THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Methods for Rapid and High Quality Acquisition of Whole Slide Images

Yilun Fan

B.Eng Honours

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2016

School of Information Technology and Electrical Engineering

ABSTRACT

In the past two decades, whole slide imaging (WSI) has been increasingly used in clinical pathology, research and education thanks to advances in technology. However, existing digital slide scanners for WSI are incapable of efficiently scanning cytology specimens in "glass-faithful quality". In particular, cervical cytology specimens used for cervical cancer screening need to be imaged at multiple focal planes and at high spatial resolution. So, the aim of this thesis is to develop novel methods for improving the scan efficiency of cytology specimens. Preliminary investigations narrowed the research into two directions: to optimise the scan parameters (e.g. the scan map, the focus profile map and the number of focal planes to be scanned) for scanning cytology specimens, and to investigate the fast slanted (specimen) scan.

Creating an accurate scan map followed by estimating the focus profile map of the specimen can significantly reduce the scan time, but previous methods are not robust to slide artefacts such as ink marker and dusts. The first research aspect of this thesis has focused on the development of a robust method that not only creates an accurate scan map but also provides a ranked list of focus candidates required for estimating the focus map. Specifically, tiles representing field-of-views to be scanned at high resolution are located on the low resolution images and evaluated with regards to whether they contain foreground objects and how good they are as focus candidates. Four metrics were proposed for the tile evaluation: threshold index (TI) that measures the image intensity; normalised auto-correlation index (NACI) that measures the spatial image similarity; auto-phase correlation index (APCI) that measures the image phase diversity; and entropy index (EI) that measures the predictability of image intensities. The experimental results indicated that while NACI, APCI and EI are effective for specimen delineation, only APCI is capable of selecting superior focus candidates and ignoring artefacts.

Knowing the three-dimensional distribution of the specimen can help to determine the minimal number of focal planes to be acquired, and hence the minimal scan time. A method has been developed to evaluate the thickness of all cell and cell clumps in cervical cytology specimens incorporating an exhaustive high-resolution scan, an over-complete wavelet transform across multi-focal planes and a cell clump segmentation method. The accuracy and usefulness of the method was evaluated and demonstrated by quantitative analysis of ten Pap stained Thin-prep cervical slides. In particular, it was found that creating a focus profile map was able to reduce the number of scanning focal planes from 25.4 to 21.4 on average. It was also shown that the distribution of cells was skewed towards the cover-slip (top of the slide), so by considering the thickness of the specimen, an improved focal map

can be produced to further reduce the required number of 1 micron spaced focal planes to 18.6.

The last research aspect of this thesis investigated the slanted scan method, which is a promising method specifically designed for rapid scan of thick (cytology) specimens. With the method, the slide is slanted at a small angle so that cross sections of the specimen can be imaged while the specimen is moved continuously. Preliminary study concluded that the slanted scan is capable of achieving higher throughput rates than that of the conventional methods with the constraints on camera frame rates, stage accuracy and illumination intensity. However, the method suffers from optical aberrations induced by the slide slant, a two-stage deconvolution method was therefore developed to correct these aberrations. Specifically, phase deconvolution is initially applied to diminish effects of the dominating coma aberration, followed by a conventional deconvolution method to improve lost image contrast. A calibration method has also been proposed to estimate the coma aberration and the point-spread-function (PSF) of the optics utilising actual cytology specimens. The efficacy of proposed algorithms was quantitatively assessed on simulated data and qualitatively on cervical cytology specimens. Results demonstrated both an improved convergence speed of the two-stage approach, especially when correcting the bend in the PSF, and a resultant image quality that was comparable to a conventionally (flat) scanned specimen.

In summary, the combination of methods proposed in this thesis have the potential to increase both overall scan efficiency and scan quality of whole slide scanners when acquiring “glass-faithful quality” images of cervical cytology specimens.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the *Copyright Act 1968* unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publication during candidature

Peer-reviewed journal articles

Fan, Y., Gal, Y., and Bradley, A. P. (2014a). An algorithm for microscopic specimen delineation and focus candidate selection. *Micron*, 66:51–62. http://www.sciencedirect.com/science/article/pii/S0968432814001127 (Fan et al. 2014a)

Fan, Y. and Bradley, A. P. (2016a). A method for quantitative analysis of clump thickness in cervical cytology slides. *Micron*, 80:73–82. http://www.sciencedirect.com/science/article/pii/S0968432815300391 (Fan and Bradley 2016a)

Fan, Y. and Bradley, A. P. (2016b). A two-stage method to correct the optical aberration induced by slide slant. *Micron*, 87:18–32. http://www.sciencedirect.com/science/article/pii/S0968432816300701 (Fan and Bradley 2016b)

Peer-reviewed conference proceedings

Fan, Y., Gal, Y., and Bradley, A. P. (2013). Performance analysis of three microscope slide scanning techniques. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6691511 (Fan et al. 2013)

Fan, Y., Gal, Y., and Bradley, A. P. (2014b). Microscopic specimen delineation using the auto phase correlation index. In *International Symposium on Biomedical Imaging (ISBI)*. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6868124 (Fan et al. 2014b)

Publication included in this thesis

The following publications were included in this thesis, of which significant portions were used in parts of different chapters (Chapter 2, 3, 4 and 5).

- **Chapter 2** - (Fan et al. 2013)
- **Chapter 3** - (Fan et al. 2014a;b)
- **Chapter 4** - (Fan and Bradley 2016a)
- **Chapter 5** - (Fan and Bradley 2016b)

For all these publications, individual contributions of all the authors are listed here.

Fan, Y., Gal, Y., and Bradley, A. P. (2013). Performance analysis of three microscope slide scanning techniques. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6691511 (Fan et al. 2013) **- incorporated in Chapter 2.**

Contributor	**Statement of contribution**
Yilun Fan	Experiment design (40%), data collection (100%), code implementation (100%), execution of experiments (100%), analysis and interpretation (40%), manuscript preparation (100%)
Yaniv Gal	Experiment design (30%), analysis and interpretation (30%), manuscript review (50%)
Andrew P.Bradley	Experiment design (30%), analysis and interpretation (30%), manuscript review (50%)

Fan, Y., Gal, Y., and Bradley, A. P. (2014b). Microscopic specimen delineation using the auto phase correlation index. In *International Symposium on Biomedical Imaging (ISBI)*. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6868124 (Fan et al. 2014b) **- incorporated in Chapter 3.**

Contributor	**Statement of contribution**
Yilun Fan	Experiment design (40%), data collection (100%), code implementation (100%), execution of experiments (100%), analysis and interpretation (40%), manuscript preparation (100%)
Yaniv Gal	Experiment design (30%), analysis and interpretation (30%), manuscript review (50%)
Andrew P.Bradley	Experiment design (30%), analysis and interpretation (30%),

	manuscript review (50%)

Fan, Y., Gal, Y., and Bradley, A. P. (2014a). An algorithm for microscopic specimen delineation and focus candidate selection. *Micron*, 66:51–62. http://www.sciencedirect.com/science/article/pii/S0968432814001127 (Fan et al. 2014a) **- incorporated in Chapter 3.**

Contributor	**Statement of contribution**
Yilun Fan	Experiment design (40%), data collection (100%), code implementation (100%), execution of experiments (100%), analysis and interpretation (40%), manuscript preparation (100%)
Yaniv Gal	Experiment design (30%), analysis and interpretation (30%), manuscript review (50%)
Andrew P.Bradley	Experiment design (30%), analysis and interpretation (30%), manuscript review (50%)

Fan, Y. and Bradley, A. P. (2016a). A method for quantitative analysis of clump thickness in cervical cytology slides. *Micron*, 80:73–82. http://www.sciencedirect.com/science/article/pii/S0968432815300391 (Fan and Bradley 2016a) **- incorporated as Chapter 4.**

Contributor	**Statement of contribution**
Yilun Fan	Experiment design (60%), data collection (100%), code implementation (100%), execution of experiments (100%), analysis and interpretation (60%), manuscript preparation (100%)
Andrew P.Bradley	Experiment design (40%), analysis and interpretation (40%), manuscript review (100%)

Fan, Y. and Bradley, A. P. (2016b). A two-stage method to correct the optical aberration induced by slide slant. *Micron*, 87:18–32. http://www.sciencedirect.com/science/article/pii/S0968432816300701 (Fan and Bradley 2016b) **- incorporated in Chapter 5.**

Contributor	**Statement of contribution**
Yilun Fan	Experiment design (80%), data collection (100%), code implementation (100%), execution of experiments (100%), analysis and interpretation (60%), manuscript preparation (100%)
Andrew P.Bradley	Experiment design (20%), analysis and interpretation (40%), manuscript review (100%)

Contributions by others to the thesis

No contributions by others.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgement

First and foremost, I would like to express my deepest appreciation to my principal supervisor Professor Andrew Peter Bradley, who guided me through this interesting research field, trained me to be a independent researcher and supported me at every moment of this thesis project. With his broadest knowledge, great patience, continuous encouragement, I am able to climb higher and higher in the last three and half years. This thesis would not have been possible without him.

I wish to express my deep gratitude to my associate advisors, Dr. Yaniv Gal and Associate Professor. Vaughan Clackson, for their generous supports and enlighting discussions to equip me with the sharpest tools to finish this thesis.

I would also like to thank all staffs at Biomedical Engineering Group who helped me during my candidature, in particular, Professor. Stephen Wilson, Dr. Karl Bertling, Dr. Andrew Smith, Dr. Yu Li, Dr. Jin Jin, and Mr.Craig Freakley for their experience sharing, encouragement and technical supports offered. Many thanks to my colleagues Dr. Mingyan li and Dr. Xia Ying, Ms. Yimeng Wang, Mr, Yaohui Wang and Ms. Yanmin Zhu for their enjoyable company and useful discussions at General South building room 528.

The financial support received from the University of Queensland International Scholarship (UQI) are gratefully acknowledged. I also appreciate the support received from the School of Information Technology and Electrical Engineering (ITEE) for my fruitful research travel to the United States in 2016.

My deepest gratitude goes to my family in China, particularly to my dad, mum, grandpa and grandma, who brought me up, given me so much, and most importantly encouraged my adventures in nature and science.

Finally, I appricate the valuable comments from two external examiniers of this thesis.

Keywords

Cervical Cylogy Screening, Whole Slide Imaging, Slide Scanner, Slanted Scan, Specimen Delineation, Cytology Specimen Thickness, Optical Aberration Correction, Deconvolution and Microscopy

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 080106, Image Processing, 55%
ANZSRC code: 090399, Biomedical Engineering not elsewhere classified, 45%

Fields of Research (FoR) Classification

FoR code: 0801, Artificial Intelligence and Image Processing, 55%
FoR code: 0903, Biomedical Engineering, 45%

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

2D two-dimensional. 20, 61, 82
3D three-dimensional. xvii, xix, 9, 17, 18, 31, 55, 56, 58–61, 76, 80, 82–84, 86–91, 93, 95, 100, 106, 108, 110

APCI auto-phase correlation index. xvi, xvii, 35–39, 42, 43, 46–54, 106, 108, 111
AUC area under of receiver operating characteristic curve. 43, 45

BF bright-field. 2, 18, 80–82, 84, 88, 90, 91, 93, 103, 104
BSFP best single focal plane. xvii, 59, 63–65, 70–72, 76

CAD computer aided diagnostics. 83
CCD charge-coupled device. 15

DOF depth-of-field. 11, 16, 23, 34, 57, 59, 65, 73, 93, 103, 107, 108
DSI dice similarity index. 43, 45, 46

EDF extended depth-of-field. xvii, 59–61, 73, 95
EI entropy index. xvi, 35, 38, 39, 42, 46, 47, 49, 50, 53, 54

FDA Food and Drug and Administration. 5
FNA fine-needle aspiration. xvi, 35, 41–43, 45, 46, 48, 50–53
FOV field-of-view. xvii, xxi, 3, 5, 6, 11, 12, 16–20, 23–28, 30, 31, 33–35, 39, 42–44, 53, 54, 59–68, 70–75, 77, 80, 91, 92, 104, 106, 108–111
FPM fourier ptychographic microscopy. 23, 24
fps frame per second. 28–30
FPSP FocalPoint Slide Profiler. 5, 6, 17

GB gigabytes. 10, 17, 18, 26, 30, 47, 57, 99

HE hematoxylin and eosin. 11, 35, 41, 54
HPV human papillomavirus. 2

IQR interquartile range. 45, 46, 67, 69, 73, 94, 102

LBC liquid based cytology. 4, 5, 9, 17, 24, 33, 55, 56, 58, 92, 106, 108, 109
lps lines per second. 26, 28–30

MB megabytes. 28, 30

MFP multiple focal planes. 26
MSE mean square error. xvi, xxi, 44, 47, 48, 52, 53
MTF modulation transfer function. 83, 84

NA numerical aperture. 24, 63, 80–82, 84
NACI normalised auto-correlation index. xvi, 35, 38, 39, 42, 43, 45–54, 106

OTF optical transfer function. xvii, 82–87, 90, 91

PPP plane parallel plate. 80, 81, 84
PSF point spread function. xviii, xix, 9, 22, 81, 82, 84, 85, 87–98, 101, 104, 107
PSNR peak signal-to-noise ratio. xviii, 90, 91, 93, 94, 102
PTF phase transfer function. 83, 84

RM review microscope. 6
RMSE root mean square error. 87, 90

SFP single focal plane. 26
SNFC sufficient number of focus candidate. xxi, 48, 53

TI threshold index. xvi, 35, 38, 42, 45–52, 106
TIS ThinPrep Imaging System. 5

WSI whole slide imaging. 1, 7, 8, 10–13, 15, 17, 18, 20, 23, 31, 56, 57, 79–81, 93, 105, 107–109, 111

1 INTRODUCTION

Pathology is the process that analyses human tissue, cell and fluid samples typically under a microscope in order to study and diagnose human diseases such as cancer. In particular, pathology has become the single most important modality for the diagnosis of cancer and more importantly for early detection of the cancer in order to improve the effectiveness of the treatment and therefore improve health outcomes. As a result, in Australia, the number of Medicare-funded pathology tests have almost doubled in the last decade, from 60.13 million in 2001 to a staggering 103.90 million in 2010 at the cost of 2 billion Australian dollars (AustralianGovernment 2013).

To increase productivity and control the cost of these pathology tests, one promising method is to digitise the traditional glass slide based pathology into digital-pathology utilising whole slide imaging (WSI). With WSI, microscopic specimens are converted into digital slides so that the management and interpretation of these specimens are undertaken in the image based environment, enabling computer aided diagnosis that has potentials to greatly increase the diagnosis productivity. However, the journey to digital-pathology is significantly harder compared to radiology going to digital because the the digitised glass slides are hundreds of times larger (gigapixels compared to megapixels), imposing greater technological challenges in terms of acquisition, storage, transmission, management and interpretation.

Existing digital slide acquisition systems and techniques for WSI are too slow and inefficient for digitising cytology specimens, because these specimens are relatively thicker and have hundreds of thousands of cells that are sparsely distributed. As a result, the computer aided diagnosis based on the WSI systems are not yet cost-efficient. This research project aims to develop methods for rapid digital slide acquisition of cervical cytology specimens. The outcomes of this research will contribute to the development of new digital slide acquisition systems with unprecedented scan speed.

The remainder of this chapter is organised as follows. Section 1.1 presents the motivation behind this study where methods have been developed to improve the productivity of cervical cytology screening. Section 1.2 discribes the current states and problems of digital slide scanners for WSI, especially when acquiring “glass-faithful quality” digital slides of cervical cytology specimens. Section 1.3 presents the objectives of the thesis and Section 1.4 covers the scope of the research described in this thesis. Finally Section 1.5 outlines the structure of this thesis.

1.1 Cervical Cytology Screening

According to the World Health Organization, in 2012, cervical cancer was the third most common cancer and had fourth highest mortality rate among women, with 527,624 new incidences and 265,672 new mortalities reported (for Research on Cancer 2015). The preventive methods of the cervical cancer include vaccination and early detection programs like the cervical cytology screening (Pap test). It is widely accepted that cervical cancer is caused by the infection of the human papillomavirus (HPV), so vaccines targeting the most common types of HPV virus (type 16, 18, 31 and 45) have been successfully developed, which can reduce the risk of the cervical cancer by around 93% (Rodríguez et al. 2013). Even after a vaccination, women still need regular cervical cytology screening, because the early treatment can effectively prevent the development of the cancer. The five-year survival rate for the early stage (IA and IB) of the cancer is around 93% (Society 2016). As a result, countries with effective screening programs, especially industrialised countries, have observed dramatic reductions in cervical cancer incidences (up to 80%) and mortality (Grabe et al. 2010; Schiffman 2007). Given the Pap test is increasingly provided in the developing world such as China and India and the fact that the test has to be repeated performed every 2-5 years, the demand for this particular screening test is overwhelming on a global scale (Grabe et al. 2010; Wright et al. 2010). Therefore, efficient screening methods need to be developed to meet the demands and to reduce the costs.

Cervical cytology screening was originally developed by Dr George Nicholas Papanicolaou in early 20th century. The principle of the test is straight forward, that is, to seek for the morphological changes of normal cervix cells towards malignant cells. Specifically, the nucleus of the cell becomes larger and more irregularly shaped, while the cytoplasm becomes smaller, so the size ratio of nucleus to cytoplasm increases. In addition, the chromatin distributions of the malignant nuclei are more coarse than that of normal cells. These cellular level changes can not be directly visualised by human eyes, so the microscopes, transmitted light bright-field (BF) microscopes to be specific, are normally used. In general, the sample cells from the transformation zone of the cervix are first collected and prepared as microscope slides for the screening. When viewed through the microscope, the cells on the slide are illuminated by light from beneath and absorb the light and hence appear darker than the background depending on the stain concentrations on the cells.

To detect any abnormalities in the specimen, all the cells on the slide should be viewed and examined under the microscope, and initially this process is done completely manually by human experts. As the technology advances, machines and computers have been developed to assist the human experts in order to improve the productivity, and the following section will focus on the technical perspectives of this development.

1.1.1 Manual Screening

The traditional, and also the historical, specimen preparation method for cervical cytology screening is the smear preparation method. Specifically, cells collected from the transformation zone of the cervix are manually smeared over the glass slide before staining and fixation. This slide is the so-called Pap smear slide, and it is reported that there are around 10,000-20,000 cells on the slide covering

a large area (25x50mm of a standard 25x75mm slide) (Malm et al. 2013). As the smear slides are prepared manually, the quality of these slides could vary significantly. The cellular materials are often distributed unevenly over the slide. In addition, the cells generally overlap and form dense cell clumps that sometimes may even prevent light from penetrating through them. There are also many foreign objects such as blood cells, mucus and inflammation. Nevertheless, this traditional preparation method is simple and cheap, at cost of around 1-2 US dollars (Bengtsson and Malm 2014).

