regional minima of the image and the boundaries of adjacent groupings are precisely located along the crest lines of the image. This is achieved by a transformation called the ‘watershed transformation’ [16].

2.3.2.3 Watershed Transformation

Watershed transformation [73] is a very powerful segmentation tool and has been widely used for numerous tasks of image segmentation. Here, the basic theory of the watershed algorithm is reviewed as follows:

Two terms ‘catchment basin’ and ‘watersheds’ are recalled firstly. Consider the topographic representation of a grey image. Now, let a drop of water fall on such a topographic surface. According to the law of gravitation, it will flow down along the steepest slope path until it reaches a minimum. The whole set of points of the surface whose steepest slope paths reach a given minimum constitutes the ‘catchment basin’ associated with this minimum. The ‘watersheds’ are the zones dividing

adjacent catchment basins. Assume that holes have been punched in each regional minimum of the surface, the surface is then slowly immersed into water. Starting from the minimum at the lowest altitude, the water will progressively flood the catchment basins of the image. In addition, dams are raised at the places where the waters coming from two different minima would merge. At the end of this flooding procedure, each minimum is surrounded by dams delineating its associated catchment basin. The whole set of dams correspond to the watersheds. This is illustrated in Fig. 2.10.

Figure 2.10: Building dams at the places where the water coming from two different minima would merge.

Unfortunately, in practice, watershed transformation often leads to over-segmentation of the image due to noise, small structure, or local irregularities in the image, which is essentially due to the presence of spurious minima. So called over-segmentation means one object is over partitioned into several parts, while under-segmentation means overlapped or clumped objects are not separated. Over-segmentation is the main problem when applying watershed algorithm to the image directly. To overcome this problem, there are two kinds of strategies: (1) marker-controlled watershed and (2) watershed followed by region merging.

Marker-Controlled Watershed. The basic idea behind the marker-controlled watershed is to transform the input image to remove the irrelevant minima in such a way that the watersheds of the transformed image correspond to meaningful object boundaries. In other words, the marker-controlled watershed extracts the markers of objects and background, and then introduces the markers into the watershed algorithm to allow the segmentation of the selected regions exclusively [99]. Many methods for marker detection have been published [5, 14–17, 75, 76], where watershed algorithms with controlled markers have the ability to perform accurate segmentation with one-to-one correspondence to the image markers. For example, ref. [106] extracts markers through histogram analysis, and ref. [107] detects markers by clustering technique. Among these marker detection approaches, image reconstruction-based approaches have been commonly used. Morphological grayscale reconstruction [5] is very useful for several image analysis tasks such as image filtering and image segmentation. The basic theory of grayscale reconstruction is given below:

Let J and I be two grayscale images defined on the same domain, taking their values in the discrete set $\{0, 1, \dots, N - 1\}$ and such that $J \leq I$ (i.e., for each pixel $p \in D_I$, $J(p) \leq I(p)$). The elementary geodesic dilation $\delta_I^{(1)}(J)$ of grey image J

$$\delta_I^{(1)}(J) = (J \oplus B) \wedge I, \quad (2.48)$$

where $\wedge$ stands for the pointwise minimum and $J \oplus B$ is the dilation of J by flat structuring element (SE) B . The grayscale geodesic dilation of size $n \geq 0$ is then given by

$$\delta_I^{(n)}(J) = \underbrace{\delta_I^{(1)} \circ \delta_I^{(1)} \circ \dots \circ \delta_I^{(1)}}_{n \text{ times}}(J). \quad (2.49)$$

The grayscale reconstruction $\rho_I(J)$ of I from J is obtained by iterating grayscale

geodesic dilations of J ‘under’ I until stability is reached, i.e.,

$$\rho_I(J) = \bigvee_{n \geq 1} \delta_I^{(n)}(J). \quad (2.50)$$

The illustration of grayscale reconstruction $\rho_I(J)$ of I from J is shown in Fig. 2.11.

Figure 2.11: Grayscale reconstruction of mask I from marker J .

Similarly, the elementary geodesic erosion $\varepsilon_I^{(1)}(J)$ of grayscale image $J \geq I$ ‘above’ I is given by

$$\varepsilon_I^{(1)}(J) = (J \ominus B) \vee I, \quad (2.51)$$

where $\vee$ stands for the pointwise maximum and $J \ominus B$ is the erosion of J by SE B . The grayscale geodesic erosion of size $n \geq 0$ is then given by

$$\varepsilon_I^{(n)}(J) = \underbrace{\varepsilon_I^{(1)} \circ \varepsilon_I^{(1)} \circ \dots \circ \varepsilon_I^{(1)}}_{n \text{ times}}(J). \quad (2.52)$$

The grayscale reconstruction $\rho_I(J)$ of mask I from marker J is obtained by iterating grayscale geodesic erosions of J ‘above’ I until stability is reached

$$\rho_I(J) = \bigwedge_{n \geq 1} \varepsilon_I^{(n)}(J). \quad (2.53)$$

For instance, h-minima transform [100], which is reconstruction-based, suppresses

all the minima whose depths are shallower than a predefined threshold level h .

Watershed Followed by Region Merging. These kind of approaches aim at merging the catchment basins of the watershed image that belong to almost homogeneous regions [99]. Some methods are still based on reconstruction technique, such as ‘waterfall algorithm’ [99]. Other commonly used methods take the output of watershed transform on the image as the starting point of a bottom-up hierarchical merging procedure, where at each step the most similar pair of adjacent regions is detected and merged [108]. Usually, the dissimilarity function [96] between two adjacent regions is based on the region homogeneity, such as mean [108–110], variance [109], second-order and third-order central moments [111] of the intensity value [108] or color value [109] of the image. Besides region homogeneity, the characteristics of edges or common boundaries between two adjacent regions are often used in dissimilarity function, such as edge magnitude [110,112], edge integrity [96], and edge orientation [96]. The joint criteria combining region homogeneity and edge have been proposed and demonstrated to give better results than using them separately [96,110,112]. Generally, the combinations of the region information and edge information are implemented in a weighted fashion [112], where the weights control the relative importance of the terms and usually are set or adjusted manually.

2.4 Review of Cell Image Segmentation

Cell image segmentation is one specific topic of the generic image segmentation, therefore the segmentation methods for generic images reviewed above can be readily applied to cell images. In addition, some cell image segmentation approaches have their own characteristics due to the nature of cell images. On one hand, cell segmentation can take advantage of some prior knowledge of the cell such as the elliptical shape of the cell, average size, average color/intensity of the cell and so on.

Integrated with the priori information, the cell segmentation algorithm can achieve improved segmentation results. On the other hand, cell segmentation has to deal with the encountered problems in cell images [145]:

- occlusion of objects
- considerable variation of object shape and /or size and/or orientation
- considerable variation of the intensity of objects from the same class, caused for example by inconsistent staining
- inhomogeneous illumination across the visual field

where the separation of overlapped cells is an open issue and has been studied actively.

Many existing approaches for cell image segmentation are based on the generic segmentation methods. For example, in [63], a watershed algorithm is applied for segmentation of clustered nuclei. And ref. [1] is based on histogram thresholding to segment the nuclei out of the background. The specific approaches for cell image segmentation are usually based on the contour or mathematical morphology characteristics of the cell. For instance, ref. [8] uses parametric fitting algorithm to fit the cell shape by minimizing a cost function. Ref. [140] optimizes the contour model of the cell by using a genetic algorithm and Tabu search. Based on the analysis of concavities in the shape of a group of cells, ref. [84] finds the dominant concave points and calculates several parameters among concave points as a criterion for segmentation. A heuristic search algorithm is proposed in [85] which combines polygonal approximation and classification of concavities. Ref. [86] gets the skeleton of cells using a distance transform and Sober operator, then retrieves the minimum distance and creates dividing lines at the minimum point corresponding to the narrow location of the overlapped nuclei.

2.5 Evaluation for Segmentation Results

The segmentation results are often evaluated through human visual inspection. Objective and quantitative segmentation evaluation is useful and plays an important role in image segmentation. The task of segmentation evaluation can be divided into two classes. One is segmentation comparison that is an inter-technique process for ranking the performance of different techniques in segmenting the same type of image. Another is segmentation characterization that is an intra-technique process for recognizing the behavior of the considered technique in segmenting various kinds of image [123]. In this study, the first class of evaluation is employed.

The first extensive survey on the evaluation of image segmentation was published in 1996 [157]. Since then, works are still going on to propose new evaluation strategies, methods, and results. The segmentation evaluation can be categorized into two classes: evaluation with and without gold standard.

2.5.1 Evaluation with Gold Standard

Manual delineations on the images by the medical experts are regarded as the gold standard, and the results of the automated segmentation technique are compared with the gold standard. Boundary error metrics and area error metrics [122], which measure discrepancy between the computer-aided segmentation results and the manual delineations, are employed commonly. It assumes that the manually delineated boundary is denoted as $\mathcal{C}_M = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{n_M}\}$ and computer-aided segmentation result as $\mathcal{C}_C = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{n_C}\}$, where each element of $\mathcal{C}_M$ or $\mathcal{C}_C$ is a point on the corresponding boundary. For each point $\mathbf{q}_j$ on boundary $\mathcal{C}_C$, the distance to the closest point in $\mathcal{C}_M$ is calculated

$$d(\mathbf{q}_j, \mathcal{C}_M) = \min_k \|\mathbf{q}_j - \mathbf{p}_k\|, \quad j = 1, 2, \dots, n_C, \quad k = 1, 2, \dots, n_M, \quad (2.54)$$

where $\| \cdot \|$ denotes 2-D Euclidean distance. Two boundary error measures are commonly used [122]. One measure is the Hausdorff distance, which is defined as the maximum of $d(\mathbf{q}_j, \mathcal{C}_M)$ over all j , and another is the mean absolute distance, which is the average of $d(\mathbf{q}_j, \mathcal{C}_M)$ over all j . The Hausdorff distance measures the worst possible disagreement between the two outlines, while the mean absolute distance estimates the disagreement averaged over the two boundaries [122].

For area error metrics, three commonly used measures are false positive (FP) volume fraction, false negative (FN) volume fraction, and true positive (TP) volume fraction [122]

$$FP = \frac{|A_c \cup A_m - A_m|}{A_m}, \quad (2.55)$$

$$FN = \frac{|A_c \cup A_m - A_c|}{A_m}, \quad (2.56)$$

$$TP = \frac{|A_c \cap A_m|}{A_m}, \quad (2.57)$$

where A_m denotes the area of the nucleus determined by manual segmentation and A_c is the area of the nucleus determined by computer-aided approaches. $\cup$ and $\cap$ denote union and intersection operations respectively.

Ref. [157] suggests that empirical methods be preferred, as there is still no general theory of image segmentation. It also points out that, in general, evaluation functions require some scaling/weighting parameters, which often have to be set on the basis of human intuition or judgment. However, some evaluation functions have been proposed that does not require user-set parameters [124].

2.5.2 Evaluation without Gold Standard

Parameter-free evaluation functions are very useful for automatic evaluation, and they turn out to correspond closely to visual judgment.

Statistical evaluation of segmentation is based on intra-region uniformity, inter-region contrast, region shape, number of miss-segmented pixels, smoothness of boundaries or accuracy of boundaries, and so on [123, 124]. The following criterion is often employed for quantitative assessment [124]

$$Q(I) = \frac{1}{1000(N \times M)} \sqrt{P} \times \sum_{i=1}^P \left[\frac{e_i^2}{1 + \log A_i} + \left(\frac{P(A_i)}{A_i} \right)^2 \right], \quad (2.58)$$

where I is the segmented image with size of $N \times M$, P is number of regions in the segmented image, e_i is the sum of Euclidean distances between color vectors of pixels in region i and color vectors attributed to region i in the segmented image, A_i is area of region i , and $P(A_i)$ is the number of regions having the exact area equal to A_i . The smaller the $Q(I)$, the better the segmentation result is.

2.6 Concluding Remarks

This chapter presents a general background on bladder inverted papilloma in p53 immunohistochemistry and its diagnosis. Then the related work and publications on generic image segmentation and cell image segmentation are reviewed respectively. From the analysis of segmentation algorithms, pixel-based segmentation and watershed-based segmentation seem very powerful and promising. Finally, evaluation of segmentation results is briefly reviewed.

Chapter 3

Supervised Learning-based Two-Step Procedure for Color Cell Image Segmentation

3.1 Introduction

As reviewed in Section 2.3.1, segmentation in the color space can be done directly in the color space through pixel classification methods such as clustering or neural network, while it also can be done by the way of converting a color image to grayscale image [5] and then thresholding the grayscale image [6,7]. In this chapter, we will show that pixel classification-based color image segmentation in color space is equivalent to thresholding grayscale images. Based on this result, a supervised learning-based C-G-T procedure for color image segmentation is developed, where the color image (C) is first converted to grayscale (G), thresholding (T) is then performed on the grayscale image to segment objects from the background. Instead of using conventional luminance-based or intensity-based fixed transform, the transform used for image conversion in the C-G-T procedure is learned off-line from the

global pixel distribution of color images while the threshold is learned on-line from local pixel distribution of the grayscale image. The combination of global and local learning makes the C-G-T procedure particularly efficient. On one hand, the globally learned transform minimizes variations among different images, which are important in the following step of overlapping or touching nuclei separation. On the other hand, the locally learned threshold provides the most suitable value for the image being segmented, exhibiting adaptability. In contrast, in pixel classification-based segmentation, the classifier is learned from global pixel distribution, and therefore lacks adaptability.

Summarily, the original contributions in this chapter are two points:

- Applying classification techniques into color image conversion. The traditional methods for color image conversion are constant weighted linear combinations. They can't achieve desired contrast when color image components change a lot over the area of a slide or from slide to slide. PCA has been used for color image conversion to generate a non-static weighted linear combination of color images to produce better contrast. Because PCA deals with data in its entirety, rather than the difference between the data classes, we do have interest to see the performance of other classification techniques, such as LDA or SVM, in color image conversion since they have good ability to separate classes.
- Demonstrating the proposed two-step segmentation method is equivalent to the pixel classification approach. Moreover, the two-step segmentation may be more adaptable since the threshold value in the second step is learned online from the current image while the corresponding threshold in the pixel classification is learned off-line from the constructed data set.

In this chapter a two-step procedure for color cell image segmentation is developed,

which is based on supervised classification technique. Section 3.1 presents the basic idea behind the proposed method. Section 3.2 then reviews the related existing algorithms for color image conversion that is the key part of the developed method. Section 3.3 begins by applying supervised classifiers LDA and SVM to color image conversion, and then the Otsu thresholding is employed to implement the two-step segmentation procedure. In Section 3.4 we demonstrate the two-step procedure is equivalent to the minimum distance classification while the two-step procedure has more adaptability. Experimental results and discussion are presented in Section 3.5. The results demonstrate the utility of the algorithm presented in this chapter on a number of cell images. Finally, section 3.5 makes concluding remarks on the developed algorithm.

3.2 A Brief Review of Color Image Conversion

As shown in the flowchart of the cell analysis system Fig. 1.1, the first step of the nuclei segmentation and separation procedure is to convert the color image into grayscale. In the literature, a few methods are available for image conversion are briefly reviewed next.

1) *Luminance, Intensity, and Value based Image Conversion*

In the luminance-based method, the grayscale information of an image is separated from the color information by converting the image from the commonly used RGB color space to YIQ space that consists of luminance, hue and saturation, denoted by Y, I and Q respectively. The luminance Y represents grayscale information of an image, while hue I and saturation Q make up color information of the image. The luminance Y is a weighted sum of the R, G, and B components

$$Y = 0.3R + 0.59G + 0.11B, \quad (3.1)$$

where the weights are calculated according to the human eye's sensitivity to the red, green, blue signals [207]. Similarly, the intensity-based conversion is based on the decomposition of a color image into hue (H), saturation (S) and intensity (I) components. The intensity component is formulated as:

$$I = \frac{R + G + B}{3} \quad (3.2)$$

Besides the intensity in HSI space, the value component of HSV color space (hue, saturation, value) is also frequently used in image conversion. The value V in HSV color space is formulated as:

$$V = \max(R, G, B) \quad (3.3)$$

The luminance, intensity and value based image conversions are motivated either by human color perception or by the physical properties of colored light. These methods are commonly used in image conversion, but they do not necessarily produce grayscale images that are optimal for segmentation purpose.