As there are huge number of cells on the slides, in general, a trained cytotechnologist is asked to initially examine all the cells for any signs of abnormalities. If such signs are found, the cytotech will put ink markers on the top surface of the glass slide around them. The pathologists will then only need to review these limited number of suspicious cells and parts to make the final diagnostic decision. A block diagram of this manual screening process is shown in 1.1. In order to interpret the cellular details such as the chromatin distribution of nuclei required by the diagnosis, high magnification (e.g. 40×) is required, with resolution close to the optical resolution limit (approximately 0.25 microns). However, the size of field-of-view (FOV) is inversely proportional to the magnification, so the manual screening is a time-consuming process. In practice, the cytotech could prescan the slide at relative quicker speed at a relative low magnification (e.g. 10×), and then switch to high magnification if something suspicious is found.

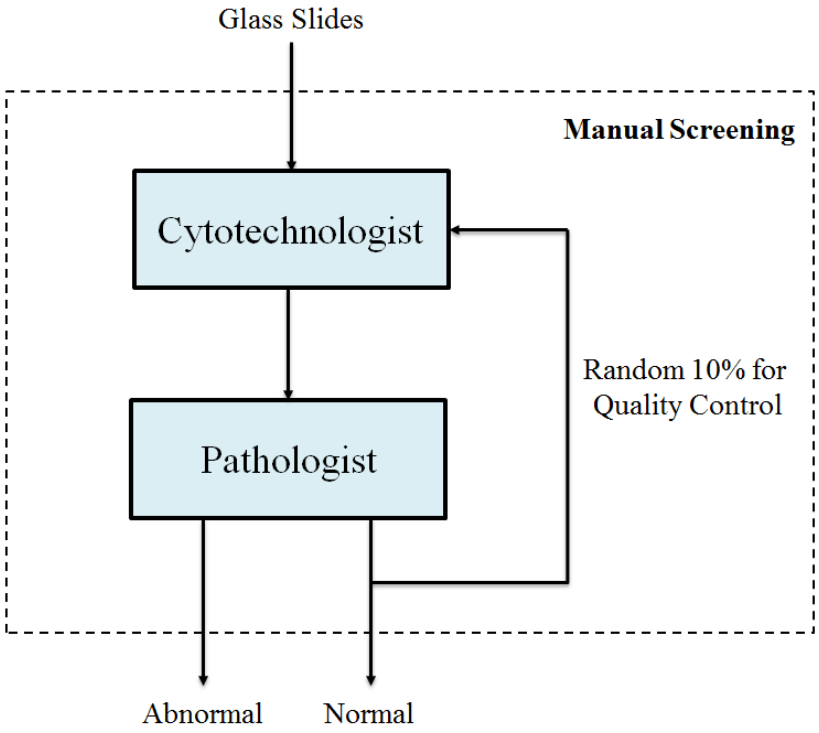


Figure 1.1: Typical block diagram of manual screening, assuming 10% of slides diagnosed as normal are re-screened for quality control.

The manual screening is also widely regarded as labour-intensive and subjective (Levi et al. 2012). The test is laborious because a huge number of cells needs to be examined but only a small fraction of them (e.g. 10-20 cells) may be visually abnormal, and thus conducting these tests could easily lead to human fatigue, habituation and lack of concentration (Malm et al. 2013). In addition, the visual

examination via a microscope is not quantitative, so the examination results will also depend on the knowledge and experience of human experts. As a result, false interpretations in the manual screening with Pap smear slides have been reported to be as high as 55% (Malm et al. 2013). Moreover, the diagnosis results from different cytotechs (inter-observer variation) and even from the same cytotech across different analysis instances (intra-observer variation) often do not agree. To reduce the false negatives (actually abnormal but diagnosed as normal), a certain proportion (normally around 10%) of slides initially diagnosed as normal are randomly selected for re-screening (as a quality control shown in Fig. 1.1) in order to increase the overall sensitivity of the testing process.

1.1.2 Machine Aided Screening

To improve the productivity of the manual screening, systems were first developed to increase the automation of the screening processes. For instance, the AcCell 2000 cytopathology system includes a motorised microscope that can automatically move the slide for a fly-over view of a specimen at low magnification and at a predefined suitable speed. When something suspicious is spotted, the user can pause the fly-over mode and then freely move the stage and change the magnification lens in order to examine the suspicious cells in greater details. After the inspection, the fly-over mode is resumed and the screening continues until the entire specimen is exhausted. This system partially reduces the workload of the cytotech and ensures the full specimen is systematically and completely viewed. However, the cytotech still need to fully examine every glass slide, and therefore the productivity improvement is limited.

1.1.3 Automated Screening

In order to actually increase the productivity of the screening, the machine should be capable of "automatically" analysing the specimens and producing quantitative results to assist human experts in the diagnostic decision making. This is generally achieved by converting the analogue signals (light intensities) into digital images and then quantitatively measuring and interpreting these digital images utilising computer based image processing techniques. For instance, the area ratio of nuclei and cytoplasm for diagnosis of the cancer can be obtained by using cells and nuclei segmentation and feature extraction techniques, and used as a feature to classify the abnormality of the specimen (Bengtsson and Malm 2014).

To automate and standardise the specimen preparation procedures and more importantly to make the appearance of specimens more consistent for consumption of both human and machines, liquid based cytology (LBC) slide preparation methods were developed. The two most popular LBC preparation methods for cervical cancer screening are Thin-prep and SurePath preparation methods. With the Thin-prep method, the collected sample specimens are submerged in a preservative liquid, and then treated normally using a centrifuge to separate the cervical cells from foreign objects (blood and mucus). The target cells are then collected using a thin membrane and transferred to a glass slide before fixation and staining. As a result, the cells are evenly distributed as a mono-layer over the slide with less overlapping, therefore, are much cleaner than the conventional smears. Moreover, the

cells are deposited in a circular area with diameter of 20mm, (or 12.55mm using Surepath), so smaller areas are examined compared to the Pap smear slides. Consequently, the LBC prepared cervical cytology slides have been widely accepted in pathology labs to replace the traditional pap smear slide. The consistent appearance of specimens is an essential step for quantitative evaluation required by the computer aided screenings. The only downside to the LBC preparation methods is associated with its cost because more material like preservative liquids and filters are required as well as complex and expensive slide preparation machines, thus increasing the cost of making a slide to multiple times that of traditional pap smear slides ($\sim$10 US dollars versus 1-2 US dollars) (Bengtsson and Malm 2014).

Apart from the development of the LBC preparation methods, the development of automated screening systems is closely related to the advances in computer technology. The earliest computer aided screening system (in 1956) used video cameras and analogue video processing circuits to measure the nuclei size, but the system had too many false positives because the effective resolution of the camera was poor (only 5 microns) (Tolles and Bostrom 1956). Since then, the resolutions of digital camera and the computing power had increased dramatically, enabling more accurate quantitative measurements of the specimen at faster speeds, and more importantly at much lower costs. The advances in computer display and storage technology promoted interactive screening programs, which were marked by the development of the first system capable of saving images of suspicious cells and displaying them to the users on computer monitors in 1994 (Mango 1994). Eventually, the first commercial product for true automated PAP screening of a full slide, Tripath, was granted approval from the Food and Drug and Administration (FDA) in the United States in 1996 (Kardos 2004). The Tripath company was later acquired by BD Diagnostics and the product was renamed the FocalPoint Slide Profiler (FPSP) (BD Diagnostic Inc., Burglington, NC). The SurePath LBC preparation method mentioned above was developed to further improve the FPSP system performance. In 2003, another automated screening system, ThinPrep Imaging System (TIS) (Hologic Inc., Boxborough, MA), was approved by the FDA; it uses the Thin-prep preparation method (Bengtsson and Malm 2014). The technical details of these two systems are described as follows:

FocalPoint Slide Profiler: The FocalPoint Slide Profiler is able to process both conventional PAP smear slides and SurePath LBC prepared slides. The slides are first scanned at low resolution ($4\times$) and the low resolution images are processed to find fields that are most likely to contain abnormal cells. Then, approximately 1000 of these fields containing single cells and cell clusters are selected for a high resolution scan at $20\times$. These high resolution FOVs are processed individually to find any signs of abnormality, and a score representing the abnormality of the entire slide is produced as well ten FOVs that are most suspicious. If any potential abnormalities are identified among these FOVs, the slide is manually screened by a cytotech (Levi et al. 2012). The computer generated score also guides the selection of slides for quality control, where at least 15% of slides with the highest scores are selected for manual screening (BD 2016). But more importantly, a maximum 25% of the slides processed by the FPSP system can be directly archived with no further review (Kardos 2004).

ThinPrep Imaging System: The ThinPrep Imaging System is designed to process Thin-prep LBC slides only. Similar to the FPSP, the FOVs of the slides are imaged and processed individually at high resolution to check any signs of abnormalities. Unlike the FPSP system, 22 locations of interests on

the slides are selected by the system and these positions are revisited under a motorised microscope that is part of the system. The microscope is called review microscope (RM), which is fitted with a motorised stage, a mouse and a touch pad for controlling microscope stage movements. If any of the 22 places is considered abnormal or suspicious, a full manual screening is initialized on the RM by the cytotech immediately. And it is reported in one study that only around 11% to 19% of slides are required to be manually screened (Halford et al. 2010). These 22 locations are selected based on a classification of several features including nuclear-to-cytoplasmic ratio, nuclear size, and nuclear staining. It is believed that the system is more sensitive than the (ThinPrep slides based) manual screening (Halford et al. 2010).

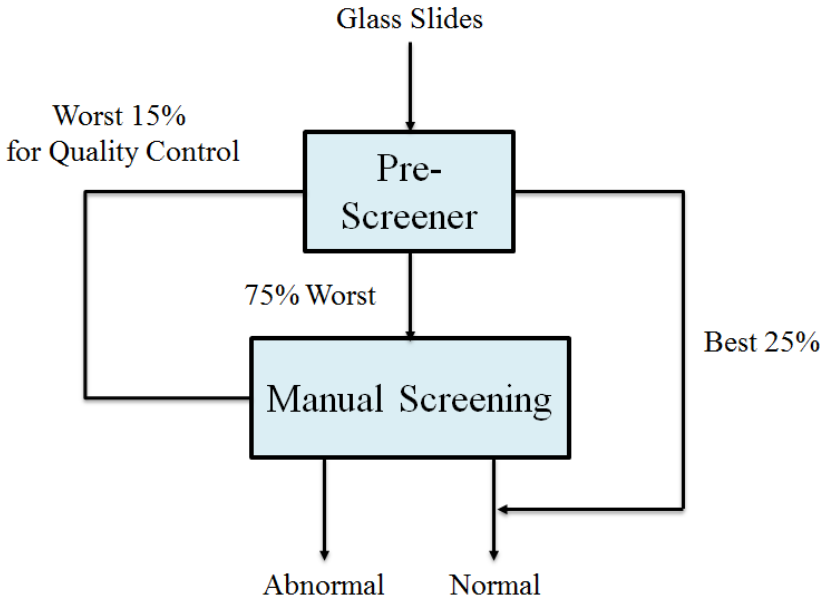


Figure 1.2: Typical block diagram of screening with automated screening machine running parallel with the manual screening. Assuming 25% of slides classified by the automated machine as normal are directly archived, 75% of slides are required to be manually screened once and 15% of slides are required to be manually screened twice for quality control.

The automated screening machines can be integrated into diagnosis work-flow in two ways, either as a pre-screener or as an adjunct screener as shown in Fig. 1.2 and Fig. 1.3 respectively. When those machines are used as pre-screeners, the glass slides are processed by the machine, and the tedious full slide manual screening is required when something suspicious is detected by the machine or if the machine failed to process the slide. As up to 25% of slides classified as normal can be directly archived without further examination with the FPSP system, at most 25% of the work-load can be directly eliminated (Levi et al. 2012). For the rest 75% of slides, the cytotech normally only needs to review a small number of FOVs (10) unless these FOVs contain suspicious cells, so the overall productivity of the screening process is significantly increased. According to some studies, up to 31% more slides can be screened per hour (Duby and DiFurio 2009; Papillo et al. 2008). However, the costs of the assisted screening are higher due to the expenditure on the machines and other infrastructure (around 8 US dollars more per test) (Papillo et al. 2008). These automated systems can also be used adjunct screeners. This is also the only option currently available to pathology labs in Australia,

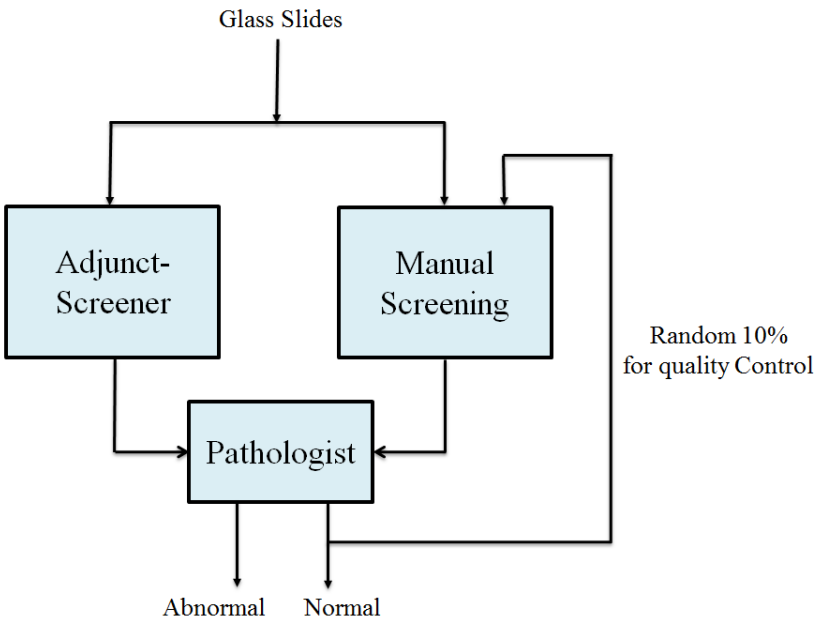


Figure 1.3: Typical block diagram of screening using automated screening machine as an adjunct screener, assuming 10% of slides diagnosed as normal are re-screened for quality control.

where the regulators dictate that each slide must be manually screened at least once (Boost 2009). The diagnostic output of each slide is a combination of results by the machine and human experts, which could reduce effects of human errors. However, the overall productivity is not improved and the costs are indeed higher compared to the manual screening alone, so the adjunct screening can be offered as an option for patients who are willing to pay extra for an additional machine-based screening.

The limitations of these automated screening machines are mainly associated with the use of glass slides. It is well known that the stain of the specimens fades under exposure to heat and light of the microscope illumination units, so the repeat use of a slide may lead to decreasing diagnostic performance (Leong and McGee 2001). Repeated handlings of slides also put them into risks of physical degradation such as scratches, breaks, and contamination of dust and fingerprints. Moreover, the storage and retrieval of glass slides in large quantities imposes logistical challenges, especially when several thousand slides are retained for a minimum five years in pathology labs.In a broader perspective, the use of glass slides makes second opinion of the case more difficult, especially by a remote consultant, because the slides could be damaged during transportation. These limitations can be overcome by digitising the entire specimen into a single image (as digital slide) utilising WSI, which has only been made available in recent years due to advances in computing processing power and the reduction in hardware costs.

1.1.4 Whole Slide Imaging Based Screening

Whole slide imaging is a relative new modality for pathology, which refers to digitalising the entire specimen into a digital slide (a single image of the specimen) and subsequently managing and interpreting the specimens in the image based environment. Compared to glass slides, the digital slides do not have risk of any degradation, they can be easily stored and managed in digital database, and they can

be efficiently distributed via computer network and, therefore, be accessed remotely in time and space. So, the digital slide and the WSI are ideal for slide archiving, quality control and reporting, and they have the potential to greatly improve work-flow efficiency given that pathologists spend considerable time on non-diagnosis related tasks like searching for and matching slides to paper works (Ghaznavi et al. 2013). The digital slides also enable remote consultation (second opinion) and remote diagnosis extending high quality health-care to regional hospitals that lack experienced pathologists (Li et al. 2007; Lu et al. 2015a).

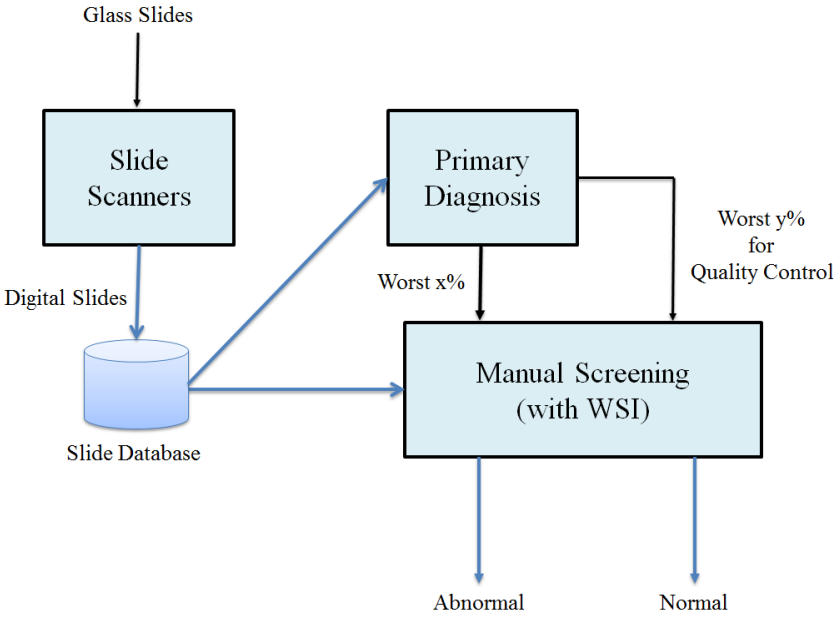


Figure 1.4: Cervical cytology screening based on whole slide imaging with automated screening machine used for primary diagnosis, assuming the automated screening machine picks x% of slides for manual screening once and y% of slides for manual screening twice.