2) *Learning-based Image Conversion*

Conversion of color image to grayscale image can be regarded as a dimension reduction problem. One approach to reduce the dimensionality is by combining features, where the linear combinations are particularly attractive because they are simple to compute and analytically tractable [121]. Principal component analysis (PCA) is the classical approach to finding effective linear transformations, where PCA seeks a projection that best represents the data in a least-squares sense. The basic theory of PCA [121] is given below:

Consider the problem of representing all of the vectors of dimension 'd' in a set of n samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ by a single vector $\mathbf{x}_0$. The squared-error criterion function $J_0(\mathbf{x}_0)$

is defined as

$$J_0(\mathbf{x}_0) = \sum_{k=1}^n \|\mathbf{x}_0 - \mathbf{x}_k\|^2. \quad (3.4)$$

We seek $\mathbf{x}_0$ to minimize J_0 . Obviously, the problem is solved by $\mathbf{x}_0 = \mathbf{m}$, where $\mathbf{m}$ is the sample mean

$$\mathbf{m} = \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k. \quad (3.5)$$

The sample mean is the zero-dimensional representation of the data set, which is simple but does not reveal any of the variability in the data. A one-dimensional representation of data is defined by projecting the data onto a line running through the sample mean

$$\mathbf{x} = \mathbf{m} + a\mathbf{e}, \quad (3.6)$$

where $\mathbf{e}$ is the unit vector in the direction of the line, and a is the scalar corresponding to the distance of any point $\mathbf{x}$ from the mean $\mathbf{m}$. If represent $\mathbf{x}_k$ by $\mathbf{m} + a_k\mathbf{e}$, then an *optimal* set of a_k is sought by minimizing the squared-error criterion function

$$\begin{aligned} J_1(a_1, \dots, a_n, \mathbf{e}) &= \sum_{k=1}^n \|(\mathbf{m} + a_k\mathbf{e}) - \mathbf{x}_k\|^2 \\ &= \sum_{k=1}^n \|a_k\mathbf{e} - (\mathbf{x}_k - \mathbf{m})\|^2 \\ &= \sum_{k=1}^n a_k^2 \|\mathbf{e}\|^2 - 2 \sum_{k=1}^n a_k \mathbf{e}^T (\mathbf{x}_k - \mathbf{m}) + \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{m}\|^2. \end{aligned} \quad (3.7)$$

Recognize $\|\mathbf{e}\| = 1$, partially differentiate with respect to a_k , and set the derivative to zero, obtain

$$a_k = \mathbf{e}^T (\mathbf{x}_k - \mathbf{m}). \quad (3.8)$$

Substitute a_k in Eq. 3.8 into Eq. 3.7, obtain

$$J_1(\mathbf{e}) = \sum_{k=1}^n a_k^2 - 2 \sum_{k=1}^n a_k^2 + \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{m}\|^2$$

$$\begin{aligned}
&= -\sum_{k=1}^n [\mathbf{e}^T (\mathbf{x}_k - \mathbf{m})]^2 + \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{m}\|^2 \\
&= -\sum_{k=1}^n \mathbf{e}^T (\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^T \mathbf{e} + \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{m}\|^2 \\
&= -\mathbf{e}^T \mathbf{S} \mathbf{e} + \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{m}\|^2, \tag{3.9}
\end{aligned}$$

where $\mathbf{S}$ is scatter matrix

$$\mathbf{S} = \sum_{k=1}^n (\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^T. \tag{3.10}$$

The best direction $\mathbf{e}$ minimizes J_1 also maximizes $\mathbf{e}^T \mathbf{S} \mathbf{e}$. Use Lagrange multipliers to maximize $\mathbf{e}^T \mathbf{S} \mathbf{e}$ subject to the constraint $\|\mathbf{e}\| = 1$. Let λ be the undetermined multiplier, differentiate

$$\mathbf{u} = \mathbf{e}^T \mathbf{S} \mathbf{e} - \lambda \mathbf{e}^T \mathbf{e}, \tag{3.11}$$

with respect to $\mathbf{e}$, and obtain

$$\frac{\partial \mathbf{u}}{\partial \mathbf{e}} = 2\mathbf{S} \mathbf{e} - 2\lambda \mathbf{e}. \tag{3.12}$$

We set the gradient vector equal to zero, obtain

$$\mathbf{S} \mathbf{e} = \lambda \mathbf{e}, \tag{3.13}$$

where $\mathbf{e}$ is an eigenvector of the scatter matrix $\mathbf{S}$. The eigenvector corresponding to the largest eigenvalue of the scatter matrix is the best projection of the data.

PCA has been applied to the image conversion problem [1, 25], and it is considered as the most efficient dimension reduction technique in statistics. However, image conversion using PCA may not be ideal because PCA maximizes variance of data on the principal axis, and consequently the grayscale image obtained exhibits maximum

variation of grey level.

3.3 Supervised Learning-based Method for Color Image Conversion

In color images a pixel is represented by a 3-dimensional vector, while in grayscale images a pixel is represented by a scalar. Converting color images into grayscale can be considered as mapping 3-dimensional vectors to scalars through the following transform:

$$y = \mathbf{w}^T \mathbf{x}, \quad (3.14)$$

where $\mathbf{w}=[w_1 \ w_2 \ w_3]^T$ denotes the transform, $\mathbf{x}$ is a vector representing a color pixel, and y is the grey level of the pixel in the grayscale image. Geometrically, the transform can be interpreted as projecting points in the 3-dimensional space onto an axis whose direction is defined by $\mathbf{w}$. Fig. 3.1 illustrates a number of object and background pixels and their projections on axes $\mathbf{w}_1$ and $\mathbf{w}_2$. It is noted that the object and background pixels are distinguishable in the 3-dimensional color space because they occupy different regions of the color space. Object and background pixels are still distinguishable when they are projected onto axis $\mathbf{w}_1$. Projecting nuclei and background pixels onto axis $\mathbf{w}_2$, however, results in overlapped distributions, indicating that object and background are hardly distinguishable in the grayscale image obtained.

To facilitate image segmentation, we believe an ideal transform or axis $\mathbf{w}$ should maximize distinctions between object and background pixels in the grayscale image resulted. If put in the framework of pattern classification, the ideal transform is such an axis on which projections of object and background pixels have the largest class separability. Note that object pixels, i.e., nuclei in this study, are considered as

Figure 3.1: Different projection directions for the same dataset ($\circ$ - object pixel, $\bullet$ - background pixel), where projection direction w_1 in (a) provides better class separability than w_2 in (b).

one class, and background pixels as another class. In the present study, we employ supervised learning-based methods to find this ideal axis. Firstly, a large number of nuclei pixels and background pixels are sampled from a variety of histological cell images of bladder inverted papilloma, and class labels, i.e., +1 and -1, are assigned to nuclei and background pixels respectively. Secondly, a class separability measure is selected as the optimization criterion. Finally, the optimization problem is solved to find the transform or axis $\mathbf{w}$. Two ways of obtaining transform $\mathbf{w}$ are introduced next.

1) Obtaining Transform $\mathbf{w}$ through LDA

To obtain good separability between samples of different classes, projections of samples on axis $\mathbf{w}$ are desired to have the following properties: i) small intra-class scatter and ii) large inter-class separation. When this concept is interpreted in image conversion, a small intra-class scatter implies small variations of the grey levels of nuclei pixels in the grayscale images obtained, while a large inter-class separation is associated with large contrast between nuclei and background in the grayscale image. To achieve both goals, it is needed to maximize the ratio of the inter-class separation to the intra-class scatter. This is exactly the goal of classical discriminant analysis [121].

Suppose there is a set of n d -dimensional samples $\mathbf{x}_1, \dots, \mathbf{x}_n$, n_1 in the subset $\mathcal{D}_1$ labeled ω_1 and n_2 in the subset $\mathcal{D}_2$ labeled ω_2 . If a linear combination of the components of $\mathbf{x}$ is defined, then the scalar dot product is obtained

$$y = \mathbf{w}^T \mathbf{x} \quad (3.15)$$

and a corresponding set of n samples $y_1, \dots, y_n$ is divided into the subsets $\mathcal{Y}_1$ and $\mathcal{Y}_2$.

If $\mathbf{m}_i$ is the d -dimensional sample mean given by

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{x} \in \mathcal{D}_i} \mathbf{x}, \quad (3.16)$$

then the sample mean for the projected points is given by

$$\begin{aligned} \tilde{m}_i &= \frac{1}{n_i} \sum_{y \in \mathcal{Y}_i} y \\ &= \frac{1}{n_i} \sum_{\mathbf{x} \in \mathcal{D}_i} \mathbf{w}^T \mathbf{x} \\ &= \mathbf{w}^T \mathbf{m}_i, \end{aligned} \quad (3.17)$$

then

$$|\tilde{m}_1 - \tilde{m}_2| = |\mathbf{w}^T (\mathbf{m}_1 - \mathbf{m}_2)|. \quad (3.18)$$

The *scatter* for projected samples labeled ω_i is defined

$$\tilde{s}_i^2 = \sum_{y \in \mathcal{Y}_i} (y - \tilde{m}_i)^2. \quad (3.19)$$

Thus, $(1/n)(\tilde{s}_1^2 + \tilde{s}_2^2)$ is an estimate of the variance of the pooled data, and $\tilde{s}_1^2 + \tilde{s}_2^2$ is called the total *within-class scatter* of the projected samples. LDA employs that linear function $\mathbf{w}^T \mathbf{x}$ for which the criterion function

$$J(\mathbf{w}) = \frac{|\tilde{m}_1 - \tilde{m}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} \quad (3.20)$$

is maximum. When the $\mathbf{w}$ maximizes $J(\cdot)$, it leads to the best separation between the two projected sets. To obtain $J(\cdot)$ as an explicit function of $\mathbf{w}$, the *scatter matrices* $\mathbf{S}_i$ and $\mathbf{S}_W$ are defined by

$$\mathbf{S}_i = \sum_{\mathbf{x} \in \mathcal{D}_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \quad (3.21)$$

and

$$\mathbf{S}_W = \mathbf{S}_1 + \mathbf{S}_2, \quad (3.22)$$

then

$$\begin{aligned} \tilde{s}_i^2 &= \sum_{\mathbf{x} \in \mathcal{D}_i} (\mathbf{w}^T \mathbf{x} - \mathbf{w}^T \mathbf{m}_i)^2 \\ &= \sum_{\mathbf{x} \in \mathcal{D}_i} \mathbf{w}^T (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \mathbf{w} \\ &= \mathbf{w}^T \mathbf{S}_i \mathbf{w}, \end{aligned} \quad (3.23)$$

therefore the sum of these scatters can be written

$$\tilde{s}_1^2 + \tilde{s}_2^2 = \mathbf{w}^T \mathbf{S}_W \mathbf{w}. \quad (3.24)$$

Similarly, the separation of the projected means obeys

$$\begin{aligned} (\tilde{m}_1 - \tilde{m}_2)^2 &= (\mathbf{w}^T \mathbf{m}_1 - \mathbf{w}^T \mathbf{m}_2)^2 \\ &= \mathbf{w}^T (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{w} \\ &= \mathbf{w}^T \mathbf{S}_B \mathbf{w}, \end{aligned} \quad (3.25)$$

where

$$\mathbf{S}_B = (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T. \quad (3.26)$$

$\mathbf{S}_W$ is the *within-class scatter matrix*, which is proportional to the sample covariance matrix for the pooled d -dimensional data. Likewise, $\mathbf{S}_B$ is the *between-class scatter matrix*. For any $\mathbf{w}$, $\mathbf{S}_B \mathbf{w}$ is in the direction of $\mathbf{m}_1 - \mathbf{m}_2$. In terms of $\mathbf{S}_B$ and $\mathbf{S}_W$, the criterion function $J(\cdot)$ can be written as

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}. \quad (3.27)$$

This expression is well known in mathematical physics as the generalized Rayleigh quotient. A vector $\mathbf{w}$ that maximizes $J(\cdot)$ must satisfy

$$\mathbf{S}_B \mathbf{w} = \lambda \mathbf{S}_W \mathbf{w}, \quad (3.28)$$

where λ is constant. If $\mathbf{S}_W$ is nonsingular, a conventional eigenvalue problem can be obtained by writing

$$\mathbf{S}_W^{-1} \mathbf{S}_B \mathbf{w} = \lambda \mathbf{w} \quad (3.29)$$

then the solution for the $\mathbf{w}$ that optimizes $J(\cdot)$ is obtained

$$\mathbf{w} = \mathbf{S}_W^{-1} (\mathbf{m}_1 - \mathbf{m}_2). \quad (3.30)$$

Thus, $\mathbf{w}$ is obtained for linear discriminant, i.e., the linear function yielding the maximum ratio of between-class scatter to within-class scatter. In the case of color image conversion, the $\mathbf{w}$ should produce grayscale images that have large contrast between nuclei and background, and small variations within nuclei and background respectively. LDA was also used in image conversion in [6], but learning of $\mathbf{w}$ is based on accumulated RGB histogram.

2) Obtaining Transform $\mathbf{w}$ through Maximizing Separating Margin

Another popular measure for class separability is the so called separating margin [27]. Separating margin reflects the difference between samples of two classes. If interpreted in the context of image conversion, separating margin reflects contrast between nuclei and background. Seeking the maximal margin hyperplane for the classification is just the goal of SVMs as reviewed in Section 2.3. Assume the data are linearly separable in the feature space. (If not, it can be solved to be linearly separable in the kernel-induced feature space.) The functional margin 1

which associates with the hyperplane $(\mathbf{w}, b)$ implies

$$\langle \mathbf{w} \cdot \mathbf{x}^+ \rangle + b = +1, \quad (3.31)$$

$$\langle \mathbf{w} \cdot \mathbf{x}^- \rangle + b = -1. \quad (3.32)$$

The geometric margin γ is equal to $1/\|\mathbf{w}\|_2$ if the functional margin is set to 1.

$$\begin{aligned} \gamma &= \frac{1}{2} \left(\left\langle \frac{\mathbf{w}}{\|\mathbf{w}\|_2} \cdot \mathbf{x}^+ \right\rangle - \left\langle \frac{\mathbf{w}}{\|\mathbf{w}\|_2} \cdot \mathbf{x}^- \right\rangle \right) \\ &= \frac{1}{2\|\mathbf{w}\|_2} (\langle \mathbf{w} \cdot \mathbf{x}^+ \rangle - \langle \mathbf{w} \cdot \mathbf{x}^- \rangle) \\ &= \frac{1}{\|\mathbf{w}\|_2}. \end{aligned} \quad (3.33)$$

Given a linearly separable training sample

$$\mathbf{S} = ((\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)), \quad (3.34)$$

the hyperplane $(\mathbf{w}, b)$ that solves the optimization problem

$$\begin{aligned} &\text{minimize}_{\mathbf{w}, b} \quad \langle \mathbf{w} \cdot \mathbf{w} \rangle, \\ &\text{subject to} \quad y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b) \geq 1, \quad i = 1, \dots, l, \end{aligned} \quad (3.35)$$

realizes the maximal margin hyperplane with geometric margin $\gamma = 1/\|\mathbf{w}\|_2$. Now transform this optimization problem into its corresponding dual problem. The primal Lagrangian is

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \langle \mathbf{w} \cdot \mathbf{w} \rangle - \sum_{i=1}^l \alpha_i [y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b) - 1], \quad (3.36)$$

where $\alpha_i \geq 0$ are the Lagrange multipliers. Differentiate L with respect to $\mathbf{w}$

$$\frac{\partial L(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^l y_i \alpha_i \mathbf{x}_i = 0 \quad (3.37)$$

and obtain

$$\mathbf{w} = \sum_{i=1}^l y_i \alpha_i \mathbf{x}_i. \quad (3.38)$$

The transform $\mathbf{w}$ in Eq. (3.30) and (3.38) can be the projection axes for color image conversion that are based on LDA and SVM respectively. The training data set is constructed by the background and nuclei pixels which are selected randomly from a large number of the studied images. The transform $\mathbf{w}$ are learned from the data set, and therefore the new weighted linear combinations of R, G, and B color components are achieved as Eq. (3.39) and (3.40) respectively:

$$Y = 0.219R + 0.582G + 0.200B, \quad (3.39)$$

$$Y = 0.313R + 0.362G + 0.326B. \quad (3.40)$$

Namely, $\mathbf{w} = [0.219 \ 0.582 \ 0.200]^T$ is the projection axis for the LDA based color image conversion and $\mathbf{w} = [0.313 \ 0.362 \ 0.326]^T$ is for the SVM based color image conversion. The weights show the different proportion that red, green, and blue signal contribute to the final image contrast. In conventional image conversion, the transform $\mathbf{w}$ is fixed, regardless of the speciality of the application images under study. For example, in luminance-based image conversion, $\mathbf{w}$ is set to $[0.3 \ 0.59 \ 0.11]^T$ in accordance with the sensitivity of human eyes to R, G and B components. However, these values do not necessarily produce grayscale images that are ideal in a specific application and in machine-based object recognition. In contrast, our learning-based methods learn the transform from images of a specific application domain in the context of machine-based pattern classification, and are therefore expected to produce grayscale images leading to better segmentation results. Note that in

the learning-based methods, transform $\mathbf{w}$ is trained in a manner of once-for-all. Once the transform is obtained in the training stage, it is fixed and is used in the same way as the conventional color image conversion without demanding additional computation.

After the grayscale image is obtained, thresholding is applied to segment nuclei out of background. In our system, Otsu's histogram thresholding [28] in Eq. 2.15 is employed.

3.4 Comparison of C-G-T procedure with Pixel Classification-based Color Image Segmentation

In our system, the segmentation of color image follows a C-G-T process: Color image (C) $\rightarrow$ Grayscale image (G) $\rightarrow$ Thresholding (T). The novelty of the proposed method is the combination of global learning and local learning, i.e., the direction of the transform is learned from global pixel distribution of color images while the threshold is learned from local pixel distribution of the grayscale image. In pixel classification-based color image segmentation, segmentation is performed directly on the color image in the color space. Next we will show that for two-class segmentation problems, i.e., segmenting objects out of background, pixel classification in color space is equivalent to the C-G-T process, but the C-G-T procedure has extra flexibility.

We first consider the minimum distance-based pixel classification. In minimum distance-based pixel classification, it is assumed that pixels of homogeneous regions in the image form clusters in the color space. In cell images, the image pixels should form two clusters, one for nuclei and another for background. Assume the cluster

centers of background and nuclei pixels are denoted by $\mathbf{c}_1$ and $\mathbf{c}_2$ respectively, a color pixel $\mathbf{x}$ is classified as a nuclei or background pixel in terms of its distances to the two centers $\mathbf{c}_1$ and $\mathbf{c}_2$:

$$d_1 = \sqrt{(\mathbf{x} - \mathbf{c}_1)^T(\mathbf{x} - \mathbf{c}_1)} = \sqrt{\mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{c}_1 + \mathbf{c}_1^T\mathbf{c}_1} \quad (3.41)$$

$$d_2 = \sqrt{(\mathbf{x} - \mathbf{c}_2)^T(\mathbf{x} - \mathbf{c}_2)} = \sqrt{\mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{c}_2 + \mathbf{c}_2^T\mathbf{c}_2} \quad (3.42)$$

The decision is made as:

$$\mathbf{x} = \begin{cases} \text{nuclei pixel} & \text{if } d_1 \geq d_2 \\ \text{background pixel} & \text{if } d_2 > d_1 \end{cases} \quad (3.43)$$

Let's first consider the case of $d_1 \geq d_2$.

$$\begin{aligned} d_1 &\geq d_2 \\ \Leftrightarrow \mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{c}_1 + \mathbf{c}_1^T\mathbf{c}_1 &\geq \mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{c}_2 + \mathbf{c}_2^T\mathbf{c}_2 \\ \Leftrightarrow -2\mathbf{x}^T\mathbf{c}_1 + \mathbf{c}_1^T\mathbf{c}_1 &\geq -2\mathbf{x}^T\mathbf{c}_2 + \mathbf{c}_2^T\mathbf{c}_2 \\ \Leftrightarrow 2\mathbf{x}^T(\mathbf{c}_2 - \mathbf{c}_1) &\geq \mathbf{c}_2^T\mathbf{c}_2 - \mathbf{c}_1^T\mathbf{c}_1 \\ \Leftrightarrow (\mathbf{c}_2 - \mathbf{c}_1)^T\mathbf{x} &\geq \frac{1}{2}(\mathbf{c}_2^T\mathbf{c}_2 - \mathbf{c}_1^T\mathbf{c}_1) \\ \Leftrightarrow \mathbf{u}^T\mathbf{x} &\geq d \end{aligned} \quad (3.44)$$

where

$$\mathbf{u} = \mathbf{c}_2 - \mathbf{c}_1 \quad (3.45)$$

$$d = \frac{1}{2}(\mathbf{c}_2^T\mathbf{c}_2 - \mathbf{c}_1^T\mathbf{c}_1) \quad (3.46)$$

Similarly, for the case of $d_1 < d_2$, we have:

$$\begin{aligned} d_1 &< d_2 \\ \Leftrightarrow (\mathbf{c}_2 - \mathbf{c}_1)^T\mathbf{x} &< \frac{1}{2}(\mathbf{c}_2^T\mathbf{c}_2 - \mathbf{c}_1^T\mathbf{c}_1) \\ \Leftrightarrow \mathbf{u}^T\mathbf{x} &< d \end{aligned} \quad (3.47)$$

The above analysis shows that the minimum distance-based pixel classification in color space is equivalent to a procedure that maps color pixel vectors to scalars through transform $\mathbf{u}$ first, and then thresholds the scalars with d .

We next consider hyperplane-based pixel classification. The hyperplane-based classifier is described by:

$$y = \text{sgn}(\mathbf{v}^T \mathbf{x} - b) \quad (3.48)$$

$\mathbf{v}^T \mathbf{x}$ actually projects color pixel vector $\mathbf{x}$ to a scalar, and $\text{sgn}(\mathbf{v}^T \mathbf{x} - b)$ is equivalent to a thresholding operation:

$$\mathbf{x} = \begin{cases} \text{nuclei pixel} & \text{if } \mathbf{v}^T \mathbf{x} - b \geq 0 \\ \text{background pixel} & \text{if } \mathbf{v}^T \mathbf{x} - b < 0 \end{cases} \quad (3.49)$$

The above analysis reveals that pixel classification-based color image segmentation in color space is equivalent to the C-G-T process, i.e., converting the color image into grayscale and then segmenting the grayscale image through thresholding. Note this is true only for two-class segmentation problems: segmenting objects out of background. *In our cell image segmentation problem, background pixels and nuclei pixels are assumed to be linearly separable in the color space.* If these pixels are nonlinearly separable, we can employ kernel-based mapping to map the color pixels to a linearly separable new space, and then use the supervised learning-based method to perform image conversion. This is not within the scope of the current study because background pixels and nuclei pixels are linearly separable in our study.

In general image segmentation problems, the classifier for pixel classification can be obtained either through off-line learning from global pixel distribution or on-line learning from local pixel distribution. In automatic cell analysis, however, on-line local learning is impractical because on-line learning of the classifier demands user intervention to define training samples, i.e., nuclei pixels and background pixels.

Global learning of classifier for pixel classification ensures computational efficiency, but the classifier learned lacks adaptability. In contrast, the C-G-T procedure combines both off-line global learning and on-line local learning. Firstly, the global learning of $\mathbf{w}$ makes the image transform computationally efficient because it is trained on a once-for-all basis. Secondly, the global learning of $\mathbf{w}$ leads to small variations of grey levels among different images. Small variation is important in our system because in the following step of separation of overlapping nuclei, small variations facilitate the pattern classification-based spurious marker merging. Thirdly, the online local learning of threshold ensures good adaptability. Note the learning of threshold is based on Otsu's histogram thresholding, which is an unsupervised learning method demanding no user intervention.

3.5 Experimental Results and Discussion

3.5.1 Performance Evaluation of Learning-based Methods for Image Conversion

The transform $\mathbf{w}$ (equation $x \ y \ z$) for image conversion is obtained through supervised learning, where SVM and LDA are both used. The generalization performance of the transform to unseen test images heavily depends on the quality of the training data. In a loose sense, a good training dataset should consist of examples that are representative of the entire data. The distribution of nuclei and background pixels in the 3-D color space are mainly determined by the color and intensity of staining. To ensure good coverage of the distribution of nuclei and background pixels, the training data should be acquired from images with a variety of color and intensity. By visual inspection, 10 representative images with various staining color and intensity are selected from our database. To ensure the generability of the method,

we selected various images with different color and intensity over area of slide and from slide to slide. 2000 background pixels and 2000 nuclei pixels are then acquired as training data from these 10 images. Moreover, the testing data are totally unseen images to make sure the reliability of testing results.

The motivation of using supervised learning-based image conversion is to maximize separability between objects and background in the grayscale image obtained. Three measures are employed to quantitatively evaluate the performance of different image conversion methods.

1) Geometric Margin

Geometric margin, a simple and intuitive measure of class separability, is defined as follows:

$$M = |m_1 - m_2| \quad (3.50)$$

where m_1 and m_2 are the means of nuclei and background pixels in the grayscale image respectively.

2) LDA

Linear discriminant in Eq. (3.27) is probably the most commonly used measure of class separability.

3) Scatter-based Class Separability Measure

Scatter is also a frequently used class separability measure [35]. Assume the within-class scatter, between-class scatter, and total scatter of the grayscale image pixels are S_W , S_B and S_T respectively, the scatter-based class separability measure is defined as in [26]:

$$J_{BW} = \frac{S_B}{S_W} \quad (3.51)$$

or

$$J_{BT} = \frac{S_B}{S_T} \quad (3.52)$$

The testing results based on the measures of class separability are listed in Table 3.1, where the larger the value of the measures, the better the class separability. From the table we can see that the learning-based methods (the first three methods) are superior than the constant weighted linear combinations (the last three methods). This may be because in real-life images the color image components vary a lot over area in slide or from slide to slide, so the constant weighted linear combinations can't produce desired image contrast. While the learning-based color image conversion can achieve better parameters according to the global pixel distribution of the specific database. In addition, the results show that there is no much difference between the three learning-based color image conversions. Theoretically, LDA based method is supposed to provide good difference between different classes, SVM based is believed to give maximal separating margin between different classes, and PCA is supposed to capture the characteristics of entire database. While in practice, their performance may be influenced by the database, such as the size or distribution of the database.

Table 3.1: Measures of Class Separability				
method	M	$J(\mathbf{w})$	J_{BW}	J_{BT}
LDA based	10.977	28.683	14.370	0.935
SVM based	10.329	26.716	13.385	0.931
PCA based	10.230	26.174	12.512	0.931
luminance-based (RGB to YIQ)	7.999	23.732	11.890	0.922
intensity-based ((R+G+B)/3)	7.938	23.527	11.787	0.922
value-based (max(R,G,B))	5.264	21.550	8.941	0.903

3.5.2 Performance Evaluation of Segmentation

The ultimate goal of introducing supervised learning-based image conversion is to facilitate image segmentation. Next, the image conversion methods are evaluated in terms of segmentation results.

First, let us evaluate the segmentation results by human visual inspection, where the results can be assessed based on the smoothness of boundaries, number of miss-segmented pixels, intra-region uniformity, inter-region contrast, and so on [123,124]. One part of the segmented images is selected and the comparison in different segmentation approaches are shown in Fig. 3.2 and 3.3, where (a) to (f) are based on color image conversion followed by thresholding, and (g) is based on minimum distance-based pixel classification. In (a), color image is converted to grayscale image by LDA. In (b), the projection for color image conversion is by SVM. The projection in (c) is obtained by PCA. And the projections in (d), (e), and (f) are respectively achieved by conventional luminance-based transformation, intensity-based transformation, and value-based transformation which are fixed linear combinations of color components R, G, and B. As mentioned above, segmentation results can be evaluated by boundaries quality, number of miss-segmented pixels and so on. For example, in the region A which is marked by red polygon in Fig. 3.2(a), the nucleus boundary is more accurate compared to the same regions in the other figures. Boundary accuracy is not only a commonly used criterion to assess segmentation results, but also provides more accurate nuclei grading in this project which is based on color and intensity inside the nuclei. Secondly, some cases of overlapping are hard to avoid due to the nature of biopsy as region B in Fig. 3.2(a). But some overlapping can be avoided if the image segmentation is proper as region C in Fig. 3.2(a) compared to other figures. It relieves the further processing for overlapping nuclei separation if such case of overlapping can be avoided. Thirdly, noise or artifact is also a criterion to evaluate segmentation results as region D in Fig. 3.2 (b). From Fig. 3.2 and

Figure 3.2: Segmentation results, where (a) to (f) are based on color image conversion followed by thresholding, and (g) is based on minimum distance-based pixel classification. The projections in (a) to (f) for color image conversion are respectively by LDA, maximizing separating margin, PCA, luminance-based transform, intensity-based transform, and value-based transform (continued overleaf).

Figure 3.3: Figure 3.2 Continued.

3.3 we can see that the first six figures are superior to the last one. This is because two-step methods are more adaptable in the step of thresholding although two-step methods and pixel classification are equivalent theoretically. Fig. 3.2 and 3.3 also show the differences among the two-step methods, for example, less overlapping cases, less noise, and more accurate boundaries. Although the differences are not that much, they are helpful for the further processing.

Next, let's give quantitative evaluations on the segmentation results. As reviewed in Section 2.5.1, the evaluation measures with gold standard, including the boundary error metrics and area error metrics [122], can be used to measure discrepancy between a computer-aided segmented image and an ideally/manually segmented image. Here two boundary error metrics including Hausdorff distance and mean absolute distance, and three area error metrics including false positive volume fraction, false negative volume fraction and true positive volume fraction are employed and their computational equations are referred to Eq. 2.54-2.57. In our study, each image contains hundreds of cells, it is impractical to evaluate the boundary error and area error of the segmented cells one by one. In the experiment, the segmentation is evaluated on 100 randomly selected cell clumps including single nucleus and overlapping nuclei. The average error measures between the computer-aided methods and the manual results are listed in Table 3.2, where the manual results are regarded as ground truth. The measures listed in Table 3.2 are the smaller the better, except that 'true positive' is the larger the better. For example, the Hausdorff distance measures the worst possible disagreement between the manually-drew boundary and the computer-aided segmentation boundary. The details of the measures were explained in Section 2.5.1. The results in Table 3.2 are consistent with Fig. 3.2 and 3.3, i.e., the two-step procedures (the first six) outperforms pixel classification-based image segmentation (the last one) and the supervised learning-based methods esp. the LDA based are better than other two-step based methods.

Table 3.2: Evaluation of computer-aided segmentation results with manual segmentation results

Method	Hausdorff				Mean absolute				True			
	distance	distance	distance	distance	Positive	Negative	Positive	Positive	False	Negative	Positive	Positive
LDA based	2.234	0.837	0.022	0.044	0.956							
SVM based	2.236	0.857	0.030	0.045	0.955							
PCA based	2.241	0.870	0.031	0.046	0.954							
luminance-based (RGB to YIQ)	2.241	0.893	0.031	0.046	0.954							
intensity-based ((R+G+B)/3)	2.245	0.893	0.033	0.048	0.952							
value-based (max(R,G,B))	2.260	0.900	0.033	0.066	0.934							
minimum distance-based pixel classification	13.416	3.594	0.400	0.143	0.900							

Besides evaluation with gold standard, the statistical evaluation of segmentation without gold standard, $Q(I)$ in Eq. 2.58, is also employed that is based on boundaries accuracy, inter-region contrast, intra-region uniformity, and so on. 8 cell images are selected randomly for testing, and the average and standard deviation of the images are computed. The results in Table 3.3 show the consistent performance of the segmentation methods compared to Fig. 3.2, 3.3 and Table 3.2, namely, the C-G-T procedure outperforms pixel classification-based image segmentation and the supervised learning-based methods esp. the LDA based are better than other two-step based methods.

Here we would like to discuss a little more about the results from the PCA and LDA based approaches. PCA and LDA are both very popular used methods for the linear combinations of variables as described in Section 3.2. LDA is to maximize the ratio of between-class scatter to that of within-class scatter [203], while PCA deals with the data in its entirety without paying any particular attention to the underlying class structure [201]. There have been a lot of publications on the comparison of PCA and LDA, where most of them focus on the face recognition problem. Generally, it is believed that algorithms based on LDA are superior to those based on PCA [202, 203]. However, ref. [201] represents that when the training data set is small, PCA can outperform LDA and also PCA is less sensitive to different training datasets. The experimental results of our project are consistent with the general opinion, i.e., LDA performs better than PCA.

3.6 Concluding Remarks

In this chapter, a supervised learning-based two-step procedure for color cell image segmentation has been developed, where color image is first projected to grayscale by a transform learned through supervised learning, and then thresholding is performed

Table 3.3: statistical assessment $Q(I)$ of segmentation results on whole image

method	image 1	2	3	4	5	6	7	8	average	SD
LDA based	6.02	6.73	13.57	4.80	2.96	3.84	6.69	5.29	6.24	3.25
SVM based	6.08	7.22	14.74	15.14	2.98	6.03	7.82	5.52	8.19	4.40
PCA based	6.17	7.44	14.80	22.18	3.25	6.14	8.03	5.72	9.22	6.21
luminance-based (RGB to YIQ)	6.26	7.56	14.91	31.57	3.35	7.97	8.84	5.74	10.78	9.04
intensity-based $((R+G+B)/3)$	6.44	7.68	15.02	36.17	3.43	8.07	9.32	5.84	11.50	10.52
value-based $(\max(R,G,B))$	6.52	7.93	15.21	47.34	3.67	9.62	9.98	6.24	13.31	14.16
minimum distance-based pixel classification	6.54	7.96	15.30	48.16	3.72	9.74	10.03	6.33	13.47	14.43

on the grayscale image to segment nuclei out of background. In the proposed C-G-T procedure, the projection for color image conversion is learned off-line from the global pixel distribution of color images, and the threshold is learned on-line from local pixel distribution of the grayscale image. The combination of global and local learning makes the proposed method efficient. Here, supervised learning based on LDA and SVM have been both discussed. Experimental results show that the proposed method provided satisfying results compared with the existing segmentation methods, and it achieved a boundary disagreement (mean absolute distance) of 0.85 while the disagreement produced by the pixel classification-based color image segmentation method is 3.59.

In this chapter we discussed the idea of projecting color image to grayscale. On one hand, the traditional mapping of RGB to grayscale (as described in section 3.2) is one of a range of projections that can be used and they are based on the optical or human eyes sensitivities. On the other hand, the color filter for contrast enhancement can do the equivalent function. For example, the Kodak Wratten color filters are used for contrast control in black and white photomicrography when using stained specimens. To enhance contrast, color filters are added to the microscope light path to absorb stained specimen color, rendering it either a darker or lighter shade of gray. Contrast can be adjusted in this manner by selectively choosing filters that absorb varying amounts of stain color [204]. Also there are some other filters for contrast enhancement [205].

The pattern classification techniques discussed here such as PCA and LDA to determining the optimal projection is obvious to the people working in the field of pattern recognition. A natural extension to the use of a three color (RGB) system would be to use hyperspectral cameras which can operate in thousands of color band. Dimension reduction is a common preprocessing step performed on many hyperspectral imaging datasets, and theoretically it is possible that a lossy opera-

tion such as dimension reduction might have a negative impact on the final image processing performance. Therefore it is crucial to choose the proper approach for the dimension reduction. Ref. [206] proposed that in many cases PCA can have a minimal impact on the data set. We do have interest to apply the powerful pattern classification approaches on hyperspectral images for the purpose of dimension reduction and other tasks such as target detection.

Chapter 4

Marker Detection for Watershed-Based Separation of Overlapped Nuclei

4.1 Introduction

In Chapter 3, a new method has been presented for segmenting nuclei from the background effectively. However, it is unable to separate overlapping or touching nuclei which is a main issue in cell image segmentation. Further processing is needed in order to obtain accurate counts of positively and negatively stained nuclei. As reviewed in the image-based segmentation methods in Section 2.3.2, a watershed algorithm which combines region growing and edge detection techniques, can be used to deal with the separation of overlapped or clumped objects [16, 63]. However, the problem of over-segmentation often occurs when watershed algorithm is directly applied to the image. As mentioned in Section 2.3, there are two main strategies to solve or alleviate this problem, where the first strategy is marker-controlled watershed and another strategy is watershed algorithm followed by region merging. These

two types of methods are based on the different ideas and they are alternative rather than complementary. The marker-controlled based algorithm is developed in Chapter 4. And a watershed and region merging based algorithm is developed in Chapter 5, which will be discussed in the next chapter.

In the literature, marker detection is usually based on photometric characteristics, i.e., grey levels of objects [16, 20, 21], or shape information [16, 22], or both [18, 22]. We observe that in cell image segmentation, the combination of different types of images provides more accurate and more reliable markers. For instance, when the staining of a nucleus is severely uneven, if only the grayscale image is considered, the nucleus is often oversegmented, even if the marker-controlled watershed algorithm is used. However, over-segmentation can be alleviated if the complemented distance image [23] which reflects shape of objects is used. On the other hand, when objects are severely overlapped or clumped, shape information of the objects is no longer reliable. In such a case, it is possible to use photometric information for correct segmentation. Based on the above considerations, we propose to combine the grayscale image and the complemented distance image to enhance watershed-based cell image segmentation.

The combination of different types of images is not straightforward. In the present study, the combination is implemented in the framework of pattern classification. The inputs of the classifier are characteristic variables extracted from valley line between two candidate markers in both the grayscale and the complemented distance images, while the output of the classifier is the decision of whether the two markers should be merged. Experimental results on a large number of histology cell images of bladder inverted papilloma demonstrate that the combination of grayscale and complemented distance images produces more reliable markers, and hence more accurate segmentation.

This chapter is organized as follows. Section 4.2 reviews the commonly used marker

detection algorithms. Section 4.3 presents the basic idea of our proposed method. Section 4.4 implements the proposed method, which begins with crucial features selection and follows the framework of classification. The experimental results are given in Section 4.5, which demonstrates the ability of the proposed method on a number of cell images. Finally, the conclusions are given in Section 4.6.

4.2 Commonly Used Approaches for Marker Detection

The root cause of the over-segmentation problem of the brute force watershed algorithm is the presence of spurious regional minima. The goal of marker detection is to identify the real markers corresponding to interested objects and remove the spurious markers. As reviewed in Section 2.3, many approaches have been proposed for marker detection [5, 14–17, 75, 76, 106, 107], most of which are based on image reconstruction techniques [5, 16]. The h-extrema transform [121] is the widely used marker detection approach based on grayscale reconstruction. The h-extrema transform includes h-maxima and h-minima, where the h-maxima transform suppresses all maxima whose depth is smaller than a given threshold level h , while the h-minima transform suppresses all minima whose depth is smaller than h . Take h-maxima for instance, the transform is achieved by performing the grayscale reconstruction of image I from image $I - h$

$$\text{HMAX}_h(I) = R_I(I - h), \quad (4.1)$$

where HMAX means h-maxima transform and R means grayscale reconstruction. The h-maxima transform is illustrated in Fig. 4.1, where grayscale reconstruction of I from $I - h$ is implemented. The regional maxima are marked by the red lines. The

illustration shows that there are ten maxima in the original image I , but only three main maxima are detected in the reconstructed image where the maxima whose depth is smaller than h are suppressed.