A possible system design of the WSI based cervical cytology screening process is shown in Fig. 1.4, where the prepared glass slides are first converted into digital slides by the slide acquisition system (slide scanner). These digital slides could be stored into a digital database so that they can be easily accessed by the cytotech and pathologist through a computer network. These digital slides can be visulized on computer systems with dedicated software that simulate the functionality of a traditional microscope such as zoom and navigation. Compared to using traditional microscopes, the navigation with the digital slide is smoother and more streamlined. More importantly computer programs can be developed to asisit diagnosis, in particularly the automated screening algorithms can be easily integrated into this digitised work-flow because the digital image of the specimen has already been created. Many previous studies had demonstrated that WSI based manual screening could give equivalent performance as that using the glass slide, but only if the glass slide has the adequate quality as discussed below (Donnelly et al. 2013).

For diagnosis using digital slides to have at least the equivalent diagnostic accuracy of that with glass slides, and for the WSI based screening to really increase productivity and to be accepted in clinical diagnosis, the WSI systems must meet two essential criteria:

1. The digital slide must be in **"glass-faithful quality"**, which means that the digital slide contains

complete and accurate information of the glass slide that is required for the interpretation of the specimen by both human and machine. In this way, processing the digital slide is equivalent to processing the glass slide, and therefore the same diagnostic decisions can be made.

2. The acquisition speed of the "glass-faithful quality" digital slide must be **fast** in order to actually increase the productivity of the diagnosis process. Given that a complete manual screening (with conventional smear slides) takes around 6 minutes, the scan time of the slide must be less than 6 minutes in order to really increase productivity (Boost 2009).

For scanning cervical cytology specimens, the digital slide image should be at multiple focal planes (as three-dimensional (3D) image stack), at a spatial resolution close to the light diffraction limit ($\sim$0.25μm/pixel) and free of optical aberrations (Bengtsson and Malm 2014). Conventionally, microscopes with high magnification objectives (normally 40$\times$, with an air objective, i.e., the medium between the objective and microscope slide is air) are required in order to interpret subtle diagnostic details in the specimen, such as the texture of cell nuclei. Therefore, when acquiring digital images of these specimens an equivalently high sampling resolution should be used. The resolving power of a microscope is commonly calculated by the Abbe distance, which describes the minimum resolvable feature size in the image plane, as shown in Eq. 1.1.

$$r = \frac{\lambda}{2NA} \tag{1.1}$$

Where, r is the Abbe distance from half-amplitude of a microscope point spread function (PSF) to its centre in image plane, λ is the wavelength of the illumination light and NA is the numerical aperture of the objective (Castleman 1996). The axial resolution of the microscope is also inversely proportional to the magnification (or NA), described as depth of focus in Eq. 1.2 (Castleman 1996).

$$depthoffocus = \frac{\lambda}{2NA^2} \tag{1.2}$$

Assuming a microscope with a 40$\times$ (NA=0.75) objective illuminated with light centered at 525 nanometer for the optimal image contrast is used for acquiring digital images of specimens, the Abbe distance of this imaging system is around 350 nanometer while the depth of focus is around 470 nanometer. Therefore, the pixel spacing of the image sensor should be smaller than 0.175 micron in order to fully sample the specimen, while the distance between imaging planes should be at around 1 micron in order to have every object within the focal depth. Unlike histology specimens that could be as thin as 4-6 microns, cervical cytology specimens are typically thicker (10 - 30 microns), because many cells overlap and form thick cell clusters even when the LBC preparation techniques are used (El-Gabry et al. 2014). Thus, acquiring only one 40$\times$ image with a single focal plane could miss details of some cellular objects and lead to poor diagnostic performance (El-Gabry et al. 2014). Moreover, secretory cells from the endocervix are often found in thick cell clusters so these clusters should be sufficiently imaged (Lee et al. 2011; Wright et al. 2013) for the specimen to be considered adequate. Lastly, the digital images of the specimen should be free from optical aberrations so that the specimen can be quantitatively measured. Therefore, the cervical cytology specimens are optimally scanned in multiple focal planes (in 3D) and at high spatial resolution, which has been proposed in a

number of previous studies (El-Gabry et al. 2014; Farahani et al. 2015; Wright et al. 2013).

In this thesis, we define "glass-faithful quality" as a whole slide image acquired with an in-plane (spatial) resolution close to the diffraction limit of a conventional bright field light microscope, which is around $0.25 \sim 0.25 \mu$m/pixel. The inter (focal) plane resolution should be less than 1μm/voxel which is close to the (half wavelength) depth of field of a microscope lens capable of imaging a spatial resolution close to the diffraction limit. However, the optimal number of focal planes that have to be acquired is dependent of the "thickness" of the specimen. This thickness issue largely remains an open question, which has not been defined in literature and so will be studied in this thesis. Specifically the "glass-faithful quality" scan of the cervical cytology specimen is defined as the generation of digital slide images of cervical cytology specimens that when the specimen is viewed on a computer screen it has the same colour, contrast, resolution and depth of field as a conventional specimen deposited on a glass slide when viewed under a conventional bright field microscope.

1.2 Whole Slide Imaging

Prior to the introduction of WSI, digital cameras were initially installed on the microscope to take "static" digital images covering only small sections of the specimen for limited research and clinical use, because scanning of entire specimen was technologically challenging at the time. The advances of technology, particularly within last two decades, have led to the development of advanced robotic microscopy and specialised commercial WSI slide scanners, enabling increasingly convenient and fast digital acquisition of the entire specimen (Farahani et al. 2015). As a result, there is a clear trend in the pathology world whereby the use of conventional glass slides and microscopes is being replaced by the WSI (Ghaznavi et al. 2013). WSI is also being used in education and research sectors, because digital slides can be easily stored, distributed and viewed, opening new learning perspectives (Paulsen et al. 2010). Specifically, specimens of standard and rare cases can be digitised as digital slides that can be conveniently shared among large groups of students without constraints in time and location, which are difficult with glass slides. More importantly, the digital slides enable quantitative evaluation and computer aided analysis of the specimen, such as counting the number of cells and automated aided diagnosis.

Despite its numerous advantages, the WSI still face many technical and practical issues in terms of costs, image quality and standards, acquisition speed, data storage and management and regulatory concerns (El-Gabry et al. 2014; Ghaznavi et al. 2013). The initial setup cost of a WSI system is high, a typical commercial slide scanner can be more than 100,000 US dollars, and additional infrastructure such as computer networks and staff training require further investments (El-Gabry et al. 2014). Even with modern slide scanners, ensuring the consistent performance of the acquisition of digital slides especially to obtain clear and sharp images is still challenging, because the scanners have to follow the subtle height variations of specimens to maintain perfect focusing and at the same time be robust to various artefacts (Ghaznavi et al. 2013; Lahrmann et al. 2013). In addition, the file sizes of digital slides are normally significantly large and could easily excess several gigabytes (GB) when cytology specimens are acquired at high resolution and at multiple focal planes (Bradley et al. 2005; Ghaznavi

et al. 2013; Wright et al. 2013). Large image file sizes are problematic not only in terms of acquisition speed where downloading this amount data from a camera already takes long time, but also in terms of data storage and data management, especially when tens of thousands of slides are required to be processed annually in pathology labs. In particular, insufficient computer processing capacity and network speed may lead to slow loading and/or freezing of images, resulting in unsatisfactory user experiences (El-Gabry et al. 2014). Furthermore, the applications of WSI on certain diagnosis are subjected to the regulation by the governments, for example, no WSI system has yet been approved for primary diagnosis of specimens stained by the hematoxylin and eosin (HE) stain (El-Gabry et al. 2014).

As the key component in WSI, the slide scanners convert a conventional glass slide into a digital slide incorporating various techniques in optics, electronics controls and computer programs. The main trend in last twenty years is to build slide scanners with higher throughput rates, which is one of essential steps for the adoption of WSI. Histology specimens are relative easy for current WSI scanners to scan, because these specimens are thin (3-4 micron), and cells are bounded together so they have relatively smooth topology and small depth variation compared to cytology specimens. Moreover, histology pathologists are examining the structures of the tissues, which can be viewed at a sufficient level of detail under modest magnification (like 20×) (Gilbertson et al. 2006). Given that the depth-of-field (DOF) of a 20× (numerical aperture = 0.4) objective lens is around 6 micron meters, which is much larger than that (∼1 micron) under 40× lens, scans of these specimens at single focal plane are generally sufficient. A digital acquisition of such a specimen can be facilitated within a few minutes with most modern WSI slide scanners, therefore the adoption of WSI in histology pathology has been relative faster than in cytology pathology (Farahani et al. 2015; Ghaznavi et al. 2013; Gilbertson et al. 2006; Rojo et al. 2006).

The acquisition of cytological specimens in "glass-faithful quality", on the other hand, requires greater efforts, because cytology specimens are optimally sampled at high spatial resolution and at multiple focal planes. In addition, cytology specimens are spatially distributed and have rough topology, so ensuring the acquired images of cells are sharp and clear is challenging. Although most commercial available slide scanners are technically capable of scanning cytology specimens in multiple focal planes, they are slow (El-Gabry et al. 2014; Rojo et al. 2006; Wright et al. 2013). As an example, in Thin-prep cervical cytology slides used for cervical cytology screening, the cells are deposited within a circular area with a diameter of $20mm$, and the size of scan area is around $314mm^2$. Assuming the specimen has a uniform thickness of 20 microns and so at least 20 focal planes are scanned. The estimated acquisition times for the specified specimen by the fastest slide scanners made by leading commercial vendors (in 2015) are listed in Table. 1.1 below, showing that the majority of these scanners have an acquisition time of more than 30 minutes per slide (3DHISTECH 2016; Hamamatsu 2016a; Leica 2016; Olympus 2016; Philips 2016; Zeiss 2016). Although the scan times of listed scanners are considerably shorter compared to the early systems reported in (Rojo et al. 2006) and (Wright et al. 2013), none of these scanners has a acquisition time less than the eight minutes in order to compete with manual screening (Levi et al. 2012; Papillo et al. 2008).

The conventional FOV scan method for slide scanners utilises an area imaging sensor and acquires

Slide Scanner	Resolution	Scan Technique	Single Plane Scan Time (seconds)	20 Planes Scan Time (minutes)
Axio Scan.Z1 **Zeiss**	20×@0.22μm/pixel	FOV	334	111
ScanScope AT2 **Lecia**	40×@0.25μm/pixel*	Line	216	72
VS120 Slide Scanner **Olympus**	20×@0.33μm/pixel	FOV	167	55
Pannoramic 250 Flash 3D **HISTECH**	41×	FOV	122	41
Intellisite Ultra Fast Scanner **Philips**	40×@0.25μm/pixel	Line	84	28
NanoZoomer-XR **Hamamatsu**	40×@0.25μm/pixel*	Line	62	20

* 20× (NA=0.75) objective lens with a 2× adapter.

Table 1.1: Selected commercial available digital slide acquisition systems as of 2015. Manufacturer names are in bold.

thousands FOV images of the entire specimen that are eventually stitched together to form an montage image of the specimen (the digital slide). Some slide scanners adopt a line scan method using a linear array imaging sensor, where the stage moves continuously from one end of the specimen to another end producing image strips of the specimen. The line scan cameras are capable of running at high frame rates and consequently achieve high data-throughput rates, which is why the fast slide scanners listed in the Table. 1.1 utilise the line scan method. However, both methods still have to require multiple passes of the specimen in order to acquire images at multiple focal planes. Recently, a novel scan method, slanted scan, was specifically designed to scan specimens at multiple focal planes (Bamford and Mayer 2009). In the slanted scan, the glass slide is slanted in a small angle (3-5 degrees) so that the cross-sections, or multiple focal planes of the specimen can be imaged at the same time. By moving the stage continuously, the specimen images at multiple focal planes can be acquired simultaneously in one pass of the specimen. Despite its great potential, the method has not been evaluated against the conventional scan methods, for example, to determine whether the method is able to acquire "glass-faithful" digital slides.

1.3 Objectives of the Research

As presented in this chapter, WSI is an enabling technology that has potentials to greatly increase the productivity of pathology tests such as the cervical cytology test. However, existing digital slide acquisition systems for WSI are inefficient in scanning "glass-faithful" digital slides of cervical cytology specimens due to the requirements of images at multiple focal planes and high spatial resolution.

The aim of this thesis is to develop methods for increasing the scan efficiency of cervical cytology specimens by focusing on the two following objectives:

- **Objective 1:** To develop methods for optimising the scanning of cytology specimens for WSI.
- **Objective 2:** To investigate whether the slanted scan method is capable of producing "glass-faithful" digital slides.

These two objectives represent two complementary directions for addressing the overall research problem in this thesis. The specific testable tasks for this thesis are consolidated in the next chapter after reviewing existing methods for the digital slide acquisition system for WSI .

2 REVIEW OF DIGITAL SLIDE ACQUISITION SYSTEMS

This chapter reviews the techniques utilised by digital slide acquisition systems for WSI, particularly for acquisitions of thick cervical cytology specimens at multiple focal planes and at high spatial resolution.

2.1 Overview of Digital Slide Acquisition Systems

The digital slide acquisition system for WSI, or the slide scanner, in essence, is an automated and computer controlled microscope. The essential components of these scanners consist of illumination units and microscopic optics lens for image formations, a digital camera for image acquisitions, a motorised stage for moving the glass slide and a computer for control and image processing. The image resolution and quality of final digital sides depend upon a combination of all the optical elements including the illumination, slide-cover-slip system, the tube lens, the objective lens, and the imaging sensor that converts the analogue signals to digital. Objective lens normally contain both intrinsic and extrinsic aberrations such as the spherical aberration and geometrical aberrations that affect the image quality, so high quality objective lens are generally utilised. The pixel size of the imaging sensor should be sufficiently small in order to meet the Rayleigh sampling criterion used in microscopy imaging (Castleman 1996). The majority of existing slide scanners utilise charge-coupled device (CCD) as the imaging sensor, because they have better photon sensitivity and low acquisition noise producing high quality images, but they have a relatively low data transfer speed (Rojo et al. 2006). Another popular type of imaging sensor is the complementary metal-oxide-semiconductor (CMOS) sensor, which generally has faster data throughput rates and is able to achieve faster frame rates. The imaging quality of these CMOS sensors have been proved greatly in recent years, so they are increasingly used in slide scanners in order to achieve faster acquisition speeds. Most commercially available scanners are fitted with a slide loader so that hundreds of slides can be digitalised with minimal human interaction (Rojo et al. 2006). To reduce the adverse effects of ambient light and dust, optics and electronic components are usually enclosed in a dark box (Altinay and Bradley 2013).

2.2 Scan Procedures in Digital Slide Acquisition Systems

The scanning process of most slide scanners consists of three typical procedures: initial specimen localisation, subsequent high resolution acquisition and final generation of the digital slide from the acquired images. The details of specimen localisation and post-processing are described as follows, and those for a high resolution scan along with the associated scan methods are covered in the next section.

2.2.1 Specimen Localisation

The first step of the acquisition process is to find where the specimen is located on the glass slide, both the lateral (XY) and axial (Z) position, which can be treated as two sub-steps. The size of the FOV through the microscope objective is generally tiny compared to the size of the specimens, so hundreds or thousands of field images (FOVs) must be acquired in order to cover the entire specimen. To this end, it is critical to know the accurate position (XY) of the specimen, and hence determine the minimum number of areas to scan to at least avoid scanning blank fields. A general method is to initially acquire a low resolution image (with the lowest available magnification lens) of the whole slide followed by a simple intensity based thresholding method such as Otsu's method to identify the foreground specimen and generate an accurate scan map of the slide (Altinay and Bradley 2011). It was reported that doing so for some histology specimens has the potential to save scanning up to 50% of the size of scan area and likewise save on the scan time and the resultant file size (Altinay and Bradley 2011). The benefits of delineating the cytology specimens could be even higher due to the fact that cytology cells are generally spatially distributed. In terms of low resolution image acquisition, some commercial slide scanners use a macro camera with the aim to reduce acquisition time because these cameras have faster frame rates as compared to the large high resolution imaging sensors (Rojo et al. 2006). Apart from using automatic specimen delineation methods, most slide scanners allow users to manually select the scan map (Rojo et al. 2006).

Finding the correct focal positions (Z) for the specimen is the second vital sub-step before the high resolution scan, because the DOF of the microscope is small, particularly at high magnification, compared to the thickness of cytology specimens. The specimen images can easily get blurred, lose high resolution details and become useless when the scanning focal planes are missed by even a few microns. Indeed, maintaining the quality of focus to ensure the final digital images are sharp and clear is regarded as one of most critical aspect of the image quality and also one of most challenging tasks for slide scanners (Cornish et al. 2012; Ghaznavi et al. 2013; Gilbertson et al. 2006; Lahrmann et al. 2013). The easiest strategy for mataining the high quality focus is to find the accurate focal depth for every point or FOV of the specimen before taking the high resolution image (Bradley et al. 2005). However, this method is extremely time-consuming with average hardware, because the focusing methods need the microscopic stage to move to multiple focal depths in order to find the in-focus focal depth, hence it is subject to the mechanical limitations of the stage (Tello-Mijares et al. 2013). Additionally, given the main imaging sensors for high resolution and large size image acquisition often have low frame rates, focusing in on a large number of different points can lead to a significant accumulation of scan

time. An alternative method is to find the focal depth of the specimen after the stage moves a certain distance if the specimen can be assumed to be reasonably flat locally, but a large number of focusing operations may be still needed if the topographic variation of the specimen is rough, especially for cytology specimens. Although various methods have been developed to increase the focusing speed, such as employing an additional fast camera specifically for the focusing (Farahani et al. 2015; Kaikai et al. 2015), a more general technique is to estimate the 3D focus profile map (*scan map*) of the entire specimen.

The focus map includes the focal depths of interesting specimen objects across the whole slide. The focus map defines what focal (Z) position the stage should be at for every spatial (XY) location of the specimen to be imaged. More specifically, if a FOV scan method (as described in the next section) is utilized, the focus map should contain the lateral positions (XY) and focal depths (Z) of each FOV. The focus map is estimated from the focal height of a limited number of points spread over the specimen, which are called focus candidates. If the topographic of the specimens is rough, more focus candidates should be sampled to accurately reflect the variations in the specimen surfaces. Conversely, few focus candidates are required when scanning smoother specimens. This approach has been employed in most commercial slide scanners for WSI as well as the FPSP system mentioned in Chapter 1 for primary diagnosis of cervical cancer (Gilbertson et al. 2006; Hilsenstein 2005; Lahrmann et al. 2013; Rojo et al. 2006). Another benefit of scanning the specimen in 3D is that specimens with rough surfaces are more likely to be adequately scanned.