Figure 4.1: h -maxima transform, where grayscale reconstruction of I from $I - h$ is implemented. The regional maxima are marked by the red lines.

4.3 The Basic Idea of the Proposed Algorithm

Marker detection is usually based on three types of images including grayscale image, gradient image, and complemented distance image. Distance image is the distance transform of a binary image. For each pixel in a binary image, the distance transform assigns a number that is the distance between that pixel and the nearest nonzero pixel of the binary image, see for example [23]. Complemented distance image is the complement of the distance image, i.e., each pixel value in the distance image is subtracted from the maximum pixel value, and the difference is used as the pixel value in the output image [29]. The performance of the marker-controlled watershed algorithm on the three types of images is evaluated using a number of cell images of bladder inverted papilloma, where the images are with the size of 1392×1040 (the same size as Fig. 2.8). Table 4.1 lists eight exemplar images only, which represent the different cases. For example, image No. 4 has severe over-segmentation problem

due to the artifact and small morphological structures, while image No. 8 has more under-segmentation problem because of the severely overlapping nuclei. Usually the segmentation performance is assessed in terms of the number of over- and under-segmentation. Over-segmentation means one object is over partitioned into several parts, while under-segmentation means overlapped or clumped objects are not separated. In the present application, each image contains hundreds of nuclei, and a large portion of the nuclei overlap or touch. Table 4.1 shows that the complemented distance image produces the least number of over-segmentations. This is because the complemented distance image has fewer minima than the grayscale and gradient images, and is less influenced by grey level unevenness of objects, as shown in the example of Fig. 4.2. We notice that there is severe grid bias on Fig. 4.2 (e), i.e., distance measure preference for square contours. The simple distance measure *Euclidean distance* was used here, actually there are some better distance measures based on Fast Marching algorithm which do not exhibit grid bias [193]. Here what we need is the regional minima extracted from the distance image (e), rather than the accurate boundary. Therefore the grid bias doesn't influence the correctness of region minima extraction. In addition, Euclidean distance is so simple and with very low computational load. Therefore we use the Euclidean distance measure for distance image implementation. A nucleus whose grey level is severely uneven is shown in Fig. 4.3. Although the h-minima transform is performed to reduce spurious markers as shown in Fig 4.3(b), the nucleus is still split into several partitions as shown in Fig 4.3(c). If marker detection is based on the complemented distance image as shown in Fig 4.3(d), the over-segmentation problem is avoided as shown in Fig 4.3(e). Even if the grey level is rather homogeneous, shape information could still be valuable. Fig. 4.4 shows an example of overlapping or touching objects, where the watershed line between two objects is quite shallow. In such a case, under-segmentation often occurs if the marker detection is performed on the grayscale image. But correct segmentation can be achieved if the complemented

Table 4.1: Comparison on accuracy of segmentation results on different types of images

image index	No. of		No. of		No. of oversegmented		No. of undersegmented	
	isolated	overlapping	or touching	grayscale image	gradient image	distance image	grayscale image	gradient image
1	235	179	1	6	1	6	13	2
2	381	235	38	45	4	2	4	7
3	486	191	23	27	9	2	8	5
4	216	264	51	64	47	0	2	3
5	152	142	22	22	0	2	1	12
6	328	189	73	89	27	2	0	15
7	448	258	3	0	0	12	15	7
8	450	194	1	0	0	9	3	10
average	321	208	27	32	11	4	6	8

distance image is used. On the other hand, grayscale image could be more reliable in segmenting severely overlapped or clumped objects as shown in Fig. 4.5. Due to severe overlapping, the convex character of nuclei becomes less obvious and some important minima of the complemented distance image are missed as shown in Fig. 4.5(c). As a result, the overlapped nuclei will not be separated. In such a case, photometric information i.e., grey level is valuable to provide reliable markers as shown in Fig. 4.5(e). Table 4.1 shows that the grayscale image is the best choice if under-segmentation is a concern.

Based on the above analysis, we propose to employ both grayscale and complemented distance images in marker detection to improve the accuracy and reliability of segmentation. The flow diagram of the proposed method is shown in Fig. 4.6, and basically it consists of four steps. Step 1 is to extract the so called 'valley line' between two minima, where the valley line represents the morphological information between the minima. Step 2 is to extract the features from the valley line that summarize the characteristics of the valley line. Step 3 is to select the discriminating features from all the collected features. In the step 4, these discriminating features are as input variables to a classification-based minima merging algorithm, and finally the minima are merged according to the meaningful objects. The details of implementation of the proposed method are given in the next section.

4.4 Detecting Markers using Both Grayscale and Complemented Distance Images

As analyzed in Section 4.3, both photometric and shape information are to be used for marker detection. However, how to combine the two types of information is not straightforward. Ref. [18] extracts seeds for the watershed algorithm based on intensity and edge information, and then implements region merging on the

Figure 4.2: Different types of images and associated minima. (a) grayscale image, (b) minima of grayscale image, (c) gradient image, (d) minima of gradient image, (e) complemented distance image, (f) minima of complemented distance image.

Figure 4.3: An example that grayscale image does not work well (severely uneven grey level). (a) grayscale image, (b) markers obtained by h-minima based on grayscale image, (c) segmentation result with markers (b), (d) marker obtained by h-minima based on complemented distance image, (e) segmentation result with marker (d).

Figure 4.4: An example that grayscale image does not work well (shallow watershed line). (a) grayscale image, (b) topographic surface of the grayscale image.

Figure 4.5: An example that complemented distance image does not work well (severe overlapping). (a) grayscale image, (b) complemented distance image, (c) minima of complemented distance image, (d) minima of grayscale image, (e) markers obtained by h-minima based on grayscale image.

Figure 4.6: Flow diagram of the proposed marker detection method

watershed image based on edge information, and finally separates the clusters of nuclei based on the shape of the cluster. In [22], the combination is implemented by performing a distance transform on a modified gradient image, but post-processing is still needed to reduce over-segmentation.

In the present study, we propose to combine the photometric and shape information in the framework of pattern classification, where characteristic variables representing the relationship between two markers in the grayscale image and the complemented distance image are used as independent inputs to the classifier, and the output of the classifier is the decision of whether the two markers should be merged.

4.4.1 Input Variables to Classifier

In mathematic morphology, a grayscale image is usually considered as a topographic surface. In human visual perception, the relationship between two markers is judged based on the distance between the two markers as well as the topographic surface

between two markers. Usually, the distance between two markers belonging to different nuclei is large, and the topographic surface exhibits significant variations. Based on the human visual perception, the information underlying the topographic surface is employed to determine whether two markers should be merged. Actually, the main information of topographic surface between two markers is contained in the valley line between the two markers. In the present study, the input variables to the classifier are features extracted from the valley line. The features are obtained in three steps:

4.4.1.1 Step 1: Extract Valley Line between Two Minima

In hydrology, the flowing water seeks the easiest downhill route, which is the one that follows the steepest slope. Valley lines can be loosely defined as the site where rain water in a terrain gathers to run downhill. In morphology, a valley line is the region whose grey levels are lower than surrounding areas [30]. The valley line can be achieved by using the method in [31], which is by looking for zero crossings of the first directional derivative taken in a direction which maximizes the second directional derivative. In the present study, the valley line between two minima is sought; hence, the minimal or geodesic path linking these two points is employed here [16].

Let's first recall the *geodesic time*. Consider a path $\mathcal{P}$ defined on the domain of a grayscale image f . The time required to cover $\mathcal{P}$ is defined as the integral of the value of f along $\mathcal{P}$ and is denoted by $\tau_f(\mathcal{P})$:

$$\tau_f(\mathcal{P}) = \int_{\mathcal{P}} f(s) ds. \quad (4.2)$$

This definition is illustrated in Fig. 4.7. Note that the designation 'time' is justified by considering units equal to the inverse of that of a speed for the image f , i.e., the

lower the grey level is, the faster the propagation is.

Figure 4.7: The time required to cover the path $\mathcal{P}$ on the image f equals the integral of the value of f along $\mathcal{P}$, where (a) is the path $\mathcal{P}$ between points p and q on an image f , and (b) is the values of f along $\mathcal{P}$.

Given two points p and q in an image f , there is a minimal path between the two points in the sense that the sum of the grey levels along these path is minimum. By definition, the sum $\tau_f(x, p) + \tau_f(x, q)$ is constant and equals $\tau_f(p, q)$ for all points x belonging to the geodesics linking p to q . Hence, the points belonging to the path can be extracted by thresholding the sum of the geodesic time function from p and q at the value corresponding to $\tau_f(p, q)$, i.e., the value of $\tau_f(p)$ at q or $\tau_f(q)$ at p .

The minima in the grayscale image to be analyzed are shown in Fig. 4.8. The valley line between minima 1 and 3, and valley line between minima 1 and 4, which associate with two different nuclei and the same nucleus respectively, are also given in Fig. 4.8. Actually, oriented gradient or tensor metrics [209] can be used to achieve good gradient of the image, which demonstrate great performance in distinguishing between light to dark and dark to light gradients and improve segmentation processing.

Figure 4.8: Extracting valley line between two minima, where valley line between minima 1 and 3, between minima 1 and 4 are shown respectively.

4.4.1.2 Step 2: Extract Characteristic Variables from the Valley Line

The curve in Fig. 4.9 shows the grey levels of pixels on a valley line, where x-coordinate denotes the position of points on the valley line in sequence, and y-coordinate is the grey value of the points. Fig. 4.9 (a) illustrates the valley line between minima 1 and 3 which belong to two different nuclei, and (b) is the valley line between minima 1 and 4 which belong to the same nucleus. The figures illustrate that the curve between two minima belonging to two different nuclei has a shape of 'hill': the grey level changes gradually from a relatively low value up to a higher value, and then down to a lower value again. The hill shape of the curve is due to the nature of cell image, i.e., the pixels of a nucleus are darker in the central region while pixels of touching region are brighter. Usually this character is not significant in the valley line between two minima in the same nucleus. As shown in Fig. 4.10, a valley line is characterized by path length, uphill height, downhill height, uphill area, downhill area, uphill gradient, and downhill gradient. Since the uphill variables including gradient, area and height are not consistently greater or smaller than downhill variables, we define the following 7 variables as features: path

length, maximum height (the greater value between uphill and downhill heights), minimum height (the smaller value between uphill and downhill heights), maximum area, minimum area, maximum gradient, and minimum gradient.

Figure 4.9: (a) Valley line in the grayscale image between minima 1 and 3, (b) valley line in the grayscale image between minima 1 and 4.

In the above, only the grayscale image is considered. As analyzed in Section 4.3, for nuclei with uneven staining, the photometric information is not adequate for correct detection of markers. To improve reliability of markers, the complemented distance image which reflects shape information of objects is incorporated into marker detection in this study.

The valley line between two minima belonging to different nuclei in the complemented distance image also exhibits a hill shape. This is because the complemented

Figure 4.10: Features for describing curve, where 7 features including path length, uphill height, downhill height, uphill area, downhill area, uphill gradient, and downhill gradient are employed.

distance is small at the central region of nuclei, and is large at the overlapping or touching region. Note, the minima are detected based on grayscale image. Similarly, 7 features can be extracted from the complemented distance image.

4.4.1.3 Step 3: Select Features based on Class Separability

The 14 variables extracted from both images are not necessarily all useful or needed in determining the relationship between two markers. Feature selection techniques are employed to find out critical variables to the decision. The goal of feature selection is to identify and select the most important and nonredundant variables from the large pool of potential variables [32]. Generally, feature selection algorithm involves two parts: (1) feature evaluation criterion and (2) search algorithm, where the evaluation criterion assesses the capacity of features to distinguish one class from another, while the search algorithm explores the potential solution space. In the present study, Mahalanobis distance class separability measure and the orthogonal sequential forward search procedure [32] are employed to evaluate the importance of the characteristic variables.

Mahalanobis distance class separability measure [35] is the commonly used evaluation criterion. For a two-class classification problem, it can be defined as

$$J(i, j) = [\mathbf{m}_i - \mathbf{m}_j]^T \mathbf{C}_{ij}^{-1} [\mathbf{m}_i - \mathbf{m}_j], \quad (4.3)$$

where $\mathbf{m}_i = [m_{1i}, m_{2i}, \dots, m_{ni}]^T$ is the mean vector of samples in n -dimensional class i . $\mathbf{C}_{ij} = \mathbf{C}_i + \mathbf{C}_j$, and $\mathbf{C}_i$ and $\mathbf{C}_j$ are the covariance matrices of classes i and j , respectively.

Sequential forward selection (SFS) is the widely used search algorithm [35]. SFS is a bottom-up procedure, i.e., it starts with an empty set, and then selects a new feature in the unselected set that maximizes the evaluation criterion, and adds this feature into the selected set until some stopping criterion is satisfied. Ref. [32] combines an orthogonal transform and SFS with the goal of reducing correlations or interactions among candidate features in the search procedure. Suppose there are N samples $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(N)$, and each sample is n -dimensional vector $\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_n(k)]^T$. Feature vector $\mathbf{x}_i$ and feature matrix $\mathbf{X}$ are defined as

$$\mathbf{x}_i = [x_i(1), x_i(2), \dots, x_i(N)]^T, \quad (4.4)$$

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(N) & x_2(N) & \dots & x_n(N) \end{bmatrix}. \quad (4.5)$$

The feature matrix $\mathbf{X}$ can be decomposed as

$$\mathbf{X} = \mathbf{Q}\mathbf{R}, \quad (4.6)$$

where $\mathbf{Q}$ is the orthogonal matrix and $\mathbf{q}_i$ is the new feature vector in the orthogonal

space, and $\mathbf{R}$ is an upper triangular matrix

$$\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n] = \begin{bmatrix} q_1(1) & q_2(1) & \dots & q_n(1) \\ q_1(2) & q_2(2) & \dots & q_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ q_1(N) & q_2(N) & \dots & q_n(N) \end{bmatrix}, \quad (4.7)$$

$$\mathbf{R} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ & \alpha_{22} & \dots & \alpha_{2n} \\ & & \ddots & \vdots \\ & & & \alpha_{nn} \end{bmatrix}. \quad (4.8)$$

In the Gram-Schmidt orthogonal transform, the orthogonal matrix $\mathbf{Q}$ is achieved by

$$\mathbf{q}_1 = \mathbf{x}_1, \quad (4.9)$$

$$\mathbf{q}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} \alpha_{ji} \mathbf{q}_j, \quad (4.10)$$

where

$$\alpha_{ji} = \begin{cases} \frac{\mathbf{q}_j^T \mathbf{x}_i}{\mathbf{q}_j^T \mathbf{q}_j} & \text{for } j = 1, 2, \dots, i-1 \\ 1 & \text{for } j = i. \end{cases} \quad (4.11)$$

The orthogonal forward selection (OFS) algorithm [32] that combines Gram-Schmidt orthogonal transform and SFS is implemented in three steps:

1. All the variables x_i ($i = 1, 2, \dots, n$) are candidates for $\mathbf{q}_1$

$$\mathbf{q}_1^{(i)} = \mathbf{x}_i. \quad (4.12)$$

Compute Mahalanobis measures J by $\mathbf{q}_1^{(i)}$. The variable $\mathbf{x}_j$ that maximizes J is identified and is added to the feature subset. Let $\mathbf{q}_1 = \mathbf{x}_j$.

2. All the remaining $n - 1$ variables are candidates for $\mathbf{q}_2$

$$\mathbf{q}_2^{(i)} = \mathbf{x}_i - \alpha_{12}^{(i)} \mathbf{q}_1 \quad 1 \leq i \leq n, \quad i \neq j, \quad (4.13)$$

where

$$\alpha_{12}^{(i)} = \frac{\mathbf{q}_1^T \mathbf{x}_i}{\mathbf{q}_1^T \mathbf{q}_1}. \quad (4.14)$$

Computer corresponding Mahalanobis measures J , and the feature maximizing J is identified and is added to the feature subset.

3. Step 2 is continued until the class separability measure J becomes stable, i.e., there is no more improvement for class separability by including more features.

In this study, 200 representative valley lines are extracted as training samples, where 100 are valley lines between markers that should be merged, and another 100 are valley lines that should not be merged. Fig. 4.11 shows that the class separability is stable until the first six features are selected, where the class separability is 0.823. Therefore the first six features are selected, including 3 variables the from grayscale image (path length, minimum gradient, minimum area) and 3 variables from the complemented distance image (maximum area, minimum gradient and maximum gradient).

4.4.2 Multiple Minima Merging based on Pattern Classification

After feature selection, only 6 variables selected are input to the pattern classifier to decide whether two minima should be merged. The pattern classifier is trained using the same training data as those used in feature selection.

Several classification schemes are investigated in our study including two linear

Figure 4.11: The relationship between class separability and number of selected features, where class separability becomes stable until the first 6 important features are selected.

classifiers, i.e., linear Support Vector Machines (SVMs) [27] (refer to Eq. 2.29) and Linear Discriminant Analysis (LDA) [121] (Eq. 3.30), and two nonlinear classifiers, i.e., SVMs with a polynomial kernel [27] and a Gaussian kernel [27] (Eq. 2.30 and Eq. 2.31). Experimental results show that SVMs with a Gaussian kernel outperforms other classifiers, and is therefore chosen in our study. The parameters of the Gaussian Kernel classifier are learned using the SVMs Toolbox [119]. The cost of constraint violation C is set to 1, and the width of the Gaussian kernel is set to 0.5. The learning phase is once-for-all, i.e., once the classifier is learned, it is fixed for test images without demanding further learning.