After creating the scan map and focus profile map, decisions need to be explicitly made in regards to the number of focal planes to be acquired in the high resolution scan. It is generally believed that the thickness of the LBC based cervical cytology specimens used for cervical cytology screening is around 20 microns, so around 20 focal planes are acquired in some studies (Bradley and Bamford 2004; Evered and Dudding 2011). Although some studies had argued that fewer number of focal planes could be sufficient for diagnosis, the exact number of focal planes required is generally decided empirically. Thus, commerical available slide scanners allow the user to select the number of focal planes to be sampled, given that they could not determine the thickness of the specimen automatically (Rojo et al. 2006). On the other hand, it is clear that the number of focal planes needs to be minimised to reduce the scan time because the size of each focal plane is huge especially at high resolution (around 8GB found in last Chapter).

2.2.2 Post-processing

After the specimen images are acquired at high resolutions, these images are stitched together to form a montage and seamless image of the specimen, which is the so-called digital slide. Various alignment methods have been proposed for stitching the images together relying on different image clues, such as the intensity of the images by cross-correlation methods, the phase of the images by phase-correlation methods and image feature based methods (Bradley et al. 2005; Kuglin and Hines 1975; Lewis 1995; Yang et al. 2013). Some methods use one or multiple of the above methods, and others extend the stitching to 3D (aligning images at multiple focal planes simultaneously) (Nozaka et al. 2013; Preibisch et al. 2009). Prior to the stitch, any intrinsic (geometric distortion) and extrinsic dis-

tortions (rotation of the camera) in the raw images should be corrected in order to effectively match the individual images (Altinay et al. 2010; Sun et al. 2005). Geometric removals are also essential for quantitive evaluation. Other improvements on the image quality of acquired images include the correction of uneven background illumination, the image fusion for removing sharp intensity changes over the overlapping areas and so on (Kayser et al. 2011; Rankov et al. 2005; Romer et al. 2003; Sun et al. 2005).

Multiple file formats are available to store these generated digital slides in digital medium, such as the Tagged Image File Format (TIFF) and the Joint Photographic Experts Group (JPEG) 2000 (Farahani et al. 2015). The JPEG 2000 format enables the effective handling of a single image larger than one giga byte as well as progressive transmission, random access and image compression (Bradley et al. 2005). Despite this, the file sizes of digital slides containing N focal planes of the specimen are approximately N times larger, in order of several GB, imposing great challenges on its storage, processing and analysis. Thus, the number of focal planes should be minimized to diminish overall storage sizes and acquisition time, given that sufficient diagnostic information ais preserved. However, if the storage is strictly limited, a possible method is to fuse the images at multiple focal planes into one image in extended depth of field with everything in focus. This has been demonstrated that to be superior than viewing only one focal plane but with limited use for only qualitative analysis (Bradley et al. 2005). The full Z stack images of cells having each individual object in best focus still needs to be saved for quantitative analysis both by human and computer aided diagnosis systems.

2.3 Scan Methods for Whole Slide Imaging

Details of four existing scan methods that have been proposed and used for WSI are presented in this section, in particular focusing on how they acquire specimens in 3D.

2.3.1 FOV Scan

The FOV scan is the most conventional and simple scan method, and therefore it has been used in many scanning systems (Bradley et al. 2005; Rojo et al. 2006). The method can be simply realised on BF microscopes when they are fitted with a computer controlled motorised microscopic stage and a digital camera. The motorised stage moves the specimen against the objective lens so that the specimen can be imaged by the digital camera. Each specimen image is a FOV image of the specimen, and the stage typically follows a serpentine scan pattern as shown in Fig.2.1 to acquire the entire specimen. Given the requirement for high resolution image acquisition of cytology specimens, the stage must stop and rest to wait any vibrations to cease before camera exposure, otherwise the final image will have motion blur degrading the image quality. During the scan, overlapping FOV images are normally taken so that the overlaps of neighbouring FOVs can be used to align the images. To acquire the specimen images in multiple focal planes, a motorised stage moves the stage along the focal axis (Z) in incremental steps.

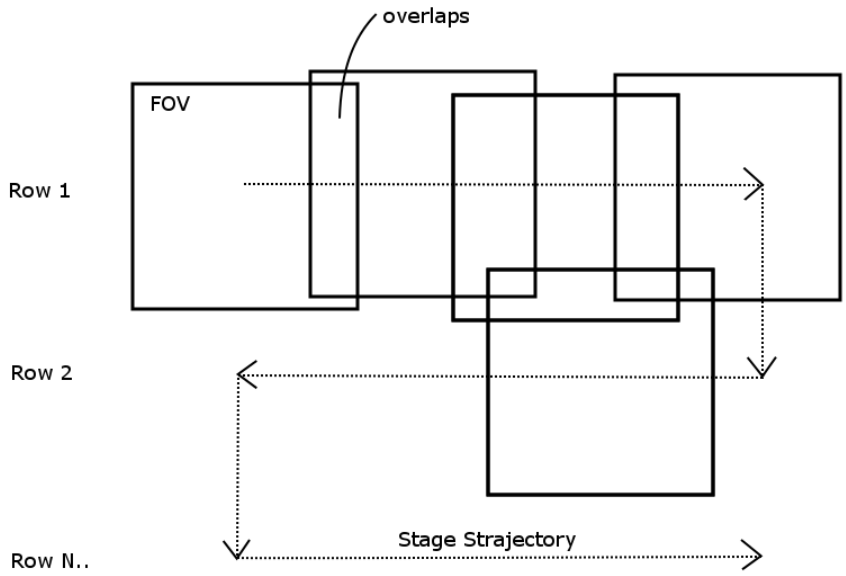


Figure 2.1: A typical FOV scan pattern for covering a large specimen.

2.3.2 Line Scan

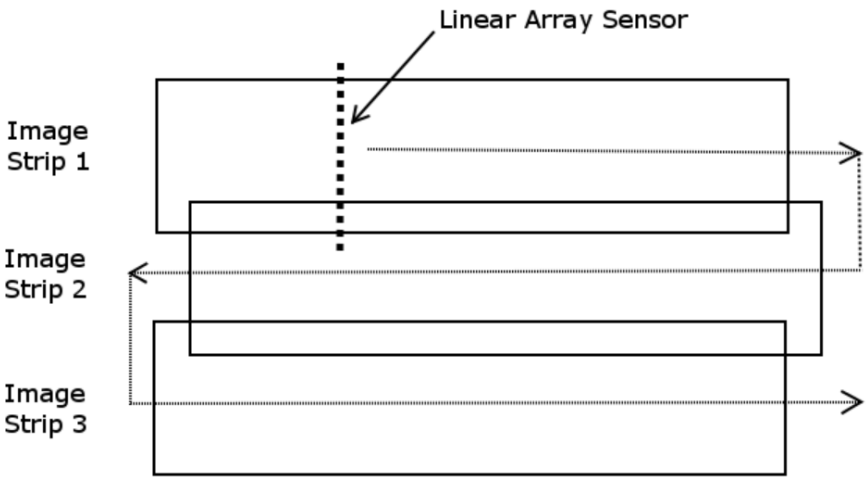


Figure 2.2: An line scan imaging sensor (left) and a typical line scan pattern for covering a large specimen (right).

Line scan cameras (linear array sensor) can also be used for image acquisition as they are able to achieve high data-throughput rates (Mayerich et al. 2008; Rojo et al. 2006). The trajectory of the microscope stage in a line scan system is illustrated in Fig. 2.2, where the long axis of the imaging sensor aligns perpendicular to the direction of stage movement so that the system acquires an image of a long strip of the specimen from one end to the other in one pass of the stage. The obtained image strips are eventually stitched together to form a whole image of the specimen. In contrast to the FOV scan, the stage moves continuously at a relatively constant velocity, and only the image overlaps in one direction (between image strips) are required. Consequently, the lines scan acquires less amount of raw data and performs one dimensional image stitches compared to the FOV scan (Yu et al. 2011). More importantly, the line scan cameras are generally faster than area scan cameras allowing greater

system data throughput rates. However, optical encoders are required on the microscope stage for synchronisation of stage movement and camera exposure. Improper synchronisation, such as false or delayed triggers, can easily result in motion blur and/or geometric distortion of the image strips, which will adversely affect the image quality. On the other hand, the line scanning reduces the geometric distortion correction from two-dimensional (2D) to only one dimensional because the image sensor is normally positioned across the optical axis of the lens. Finally, each image strip and each pass of the stage only sample one depths of the specimen, so multiple passes of the stage are required to sample thick cervical cytology specimens.

2.3.3 Lens Array Scan

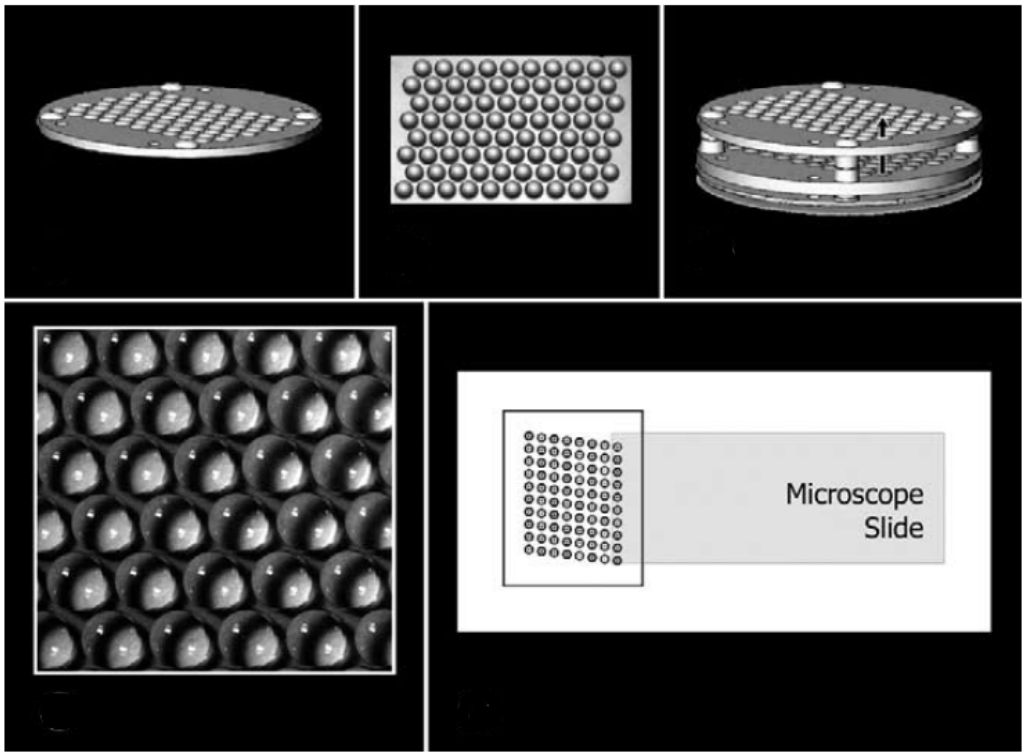


Figure 2.3: Top row: Schematic of the lens array. Bottom left: Close look of the lens. Bottom right: Relative position of the lens array to the glass slide at the start of the scan. Images reproduced from (Weinstein et al. 2004) without permission.

In conventional slide scanners, only one objective lens can be used at once so that only one FOV (or one line in a line scan) can be acquired in one camera exposure. It is clear that the scan speed can be increased dramatically if multiple objective lenses can be used to take multiple FOVs simultaneously. However, the relative size of microscope objective lenses compared to the size of a glass slide does not allow such design. For instance, a conventional 40× Zeiss objective lens has a diameter of 30mm, which is already larger than the diameter of the specimen deposition area of the Thin-prep cervical cytology specimen (Zeiss 2015). In order to acquire multiple FOVs at the same time, miniature objective lenses must be designed. Recently, Ronald et al. developed a lens array imaging system for WSI consisting of 80 miniature objective lenses so that 80 FOVs can be imaged simultaneously as shown in Fig. 2.3 (Weinstein et al. 2004). A 24 mega-pixel imaging sensor is used to acquire images.

The acquisition time for a specimen of 2.25cm2 is 58 seconds, at resolution of $\sim$0.47μm/pixel, which is approximately equivalent to 20$\times$ magnification (Weinstein et al. 2004).

The main limitation associated with the lens array scan method is its poor image resolution. Conventional high quality objective lenses are compound lenses that consist of a number of optic lens elements necessary to correct various optical aberrations, such as chromatic and spherical aberrations, for achieving optimal image resolution and image quality. In the case of the lens array, on the other hand, there is limited space in which to pack several optical lenses so the resolution of system is almost the double that of the diffraction limit, which is not acceptable for scanning cervical cytology specimens.

2.3.4 Slanted Scan

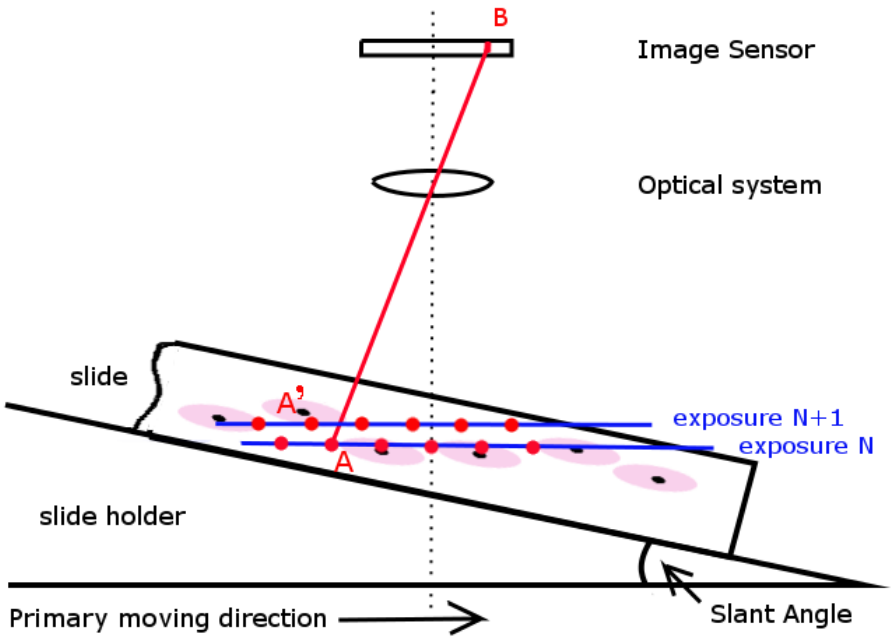
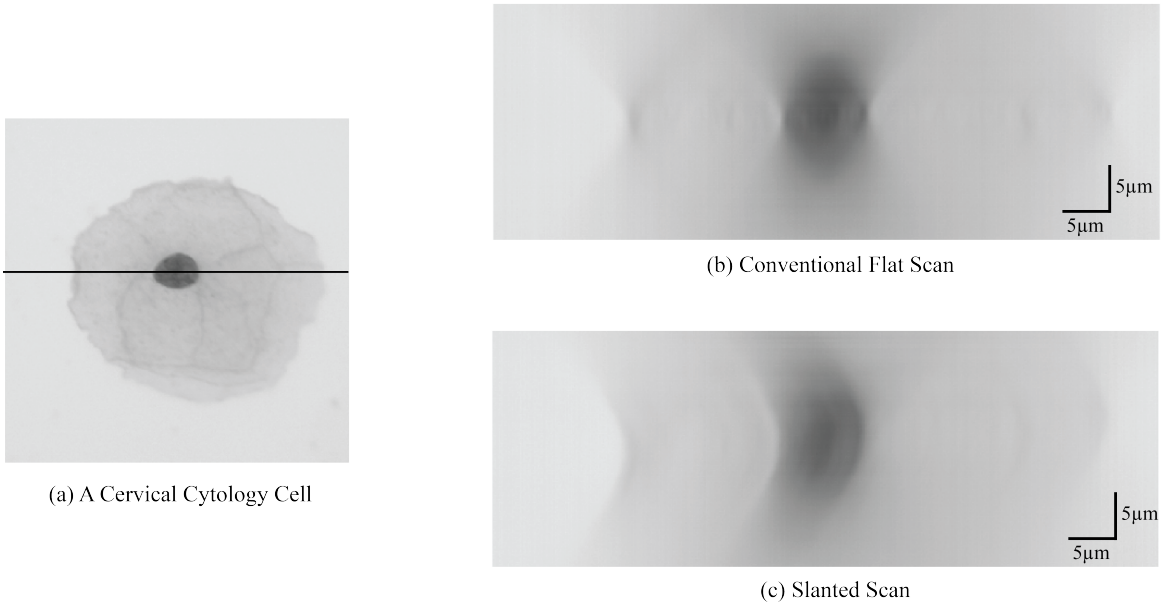


Figure 2.4: The schematic of slanted scan method.

Recently, a novel slide scan method, the slanted scan, has been specifically designed to acquire digital slide of thick specimens at multiple focal planes (Bamford and Mayer 2009). The microscope slide is slanted at a small angle (around 3-5 degrees) so that multiple depths of the specimen can be imaged simultaneously in a single pass of the stage. The principle of the method is illustrated in Fig. 2.4. As the slide is slanted an angle with respect to the optical axis, the cross-section of the specimen (several focal planes shown as red dots) can be imaged in one exposure. For example, image line A' contains information from specimen depth A. Next, by moving the stage perpendicular to the line imaged at A (perpendicular to the paper surface), multiple image lines from different depths of the specimen can be simultaneously acquired, thereby forming images of cross sections of the specimens at multiple focal planes. The acquisition can be done by using several linear array sensors at appropriate positions, or multiple lines can be scanned via an area scan camera. To acquire the whole specimen, the stage follows a trajectory similar to that of a line scan system but only one pass

of the specimen is required. Overlaps are only required between image strips, and these these image strips eventually stitched together to form complete images of the specimen in each focal planes.

The main advantage of the slanted scan is its ability to scan a section different focal planes of the specimen so that only a single pass of the stage is required. Rather than slanting the specimen, an alternative approach is to slant the image sensor, e.g. by placing several linear imaging sensors at different focal distances in the imaging plane, which had been attempted earlier (Nordin 1989). However, this method requires the sensor to be fixed at accurate slant locations prior to the scan. So it would be difficult for it to adjust the slant angle if specimens of different thickness need to scanned, impeding the robustness and applicability of the system. However, with the specimen slanted, one can make several interchangeable slide holders for different slant angles, and meanwhile the imaging system can be still used to perform conventional flat scans. To this end, a conventional slide scan system can be easily converted into a slanted scan system with minimal hardware modifications. Apart from this, when scanning a Z stack of images of the specimen, cells are more likely to be acquired in focus, leading to a higher success rate of slide digitalisation.