The algorithm of minima-merging using both photometric and shape information is described in Table 4.2. Refer to Fig. 4.8 for an instance. There are 5 minima in the grayscale image, and the merging procedure starts from minimum 1. The minimum 4 is the nearest to minimum 1, and hence the relationship between minima 1 and 4 is first studied. This includes valley line extraction, feature extraction, and relationship determination through classification. The result is that the two markers should be merged. In the second iteration, the nearest minimum to minima 1-4 cluster is

Table 4.2: Algorithm of minima-merging

```

% Initialization:

Look-up table of distance  $D$ 

Current table of distance  $D_C = D(1, 1 : N)$  % starts with minimum 1

Label of minima  $L_M = [1, 0, \dots, 0]$  % minimum 1 is 1, other are 0

Number of nuclei  $K=1$ 

% Processing:

while  $all(L_M) \neq 1$  % loop ends when all the minima are processed

     $i, j = argmin(d_{ij}), d_{ij} \in D_C$  % get nearest two minima i and j

    Study the relationship between minima i and j:

        Compute valley line on the grayscale image between minima i and j

        Compute valley line on the complemented distance image

        Get six features from two curves

        Input features to classifier, and output:

            1 % minima i and j should be merged

                 $d_{ij} = LargeValue, d_{ji} = LargeValue$  % update  $D$ 

                 $L_M(1, i) = K, L_M(1, j) = K$  % assign i, j the same label  $K$ 

                 $D_C = D(find(L_M == K), 1 : N)$  % update  $D_C$ 

            0 % minima i and j should not be merged

                 $K = K + 1$ 

                 $L_M(1, j) = K$  % a new nucleus

                 $D_C = D(j, 1 : N)$  % update  $D_C$ 

end

```

searched, and minimum 3 is found to be closest to minimum 1. The relationship between minima 3 and 1 is studied, and the result is that minimum 3 should not be merged with minima 1-4 cluster. Therefore minimum 3 is considered as a new marker. The iterative procedure is continued until all the minima are processed. The clustered minima are considered as markers, based on which watershed algorithm is applied to separate overlapping nuclei.

4.5 Experimental Results and Discussion

As introduced in Chapter 1, in the present automatic cell analysis system, nuclei are first segmented from the background using supervised learning based method as described in Chapter 3, and then the proposed marker detection algorithm is employed to guide the watershed transform for separating the overlapped or clumped nuclei. The user interface of the automatic cell analysis system is given in Fig. 4.12, where the user can choose the image to be analyzed, and then the percentages of nuclei in each grade as well as the processed image are shown. The centers of each nucleus in the segmented image are marked by a red cross.

To merge or not merge two markers is decided by the classifier, which is learned through supervised learning. As discussed in Section 4.3, the training data should cover the distribution of entire data as comprehensive as possible to achieve good generalization. In marker merging, the major factors of concern are nuclei shape, size, and degree of overlapping. Therefore the training dataset for marker merging classifier should consist of nuclei with various shapes, sizes and degrees of overlapping. In the experiment, 200 training samples (nuclei) were acquired from the 10 images used in Section 3.5.

The segmentation results of part of an exemplar testing image are shown in Fig. 4.13. (a) is the grayscale cell image converted from color cell image using the

Figure 4.12: User interface of automatic cell analysis system.

Figure 4.13: Comparison of segmentation results. (a) grayscale cell image, (b) segmentation result by using watershed algorithm directly. (c), (d) and (e) segmentation results by marker-controlled watershed, where marker detection is performed on grayscale image, gradient image, and complemented distance image, respectively, all using h-minima algorithm. In (f), markers are detected by the proposed supervised minima-merging algorithm (continued overleaf).

Figure 4.14: Figure 4.13 Continued.

method as described in Chapter 3. (b) is segmentation result by using watershed algorithm directly. (c), (d) and (e) are segmentation results of marker-controlled watershed, where marker detection is performed on grayscale image, gradient image, and complemented distance image, respectively, all using h-minima algorithm. In (f), markers are detected by the proposed supervised minima-merging algorithm. The figures illustrate that the proposed marker detection method can separate overlapped nuclei very well corresponding to the real nuclei.

A large number of cell images have been applied to test the proposed method. The same exemplar testing images as used for marker detection on the different kinds of images (Table 4.1) are employed to the proposed method for comparison. A comparison between Table 4.1 and 4.3 shows that the combination of grayscale image with complemented distance image produces improved results of 0.4% over-segmentation and 0.2% under-segmentation, compared with 4.4% over-segmentation and 1.1% under-segmentation of the popular reconstruction-based method. The evaluation measure $Q(I)$ is also applied to the 8 exemplar testing images, and the results in Table 4.4 show the good performance of the proposed method again.

Computational time is also considered in the study. The image segmentation is performed on NEC PC with Intel Pentium 4, 2.4GHz, 512MB total memory, and 512KB Cache RAM. All the programs are implemented in Matlab language. H-minima transform applied on grayscale images are used to compare the execution time with our proposed method. In addition, there is no much difference on execution time among using h-minima on grayscale, complemented, or gradient images in our project. This is because what h-minima does is to reconstruct a marker image from the mask image and image type doesn't have much influence on the transform. As Table 4.5 shows, the average time needed by the proposed method is longer than that of h-minima, however, the increase of computation brings improvement in segmentation accuracy.

Table 4.3: Comparison on accuracy of the proposed method		
image index	No. of oversegmented	No. of undersegmented
1	1	2
2	4	0
3	0	1
4	2	0
5	2	1
6	8	3
7	1	2
8	0	1
average	2	1

4.6 Concluding Remarks

In this chapter, a new marker detection algorithm has been presented for watershed-based separation of overlapping or touching nuclei. The proposed method makes use of both photometric and shape information of the nuclei, and combines these two kinds of information in the framework of pattern classification. The accuracy of the segmentation results provided by the proposed method is improved compared with the commonly used marker detection algorithm (h-minima). The computation load of the proposed method may be reduced to make the algorithm more efficient through some attempts, for example, grouping the regional minima by means of fuzzy clustering technique instead of by using the analysis on valley line.

Table 4.4: Statistical assessment $Q(I)$ of segmentation results.

image	h-minima			proposed
index	grayscale	gradient	complemented	method
	image	image	distance image	
1	17.7	19.3	11.5	6.8
2	16.4	18.6	11.0	7.6
3	17.1	18.4	11.5	7.3
4	22.8	24.7	14.4	9.7
5	9.8	10.6	7.2	6.4
6	41.6	45.7	36.9	9.6
7	19.8	22.3	17.4	10.6
8	27.2	29.6	22.5	19.5
average	21.6	23.7	16.6	9.7

Table 4.5: Comparison on Execution Time (in seconds)

image index	h-minima	the proposed method
1	30	41
2	39	33
3	31	66
4	16	19
5	24	24
6	30	42
7	48	56
8	26	58
average	31	42

Chapter 5

Region Merging for Watershed-Based Image Segmentation

5.1 Introduction

As mentioned in previous chapters, basically there are two types of methods to address the issue of over-segmentation in the watershed-based image segmentation. The first one is the marker-controlled watershed algorithm, which extracts markers of objects and background and then introduces the markers into the watershed algorithm to allow the segmentation of the selected regions exclusively [99]. Due to the complexity of real images, however, it is not always possible to associate obvious geometric or photometric characteristics with the objects to be segmented, and therefore accurate marking is sometimes difficult or even impossible [99].

The second type of approaches aims at merging the catchment basins of the watershed image that belong to almost homogeneous regions [99]. Some of the approaches

are based on image reconstruction, such as the waterfall algorithm, see [99] for example and the references therein. Other methods in this category take the output of watershed transform of the image as the starting point of a bottom-up hierarchical merging procedure, and merge the most similar pair of adjacent regions at each step [108]. Usually, the dissimilarity function [96] of the edge between adjacent regions is based on the region homogeneity, such as mean [108–110], variance [109], second-order and third-order central moments [111] of the intensity value [108] or color value [109] of the image. Besides region homogeneity, the characteristics of edge or common boundary between two adjacent regions are also used in dissimilarity function, such as edge magnitude [110, 112], edge integrity [96], and edge orientation [96]. The joint criteria combining region homogeneity and edges have been proposed and demonstrated to give better results than using them separately [96, 110, 112]. Generally, the combinations of the region information and edge information are implemented in a weighted fashion [112], where the weights control the relative importance of the terms and usually they are set or adjusted manually.

In mathematic morphology, an image can be represented as a topographic surface. It is easy to tell whether a pair of adjacent regions that should be merged or should not be merged through human visual inspection of the topographic surfaces of the two regions. Motivated by this, we propose to use topographic characteristic variables to define dissimilarity function. In addition, the topographic characteristic variables are combined and processed in the framework of pattern classification, i.e., the characteristics variables extracted from topographic surface are used as input variables to the classifier and the output of the classifier is the decision value. In this thesis, region merging based on both supervised classification and unsupervised classification are discussed. The strength of the proposed method is that the characteristic variables are derived based on human visualization, and is expected to provide better results because humans are effective processors of visual information. In addition, the characteristic variables are combined naturally through classifica-

tion techniques, and hence the weights of the variables are learned automatically rather than the manual adjustment of weight parameters as done in [96,108,112,113].

This chapter is organized as follows. Section 5.2 reviews the existing region merging approaches. Section 5.3 analyzes the topographic characteristics of the adjacent regions. Section 5.4 presents the proposed supervised-based region merging algorithm, which makes use of the topographic information of the adjacent regions. Section 5.5 applies the similar idea to the framework of unsupervised classification. The experimental results and discussion are given in Section 5.6. The supervised-based region merging demonstrated good ability to control the over-segmentation and under-segmentation problems. Finally, the concluding remarks are given in Section 5.7.

5.2 Brief Review of Region Merging

Watershed-based region merging basically contains several issues: image smoothing, gradient image generation, representation of image partition, merging sequence, dissimilarity function, and merging stop criterion. Let's review them briefly.

Prior to the image partitioning using a watershed algorithm, the original image needs to be smoothed to avoid severe over-segmentation. The commonly used image reconstruction-based filtering [100,101] can be employed to filter out the noise or minor structure in the image, where *opening by reconstruction* is employed first to filter the spurious maxima, and then *closing by reconstruction* is used to filter the spurious minima. Opening by reconstruction means reconstructing the original image f from its opening $\gamma_B(f)$, where opening is defined as the erosion of f by the structuring element (SE) B followed by the dilation with the transposed SE $\tilde{B}$:

$$\gamma_B(f) = \delta_{\tilde{B}}[\varepsilon_B(f)]. \quad (5.1)$$

The reconstruction of original image from its opening refers to Eq. 2.50 and 2.53. Similarly, closing by reconstruction means reconstructing the original image f from its closing $\phi_B(f)$, where closing is defined as the dilation of f with a SE B followed by the erosion with the transposed SE $\tilde{B}$:

$$\phi_B(f) = \varepsilon_{\tilde{B}}[\delta_B(f)], \quad (5.2)$$

where ε_B is the erosion transform with the structuring element B , and $\varepsilon_{\tilde{B}}$ is the dilation transform with the transposed structuring element $\tilde{B}$. The reconstruction of original image from its closing refers to Eq. 2.50 and 2.53. This noise suppression allows a more accurate calculation of the gradient image and reduction of the number of the detected false edges [108]. Because watershed transform is usually applied to the gradient magnitude of the original image [114], a better gradient image is helpful for a more accurate final segmentation and for alleviating over-segmentation.

After the image smoothing, the gradient image is computed. The employed morphological gradient, also called *Beucher gradient*, is defined as the arithmetic difference between the dilation δ and the erosion ε by the SE B [100]:

$$\rho_B = \delta_B - \varepsilon_B. \quad (5.3)$$

A watershed transform is applied to the gradient image, and it produces an initial image partition of a large number of disjointed regions. Such initial over-segmentation is due to the high sensitivity of the watershed algorithm to the gradient image, and the over-segmentation can be reduced by merging the regions belonging to almost homogeneous region.

Conventionally, image partition is represented by the data structure *region adjacency graph* (RAG) [115]. A RAG of K -partition is defined as an undirected graph, $G = (V, E)$, where $V = 1, 2, \dots, K$ is the set of nodes and $E \subset V \times V$ is the set of edges.

Each region is represented by a graph node, and between two regions (nodes) $i, j \in V$ there exists the edge (i, j) if the regions are adjacent. A cost is assigned to each graph edge expressing the dissimilarity between the two adjacency regions [108]. A fast nearest neighbor graph (NNG) was proposed instead of RAG to speed up the merging procedure [108].

For an initial image partition obtained from the watershed transform and a given criterion or dissimilarity function, there exists an ordered sequence of region merging that optimizes the criterion. Finding a globally optimal merging sequence is a very difficult task due to the prohibitively large search spaces. Therefore, a sub-optimal sequence is often used in practice: for each merging step, the pair of adjacent regions in the current image partition is sought that optimizes the criterion. Such a merging sequence cannot guarantee that the final result is optimal with respect to the criterion. However, given the usual quality of the initial image partition yielded by watershed transform, the occurrence of wrong merges is not significant [112].

Usually, the dissimilarity functions make use of the characteristics of region homogeneity and/or edge of the adjacent regions. They are summarized briefly in the following:

1) Mean-based Dissimilarity Function

Dissimilarity functions are often defined based on the difference of intensity mean between the two adjacent regions [108, 112]:

$$\delta(R_M^i, R_M^j) = \frac{\|R_M^i\| \cdot \|R_M^j\|}{\|R_M^i\| + \|R_M^j\|} [\mu(R_M^i) - \mu(R_M^j)]^2, \quad (5.4)$$

where $\mu(R_M^i)$ and $\mu(R_M^j)$ are the mean of two adjacent regions labelled i and j of a certain image partition M respectively, and $\|R_M^i\|$ and $\|R_M^j\|$ are numbers of pixels contained in the two regions.

2) Mean and Variance-based Dissimilarity Function

Ref. [109] calculates dissimilarity by measuring the Euclidean distance of the mean between two adjacent regions and the new variance after the two regions are merged. When either of the two components (mean and variance) is smaller than its corresponding threshold, the two regions are considered as similar and are merged.

3) Edge-based Dissimilarity Function

In the edge-based dissimilarity function, the information on the edge between the two adjacent regions is taken into consideration for merging criterion. An obvious choice is the average gradient magnitude of the edge $E_{i,j}$ of the adjacent regions R_M^i and R_M^j [112]

$$\overline{\|\nabla E_{i,j}\|} = \frac{1}{\|E_{i,j}\|} \cdot \sum_{(x,y) \in E_{i,j}} \|\nabla I(x,y)\|, \quad (5.5)$$

where $\|E_{i,j}\|$ is the number of pixels constituting the edge. Usually the modified edge criterion, i.e., edge with extra factor, that counts the influence of the length of edge is used [112]

$$\delta_{AGE}(R_M^i, R_M^j) = \overline{\|\nabla E_{i,j}\|} \cdot \frac{\min(\|E_i\|, \|E_j\|)}{\|E_{ij}\|} \quad (5.6)$$

4) Region Homogeneity and Edge-based Dissimilarity Function

Combined criteria have been proposed to integrate the information of region homogeneity and edge in a weighted fashion, such as [112]

$$\delta_C(R_M^i, R_M^j) = \alpha \cdot \delta_R(R_M^i, R_M^j) + \beta \cdot \delta_E(R_M^i, R_M^j), \quad (5.7)$$

where R_M^i, R_M^j represent two adjacent regions labelled i and j of a certain image partition M , δ_R represents the difference between the two regions such as difference of intensity mean in Eq. 5.4, δ_E is the edge information between the two regions such as the average gradient magnitude in Eq. 5.6, and δ_C is the combined criterion. The weight parameters α, β are predefined constants [110, 112], or adjustable

coefficients corresponding to the stage of region merging [96]. The combined dissimilarity functions take advantage of information from regions and edges, and produce more visually appropriate final segmentation results than using the information separately [96, 112].

The iterative hierarchical region merging can be terminated by allowing user to manually control the stopping point of the merging process [116]. Merging can also be stopped by some predefined criteria, such as the number of regions left in the final image partition [96, 112], the number of iterations [113], or a threshold for dissimilarity function [108].

The proposed region merging algorithm in the present study mainly focuses on the improvement of dissimilarity function and stopping criterion. Firstly, the dissimilarity function is defined on the basis of topographic characteristics of two adjacent regions. These characteristics are summarized based on human visual inspection, which are expected to provide more accurate information because the human eye is more effective in visual information processing. And the characteristics extracted from region and edge are combined and processed in the framework of pattern classification rather than in a weighted fashion set manually. Secondly, the stopping criterion is learned through classification, rather than a manually defined constant. The major issues in the proposed method are: characteristics extraction and classification, details of which are presented next.

5.3 Topographic Characteristics of the Two Adjacent Regions

The topographic surface of two adjacent regions in a watershed image are illustrated in Fig. 5.1, where (a) is two regions that should be merged while (b) is two regions

Figure 5.1: The topographic representation of the two adjacent regions. (a) two regions that should be merged, (b) two regions that should not be merged.

that should not be merged. The regional minima, watershed lines, and catchment basins are illustrated in Fig. 5.1. To better describe the characteristics of adjacent regions, we introduce three new concepts, namely, ‘common boundary’, ‘partial catchment basin’ and ‘hump’ as shown in Fig. 5.1. The common boundary is defined as the watershed line between the two adjacent regions being studied. The partial catchment basin is the set of pixels within the catchment basin that are lower than or equal to the common boundary. With reference to the definition of catchment basin [98], the hump is defined as follows:

Let I be a grayscale image, the hump $H(M_i, M_j)$ associated with minima M_i and M_j is the set of pixels p such that a water drop at the common boundary falling at p flows down along the relief, following a certain descending path, and eventually reaches M_i or M_j .

By carefully studying the topographic illustration in Fig. 5.1 (a)(b), we summarize the following difference between case ‘should be merged’ and ‘should not be merged’ for the two adjacent regions from a human visual inspection point of view.

1) *Edge*: the common boundary in the ‘merge’ case is low, i.e., its intensity is small, while the intensity of the common boundary in the ‘not merge’ case is comparatively large.

2) *Region, catchment basin*: the partial catchment basin is much smaller than the whole catchment basin in the ‘merge’ case, i.e., the volume (intensity sum) and the intensity mean of the partial catchment basin are smaller than those of whole catchment basin in the ‘merge’ case comparing with the ‘not merge’ case.

3) *Region, hump*: the volume and height of the hump are smaller in the ‘merge’ case than those in the ‘not merge’ case.

Based on above observation, 24 features are defined which are extracted from common boundary, small catchment basin (catchment basin with fewer pixels), large

catchment basin (catchment basin with more pixels), and hump respectively as shown in Table 5.1 to describe the difference of characteristics of the two adjacent regions. The features include intensity mean of the common boundary, length of the common boundary; volume, intensity mean, and intensity variance of the partial catchment basin, volume, mean, and variance of the whole catchment basin, the ratio of volume, mean, and variance of partial catchment basin to whole catchment basin; volume of hump, area of hump, small height of hump (the intensity difference between the common boundary and the shallower minimum), and large height (the intensity difference between the common boundary and the deeper minimum).

The features obtained that describe the characteristics of the two adjacent regions need to be classified into the category ‘merge’ or ‘not merge’. Data classification can be performed in two modes: supervised and unsupervised classification. The supervised classification is classifying the new data into classes where the class patterns are learned from the training samples with category label, while the unsupervised classification is naturally partitioning homogenous data into clusters without explicit supervision. The supervised-based and unsupervised-based region merging are discussed in the next two sections respectively.

5.4 Supervised Learning-based Region Merging

5.4.1 Feature Selection on the Characteristic Variables

In the context of supervised classification problems, a good choice of features is a key for building compact and accurate classifiers [117], where feature selection is the task of choosing a small set out of a given set of features that capture the significant properties of the data. The 24 characteristic variables extracted from the two adjacent regions are based on observation and may not necessarily all useful

in determining the relationship between two regions, therefore, feature selection techniques are employed here to find out critical variables to the decision.

As for the training dataset, it should cover the distribution of entire data to ensure a generalization performance for the unseen data. The distribution of characteristic variables describing the relationship between two adjacent regions are mainly determined by the properties of intensity within the regions such as intensity mean or variance, and the size of the region as well, therefore 200 representative pairs of adjacent regions with various intensities and sizes are selected as training samples, where 100 pairs should be merged and another 100 pairs should not be merged.

In the procedure of feature selection, Mahalanobis distance class separability measure and orthogonal sequential forward search procedure [32] are employed to evaluate the importance of the characteristic variables (refer to Eq. 4.3-4.11). Finally 15 features are selected, including 1 variable from common boundary (length), 5 variables from small catchment basin (volume of partial catchment basin, volume of whole catchment basin, ratio of volume, variance of partial catchment basin, ratio of variance), 5 variables from large catchment basin (volume of partial catchment basin, ratio of volume, ratio of mean, variance of partial catchment basin, ratio of variance), and all the 4 variables from hump.

5.4.2 Region Merging based on Supervised Classification

After feature selection, 15 variables selected are input to the pattern classifier to decide whether the two adjacent regions should be merged or not. The classifier is trained using the same training dataset as those used in the feature selection.

Several supervised classification schemes are investigated in the present study including two linear classifiers, i.e., linear Support Vector Machines (SVMs) [27] (refer to Eq. 2.29) and Linear Discriminant Analysis (LDA) [121] (Eq. 3.30), and two non-

linear classifiers, i.e., SVMs with a polynomial kernel [27] and a Gaussian kernel [27] (Eq. 2.30 and Eq. 2.31). Experimental results show that SVMs with a Gaussian kernel outperforms other classifiers, and is therefore chosen in the present study. The parameters of the Gaussian Kernel classifier are learned using the SVMs Toolbox [119]. The cost of constraint violation C is set to 1, and the width of the Gaussian kernel is set to 0.5.

The algorithm of supervised-based region merging is described in Table 5.2, which is programmed in Matlab 6.5. Firstly, a table of dissimilarity value D is built for each pair of adjacent regions in the watershed image, where the dissimilarity value is the decision value output from the classifier. The 15 features selected from the two adjacent regions are input to the SVMs classifier with a Gaussian kernel, and the output of the classifier is whether the two regions should be merged. The output pattern label '+1' means 'not merge', and label '-1' means 'merge'. The smaller the decision value, the two regions have higher potential to be merged. We search the least dissimilarity d_{least} in the table D , i.e., the highest merging potential for the two adjacent regions, merge these two regions and modify the table D accordingly if $d_{least} < 0$; stop the merging if $d_{least} \geq 0$.

5.5 Unsupervised Learning-based Region Merging

Obviously, supervised and unsupervised learning approaches have their own pros and cons respectively. Supervised machine learning is likely to provide more accurate classification result than unsupervised does [208], while the retaining is needed if the learning model is to be applied objects that are too different with the training data. In addition, supervised learning may have the problem of over-training and highly dependent on the quality of training data. Therefore we do have interest to

Table 5.2: The Algorithm of The Supervised Learning-based Region Merging

% Initialization:
Look-up table of dissimilarity value D (for each pair of adjacent regions in the watershed image)
% Processing:
while $Flag == 1$ % loop ends when the least dissimilarity value $d_{least} \geq 0$
Find the pair regions with the least dissimilarity value d_{least}
if $d_{least} < 0$ % the pair regions should be merged
i) The pair regions are set with a new label and the common boundary between them is filled with the same label %merge pair regions
ii) Remove d_{least} from table D , find the regions connected to the merged pair regions, and update their dissimilarity value %modify D
iii) $Flag=1$;
else
$Flag=0$;
end
end

try the region merging algorithm on the unsupervised learning approaches.

Basically, unsupervised classification approaches are based on clustering techniques, where the k-means clustering technique is an elementary but effective tool [121] (refer to Eq. 2.16-2.18). Therefore, the unsupervised-based region merging in the present study is based on k-means clustering technique. Without performing feature selection, all the 24 features from the two adjacent regions are employed and are clustered into two categories by k-means clustering, where $k = 2$, one category for ‘merge’ and another for ‘not merge’.

The algorithm for the unsupervised-based region merging is given in Table 5.3. The features F of all the pairs of adjacent regions in the watershed image are obtained to describe the relationship between the two adjacent regions, and are grouped into two clusters. Here the cluster with smaller features is class ‘merge’ and the cluster with larger features is considered as ‘not merge’, because the features associated with the ‘merge’ case are small comparing with the features in the ‘not merge’ case as discussed in Section 5.3. The distances of the features to the cluster centers ‘merge’ and ‘not merge’ are saved in D_M and D_{NM} respectively. Search the pair regions with the least distance d_{Mleast} in the class ‘merge’, merge the pair regions, and modify F , D_M and D_{NM} accordingly. Here the features corresponding to the pair regions that were merged already, i.e., the ‘processed’ features, are still in the features dataset F and are clustered in the iterative process of region merging, with the goal to avoid too few ‘merge’ features in the clustering procedure as the regions merge. The iterations are terminated when all the features in the class ‘merge’ are ‘processed’.

Table 5.3: The Algorithm of Unsupervised Learning-based Region Merging

<pre>% Initialization: Get features F (for each pair regions in the watershed image) Cluster F into 2 classes by k-means clustering, and get distances D_M, D_NM to cluster centers 'merge' and 'not merge' respectively % Processing: while Flag == 1 % loop ends when the pair regions in the class 'merge' are all processed i) Find the pair regions with the least distance d_Mleast in the class 'merge' ii) The pair regions are set with a new label and the common boundary between them is filled with the same label %merge regions iii) Find the regions connected to the merged regions, get the new features and add them in F, and mark the features corresponding to the merged pair regions with 'processed' %modify F iv) Cluster F into 2 classes, and update distance table %modify D_M, D_NM if the pair regions in the class 'merge' are all 'processed' Flag=0; else Flag=1; end end end</pre>

5.6 Experimental Results and Discussion

5.6.1 Performance Evaluation of Supervised Learning-based Region Merging

The motivation of proposing the supervised-based region merging for watershed image is to improve the accuracy of the final segmentation results. Therefore, the performance of the proposed method is evaluated by the final segmentation results using a number of cell images, and is compared with commonly used methods such as marker-controlled watershed by h-minima marker detection [99], waterfall [99], and region merging approaches with different dissimilarity functions as reviewed in Section 5.2. The segmentation results on a small part of one testing image by different methods are shown in Fig. 5.2 and 5.3, where (a) is the original grayscale image, (b) is segmentation result by using watershed algorithm directly, (c) is result by marker-controlled watershed with h-minima for marker detection, (d) is result by waterfall transform, (e) is result by region merging with dissimilarity function based on intensity mean, (f) is result by region merging with dissimilarity function based on intensity mean and variance, (g) is result by region merging with dissimilarity function based on edge, (h) is result by region merging dissimilarity function based on intensity mean and edge, and (i) is result by the proposed supervised-based region merging approach. The images show that watershed transform followed by the proposed supervised-based region merging outperforms other approaches.

The segmentation results of 8 exemplar testing images are listed in Table 5.4. Compare to the existing methods, the proposed supervised-based region merging algorithm significantly reduces the numbers of over-segmentation and under-segmentation. Furthermore, it outperforms other region merging-based algorithms. This is mainly because of the supervised learning.

Table 5.4: Comparison on accuracy of the proposed supervised learning-based approach with commonly used methods

image index	No. of isolated	No. of overlapped	h-minima		waterfall	mean		mean & variance		edge with extra factor		mean & edge with extra factor		proposed supervised-based		
			over-segment no.	under-segment no.		o-s	u-s	o-s	u-s	o-s	u-s	o-s	u-s	o-s	u-s	
1	235	179	1	2	20	3	7	8	4	7	5	7	2	7	1	1
2	381	235	4	7	21	1	6	6	3	6	8	3	4	4	2	0
3	486	191	9	5	27	1	5	1	4	4	7	5	1	5	0	0
4	216	264	47	3	50	4	40	5	26	8	45	12	37	7	2	1
5	152	142	0	12	23	3	7	7	6	9	7	6	5	11	2	1
6	328	189	27	15	32	1	19	1	14	7	31	3	13	3	3	2
7	448	258	0	7	36	1	10	3	5	3	5	4	11	5	0	2
8	450	194	0	10	17	1	10	1	7	5	20	2	2	4	1	1
average	321	208	11	8	28	2	13	4	9	6	16	5	9	6	1	1

Statistical evaluation of segmentation $Q(I)$ [124] (Eq. 2.58) is also employed to the 8 exemplar testing images, and the results are listed in Table 5.5. From Table 5.5 we can see that generally speaking all the region merging-based methods did better job than the h minima transform and waterfall transform. Among the region merging-based methods, the proposed supervised region merging algorithm performed best. Note that the testing images used here are as the same as used in Chapter 4. The comparison of over-segmentation and under-segmentation (Table 4.1, 4.3 and 5.5) shows that watershed followed by the proposed supervised-based region merging outperforms the marker-controlled watershed with the marker detection approach presented in Chapter 4. This can be compared with the evaluation $Q(I)$ as shown in Table 4.4 and 5.6.

Table 5.5: Statistical assessment $Q(I)$ of segmentation results on cell images (the proposed supervised learning-based region merging approach)

image index	h minima	waterfall	mean mean	mean& variance	edge with extra factor	mean&edge with extra factor	the proposed (supervised)
1	11.5	11.9	6.3	6	6.4	6.3	6.1
2	11	11.2	7.9	7.1	12.6	7.2	7.4
3	11.5	10.7	6.8	9.1	8.0	7.4	6.0
4	36.9	23.3	21.1	20.4	26.5	21.4	19.5
5	14.4	13.9	8.0	7.2	8.0	8.0	6.2
6	22.5	22.5	14.8	14.6	16.6	15.1	10.3
7	17.4	22.3	13.6	9.5	14.7	10.9	9.5
8	7.2	11.3	7.7	7.6	7.9	7.8	6.3
average	16.6	15.9	10.8	10.2	12.6	10.5	8.9

5.6.2 Performance Evaluation of Unsupervised Learning-based Region Merging

The proposed unsupervised learning-based region merging is also applied to the cell image segmentation. The segmentation result of part of a testing image is shown in Fig. 5.5, and the statistical assessment $Q(I)$ is listed in Table 5.6, where the testing images are as the same as those used in the evaluation of the proposed supervised learning-based region merging. Compare Table 5.6 with Table 5.5 we can see that unsupervised-based region merging provided worse results than the supervised-based region merging did, and even worse than the traditional methods. The possible reason is that other methods are trained by supervised classification or are set to some specific parameters through experiments such as the threshold h in the h -minima transform. The generability of the unsupervised region merging algorithm needs further study as the project goes to real application and data set includes more different cases.

5.7 Conclusion

In the present study, a new region merging approach has been proposed for watershed-based image segmentation. First, the features from the pair of adjacent regions are extracted that contain the information of region and edge. And then the features are put into the framework of pattern classification to decide whether the two regions should be merged or not. The iterative hierarchical region merging is terminated automatically through the learning of pattern classification. The supervised-based and unsupervised-based region merging are both discussed. Experimental results for the cell image segmentation show that the proposed supervised-based region merging algorithm reduced over-segmentation and under-segmentation effectively compared with reconstruction-based method and other commonly used region merging meth-

Table 5.6: Statistical assessment $Q(I)$ of segmentation results on cell images (the proposed unsupervised-based region merging approach)

image index	the proposed (unsupervised region merging)
1	12.3
2	11.5
3	11.7
4	36.2
5	9.0
6	23.2
7	18.6
8	14.7
average	17.2

ods.

Figure 5.2: Comparison of the segmentation results, (a) the original grayscale image, (b) segmentation result by using watershed algorithm directly, (c) result by marker-controlled watershed with h-minima for marker detection, (d) result by waterfall transform, (e) result by region merging with dissimilarity function based on intensity mean, (f) result by region merging with dissimilarity function based on intensity mean and variance, (g) result by region merging with dissimilarity function based on edge, (h) result by region merging dissimilarity function based on intensity mean and edge, and (i) result by the proposed supervised-based region merging approach (continued overleaf).

Figure 5.2 continued.

Figure 5.2 continued.

Figure 5.3: The segmentation result of the cell image by the proposed unsupervised-based region merging approach.

Chapter 6

Conclusions and Recommendations

6.1 Conclusions

Cell analysis is prominent and usually the only sure way among the medical diagnostic approaches. In the recent years, digital image processing techniques have been applied to automate cell analysis with the goal of providing accurate and efficient diagnostic results. This thesis presents an in-depth research on the automatic cell analysis, especially for the bladder inverted papilloma with staining in p53 immunohistochemistry in which the main work focuses on color cell image segmentation.

Pattern classification techniques are very powerful and promising, and have been applied to the image segmentation, named as pixel classification or pixel-based segmentation. In this thesis, we demonstrate that pixel classification-based color image segmentation in color space is equivalent to performing segmentation on grayscale image through thresholding. Based on this result, a supervised learning-based 2-step procedure for color cell image segmentation has been proposed. Here color image

is first mapped to grayscale via a transform learned through supervised learning, thresholding is then performed on the grayscale image to segment objects out of background. The transform used for image conversion is learned off-line from the global pixel distribution of color images while the threshold is learned on-line from local pixel distribution of the grayscale image. The combination of global and local learning provides high efficiency.

The separation of overlapped or clumped nuclei is another important issue in the cell image segmentation, which can not be solved by the procedure mentioned above (histogram thresholding). The watershed algorithm, which combines the advantages of region growing and edge detection, is a powerful tool for the image segmentation, especially for the case of objects separation. However, there is the problem of over-segmentation when applying watershed algorithm to the image directly. Basically, there are two ways to overcome the over-segmentation problem, including (1) marker-controlled watershed and (2) watershed followed by region merging. Based on the first concept, a new marker detection algorithm for watershed-based separation of overlapped nuclei has been proposed, which employs both photometric and shape information and combines the two naturally in the framework of pattern classification to provide more reliable markers. The problem of over-segmentation and under-segmentation have been alleviated largely by the proposed marker detection approach.

Based on the second concept, a new region merging method for watershed-based image segmentation has been proposed. The topography-based characteristics are extracted from the adjacent regions, including both region and edge information, and are processed in the framework of pattern classification, rather than the traditional weighted fashion with manually controlled parameters. The iterative hierarchical region merging is terminated automatically through classification. Both supervised-based and unsupervised-based region merging have been discussed.

In summary, three newly developed algorithms for cell image segmentation as mentioned above all made use of pattern classification techniques, and did some attempts in the field of image conversion, marker detection, and region merging, respectively.

This project is collaborated with Department of Pathology, Singapore General Hospital. Currently, it is in the stage of research and it is early for the clinical application. More efforts are needed to ensure the accuracy, low computational load and result reproducibility for the automatic papilloma diagnosis in the daily application.

6.2 Recommendations for Further Research

Image segmentation is fundamental and crucial for various image processing tasks, such as computer vision, object recognition, image compression, etc. [96]. Therefore, it is very useful to explore efficient and accurate segmentation algorithms. Some possible works are listed below:

- More pattern classification techniques can be tried in the field of image segmentation, not only in the procedure of pixel classification but also in the step of feature selection, feature classification, and so on. Besides the classification methods used in the thesis, some other powerful approaches can be studied. For example, for the case of supervised classification, neural networks (NN) have good performance for classifying the nonlinear separable data patterns. For the case of unsupervised classification, fuzzy clustering, knowledge-based clustering, etc. can be used instead of traditional clustering to improve the classification ability and consequently the segmentation results.
- Pixel classification incorporated with spatial information can be developed. Classification-based image segmentations usually categorize pixels into the corresponding categories based on the color features, while the spatial infor-

mation is used seldom. Therefore some morphology-based or statistical-based algorithms can be introduced into the pixel classification to enhance its segmentation ability.

- Model-based segmentations, esp. the deformable modelling-based segmentation can be developed. Deformable templates are very powerful segmentation tools which are widely used in the segmentation of Magnetic Resonance Images (MRI) and are also applied for cell image segmentation [158]. Deformable modelling is esp. suitable for the segmentation of the objects with non-rigid boundaries. Therefore, a deformable modelling-based algorithm can be developed to deal with the cell segmentation.
- Feature definition, extraction, and selection based on the priori knowledge can be done to describe the characteristics of the image and discriminate the difference among the different objects.
- In addition, image filtering techniques should be studied to remove the noise or artifacts in the images as the preprocessing for the image segmentation with the goal of achieving better segmentation results. Mathematical morphology-based filtering is very powerful. It is a nonlinear technique so it can remove the noise while keeping other characteristics such as edges unchanged. Therefore a morphological filtering should be developed.

In closing, research effort has been undertaken on developing more efficient and more reliable techniques to automate medical image analysis. These attempts not only alleviate the load of medical practitioners, but also pave the way for the research and development of new medical diagnosis methods and imaging techniques.

Author's Publications

Journal Papers

1. K.Z. Mao, Peng Zhao, Puay-Hoon Tan, Supervised learning-based cell image segmentation for p53 immunohistochemistry in bladder inverted papilloma, *IEEE Transactions on Biomedical Engineering*, vol. 53, pp 1153-1163, 2006.
2. Peng Zhao and K.Z. Mao, Region merging using characters of catchment basins and classification techniques for watershed-based image segmentation, in preparation.