(a) A Cervical Cytology Cell

(b) Conventional Flat Scan

(c) Slanted Scan

Figure 2.5: The side image of a cervical cytology cell (a) without (b) and with (a) slide slant. The slide slant angle is approximately 4 degrees. Notice the shape of cell nucleus is nearly circular so the shapes of nucleus blur resemble the point-spread-function of the imaging system.

Despite the slanted scan being capable of simultaneously scanning multiple depths of the specimen, the slide slant induces optical aberrations resembling that of a lens coma. The main artefacts include movement of the object blur at different depths (bending of the PSF) and the reduction of the image contrast compared with images acquired in the conventional flat orientation (Arimoto and Murray 2004; Bamford and Mayer 2009) as shown in Fig. 2.5. The existence of these aberrations means that the digital images acquired with the method are not "glass-faithful" yet. So, these artefacts must be corrected in order for the slanted scan method to be capable of producing "glass-faithful quality" digital slides and more importantly to compete with the conventional scan in terms of image quality. In addition, the optimal slide slant angle should be investigated, which is also relevant to the thickness

of the specimen.

2.3.5 Fourier Ptychographic Microscopy

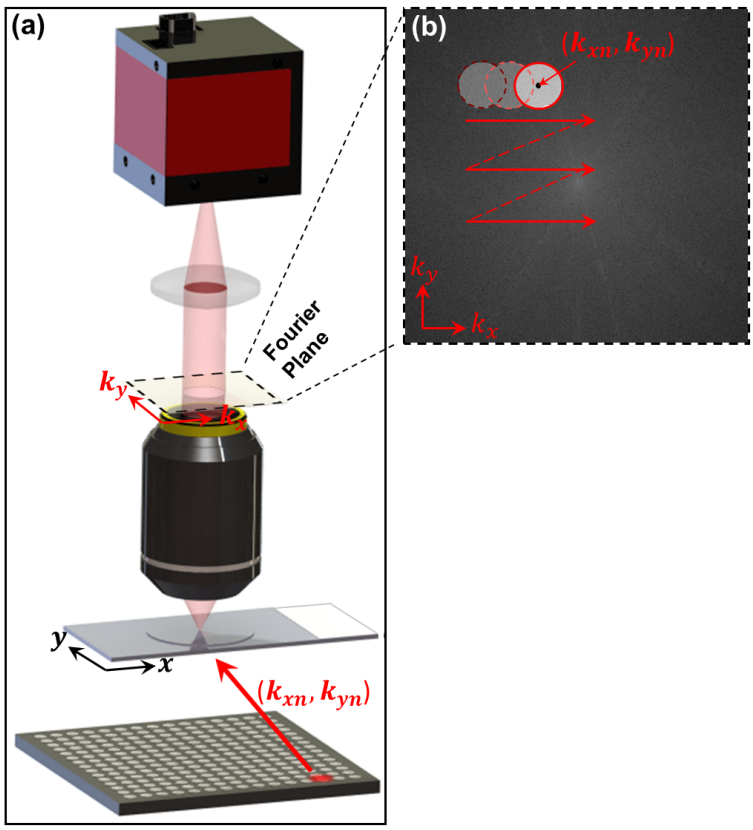


Figure 2.6: Fourier Ptychographic setup and imaging procedure. (a) An LED array placed beneath a conventional microscope illuminates the sample with unique wavevector k. (b) The object's finite spatial frequency support, defined by the microscope's NA in the Fourier domain (red circle), is imposed at offset locations to reflect each unique LED illumination angle. The Fourier transform of many shifted low-resolution measurements (each circle) are stitched together to recover a complex sample spectrum with resolution extending well beyond the objective lens cutoff frequency. Figure reproduced from (Zheng 2016b) without permission.

fourier ptychographic microscopy (FPM) is a super-resolution technique capable of greatly increasing the resolution of a objective lens by incorporating angular varying illuminations and a phase retrieval algorithm based image reconstruction method as illustrated in illustrated in Fig. 2.6 (Zheng et al. 2013a). As high resolution images can be obtained from a lower resolution objective lens, more celluar objects and hence information can be captured in a single FOV leading to acquisitions of fewer number of them, which could result short digital slide acquisition time if it is used for WSI. Specifically, it had been demonstrated that a FPM system fitted with a 2× objective lens was able to produce a digital slide at resolution equivalent to using a 20× objective lens, where the FOV of the system is the same as that of the 2× lens, which is much larger than that of 20× lens (Zheng et al. 2013a). More importantly, the reconstructed images have the same DOF as that of 2× objective lens, which is considerably larger than that of a 20× ojective lens. So only one FOV image may be needed to cover the all focal planes of cervical cytology specimens if the FPM extend the final resolution close

to the diffraction limit. Given the fewer number of FOV images are required, the overall acquisition time could be small with the FPM. However, hundreds of images are acquired to produce each high resolution FOV image, and around 40% of overlaps are needed between images for the complex and computational expensive reconstruction algorithms, the actual acquisition time reduction is limited (Ou et al. 2014; Zheng et al. 2013a;b). More complicating, the method is currently applicable to microscopes with coherent illuminations only, where ring artefacts appear in the images at cell edges due to the attenuation of the high frequency signals, which is distinct to that in a spatially incoherent system (Goodman 2005). Apart from this, there are also noticeable artefacts on the reconstructed images especially in the blank area, which could be contaminations on the light path brought into focus due to the extended depth of field in FPM (personal communication) (Zheng 2016a). Therefore, despite its great potentials, the FPM method is not further studied in this thesis.

2.4 Performance Analysis of the FOV Scan, Line Scan and Slanted Scan

In this section, the performance of the FOV scan, the line scan and the slanted scan as well as their system requirements are analysed and compared by estimating the scan time of a typical LBC cervical cytology slide for cervical cytology screening. The lens array scan system is not included in this comparison because it failed to meet the requirement for spatial resolution needed for digitisation of cytology specimens.

Settling Time (T_m)	200ms
Settling Time (T_{mz})	100ms
FOV Size	2048x2048
Imaging Resolution	0.2μm/pixel
Deposition	Circular with diameter of 20mm
Scan Area	$3.14mm^2$
FOV Overlaps	20%
Line Scan Overlaps	20%
Acquiring Number of Focal Planes	20

Table 2.1: Scan configurations.

Where specific parameters (settling, exposure times etc) are required we have estimated these based on commercially available hardware: specifically an Olympus BX40 microscope with 40$\times$ objective (0.75numerical aperture (NA)), an AcCell2000 motorised stage and a high resolution camera with 7.4μm pixel size (SPOT Insight). This system produces a spatial resolution of 0.2μm/pixel (notice not the effective resolution), which is close to the diffraction limit of the optical microscopes. The scan target was Thin-prep prepared cervical cytology slides, where the cells are deposited inside an circular area with diameter of 20mm. It is also assumed that specimen has a constant thickness of 20μm and so 20 focal planes are acquired if the focal distance between the slices was set to approximately the focal depth of the system (1μm). These parameters meet the requirement for digital slide to be "glass-faithful quality" proposed in Chapter 1. The assumed overlap between adjacent FOVs in FOV scan

and image strips in line scan is 20%. The scanning parameters are summarized in Table. 2.1, and these parameters were the same as an experimental scanning system specially designed to scan the Thin-prep slides at high quality (images at multiple focal plane and high spatial resolution) in (Bradley et al. 2005). It is assumed that the scan map and the focus profile map of the specimen have already been obtained, say using a method such as (Hilsenstein 2005), so that there are no extensive focusing during the high resolution scan. The equations proposed in this thesis only estimate the time for high resolution acquisition, not only because it generally account for the most part of acquisition but also based on fact that the post-processing generating the final digital slide could run concurrently with the scan or during the acquisitions of other slides. Finally, the time estimation equations consider only the acquisition of monochrome images, but these can be easily extended to a colour acquisition assuming that each colour plane is acquired in a separate pass, i.e., taking three times longer. Alternatively, colour images can be acquired in the same time assuming there are separate Red-Green-Blue (RGB) sensors.

2.4.1 Scan Time Estimation

FOV Scan: The time spent between two FOVs depends on microscope stage movement time (T_m) and the camera operation time (T_c). Here T_m consists of the stage acceleration time (T_a), cruising time (T_{cr}), deceleration time (T_d), settling time (T_s) and a variable waiting time (T_{ws}), i.e., $T_m = (T_a + T_{cr} + T_d + T_s + T_{ws})$. A typical camera operation cycle starts from exposure of a FOV (T_e), which is then downloaded (T_r) and lastly, waiting for stage to move to the next FOV (T_{wc}), i.e., $T_c = (T_r + T_e + T_{wc})$. To keep image contrast constant it is assumed that illumination intensity is inversely proportional to exposure time.

For an FOV system to achieve maximum scan speed, the stage and the camera operations are synchronised so that the camera starts exposure immediately after the settling of the stage for the minimum of T_{ws} and T_{wc}. The shortest possible scan time for a single focal plane scan ($T^{SFP}_{FOVScan}$) is shown in Eq. 2.1, where (N_{FOV}) is number of FOVs required to image a single focal plane of the whole specimen. The N_{FOV} is found by counting the total number of FOVs required to cover over the entire specimen area (scan map) one by one while maintaining the desired overlap area, as demonstrated in (Altinay and Bradley 2011). The minimum of T_m and T_c is around 200ms and 295ms for the Olympus platform.

$$T^{SFP}_{FOVScan} = \max(T_m + T_e, T_c) N_{FOV} \tag{2.1}$$

When multiple focal planes (N_f) are acquired the scanner needs to capture multiple images at each FOV. Here we include the time for the stage to change its focal position ($T_{mz} \approx$ 100ms, smaller than T_m), stage running time in the focal axis (T_{rz}), settling time (T_{sz}) and waiting time (T_{wz}), i.e. $T_{mz} = (T_{rz} + T_{sz} + T_{wz})$. Finally, the scan time for multi-focal plane virtual slide is shown in Eq. 2.2.

$$T^{MFP}_{FOVScan} = \left(\max(T_m + T_e, T_c) + (N_f - 1)\max(T_{mz} + T_e, T_c)\right) N_{FOV} \tag{2.2}$$

Line Scan: In a line scan system the microscope stage moves exactly one (projected) pixel width

between each exposure of the line scan camera. Therefore, the data acquisition time is defined by the total number of lines required divided by the scan rate of the image sensor measured in lines per second (lps). The time for stage to move between two image strips ($T_{sr} = T_m \approx 200$ms) is also considered. It is assumed that the stage moves at a constant speed from the start to the end of the strip while following the focus profile map, so the T_m and T_{mz} are not considered. For scans of multiple focal planes, multiple single focal plane scans are performed sequentially. The scan time for a single focal plane (SFP) and multiple focal planes (MFP) are given as below, where N_{ppr} is the number of pixels per image strip and N_r represents the number of image strips (N_{ppr} could be different for different strips). The N_r is found by counting the number of rows of FOVs in FOV scan, because the image strips are having the same overlap area.

$$T^{SFP}_{LineScan} = \frac{N_{ppr}N_r}{fps} + (N_r - 1)T_{sr} \tag{2.3}$$

$$T^{MFP}_{LineScan} = T^{SFP}_{LineScan}N_f \tag{2.4}$$

Slanted Scan: The total scan time of a slanted specimen system is estimated by Eq. 2.5. This equation is the same as Eq. 2.3 as the slanted scan imaging sensor is equivalent to multiple line scans simultaneously acquiring images at multiple depths of the specimens. Therefore, the total number of image lines acquired in each focal plane is the same as from a single line scan in a single focal plane, neglecting the geometry deformation of the pixels in the scan plane due to the slant angle (that is small given $cos(3°) \approx 0.9986$).

$$T^{MFP}_{SlantedScan} = \frac{N_{ppr}N_r}{fps} + (N_r - 1)T_{sr} \tag{2.5}$$

2.4.2 Results

An FOV scan system needs to acquire over 67,260 FOVs in order to produce one digital slide of a Thin-prep cervical cytology specimen at resolution of 0.2μm/pixel with 20 focal planes. The monochrome version of the raw data size for the FOV scan is around 282.11GB (8 bit grayscale). However, the line scan and slanted scan system need only acquire 223.67GB of raw data for the same slide, due to reduced necessity of overlaps between acquired images, a reduction of 21%.

Next, we select a number of area scan and line scan cameras available on the market and estimate the scan time and data throughput rate of the system using the equations proposed in Section. 2.4.1, as if they were attached to the Olympus system. The selected cameras have the same or smaller pixel size providing at least the same spatial resolution as the SPOT camera. We also assume the Olympus stage has sufficient speed and precision to accomplish the selected line scan and slanted scan acquisitions.

Table. 2.2 outlines the selected cameras and their key specifications as well as the scan times and average system data throughput rates (Hamamatsu 2016b; Turnkey 2016). It is assumed that the ORCA camera runs with the same exposure time as that of SPOT camera used in Olympus system to simplify the simulation. For line scan cameras, the maximum possible exposure times for achieving the maximum frame rate were used, assuming they have negligible data download times. Fig. 2.7

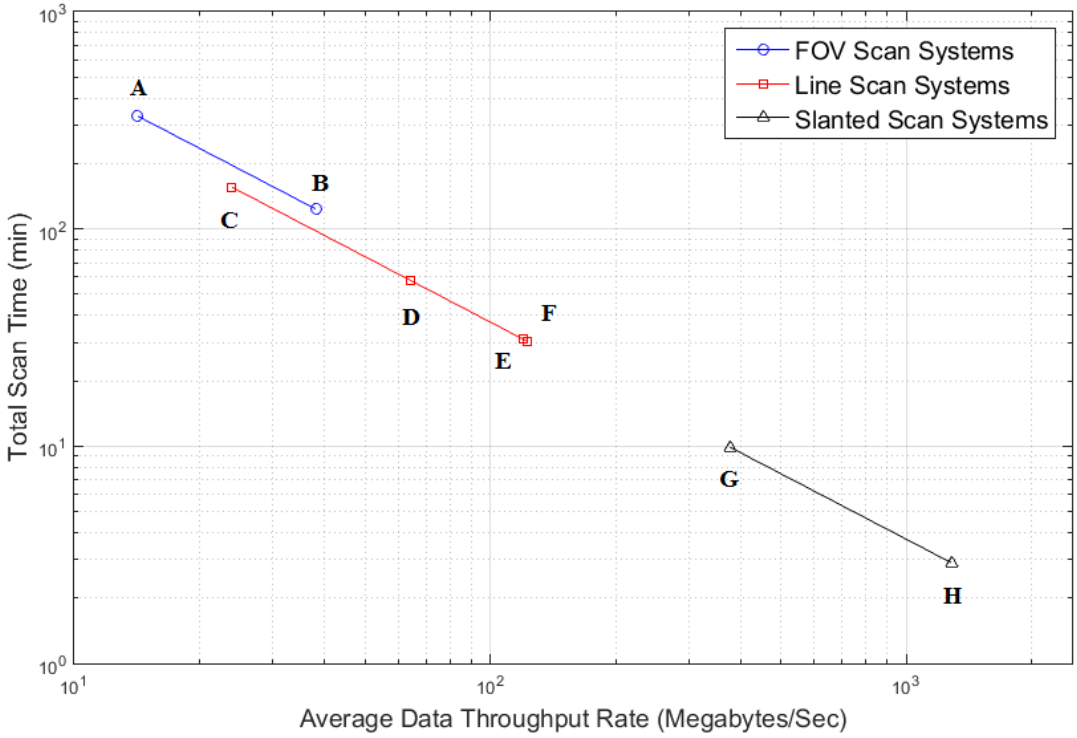


Figure 2.7: Average system throughput rate versus scan time.

Pt.	System configuration	Scan Mode	Camera Speed (fps or lps)	Exposure Time	Max Camera Throughput Rate(MB/s)	Scan Time (s)	Average System Throughput Rate (MB/s)
A	SPOT Insight T_c = 295ms	FOV	3.39	5ms	14.22	19,842	14.22
B	ORCA-Flash4.0 V2	FOV	100	5ms	419.43	7,399	38.13
C	Piranha 2 P2-2X-06K40	Line	12K	83.3μs	24.58	9,361	23.89
D	Piranha ES-8k-34 kHz	Line	34K	29.41μs	69.63	3,472	64.42
E	Piranha ES-8k-68 kHz	Line	68K	14.71μs	139.26	1,866	119.86
F	Piranha4-8k-70 kHz	Line	70K	14.29μs	143.36	1,820	122.89
G	Custom camera*	Slanted	12K	83.3μs	491.52	468	377.87
H	Custom camera*	Slanted	34K	29.41μs	1,392.64	174	1,288.37

*only 20 lines read out.

Table 2.2: Simulated scan times and data rates for various scan modes and systems.

shows comparisons of scan times and data throughput rates of the simulated systems (A-H) in Table. 2.2 as a graph.

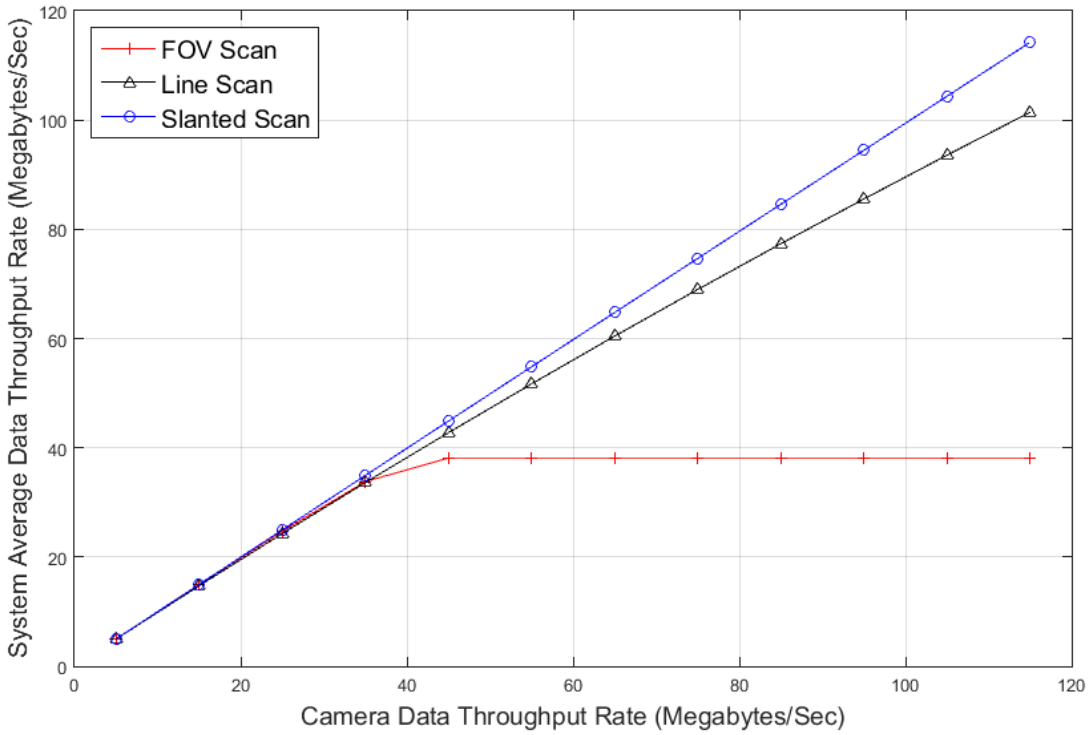


Figure 2.8: Camera throughput Vs System throughput.