Conference Papers

1. K.Z. Mao, Peng Zhao, Puay-Hoon Tan, Learning-based method for p53 immunohistochemically stained cell image segmentation, 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Shanghai, China, Sept. 1-4, 2005.
2. Peng Zhao, K.Z. Mao, Puay-Hoon Tan, Marker detection for watershed algorithm using both intensity and shape information: applied to cell image segmentation, The 2005 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences, Las Vegas, USA, Jun. 20-23, 2005.
3. Peng Zhao, K.Z. Mao, T.S. Koh, P.H. Tan, Automatic cell analysis for P53 immunohistochemistry in bladder inverted papilloma, IEEE EMBS Asian-Pacific Conference on Biomedical Engineering, Japan, Oct. 20-22, 2003.
4. K.Z. Mao, Peng Zhao, T.S. Koh, P.H. Tan, Overlapping/touching cell nuclei segmentation based on analysis of perpendicular distance curve, IEEE EMBS Asian-Pacific Conference on Biomedical Engineering, Japan, Oct. 20-22, 2003.

Bibliography

- [1] P. Ranefall, K. Wester, and E. Bengtsson, "Automatic quantification of immuno-histochemically stained cell nuclei using unsupervised image analysis," *Analytical cellular pathology*, vol. 16, pp. 29-43, 1998.
- [2] J. Angulo and G. Flandrin, "Microscopic image analysis using mathematical morphology: application to haematological cytology," *Science, Technology and Education of Microscopy: An overview*, Formatex Eds., Spain, pp. 304-312, 2003.
- [3] N. Vandenbroucke, L. Macaire, and J.G. Postaire, "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis," *Computer Vision and Image Understanding*, vol. 90, pp. 190-216, 2003.
- [4] P. Ranefall, L. Egevad, B. Nordin, and E. Bengtsson, "A new method for segmentation of color images applied to immunohistochemically stained cell nuclei," *Analytical Cellular Pathology*, vol. 15, pp. 145-156, 1997.
- [5] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Trans. Image Processing*, vol. 2, pp. 176-201, 1993.
- [6] T. Würflinger, J. Stockhausen, D.M. Ebrecht, and A. Böcking, "Robust automatic coregistration, segmentation, and classification of cell nuclei in multi-modal cytopathological microscopic images," *Computerized Medical Imaging and Graphics*, vol. 28, pp. 87-98, 2004.
- [7] H. Stern and B. Efros, "Adaptive color space switching for tracking under varying illumination," *Image and Vision Computing*, vol. 23, pp. 353-364, 2005.
- [8] H.-S. Wu, J. Barba, and J. Gil, "A Parametric fitting algorithm for segmentation of cell images," *IEEE Trans. Biomedical Engineering*, vol. 45, pp. 400-407, 1998.

- [9] T.Z. Jiang, F.G. Yang, and Y. Fan, "A parallel genetic algorithm for cell image segmentation," *Electronic Notes in Theoretical Computer Science*, vol. 46, 2001.
<http://www.elsevier.nl/locate/entcs/volume46.html>
- [10] L.P. Shen, X.Q. Song, M. Iguchi, and F.J. Yamamoto, "A method for recognizing particles in overlapped particle images," *Pattern Recognition Letters*, vol. 21, pp. 21-30, 2000.
- [11] V. Grau, A. U. J. Mewes, M. Alcañiz, R. Kikinis, and S. K. Warfield, "Improved watershed transform for medical image segmentation using prior information," *IEEE Trans. Medical Imaging*, vol. 23, pp. 447-458, 2004.
- [12] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, 1991.
- [13] C.D. Ruberto, A. Dempster, S. Khan, and B. Jarra, "Analysis of infected blood cell images using morphological operators," *Image and Vision Computing*, vol. 20, pp. 133-146, 2002.
- [14] C.Pan, C.X.Zheng, and H.J.Wang, "Robust color image segmentation based on mean shift and marker-controlled watershed algorithm," *Int. Conf. Machine Learning and Cybernetics*, Xian, China, pp. 2752-2756, 2003.
- [15] Z.G. Jiang, F.Y. Xie, and X.K. Zhou, "An image segmentation method for immune cell image analysis," *Int. Society for Optical Engineering*, vol. 4875, pp.610-613, 2002.
- [16] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, 1999.
- [17] F. Meyer, "Levelings, image simplification filters for segmentation," *J. of Mathematical Imaging and Vision*, vol. 20, pp. 59-72, 2004.
- [18] C. Wählby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson, "Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections," *J. of Microscopy*, vol. 215, pp. 67-76, 2004.
- [19] P.S. Umesh Adiga and B.B. Chaudhuri, "An efficient method based on watershed and rule-based merging for segmentation of 3-D histo-pathological images," *Pattern Recognition*, vol. 34, pp. 1449-1458, 2001.

- [20] J. Serra and P. Soille, "Watershed, hierarchical segmentation and waterfall algorithm," *Mathematical Morphology and Its Applications to Image Processing*, 69-76, 1994, Kluwer Academic Publishers.
- [21] H. Gao, W.C. Siu, and C.H. Hou, "Improved techniques for automatic image segmentation," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, pp. 1273-1280, 2001.
- [22] E.R. Dougherty, *Mathematical Morphology in Image Processing*, Marcel Dekker, 1993.
- [23] H. Breu, J. Gil, D. Kirkpatrick, and M. Werman, "Linear time euclidean distance transform algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 529-533, 1995.
- [24] B.K. Nayak and B.R. Das, "Mutation and methylation status of p53 gene promoter in human breast tumors," *Tumor Biology*, vol. 20, pp. 341-346, 1999.
- [25] C. Macaulay, H. Tezcan, and B. Palcic, "Adaptive color basis transformation. An aid in image segmentation," *Anal. Quant Cytol. Histol.*, vol. 11, pp. 53-58, 1989.
- [26] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1973.
- [27] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, 2000.
- [28] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Systems, Man, and Cybernetics.*, vol. 9, no. 1, pp. 62-66, 1979.
- [29] Handbook of Matlab (Release 13), Imaging Processing Toolbox.
- [30] J. Serrat, A. López, and D. Lloret, "On ridges and valleys," *15th Int. Conf. Pattern Recognition*, Spain, pp. 59-66, 2000.
- [31] R. Haralick, "Ridges and valleys on digital images," *Computer Vision, Graphics, and Image Processing*, vol. 22, pp. 28-38, 1983.
- [32] K.Z. Mao, "Orthogonal forward selection and backward elimination algorithms for feature subset selection," *IEEE Trans. Systems, Man and Cybernetics*, Part B, vol. 1, pp. 629-634, 2004.

- [33] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, Wiley-Interscience Publication, 2001.
- [34] OSU Support Vector Machines (SVMs) Toolbox version 3.00, Feb. 2002, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [35] P.A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*, Prentice/Hall Int., 1982.
- [36] A. Madabhushi and D.N. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," *IEEE Trans. Medical Imaging*, vol. 22, No. 2, 2003.
- [37] M. Borsotti, P. Campadelli, and R. Schettini, "Quantitative evaluation of color image segmentation results," *Pattern Recognition Letters*, vol. 19, pp. 741-747, 1998.
- [38] Y.J. Zhang, "A review of recent evaluation methods for image segmentation," *Int. Symposium on Signal Processing and its Application (ISSPA)*, Kuala Lumpur, Malaysia, 13-16 August, 2001.
- [39] URL: http://www.urologyinstitute.com/html/bladder_cancer.html
- [40] URL: http://www.wellnesspartners.com/utidir/link_directory/dynamic-frameset.html
- [41] URL: <http://www.cancerindex.org/clinks3d.html>
- [42] URL: <http://www.matriotech.com/bladdercancer.html>
- [43] URL: <http://www.cancer.gov/cancerinfo/wyntk/bladder>
- [44] URL: <http://www.drrajmd.com/bladdercancer/bladdercancer.html>
- [45] URL: http://teachmefinance.com/Scientific_Terms/inverted_papilloma.html
- [46] URL: <http://www.urologychannel.com/bladdercancer/index.shtml>
- [47] URL: <http://www.meb.uni-bonn.de/cancer.gov/CDR0000062908.html>
- [48] URL: <http://www.emedicine.com/MED/topic2344.htm>
- [49] URL: http://hcd2.bupa.co.uk/fact_sheets/html/bladder_tumors.html
- [50] URL: http://www.medicinenet.com/bladder_cancer/article.htm

- [51] URL: http://www.urologyinstitute.com/html/bladder_cancer.html
- [52] URL: <http://kidney.niddk.nih.gov/kudiseases/pubs/yoururinary/>
- [53] URL: <http://www.vghtpe.gov.tw/~cmj/5701/570110.htm>
- [54] URL: <http://www.bioscience.org/atlas/tumpath/kidurin/bladder/4/synopsis.htm>
- [55] S. Urakami, M. Igawa, H. Shirakawa, H.S.Ina, and T. Ishibe, "Biological characteristics of inverted papilloma of the urinary bladder," *British Journal of Urology*, vol. 77, pp. 55-60, 1996.
- [56] URL: <http://www.bostwicklaboratories.com/digital/bladder/bladder2.html>
- [57] URL: <http://www.towerurology.com/handler.cfm?event=practice,template&cpid=10091>
- [58] URL: http://pathology2.jhu.edu/bladder_cancer/diagnoss.cfm#pathology
- [59] L.M. McShane, R. Aamodt, C.C. Cardo, R. Cote, D. Faraggi, Y. Fradet, H.B. Grossman, A. Peng, S.E. Taube, and F.M. Waldman, "Reproducibility of p53 Immunohistochemistry in Bladder Tumors," *Clinical Cancer Research*, vol. 6, pp. 1854-1864, 2000.
- [60] A. Ponticiello, E. Barra, U. Giani, M. Bocchino, and A. Sanduzzi, "p53 immunohistochemistry can identify bronchial dysplastic lesions proceeding to lung cancer: a prospective study," *European Respiratory Journal*, vol. 15, pp. 547-552, 2000.
- [61] N. Malpica, A. Santos, A. Tejedor, A. Torres, M. Castilla, P. Garcí-Barreno, and M. Desco, "Automatic quantification of viability in epithelial cell cultures by texture analysis," *Journal of Microscopy*, vol. 209, pp. 34-40, 2003.
- [62] J. Boezeman, R. Raymakers, G. Vierwinden, and P. Linssen, "Automatic Analysis of Growth Onset, Growth Rate and Colony Size of Individual Bone Marrow Progenitors," *Cytometry*, vol. 28, pp. 305-310, 1997.
- [63] N. Malpica, C.O. de Solózano, J.J. Vaquero, A. Santos, I. Vallcorba, J.M.García-Sagredo, and F. del Pozo, "Applying Watershed Algorithms to the Segmentation of Clustered Nuclei," *Cytometry*, vol. 28, pp. 289-297, 1997.

- [64] W.Y. XU-VAN OPSTAL, C. RANGER, O. LEJEUNE, P. FORGEZ, H. BOUDIN, J.C. BISCONTE, and W. ROSTENE, "Automated Image Analyzing System for the Quantitative Study of Living Cells in Culture," *Microscopy Research and Technique*, vol. 28, pp. 440-447, 1994.
- [65] D. YOUNG, C. A. GLASBEY, A.J. GRAY, and N.J. MARTIN, "Towards automatic cell identification in DIC microscopy," *Microscopy Research and Technique*, vol. 192, pp. 186-193, 1998.
- [66] J. Lindblad, C. Wählby, E. Bengtsson, and A. Zaltsman, "Image Analysis for Automatic Segmentation of Cytoplasms and Classification of Rac1 Activation," *Cytometry*, vol. 57A, pp. 22-33, 2004.
- [67] Y. Chen, K. Biddell, A. Sun, P. Relue, and J. Johnson, "An automatic cell counting method for optical images," in The First Joint Conference of the Biomedical Engineering Society and the IEEE Engineering in Medicine and Biology Society, (Atlanta, Ga, USA), p. 819, 1999.
- [68] Q. Liao and Y. Deng, "An accurate segmentation method for white blood cell images," in 2002 IEEE International Symposium on Biomedical Imaging, (Washington, DC, USA), pp. 245248, 2002.
- [69] H. Cheng, X. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, pp. 22592281, 2001.
- [70] W. Skarbek and A. Koschan, "Color image segmentation - a survey," tech. rep., Technical University of Berlin, October 1994.
- [71] K. Fu and J. Mui, "A survey on image segmentation," *Pattern Recognition*, vol. 13, pp. 316, 1981.
- [72] B. Sankur and M. Sezgin, "Image thresholding techniques: a survey categories," tech. rep., Bogazici University Electric-Electronic Engineering Department, Eebek, Istanbul, Turkey.
- [73] L. Vincent and P. Soille, "Watershed in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 13, pp. 583598, 1991.
- [74] F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, pp. 2146, 1990.

- [75] J. Fan, D. Yau, A. Elmagarmid, and W. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. on Image Processing*, vol. 10, pp. 14541466, 2001.
- [76] N. Malpica, C. Solórzano, J. Vaquero, A. Santos, I. Vallcorba, J.M.G.-Sagredo, and F. Pozo, "Applying watershed algorithm to the segmentation of clustered nuclei," *Cytometry*, vol. 28, pp. 289297, 1997.
- [77] N. Sweeney and B. V. Sweeney, "Efficient segmentation of cellular images using gradient-based methods and simple morphological filters," in Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (Chicago, Illinois, USA), pp. 880882, 1997.
- [78] Y. Chen, K. Biddell, A. Sun, P. Relue, and J. Johnson, "An automatic cell counting method for optical images," in The First Joint Conference of the Biomedical Engineering Society and the IEEE Engineering in Medicine and Biology Society, (Atlanta, Ga, USA), p. 819, 1999.
- [79] P. Spyridonos, P. Ravazoula, D. Cavouras, and G. Nikiforidis, "An automatic classification system of urine bladder tumors employing morphological and textural nuclear features," in IEEE International Conference on Image Processing, (Thessaloniki, Greece), pp. 853856, 2001.
- [80] G. Ongun, U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac, and S. Beksac, "Feature extraction and classification of blood cells for an automated differential blood count system," in INNS-IEEE International Joint Conference on Neural Networks, (Washington, DC, USA), pp. 24612466, 2001.
- [81] Q. Liao and Y. Deng, "An accurate segmentation method for white blood cell images," in 2002 IEEE International Symposium on Biomedical Imaging, (Washington, DC, USA), pp. 245248, 2002.
- [82] F. Yang and T. Jiang, "Cell image segmentation with kernel-based dynamic clustering and an ellipsoidal cell shape model," *Journal of Biomedical Informatics*, vol. 34, pp. 6773, 2001.
- [83] R. Pohle and K. Toennies, "Segmentation of medical image using adaptive region growing," tech. rep., Otto-von-Guericke University.
- [84] G. Fernández, M. Kunt, and J.-P. Zrýd, "A new plant cell image segmentation algorithm," algorithm, in 8th International Conference on image analysis and processing, (San Remo, Italy), pp. 229234, 1995.

- [85] W. Wang, "Binary image segmentation of aggregates based on polygonal approximation and classification of concavities," *Pattern Recognition*, vol. 31, pp. 15031524, 1998.
- [86] N. Sweeney and B. V. Sweeney, "Efficient segmentation of cellular images using gradient-based methods and simple morphological filters," algorithm, in Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (Chicago, Illinois, USA), pp. 880882, 1997.
- [87] K. Rao and A. Dempster, "Modification on distance transform to avoid over-segmentation and under-segmentation," in Proceedings of the 4th EURASIP-IEEE Region 8 International Symposium on Video/Image Processing and Multimedia Communications, (Zadar, Croatia), pp. 295 299, 2002.
- [88] URL:<http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/backgroundsubtraction>.
- [89] N. Theera-Umpon and P. Gader, "Training neural networks to count white blood cells via a minimum counting error objective function," in Proceedings of the 15th International Conference on Pattern Recognition, (Barcelona, Spain), pp. 299302, 2000.
- [90] M. Beksac, M. Beksac, V. Tipi, H. Duru, M. Karakas, and A. Cakar, "An artificial intelligent diagnostic system on differential recognition of hematopoietic-cells from microscopic images," *Cytometry*, vol. 30, pp. 145150, 1997.
- [91] D. Bourgin, "David Bourins color spaces FAQ," URL: <http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/backgroundsubtraction>.
- [92] T. Carron and P. Lambert, "Color edge detector using jointly hue, saturation and intensity," in 1994 International Conference on Image Processing, (Austin, Texas, USA), pp. 977981, 1994.
- [93] C. Zhang and P. Wang, "A new method of color image segmentation based on intensity and hue clustering," in 15th International Conference on Pattern Recognition, (Barcelona, Spain), pp. 613616, 2000.
- [94] J. Song, "Hue-grayscale collaborating edge detection and edge color distribution space," URL:<http://www.cse.cuhk.edu.hk/lyu/staff/SongJQ/Edge.ppt>.

- [95] L. M. N. Vandenbroucke and L.-G. Postaire, "Color pixels classification in an hybrid color space," in 1998 International Conference on Image Processing, (Chicago, Illinois, USA), pp. 176180, 1998.
- [96] S. E. Hernandez, K.E. Barner, and Y. Yuan, "Region merging using homogeneity and edge integrity for watershed-based image segmentation," *Optical Engineering*, vol. 44, 017004, 2005.
- [97] J.B.T.M. Roerdink and A. Meijster, "The watershed transform: definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, pp. 187-228, 2001.
- [98] L. Vincent and P. Soille, "Watershed in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 583-598, 1991.
- [99] S. Beucher, "Watershed, hierarchical segmentation and waterfall algorithm," in *Proc. Int Symposium on Mathematical Morphology (ISMM)*, pp. 69-76, Kluwer, 1994.
- [100] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlas, 1999.
- [101] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Trans. Image Processing*, vol. 2, pp. 176-201, 1993.
- [102] K. Stapor, A. Switonski, R. Chrastek, and G. Michelson, "Segmentation of fundus eye images using methods of mathematical morphology for glaucoma diagnosis," in *Int. Conf. Computational Science (ICCS)*, 6-9 June, 2004, Krakow, Poland.
- [103] G. Bueno, O. Musse, F. Heitz, and J.P. Armspach, "Three-dimensional segmentation of anatomical structures in MR images on large data bases," *Magnetic Resonance Imaging*, vol. 19, pp. 73-88, 2001.
- [104] V. Grau, A.U.J. Mewes, M. Alcañiz, R. Kikinis, and S.K. Warfield, "Improved watershed transform for medical image segmentation using prior information," *IEEE Trans. Medical Imaging*, vol. 23, No. 4, pp. 447-458, 2004.
- [105] C. Wählby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson, "Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections," *J. of Microscopy*, vol. 215, pp. 67-76, 2004.

- [106] Z.G. Jiang, F.Y. Xie, and X.K. Zhou, "An image segmentation method for immune cell image analysis," *Int. Society for Optical Engineering*, vol. 4875, pp. 610-613, 2002.
- [107] O. Lezoray and H. Cardot, "Bayesian marker extraction for color watershed in segmenting microscopic images," in *Proc. 16th Int. Conf. Pattern Recognition*, Quebec, Canada, 2002, vol. 1, pp. 739-42, 2002.
- [108] K. Haris, S.N. Efstratiadis, N. Maglaveras, and A.K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Trans. Image Processing*, vol. 7, pp. 1684-1699, 1998.
- [109] C.P. Chau and W.C. Siu, "Generalized hough transform using regions with homogeneous color," *Int. J. of Computer Vision*, vol. 59, no. 2, pp. 183-199, 2004.
- [110] E. Navon, O. Miller, and A. Averbuch, "Color image segmentation based on adaptive local thresholds," *Image and Vision Computing*, vol. 23, pp. 69-85, 2005.
- [111] J.B. Kim and H.J. Kim, "Multiresolution-based watersheds for efficient image segmentation," *Pattern Recognition Letters*, vol. 24, pp. 473-488, 2003.
- [112] R.L.V.P.M. Pires, P.D. Smet, and W. Philips, "Watershed segmentation and region merging," in *Proc. of the SPIE - The Int. Society for Optical Engineering*, vol. 5308, no. 1, p 1127-35, 2004.
- [113] M. Nieniewski, "Extraction of diffues objects from images by means of watershed and region merging: example of solar images," *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 34, pp. 796-801, 2004.
- [114] J.M. Gauch, "Image segmentation and analysis via multiscale gradient watershed hierarchies," *IEEE Trans. Image Processing*, vol. 9, pp. 69-79, 2000.
- [115] X.L. Wu, "Adaptive split-and-merge segmentation based on piecewise least-square approximation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, pp. 808-815, 1993.
- [116] S.C. Orphanoudakis, G. Tziritas, and K. Haris, "A Hybrid Algorithm for the Segmentation of 2D and 3D Medical Images," *Proc. Information Processing in Medical Imaging (IPMI)*, pp. 385-386, Brest, France, June 25-30, 1995.

- [117] R.G. Bachrach, A. Navot, and N. Tishby, "Margin Based Feature Selection - Theory and Algorithms," *Proc. 21st Int. Conf. Machine Learning*, Banff, Canada, 2004.
- [118] I. Kononenko, "Estimating Attributes: Analysis and Extensions of RELIEF," *Proc. 7th European Conf. Machine Learning*, Catania, Italy, 1994.
- [119] OSU Support Vector Machines (SVMs) Toolbox version 3.00, Feb. 2002, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [120] C.W. Hsu, C.C. Chang, and C.J. Lin, "A practical guide to support vector classification," <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [121] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, Wiley-Interscience Publication, 2001.
- [122] A. Madabhushi and D.N. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," *IEEE Trans. Medical Imaging*, vol. 22, No. 2, 2003.
- [123] Y.J. Zhang, "A review of recent evaluation methods for image segmentation," *Int. Symposium on Signal Processing and its Application (ISSPA)*, Kuala Lumpur, Malaysia, 13-16 August, 2001.
- [124] M. Borsotti, P. Campadelli, and R. Schettini, "Quantitative evaluation of color image segmentation results," *Pattern Recognition Letters*, vol. 19, pp. 741-747, 1998.
- [125] L. M. N. Vandenbroucke and L.-G. Postaire, "Color pixels classification in an hybrid color space," in *1998 International Conference on Image Processing*, (Chicago, Illinois, USA), pp. 176180, 1998.
- [126] J. Gauch and S. Pizer, "The intensity axis of symmetry and its application to image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 753770, 1993.
- [127] J. Zou, H.-H. Chang, and H. Yan, "Shape skeletonization by identifying discrete local symmetries," *Pattern Recognition*, vol. 34, pp. 18951905, 2001.
- [128] S. He, R. Dai, B. Lu, C. Cao, H. Bai, and B. Jing, "Medial axis reformation: a new visualization method for ct angiography," *Academic Radiology*, vol. 8, pp. 726733, 2001.

- [129] B. Morse, S. Pizer, and A. Liu, "Multiscale medial analysis of medical images," in *in Proceedings of the 13th International Conference of Information Processing in Medical Imaging*, (Berlin, Germany), pp. 112131, 1993.
- [130] G. Rubin, D. Paik, P. Johnson, and S. Napel, "Measurement of the aorta and its branches with helical ct," *Radiology*, vol. 206, pp. 823829, 1998.
- [131] M. Brady and H. Asada, "Smoothed local symmetries and their implementation," *International Journal on Robotics Research*, vol. 3, pp. 3661, 1984.
- [132] H. Rom and G. Medioni, "Hierarchical decomposition and axial shape description," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 973981, 1993.
- [133] C. Lantuéjoul and F. Maisonneuve, "Geodesic methods in quantitative image analysis," *Pattern Recognition*, vol. 17, pp. 177187, 1984.
- [134] URL:http://www.hsc.stonybrook.edu/som/urology/urology_cp_bladdercancer.cfm
- [135] URL:<http://www.craighospital.org/SCI/METS/bladderCancer.asp>
- [136] URL:<http://www.medterms.com/script/main/art.asparticlekey=2466>
- [137] L.G. Hefter and I.S. Young, "Inverted papilloma of bladder," *Urology*, vol. V, No. 5, pp. 688690, 1975.
- [138] URL:<http://http://www.webpathology.com/case.cfm?case=54>
- [139] C.W. Cheng, L.W. Chan, C.K. Chan, C.F. Ng, H.Y. Cheung, S.Y.D. Chan, W.S. Wong, and K.F. To, "Is surveillance necessary for inverted papilloma in the urinary bladder and urethra?," *ANZ Journal of Surgery*, vol. 75, pp. 213217, 2005.
- [140] T.Z. Jiang and F.G. Yang, "An evolutionary tabu search for cell image segmentation," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 32, pp. 675678, 2002.
- [141] K. Jiang, Q.M. Liao, and Y. Xiong, "A novel white blood cell segmentation scheme based on feature space clustering," *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 10, pp. 1219, 2006.
- [142] C. Wahlby, "Algorithms for applied digital image cytometry," *PhD thesis*, Uppsala University, Sweden, 2003.

- [143] J. Angulo and G. Flandrin, "Microscopic image analysis using mathematical morphology: application to haematological cytology," *Science, technology and education of microscopy: an overview*, pp. 304312.
- [144] F. Ortiz, F. Torres, E.D. Juan, and N. Cuenca, "Color mathematical morphology for neural image analysis," *Real-time imaging*, pp. 455465, 2002.
- [145] T.W. Nattkemper, "Automatic segmentation of digital micrographs: a survey," *MEDINFO 2004*, 2004.
- [146] L. Lucchese and S.K. Mitra, "Color image segmentation: a state-of-the-art survey," URL: http://web.engr.oregonstate.edu/~luca/proc_insa_01.pdf.
- [147] J. Vergés-Llahí, "State-of-the-art survey on color segmentation methods (Chapter 3 of thesis)," URL: http://www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0906105-154139/04Jvl04de11.pdf
- [148] M. Celenk, "A color clustering technique for image segmentation," *Graphical Models Image Process.*, vol. 52, pp. 145170, 1990.
- [149] S.D. Zenzo, "A note on the gradient of a multi-image," *Comput. Vision Graphics Image Process.*, vol. 33, pp. 116125, 1986.
- [150] A. Cumani, "A note on the gradient of a multi-image edge detection in multispectral images," *Comput. Vision Graphics Image Process.*, vol. 53, pp. 4051, 1991.
- [151] K.N. Plataniotis, A.N. Venetsanopoulos, *Color Image Processing and Applications*, Springer, 2000.
- [152] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, pp. 146165, 2004.
- [153] H.X. Xu, G.X. Zhua, J.W. Tian, X. Zhang, and F.Y. Peng, "Image segmentation using support vector machine," *MIPPR 2005: Image Analysis Techniques*, vol. 6044, 60441, 2005.
- [154] C. Kotropoulos and I. Pitas, "Segmentation of ultrasonic images using support vector machines," *Pattern Recognition Letters*, vol. 24, pp. 715727, 2003.

- [155] J.C. Fu, S.K. Lee, S.T.C. Wong, J.Y. Yeh, A.H. Wang, and H.K. Wu, "Image segmentation feature selection and pattern classification for mammographic microcalcifications," *Computerized Medical Imaging and Graphics*, vol. 29, pp. 419-429, 2005.
- [156] R. Nevatia, "A color edge detector and its use in scene segmentation," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-7, no. 11, pp. 820-826, 1977.
- [157] Y.J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition*, vol. 29, no. 8, pp. 1335-1346, 1996.
- [158] A. Garrido and N. Pérez de la Blanca, "Applying deformable template for cell image segmentation," *Pattern Recognition*, vol. 33, pp. 821-832, 2000.
- [159] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Tran. on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [160] P.J. Sullivan, R.P. Campagnoli, and K.S. Latimer, "Evaluation of Proliferation Marker (PCNA, Ki-67, and p53) Expression in Avian Hepatic Tissues," URL: <http://www.vet.uga.edu/IVCVM/1998/sullivan/sullivan.htm>
- [161] M.K. Ng, "A note on constrained k-means algorithms," *Pattern Recognition*, vol. 33, pp. 515-519, 2000.
- [162] E. Kunze, A. Schauer, and M. Schmitt, "Histology and histogenesis of two Different types of inverted urothelial papillomas," *Cancer*, vol. 51, pp. 348-358, 2006.
- [163] I. Potts and E. Hirst, "Inverted papilloma of the bladder," *Journal of Urology*, vol. 90, pp. 175-179, 1963.
- [164] S.K. Kopparapu, P. Mudalige, and P. Corke, "A multiresolution based image segmentation," *IEE Image Processing and Its Applications*, pp. 567-571, 1999.
- [165] J.B. Kim and H.J. Kim, "Multiresolution-based watersheds for efficient image segmentation," *Pattern Recognition Letters*, vol. 24, pp. 473-488, 2003.
- [166] F.G. Yang and T.Z. Jiang, "Pixon-based image segmentation with Markov random fields," *ACCV2002: The 5th Asian Conference on Computer Vision*, 2002, Melbourne, Australia.

- [167] Markov Random Field Theory, URL: <http://www.fmrib.ox.ac.uk/analysis/techrep/tr00yz1/tr00yz1/node4.html>
- [168] Markov Random Field Segmentation, URL: <http://splweb.bwh.harvard.edu:8000/pages/papers/elena/node3.html>
- [169] J.X. Sun and D.B. Gu, "Bayesian image segmentation based on a inhomogeneous hidden markov random field," *Proceedings of the 17th International Conference on Pattern Recognition*, 2004, Melbourne, Australia.
- [170] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Tran. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.
- [171] I. Papila and O. Ersoy, "Multiscale segmentation of remotely sensed images using pairwise Markov chains," *IEEE*, pp. 2123-2126, 2004.
- [172] C. Kervrann and F. Heitz, "Statistical deformable model-based segmentation of image motion," *IEEE Tran. on image processing*, pp. 583-?, 1999.
- [173] V. M-Yedid, S. Tilie and J-C O-Marin, "Color image segmetation based on Markov random field clustering for histological image analysis," *IEEE*, pp. 796-798, 2002.
- [174] C.A. Bouman and M. Shapiro, "A multiscale random field model for bayesian image segmentation," *IEEE Tran. on image processing*, vol. 3, No. 2, pp. 162-177, 1994.
- [175] J.Q. Liu and Y-H Yang, "Multiresolution color image segmentation," *IEEE Tran. on image processing*, vol. 16, No. 7, pp. 689-700, 1994.
- [176] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321-332, 1988.
- [177] B. Appleton, "Globally optimal geodesic active contours," *Journal of Mathematical Imaging and Vision*, vol. 23, pp. 67-86, 2005.
- [178] R. Malladi, J.A. Sethian, and B.C. Vemuri, "Shape modeling with front propagation: a level set approach," *IEEE Tran. on Pattern Analysis and Machine Intelligence*, vol. 17, No. 2, pp. 158-175, 1995.

- [179] S. Osher and J.A. Sethian, "Fronts propagating with curvature-dependent speed: algorithm based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12-49, 1988.
- [180] D. Adalsteinsson and J.A. Sethian, "A fast level set method for propagating interfaces," *Journal of Computational Physics*, vol. 118, pp. 269-277, 1995.
- [181] P.O. Persson, "The level set method," Lecture notes, MIT 16.920J/2.097J/6.339J Numerical methods for partial differential equations, URL: <http://www-math.mit.edu/18086/levelsetnotes.pdf>
- [182] J.A. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, pp. 1591-1595, 1996.
- [183] R. Malladi and J.A. Sethian, "Level set and fast marching methods in image processing and computer vision," *International Conference on Image Processing 1996*, vol. 1, pp. 489-492, 1996, Switzerland.
- [184] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, "Fast geodesic active contours," *IEEE Tran. Image Processing*, vol. 10, No. 10, pp. 1467-1475, 2001.
- [185] L.D. Cohen and R. Kimmel, "Global minimum for active contour models: a minimum path approach," *International Journal of Computer Vision*, vol. 24, No. 1, pp. 57-78, 1997.
- [186] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours," *Signal Processing*, vol. 71, pp. 203-213, 1998.
- [187] E. Sifakis, C. Garcia, and G. Tziritas, "Bayesian level sets for image segmentation," *Journal of Visual Communication and Image Representation*, vol. 13, pp. 44-64, 2002.
- [188] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, No. 1, pp. 61-79, 1997.
- [189] D. Adalsteinsson and J.A. Sethian, "A fast level set method for propagating interfaces," *Journal of Computational Physics*, vol. 118, pp. 269-277, 1995.
- [190] J. Xin and Z. Yan, "The application of watershed algorithm in the initial contour selection of mumford-shah model," *2005 Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 1773-1776, 2005, China.

- [191] S.R. Gunn, M.S. Nixon, "A robust snake implementation: a dual active contour," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, No. 1, pp. 63-68, 1997.
- [192] B. Appleton, "Optimal geodesic active contours: application to heart segmentation," *APRS Workshop on Digital Image Computing*, pp. 27-32, 2003, Australia
- [193] B. Appleton, "Globally minimal contours and surfaces for image segmentation," PhD thesis, 2004, University of Queensland, Australia
- [194] J. A. Sethian and A. Vladimirsky, "Ordered upwind methods for static Hamilton-Jacobi equations," *Proc. Nat. Acad. Sciences*, vol. 98, No. 20, pp. 11069-11074, 2002.
- [195] H.X. Xu, G.X. Zhu, J.W. Tian, X. Zhang, F.Y. Peng, "Image segmentation using support vector machine," *Proc. SPIE*, vol.6044, pp.536-544, 2005, China.
- [196] M. Bonneville, J. Meunier, Y. Bengio, J.P. Soucy, "Support vector machines for improving the classification of brain pet images," *Proc. SPIE Medical Imaging*, vol.3338, pp.264-273, 1998.
- [197] S. Wang W.Y. Zhua, and Z.P. Liang, "Shape Deformation: SVM Regression and Application to Medical Image Segmentation," *8th IEEE International Conference on Computer Vision*, vol.2, pp.209-216, 2001, Canada.
- [198] Y.Q. Zhan and D.G. Shen, "Deformable Segmentation of 3-D Ultrasound Prostate Images Using Statistical Texture Matching Method," *IEEE Trans. Medical Imaging*, vol.25, No. 3, pp.256-272, 2006, Canada.
- [199] H.F. Ng, "Automatic thresholding for defect detection," *Proceedings of the Third International Conference on Image and Graphic*, pp.532-535, 2004, China.
- [200] X.D. Yang and V. Gupta, "An improved threshold selection method for image segmentation," *Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp. 531-534, 1993, Canada.
- [201] A.M. Martinez and A.C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, No. 2, pp. 228-233, 2001.
- [202] K. Delac, M. Grgic, and S. Grgic, "Statistics in face recognition: analyzing probability distributions of PCA, ICA, and LDA performance results," *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, pp. 289-294, 2005, Croatia.

- [203] C.Y. Xie, M. Savvides, and B.V.K.V. Kumar, "Comparison of feature space methods for face recognition," *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, 2006, USA.
- [204] "Filter control of specimen contrast in black & white photomicrography," URL: <http://www.olympusmicro.com/primer/java/photomicrography/contrastfilters/index.html>
- [205] M. Møller, O.V. Glistrup, and W. Olsen, "Contrast enhancement of the brownish horseadish peroxidase-activated 3,3'-diaminobenzidine tetrahydrochloride reaction product in black and white photomicrography by the use of interference filters," *The Journal of Histochemistry and Cytochemistry*, Vol. 32, No. 1, pp. 37-42, 1984.
- [206] M.D. Farrell and R.M. Mersereau, "On the impact of PCA dimension reduction for hyperspectral detection of difficult targets," *IEEE Geoscience and Remote Sensing Letters*, Vol. 2, No. 2, pp. 192-195, 2005.
- [207] D.J. Lee, "Color Space Conversion for Linear Color Grading," *Proceedings of SPIE 2000*, Vol. 4197, pp. 358-366, 2000.
- [208] P. Chaovalit and L. Zhou, "Movie Review Mining: a Comparison between Supervised and Unsupervised Classification Approaches," *Proceedings of the 38th International Conference on System Sciences*, pp. 112c, 2005, USA.
- [209] J.A. Sethian and A. Vladimirovsky, "Ordered upwind methods for static Hamilton-Jacobi equations," *Proceedings of the National Academy of Sciences*, vol. 98, no. 20, pp. 11069-11074, 2001, USA.
- [210] J. Cornelis, E. Nyssen, A. Katartzis, L. Kempen, P. Boekaerts, R. Deklerck, and A. Salomie, "Statistical models for multidisciplinary applications of image segmentation and labelling," *Proceedings of ICSP 2000*, vol. 3, pp. 2103-2110, 2000.
- [211] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society: Series B*, vol. 36, pp. 192-326, 1974.
- [212] N.R. Pal and S.K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, 1993.