The results suggest that the existing Olympus system performing an FOV scan takes the longest time among all systems at around 19,842 seconds (Pt.A), or close to 6 hours, which is close to the actual scan times in early experiments (Bradley et al. 2005). Obviously, this is too long to be clinically practical. When a faster area scan camera is used, the scan time is reduced to 7,399 second (Pt.B), but the increased camera throughput rate is not reflected by the increase in system throughput rate. The relationship between camera throughput rate and final system throughput rate is demonstrated in Fig. 2.8. The graph shows that the system throughput rate of FOV scan systems hits a limit at 38.13megabytes (MB)/s despite increasing camera throughput. This insists that the maximum scan time of the current Olympus system configuration is around 7399s (Pt.B), or around 2 hours, which is impractical.

However, the line scan and slanted scan systems are able to break the 38.13MB/s throughput limit, and their system throughput rates keep pace with the increasing camera data rates. Specifically, all but one of the line scan systems have shorter scan times than the fastest FOV scan system (Pt.B). The 12K line scan system (Pt.C) has a faster frame rate, but the scan time is longer than the fastest FOV scan system due to slower system throughput rate. The fastest selected line scan camera is the 'Piranha4-8k-70kHz', which has a readout clock of 70K frame per second (fps), showing the fastest line scan system that acquires raw data in 1,820s, or 30 minutes (Pt.F). But, the 70K lps camera would have maximum of 83.3μs for each exposure, under such a short period of time, the system either requires an expensive light source or reduced signal to noise ratio hence reduced image quality.

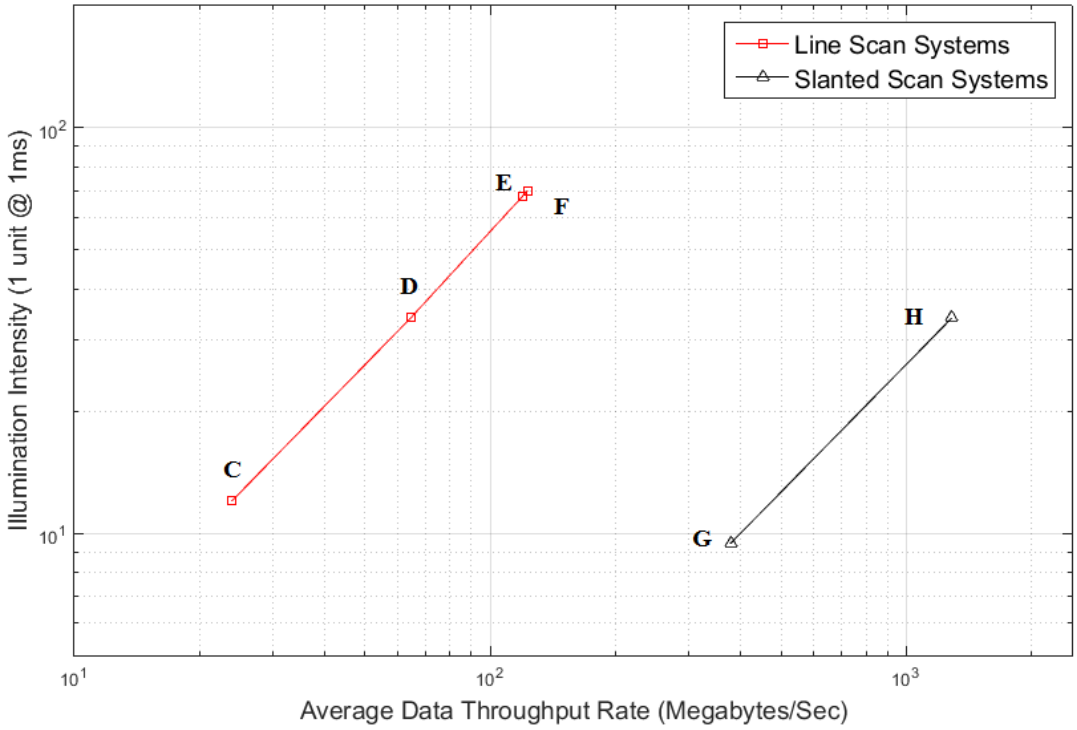


Figure 2.9: Camera throughput Vs Illumination intensity.

The camera used in slanted scan systems can either be a combination of several line scan cameras, or an area scan sensor with only a subset of lines read out (Bamford and Mayer 2009). In this way, the frame rate can be increased to close to that of a line scan camera. The two slanted scan systems are assumed to have two area imaging sensors, that only read out 20 evenly spaced lines at the same frame rates to those in 12K lps and 34K lps line scan system. Comparing 12K fps slanted scan system (Pt.G) and 70K lps line scan system (Pt.H), the slanted scan system has nearly a six times longer exposure time. Thus, to obtain the same image quality, system F needs a six times brighter illumination source regardless other image quality effects. Fig. 2.9 illustrates the estimated illumination intensities required for the line scan and slanted scan systems. Given the 12K lps line scan system (Pt.C) and the 12K fps slanted system (Pt.G) run at the same clock speed, they will to produce images of the same quality under the same illumination intensity. However, the slanted scan system has a 20 times higher data throughput rate than the line scan system. The fastest line scan system (Pt.F), on the other hand, requires an increase of more than 6 times (72/12) in light intensity as compared to the 12K fps slanted scan system (Pt.G). But this only achieves a 26.4% increase in the data throughput rate compared to the line scan system (Pt.G). The 34K fps slanted scan system has the shortest estimated scan time among all systems, which is able to digitise a Thin-prep cervical cytology specimen slide with multi-focal planes and high resolution in around 3 minutes. Thus, the slanted scan technique has the potential to achieve the fastest scan time among the three slide scan techniques under the constrains of exposure time and mechanical system limits. This is due to a system design that combines fast readout speed and acquisition of multiple focal planes in a single scan.

Despite being the fastest scan system, the slanted scan system requires the highest precision hard-

Stage Accuracy	FOV scan	Line Scan	Slanted Scan
Primary direction	20μm	0.2μm	0.2μm
Secondary direction	20μm	20μm	20μm
Focusing direction	1μm	1μm	0.2μm

Note: Assumes focal depth of 1μm and a pixel resolution of 0.2μm.

Table 2.3: Microscope stage accuracy requirements.

ware. Table. 2.3 shows the minimum requirements for microscope stage positional accuracy. However, the primary direction of movement for the FOV scan and secondary moving direction for all other techniques requires the least positional accuracy as these images can be aligned digitally. The primary direction of movement for the line and slanted scan technique require pixel-resolution accuracy due to the synchronous movement of the microscope stage and the camera. The slanted scan technique requires pixel resolution accuracy in the focus direction in order to correctly track the specimen slant angle.

The maximum camera data throughputs shown on Table. 2.2 should also be used for selection of inter-connectors between the camera and the storage system. For instance, a Universal Serial Bus (USB) 2.0 port (maximum transfer speed 60MB/s) and early generation FireWire (IEE 1394 400, maximum transfer speed 49MB/s) are sufficient for the Olympus system and the slow 12K lps line scan system (Pt.C). But, the 100 fps area scan camera, 12K fps and 34K fps slanted scan systems have much a higher data throughput rate, therefore they require faster data connectors, such as newer generation USB (625MB/s for 3.0 protocal and 1250MB/s for 3.1 protocal) or Thunderbolt connectors (more than 1250MB/s) (Network 2016; Tips 2016).

Rather than turning the illumination light source on constantly in the conventional FOV scan, pulsed light can be alternatively used to reduce the motion blur. In this way, the stage is able to move continuously if the light pulses are strong and short enough to ensure camera exposure while minimising the motion blur. Consequently, the scan speed is no longer limited by the stage mechanics only but also the camera frame rates and the intensity of the light source, which can be considerably faster than the conventional FOV scan method. For instance, the 3DHISTECH Pannoramic 250 Flash, one of the fastest commercial systems adopting the pulsed FOV scan method, is three times faster (2,460s vs 7,399s) than the fastest FOV scan system (Pt.B in Table. 2.2) proposed in our analysis. But, this scan speed is still considerably slower than some line scan systems and all estimated slanted scan systems, because the line scan cameras can run at greater data throughput rates.

It also worth mentioning that the shape of scan maps used for estimating scan times of Thin-prep slides is circular rather than the square scan map used in our preliminary study and it is closer to the actual deposition area of the specimen (Fan et al. 2013). As a result, the amount of data acquired and the scan times presented here are less than those calculated in the preliminary study (Fan et al. 2013), for example, a smaller data file of 83.3GB (via the FOV scan) and a shorter scan time of 49 seconds for the slanted scan system (Pt.H), highlighting the importance of generating an accurate scan map before the high resolution scan, especially for scanning cytology specimens in “glass-faithful quality”.

2.5 Summary

In this chapter, techniques for the acquisition of digital slides of cervical cytology specimens in "glass-faithful quality" for WSI are presented and analysed. Specifically, it is shown that initially generating accurate scan maps of the specimen has great potential to minimise the size of scan areas and the amount of data to be scanned and ultimately the slide scan time. Scan times can be further reduced by estimating the focus profile maps of the specimen before the high resolution scan, which is also important for slide scanners to produce in-focus images and to ensure the final digital slide is of high quality. Moreover, the actual thickness of cytology specimens, in this study the cervical cytology specimen, should be evaluated in order to determine the minimal number of focal planes and data to be acquired and consequently the minimal scan time. This vital parameter is currently determined empirically in this chapter as well as in previous studies. Furthermore, the thickness of the cervical cytology specimen will determine the appropriate angle the slide needs to be slanted if the slanted scan method is utilised.

The equations for estimating scan times of three existing scan methods, namely FOV scan, line scan and slanted scan, have been proposed to quantitatively evaluate the performance of these methods in the acquisition of "glass-faithful quality" digital slides from cytology specimens. The simulation results indicated that the throughput of the conventional FOV scan technique (no pulsed illumination) is limited by the settling time of the mechanical stage, while the line scan and slanted scan techniques overcome this limitation due to the fact that the stages move continuously. The simulation results are in accordance with the current trend among commercial slide scanners where faster scanners are using the line scan method as shown in Table. 1.1. However, the line scan method is limited by the camera exposure time because the faster frame rates result in shorter exposure times that could lead to higher image acquisition noise and poor image quality. The slanted scan has the potential to achieve faster system throughput rates and shorter scan times than the line scan systems within the constraints of camera frame rate and illumination. Assuming a slanted scan is scanning 20 focal planes and running at the same frame rate as the NanoZoomer-XR system mentioned in Chapter 1, the slanted scan has potential to acquire a "glass-faithful quality" cervical cytology specimen in around 1 minute, which is significantly faster than the existing systems. But, the slide slant induces optical aberrations, which should be corrected so that the digital slides acquired with the scan method are in "glass-faithful quality".

2.6 Aims of the Thesis

The objectives proposed in Chapter 1 are consolidated into following aims:

- **Aim 1:** To develop methods for improving the scan efficiency of cytology specimens.
  - **Aim 1.1:** To develop methods to find accurate 3D positions of cells in cytology specimens; The initial step of digitising a glass slide is to generate a 3D map of the specimen on the slide. An accurate scan map of the specimen can substantially reduce the amount of data to be acquired by the slide scanner particularly when multiple focal planes are

required, and hence the total acquisition time, while an accurate focus profile map of the specimen ensures that all the specimen images are sharp and clear. Conventional methods for specimen delineation and selection of focus candidates used for estimating the focus profile map are independent and not robust in relation to slide artefacts like dust, so a robust method needs to be developed to generate accurate scan map and good quality focus candidates. The proposed method needs to be carefully validated with representative glass slides that contain artefacts.

  - **Aim 1.2:** To evaluate the thickness of the cytology specimen.
    Once the scan map and the focus profile map of the specimen is generated, the number of focal planes to be scanned is determined depending on the thickness of the specimen. As the existing scans empirically estimate the thickness of the specimen, a quantitative method is required to calculate the accurate number of focal planes to be used, because the amount of data and time spend on each single focal plane is considerable especially when scanning cytology specimens to achieve a "glass-faithful quality" digital result. The proposed method only needs to determine the thickness of cell clumps because the full height of the clump should be acquired. The effectiveness of the method will be evaluated as well as a demonstration of use of the method in terms of potentially improving the scan speed.

- **Aim 2:** To investigate and correct the optical aberrations due to the slide slant in the fast slanted scan method.
  Although the slanted scan is shown to have the potential to achieve fast acquisition speed, the images acquired via the method must be of high quality and free of aberrations in order for the slanted scan method to be considered as an alternative to the conventional flat scan methods. The preliminary study of the scan method in this chapter found that the slide slant induces optical aberrations that cause adverse artefacts and degrade image resolution. Therefore, aberration correction methods need to be proposed, which should be effective, convenient and evaluated with real specimens against using conventional scan methods.

3

A TILE BASED METHOD FOR SPECIMEN DELINEATION AND FOCUS CANDIDATES SELECTION

This chapter presents the work to attain **Aim 1.1**, which is to develop a robust method to find the accurate spatial location of the cytology specimen on the glass slide. To this end, a tile based specimen delineation method was developed that not only provides an accurate scan map but also provides a ranked list of FOV as focus candidates needed for estimating the focus profile map. The method first divides the low resolution image of the slide into tiles representing FOVs at high resolution. These tiles are then evaluated on whether they belong to the foreground and how good are they as focus candidates. The method is used as the first step to digitalising the slide followed by the estimation of the focus profile map of the specimen, the high resolution scan and the generation of the final digital slide. The proposed method is applicable to not only cytology specimens but also to histology specimens.

The development of this novel method has been described in a conference article accepted by the IEEE International Symposium on Biomedical Imaging (ISBI), 2014, and a journal article accepted by *Micron*. The conference paper included preliminary work, where the framework of the method was proposed and two tile-evaluating metrics were evaluated on LBC cervical cytology cells. The journal article, which is presented in this chapter, extended the number of evaluating metrics to four and included evaluations of additional specimen types including fine needle aspiration prepared cytology slides and histology slides.

Related publication details

Fan, Y., Gal, Y. and Bradley, A.P., 2014, April. Microscopic specimen delineation using auto-phase correlation index. In Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on (pp. 1336-1339). IEEE.

Manuscript revision history	
Submitted to the *ISBI* conference	9 August 2013
Accepted	27 November 2013
Presented (traditional poster) in the ISBI conference	2 May 2014

Fan, Y., Gal, Y. and Bradley, A.P., 2014. An algorithm for microscopic specimen delineation and focus candidate selection. Micron, 66, pp.51-62.

Manuscript revision history	
Submitted to *Micron*	16 December 2013
Interim decision	
Revision submitted	6 April 2014
Accepted	22 May 2014
Published/Available online	1 June 2014

3.1 Introduction

The acquisition time of high-resolution microscope slide images is normally prohibitively long, because the large number (many thousands) of FOV (long strips required in line scan) and the requirement to maintain focus over the whole specimen so that the final digital slide image is clear and sharp (Bradley et al. 2005; Rojo et al. 2006). To reduce acquisition time, it is typical to initially delineate the specimen from a low-resolution scan to minimise acquisition of blank FOVs, and then to construct a *focus map*, using a small subset of FOVs, that the system can follow rather than focusing on each high-resolution image (Altinay and Bradley 2011; Rojo et al. 2006). The most straightforward approach to find the region of interest (specimen) from the background (blank areas) is to use an intensity thresholding followed by spatial smoothing (Altinay and Bradley 2011). This *scan map* can then be used to guide the acquisition of high-resolution FOVs over the specimen (Altinay and Bradley 2011; Rojo et al. 2006). Once the scan map is constructed, a focus map is formed by interpolating the focal positions measured from a limited number of FOVs within the specimen's scan map (Gilbertson et al. 2006; Hilsenstein 2005; Lahrmann et al. 2013). Typically, these focus candidates are either randomly selected or follow a predefined pattern (Hilsenstein 2005).

This has led to the development of *robust* methods that reject outliers when interpolating the focus map. In (Hilsenstein 2005), outlier focus candidates are removed by smoothing the topgraphic of the focus map. A more sophisticated method is to image image every focus candidate at high resolution and find whether they contain real cells or contaimnations, such as methods developed in (Lahrmann et al. 2013; Lopez et al. 2013), where classfication methods were developed to distinguish cells from foreign objects based on image features of pre-classified images. In this way, one focus candidates containg cells are used for focus map interpolation to ensure the quality of it.In the absence of such techniques the interpolated focus map has the potential to be biased towards the artefacts (say) on top of the slide and hence the images acquired become de-focused and appear blurred. This problem becomes worse as the magnification increases, because the DOF of the objective lens reduces. For example, at 40$\times$ magnification the DOF is around one micron (Castleman 1996), while dust artefacts are around two hundreds of microns away from the specimen, depending the thickness of the coverslip. Another solution is to acquire a large Z-stack of images at more focal planes so that cells are guaranteed to be included, but this further increase scan times (Bradley et al. 2005). Therefore, in

this chapter we propose a combined scan and focus map algorithm that aims to be more effective and efficient by not including these outliers in either the scan or focus map.

Compared to the methods developed by (Lahrmann et al. 2013; Lopez et al. 2013), the algorithm proposed in this chapter acquires and analyses a low-resolution image of the entire microscope slide to produce both a specimen scan map and an ordered list of focus candidates. The tile-based method locates potential (i.e. yet to be acquired) high-resolution FOV images (tiles) of the specimen in the low-resolution image, so that they can evaluated to determinate whether they contain cellur objects, whether they are in focus, whether contain contaminations such as dusts, even possibly the amount of cells in the FOV. In this way, these tiles can be classfied for whether they are included in the scan map and be ranked to guide the selection of focus candidates. Therefore, the tile-based allows the generation of scan maps and focus candidates simuoutenously.

In this thesis, we investigated and compared four different tile evaluation metrics measuring distinct property of an image for the proposed tile-based specimen delineation method. These metrics are: threshold index (TI) that measures the percentage of foreground to background pixels in each tile; normalised auto-correlation index (NACI) that measures spatial image similarity within each tile; auto-phase correlation index (APCI) that measures the phase diversity of each tile; and entropy index (EI) that measures the predictability of image intensities within each tile. Among them, the APCI is a novel metric proposed in this thesis research. The four metrics are evaluated on forty pathology slides including: 16 Thin-prep cervical cytology and 9 breast fine-needle aspiration (FNA) specimens, both PAP stained, as well as 15 HE stained histology neuroanatomy slides.

This chapter is structured as follows: Section. 3.2 gives details the tile-based specimen delineation method and the four tile evaluation metrics; Section. 3.3 describes the slides analysed and the experimental methodology; Section 3.4 illustrates the results for both slide delineation and the evaluation of focus candidates; Section. 3.5 discusses the results and compares the advantages and disadvantages of each metric.

3.2 Methodology

The flow chart of the tile-based specimen delineation method is shown in Fig. 3.1. The tile-based specimen delineation method includes five steps, namely, low-resolution acquisition of the whole slide image, cover-slip boundary detection, tile evaluation, scan map segmentation and focus candidate selection.

3.2.1 Low-resolution Scan

The algorithm starts with image acquisition of the whole slide under low magnification. The initial low-resolution scan is relatively fast either from a stand-alone macro camera, used in most commercially available whole slide scanners, or from a low magnification objective of motorised microscope (Rojo et al. 2006). The low-resolution FOV images are then stitched together to form a single image of the whole slide (Altinay and Bradley 2011).

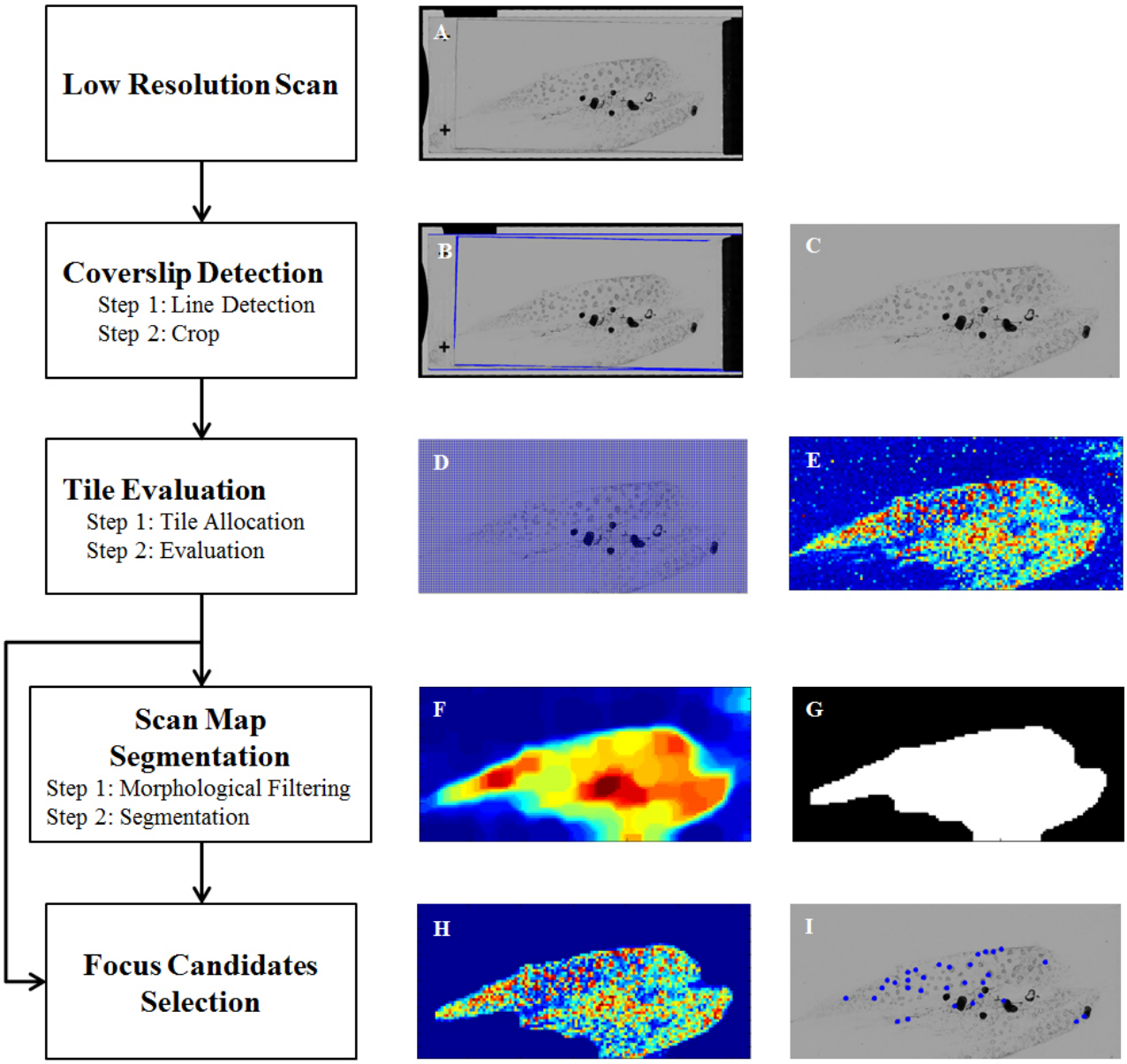


Figure 3.1: Flow chart (left) and example images of the tile-based specimen delineation method. A: Low-resolution image of a slide; B: Detected lines (green); C: Cropped slide image without the cover-slip; D: Tile allocation on the low-resolution slide(each blue box represents a high-resolution tile); E: Tile matrix (showing the APCI metric); F: Tile matrix after morphological filtering; G: Scan map (white is foreground); H: Ranked tiles across the scan map; I: Positions of 30 selected focus candidates (blue points).

3.2.2 Cover-slip Boundary Detection

Normally, a cover-slip is glued on the top of the glass microscope slide to protect the specimen and an identification label is put on the side. On the low-resolution slide image, the edges of the cover-slip and the glass slide normally have clear and strong boundaries allowing relatively easy detection. Therefore, the cover-slip boundary is detected and only the image area *inside* the cover-slip is processed at the next step. This not only reduces the effects of artefacts such as cover-slip edges, glue and labels, but also reduces the number of tiles to be analysed. First, we apply a Sobel edge detector (Gonzalez and Woods 2008) to the low-resolution image. Next, we apply a Hough transform (Gonzalez and

Woods 2008) to the edge image and select the strongest peaks from the Hough matrix. The locations of lines corresponding to these peaks are found which, because these are the strongest lines in the image, relate to the locations of the edges of the slide and cover-slip, e.g., as shown in Fig. 3.1.B. Finally, we remove the image areas outside of these lines and process the image representing the enclosed area (under the cover-slip).

3.2.3 Tile Evaluation

We first locate the tiles to be processed in the low-resolution image produced from above step. The size of these tiles depends on the difference in magnification between the initially acquired low-resolution image and the high-resolution image yet to be acquired. The tiles are positioned adjacent to each other so that they form a mosaic of the whole low-resolution image, such as shown in Fig.3.1.D. The positions of each tile are determinate based on the initial position of the first tile and any overlap (for stitching) between tiles (Altinay and Bradley 2011).

After tile allocation, all of the tiles are evaluated individually by the proposed metrics, and their values are placed into a *tile matrix* based on their position in the image, an example is shown in Fig. 3.1.F. The tile matrix does not contain the actual positions of the tiles, rather they can be found based on the geometry and physical position of the slide.

The following sub-sections describe the four tile evaluation metrics studied in this paper, namely thresholding index (TI), normalised auto-correlation index (NACI), auto-phase correlation index (APCI) and entropy index (EI). Among the equations below, we let I represent a low-resolution image tile, which has dimensions of M by N, where normally $M = N$, $I(m, n)$ represents a pixel intensity at position (m, n) of the image I, where $0 \leq m \leq M$ and $0 \leq n \leq N$. A number of cytology images, their corresponding values given by the metrics and processing time are shown in Fig. 3.2 and Table 3.3. In Table 3.3, the threshold index of images B, C and D is calculated using the threshold value obtained over the whole low-resolution image of the specimen, e.g., A in Fig. 3.2.

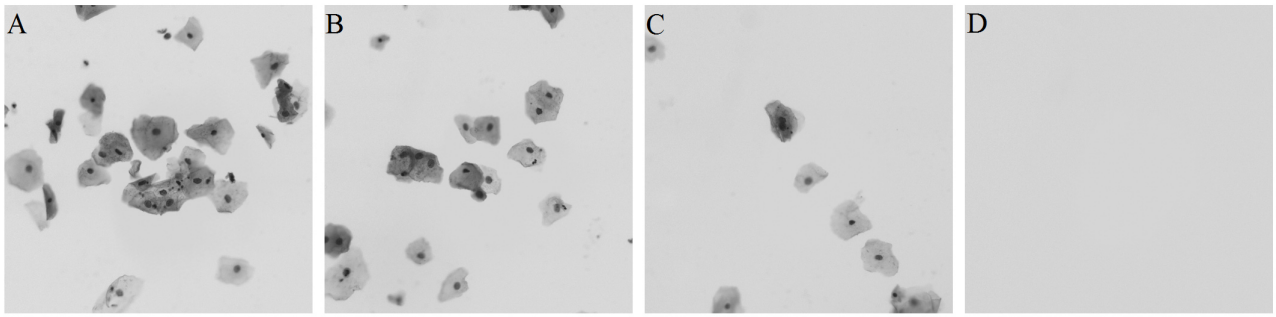


Figure 3.2: Sample cytology images at magnification of $40\times$.

Thresholding Index (TI)

The most straight forward way to tell whether a tile is empty or contains cellular objects is by measuring the image intensity. This is effectively the conventional approach to specimen delineation, which assumes that darker pixels relate to specimen and brighter pixels to background. Therefore, we initially apply the classical Otsu's thresholding method (Otsu 1975) to the low-resolution image

	A	B	C	D	Average Time (s)
TI	0.1385	0.0737	0.0289	0	0.0025
NACI	0.9991	0.9984	0.9964	0.2082	0.4328
APCI	0.7569	0.6792	0.5300	0.1028	0.0362
EI	0.8190	0.8020	0.7450	0.6543	0.0031

Table 3.3: Evaluation of images in Fig. 3.2 and average computation time over 1000 512x512 images.

obtained from the previous step of the algorithm 3.2.2. This produces a binary image that, to a first approximation, separates foreground from background. Next we calculate the threshold index of each tile as the percentage of foreground pixels in the tile,

$$TI = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (I(m,n) > t) \tag{3.1}$$

Where t is a global threshold value found by Otsu's method in order to distinguish specimen from background. Due to its connection with the traditional specimen delineation methods, we treat the TI as the baseline for comparison.

Normalised Auto-Correlation Index (NACI)

It is well known that cross correlation is able to find the similarity between two signals and to co-register two images that have some degree of overlap (Lewis 1995). However, the method fails when the two images contain objects with strong intensities or an uneven illumination gradient, because these high intensity areas result in high correlation measurements as their signal power is strong (Lewis 1995). Therefore, normalised cross-correlation is typically implemented as it is more robust (Lewis 1995). We also know that the auto-correlation is able to highlight repeated or periodic signals. Thus, if we correlate two images of the same scene but with different background noise, the auto-correlation index will estimate the power of the signal with respect to the power of the (random) underlying noise (Gonzalez and Woods 2008). Here, we acquire two low-resolution images I_1 and I_2 of each slide with the same exposure and illumination settings, but with different realisations of the background noise. Therefore, NACI can be used to directly compare and measure the amount of specimen in each tile. As we only require the complete correlation of two tile images, NACI is as follows:

$$NACI = \frac{1}{MN} \left(\frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (I_1(m,n) - \mu_1)(I_2(m,n) - \mu_2)}{\sigma_1 \sigma_2 + \epsilon} \right) \tag{3.2}$$

Where μ_1, μ_2 are the mean value and σ_1, σ_2 are the standard deviation of the image I_1, I_1, and ϵ is a small constant to prevent division by zero.

As shown in Fig. 3.2 and Table. 3.3, the more cellular objects in an image, the higher NACI. In contrast, if the tile is empty, such as image D in Fig. 3.2, a low NACI value is obtained.

Auto-Phase Correlation Index (APCI)

As an alternative to cross-correlation, phase correlation can be used for robust registration of im-

ages (Kuglin and Hines 1975). Unlike cross-correlation, phase correlation considers the phase of the images only, effectively equalising image spectral magnitudes (Kuglin and Hines 1975). The principle of APCI can be explained in the same way as NACI. The two realisations of each image are correlated in the phase domain and APCI reflects the degree of coherent phase of the specimen in the tile compared to the incoherent phase of the background noise. For example, the cytology images in Fig. 3.2 contain more cells and have more edges (of nuclei and cytoplasm) and hence are assigned higher APCI values.

Unlike NACI, APCI is conveniently calculated in the phase domain by normalising the power spectrum of the two images (Fan et al. 2014b):

$$APCI = \frac{1}{MN}\left|\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left(\frac{\hat{I}_1(m,n)\hat{I}_2(m,n)^*}{|\hat{I}_1(m,n)\hat{I}_2(m,n)^*|}\right)\right| \tag{3.3}$$

Where $\hat{I}_1$ and $\hat{I}_2$ are Fourier transforms of the image I_1 and I_2, and * indicates complex conjugate. To minimise boundary effects we apply a Tukey window (Gonzalez and Woods 2008) with a radius of 95% of M and N.

Entropy Index (EI)

The entropy of an grayscale image calculates the probability of occurrence for each intensity level to reflect how unpredictable or random an image is, as shown in Eq. 3.4. In this way, an blank tile or FOV has a low entropy value because the intensity is highly predicable, while a tile containing cellular objects has larger entropy as it is more unpredictable. All three metrics proposed above have an operating range of $\{0, 1\}$, and so we normalise entropy (3.4) in the same way, as shown in (3.5). As shown in Fig. 3.2 and Table. 3.3, EI successfully ranks the four images based on the number of cells, although the empty tile is given a value close to those with specimen.

$$Entropy = -\sum_i (p_i \log(p_i)) \tag{3.4}$$

$$EI = 1 - \frac{1}{Entropy + 1} \tag{3.5}$$

Where p_i is the probability of the intensity value i in image I. Here, image I has eight bits of precision and so has 256 distinct grey levels.

3.2.4 Scan Map Segmentation

The tile matrix from the tile evaluation process above is processed with morphological filters and then the filtered tile matrix is segmented to produce the final scan map.

Morphological Filtering

The morphological filters are used to increase the sensitivity of specimen delineation at the cost of reduced specificity. In practice, an ideal scan map should consist of tiles that encompass the complete specimen, with the minimum amount of background and other outliers. In this study, we assume the

slides are prepared so that the specimens appear in one or two main areas, either as a tissue slice or a contiguous deposition (smear) of cells. Therefore, the morphological filter may remove isolated tiles containing isolated cells or small tissue fragments. However, we also believe that the scan map should be efficient for the high-resolution scan. For example, for a system using a line scan sensor, it is faster and easier for the sensor to scan over a small area of background, which means the scan map should not contain small holes. To meet these goals, we apply a number of grey scale morphological operators (image opening and image closing with a disk structure of size 3 followed by hole closing) to the tile matrix (Gonzalez and Woods 2008). The combination of grey scale morphology followed by a binary threshold has been shown to be preferable to the initial application of a threshold followed by binary morphology (Bradley and Stentiford 2003).

Segmentation

After the application of the morphological filters, a threshold value needs to be found to distinguish the tiles that contain foreground (specimen) from tiles that contain primarily background. The classical Otsu thresholding method is an effective way of automatically selecting a threshold value by analysing the distribution of tile index values. The method assumes that each object distribution follows a Gaussian (normal) distribution. In this way, if the intensities of foreground and background distributions overlap, the Otsu's method picks a threshold value around the centre of the overlap (Gonzalez and Woods 2008). However, in our application, the foreground (specimen) distribution is both more important and more variable than the background, as a user would prefer to acquire the entire specimen and tolerate a degree of background. Therefore, we implemented a modified version of the Otsu method to bias the threshold away from the specimen distribution in order to increase sensitivity.

First, the classical Otsu's method is applied to the tile matrix and an initial threshold (k_1) that separates the two intensity classes (foreground and background) is found. We assume a bimodal distribution because we want the metric to detect (foreground) specimen and additional classes may represent artefacts. Next, we find the mean (μ) and standard deviation (σ) of the foreground class, which is defined to have higher index. We place a *biased* threshold (k_2) two standard deviations below the mean of the foreground distribution (k_2=μ-2σ). Assuming a Normal distribution, this assigns 97.5% of the foreground distribution to the foreground class. The two standard deviation distance is obtained heuristically using a number of different slide types and metrics with the aim of maximising sensitivity. The final threshold (k) is then taken as the minimum of k_1 and k_2. Finally, the scan map is obtained using the threshold k as shown in Fig. 3.1.G.

3.2.5 Focus candidate selection

The tile matrix produced from Section 3.2.3 also provides a ranking of all tile indexes with regard to the extent of objects in the tile. Obviously, in the Fig. 3.2, the image tile A is a good focus candidate and should be selected with priority as a focus candidate, while the tile D is a poor focus candidate and should not be selected.

Here we utilised a simple greedy search method to select a subset of focus candidates based on their tile index value and spatial location. First, we ranked all tiles in the tile matrix in descending

order. Next, we selected the highest valued tile as the first focus candidate, and iterated down through the ranking. A tile is selected as a focus candidate if the tile belongs to the scan map (obtained from Section 3.2.4) and if the position of the tile is a sufficient distance away from a previously selected focus candidate. This process is repeated until the required number of focus candidate is found or the list of tiles is exhausted.

3.3 Experimental Methodology

3.3.1 Experimental Slides

The experimental slides include sixteen cervical cytology slides, prepared using the Thin-prep monolayer slide preparation technique with PAP stain, nine slides consisting of PAP stained smears from FNA for breast cancer diagnosis, and fifteen histology slides of HE stained neuroanatomy specimens. To examine the robustness of the proposed method, we purposely selected slides with a range of conditions that are seen in routine clinical practice, including a varying degree of stain intensity, age and cell density. Particularly, some old slides are contaminated by dust and glue artefacts, which would be digitized for archive and research. Some slides are from post-diagnosis and so contain ink markers from previous manual diagnosis (as may be scanned for quality assurance purposes). All slides are standard microscope slides with dimensions of 75 by 25mm, thickness of around 1mm and with cover-slips around 40 by 20mm. The sample slides are summarised in Table. 3.4.

	Thin-prep	FNA	Histology	Total
With Artefact	10	1	5	16
Without Artefact	6	8	10	24
Total	16	9	15	40

Table 3.4: The number of slides in each specimen type and condition.

3.3.2 Digital Slide Acquisition System

A fully functional digital slide acquisition platform is used in this study. The system consists of a motorised stage, an Olympus BX40 microscope with $4\times$ and $40\times$ objective lens, and a four megapixel SPOT Insight camera with a pixel size of 7.4 µm and 100% fill factor. The maximum resolution produced by the system is 0.185 µm/pixel under the $40\times$ objective.

3.3.3 Scan Map Ground truth

On the Thin-prep slides, the cellular objects are approximately evenly spaced inside a circular area, which has a measured diameter of around 20mm. Therefore, a circle of diameter of 5400 pixels (≈20,000/3.7 @ 3.70 µm/pixel image resolution) was manually placed to represent the specimen ground truth. The FNA and Histology slides are arbitrarily shaped and so a simple linear polygon was constructed around the cellular objects of interest as the ground truth. This was done manually

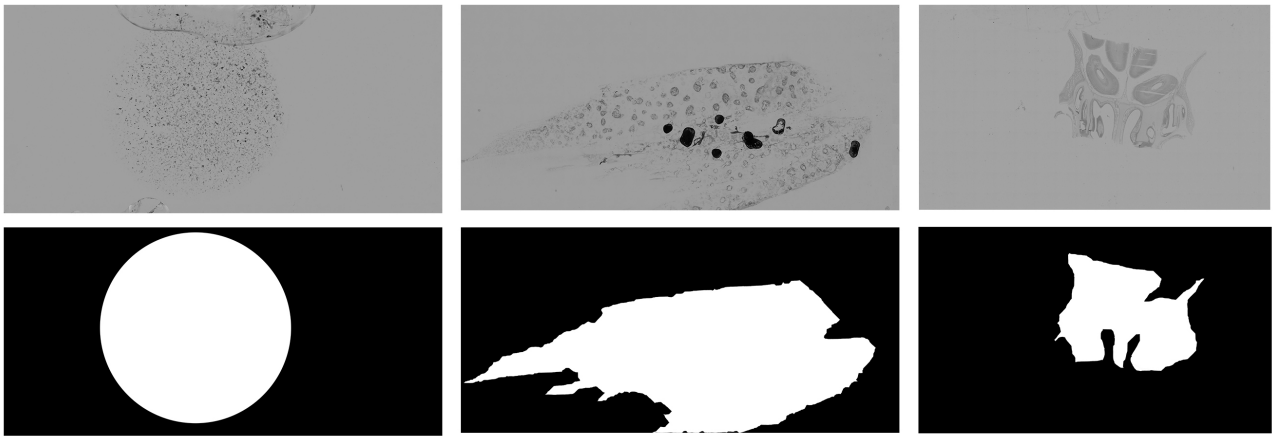


Figure 3.3: Two low-resolution images of a Thin-prep slide (top left), a FNA slide (top middle) and a Histology slide (top right) and their ground truth at bottom (white is foreground).

using the Photoshop software (Adobe Systems Inc., San Jose, CA, USA). Examples of the scan map ground truth can be seen in Fig. 3.3.

3.3.4 Experimental Settings

This section gives details of specific parameter settings for the proposed method.

Low-resolution Scan

All glass slides were imaged using the $4\times$ objective so that the low-resolution FOV images had a spatial resolution of 3.70 µm/pixel (as $2\times$ pixel binning was used on the camera which effectively gives a pixel size of 14.8 µm). Before acquisition, each slide was manually focused and the focal axis of the stage was fixed during the scan. This is a plausible approach at low-resolution as the focal depth under the $4\times$ objective is larger than the thickness of our specimens (around 20 µm) and smaller than the thickness of a slide.

In addition, for the calculations of NACI and APCI we require two images for each tile. Therefore, during acquisition of low-resolution images, we acquired two images at each FOV with the same exposure time and the illumination. As the number of images and the image exposure time are much smaller than for high-resolution scan, the increase in scan time due to the acquisition of the second image is insignificant. For example, with our system, each whole slide needs around 200 FOVs to be acquired and each exposure takes around 0.2 seconds, so the total time due to the second acquisition is around 40 seconds, which is acceptable and could be further reduced by utilising faster cameras. Two images of the slide were then produced: both of them were used for calculating the NACI and APCI, and the average image was used for calculating TI and EI.

Cover-slip Boundary Detection

To reduce processing time, we searched for lines only in the horizontal and vertical directions with a five degree tolerance and a step size of 0.2 degree. We down-sampled the original image by a factor of eight to further reduce processing time without sacrificing accuracy. The number of peaks examined in Hough matrix was 50, while the minimum line length was 225 pixels, which is around one third of the length of the short edges of the cover-slip in the down-sampled image. After all lines meeting

these criteria were found, we extended them to intersect with the edge of the image. The four most inward lines closest to the centre of the image then define the cover-slip. With the above parameters, the cover-slip of all experimental slides (40) were successfully identified. The above parameters were selected heuristically for the image resolution and the cover-slip sizes used in this study.

Tile Allocation

At high magnification, each high-resolution FOV captures a spatial area of 379 µm by 379 µm, so the size of the tile projected on to the low-resolution image is 103 by 103 pixels (≈379/3.7). For simplicity, we assume that no overlap is required between two high-resolution FOVs and hence the tiles. We put the first tile at the top left corner of the low-resolution image. For calculation of the APCI and NACI metrics, we tile the second slide image in the same way as the first image.

Morphological Filtering

The sizes of the structuring elements are first guided by the composition of the specimen such as cell density and gaps between tissue parts. This information is then transferred to the final sizes based on the image resolution, size of the FOV and overlap between tiles. Finally these parameters are fine-tuned with experiments of sample slides to maximise the segmentation performance.

First, we smooth the tile matrix via a circular averaging filter with size of three tiles because the values of tiles vary greatly even they all represent foreground, some may form distinct groups, which makes classification challenging. Next, we applied an image opening on the tile matrix with a disk structuring element (assuming the specimen is isotropic) of size three (tiles) in order to remove isolated and small-valued tiles that are less likely to be specimen. Then, we applied an image closing on the tile matrix with a disk structuring element of size five (tiles) to expand the potential specimen area and fill small holes. The same set of morphological filters (proposed in 3.2.4) were applied to all four tile evaluation metrics. Though the above filters were initially tuned to the Thin-prep slides, they all showed good performance on processing the FNA and histology slides.

3.3.5 Scan Map Segmentation Evaluation

We measure the Area Under of each Receiver operating characteristic (ROC) curve (AUC) to estimate the ranking performance of the four metrics (Bradley 1997). As area under of receiver operating characteristic curve (AUC) measures the probability of correct ranking a foreground pixel higher than a background pixel it is independent of the threshold chosen to delineate (segment) the specimen. For all generated scan maps, we calculate the dice similarity index (DSI) (Gonzalez and Woods 2008) which compares the obtained scan maps with the ground truth described in Section. 3.3.3 using the formula below:

$$DSI = \frac{2|A \cap B|}{|A| + |B|} \tag{3.6}$$

where A and B are two binary vectors of ground truth and generated scan maps, with 1 to be the specimen and 0 to be the background.

3.3.6 Focus Map Evaluation

In order to evaluate the quality of the selected focus candidates, we attempt to construct the focus map in the same way as in real digital slide acquisition. That is by interpolating the focus map across the specimen area using a subset of selected focus candidates. Then, the focus maps are quantitatively compared with the ground truth as the number of focus candidates used for focus map interpolation is increased. This approach highlights the effectiveness of each metric at selecting focus candidates, i.e. whether it produce a better ranking, and so fewer candidates are required to build an accurate focal map.

Focus Map Ground-truth Generation

We first re-locate each tile from the tile allocation procedure at high magnification ($40\times$) and run a gready focus searching algorithm to find the focal depths of all FOVs. To ensure these FOVs match exactly the tiles found from processing low resolution slide images, the mechnical stage for the microscope needs to have high positional accuracy, i.e., close to the pixel resolution of low resolution images. The focus metric used was normalised variance and for each tile, and the fast hill-climbing search method for focusing, which samples a range of focal depths with progressively smaller focal steps till a maximum focus value is found (Sun et al. 2004; Tello-Mijares et al. 2013). Next, we manually classify inlier FOVs that had cervical cells and outlier FOVs that contained glue, ink marker artefacts or were focused on the top of the cover-slip. The ground truth of the focus map excludes all these outlier tiles and includes all inlier tiles containing specimen.

Focus Map Construction

After we obtain the raw focal positions and focus map ground truth for each slide, we interpolate the focus map using the selected focus candidates and its focus information, and compare it with the ground truth using mean square error (MSE). Specifically, we used a third order polynomial surface fitting in both directions for generating the focus map, while minimising the effect of any remaining outliers with the robust regression option enabled.

The optimal distance between two focus candidates depends on the topographic variation of the specimen's focal plane. For instance, in general, the histology slides are of a more uniform depth than the cytology slides, hence it requires a smaller number of data points for accurate interpolation. In this study, we set the minimum distance to be three FOVs, which may be a sub-optimal distance, but is sufficiently small for all three slide-types. A smaller distance will result in an increase in the number of focus candidates sampled, which may not be desirable in practice.

We also implement a benchmark focus candidate selection algorithm based on a random selection of tiles (i.e. ignoring the tile evaluation metric) for comparison. Specifically, we randomly select a tile as a candidate and feed it to the greedy search process rather than using an ordered tile list. The ground truth scan map of the slide is used as the constrained map for the focus candidates in this case. We run the random selection of focus candidates 200 times and present the averaged results.

Since acquisition of the full focus map and focus map interpolation are time consuming (may take more than 30 hours), we selected eight slides and sampled a rectangle area of the slide containing a combination of the specimen and artefacts.

3.4 Results

3.4.1 Scan Map Segmentation

Post-processing	Slide Type	Area Under ROC Curve			
		TI	NACI	APCI	EI
Before	Thin-prep	0.818	**0.903**	0.832	0.836
	FNA	0.875	**0.955**	0.919	0.950
	Histology	**0.993**	0.989	0.985	0.991
	Average	0.895	**0.949**	0.912	0.926
After	Thin-prep	0.958	**0.970**	0.963	0.947
	FNA	0.955	0.983	0.978	**0.990**
	Histology	**0.997**	0.996	0.996	0.996
	Average	0.970	**0.983**	0.979	0.978
Before	Artefacts	0.849	**0.910**	0.854	0.861
	No Artefacts	0.904	**0.951**	0.894	0.908
After	Artefacts	0.942	**0.961**	0.943	0.928
	No Artefacts	0.970	**0.993**	0.977	0.977

Table 3.5: Area under the ROC curve before and after post-processing for the four tile evaluation metrics. The highest value among the four is shown in **bold**.

The AUC of the four metrics are shown in Table 3.5. Averaged over all slide types NACI achieved the highest AUC both before and after post-processing. Despite TI showing good performance when processing histology slides, the metric has the lowest AUC on average. Overall, all four of metrics achieved good overall performance with an average AUC of 0.97 after post-processing. When looking at the AUC based on the existence of artefacts, NACI again achieves the best results, both with/without artefacts and before/after post-processing. The performance of all four metrics are affected by the presences of artefacts. In particular, there was a 3-4% improvement in AUC when processing slides without artefacts.

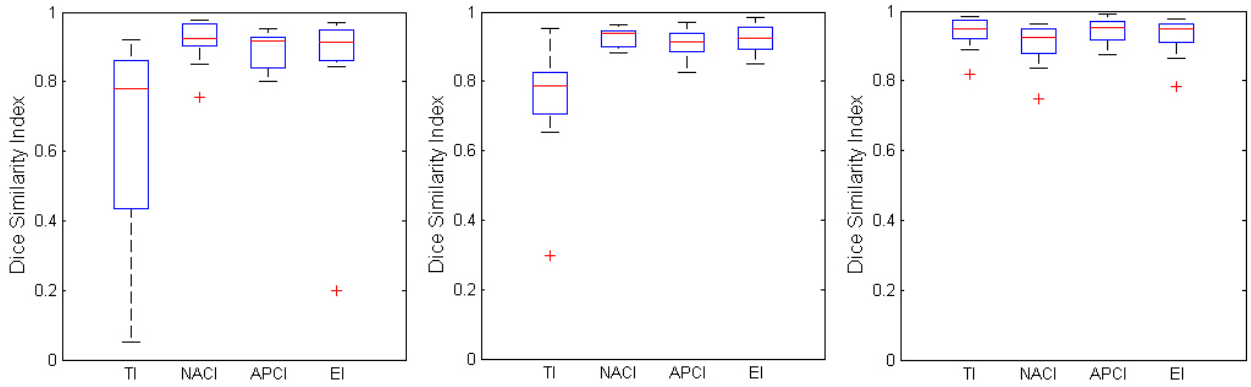


Figure 3.4: Dice similarity index of Thin-prep slides (left), FNA slides (middle) and Histology slides.

The evaluation results of the scan map segmentation obtained after the tile matrices were thresholded are illustrated in Fig. 3.4 and Fig. 3.5. The TI metric achieved the poorest segmentation results for Thin-prep and FNA specimens, as shown with the lowest median DSI and largest interquartile

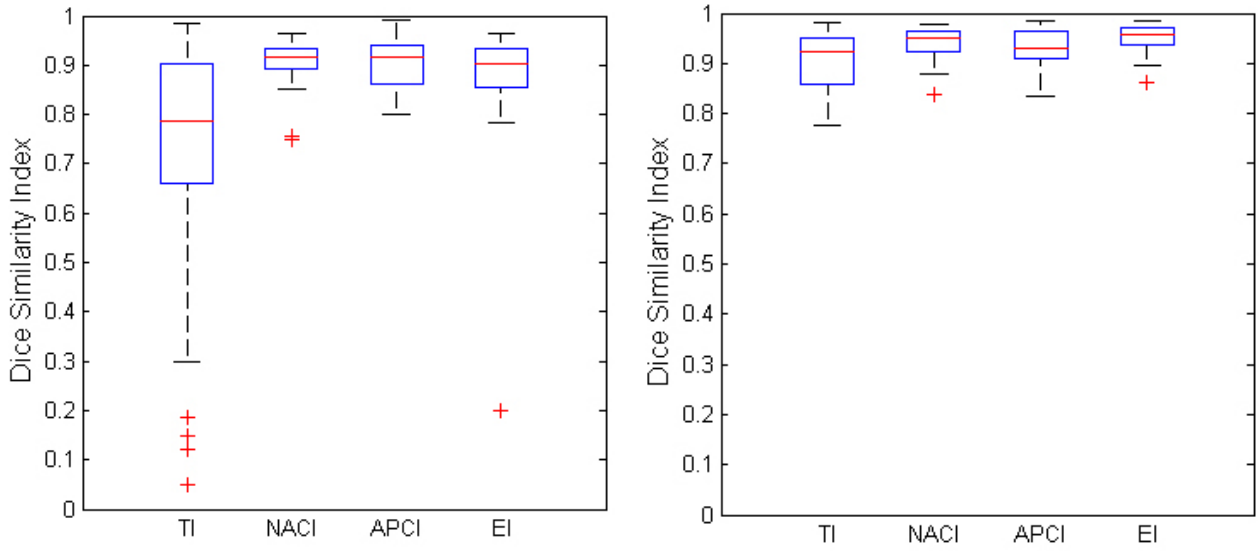


Figure 3.5: Dice similarity index of slides with artefacts (left) and without artefacts.

range (IQR). However, the same metric achieved a competitive DSI (0.9502) when processing the Histology specimens, which is considerably better than that of Thin-prep (0.7782) and FNA (0.7854) specimens. In contrast, the scan maps generated using the other three metrics showed consistent high quality regardless of specimen type. Although the three metrics (APCI, NACI and EI) achieved similar DSI scores, APCI received the lowest False Positive Rate (FPR): 3.94% (APCI), 6.51% (NACI) and 6.02% (EI) for the Thin-prep slides; 0.97% (APCI), 3.21% (NACI) and 1.17% (EI) for the FNA slides; and 1.40% (APCI), 3.58% (NACI) and 2.16% (EI) for the Histology slides.

When all slides are grouped as to the presence of slide artefacts, TI produces the worst performance among the four metrics. Specifically, its performance is affected significantly by the existence of artefacts, producing a decline in median DSI (from 0.9224 and an IQR 0.0043 without artefacts to a DSI of 0.7854 and IQR 0.0904 with artefacts). However, the other three metrics are less affected by the presence of artefacts, with their corresponding DSI values being maintained above 0.90.

False Negative Rate (%)				
	APCI	NACI	EI	TI
Original	5.44	2.33	5.37	13.00
Modified	4.14	1.54	3.32	10.37
Dice Similarrity Index				
	APCI	NACI	EI	TI
Original	0.9045	0.9139	0.8590	0.7123
Modified	0.9155	0.9170	0.9024	0.7790

Table 3.6: Segmentation results of the original and modified Otsu threshold.

The modified Otsu method also proved effective. Averaged across all experimental slides, the FNR decreased consistently across all metrics while DSI also increased, this is shown in Table. 3.6.

Regarding the computational complexity of the four metrics, we estimated average computation time of the tile evaluation process for all forty slides. All processing was performed in the MATLAB

environment using a single Intel i7-3770 3.4GHz processor with 8 GB of RAM. The NACI takes considerably longer than the of other three metrics, at 95.86 seconds compared with 14.35 seconds for APCI, 0.98 seconds for EI and 0.43 seconds for TI. This result is consistent with the complexity of each metric as defined in Section 3.2.3.

3.4.2 Focus Map Evaluation

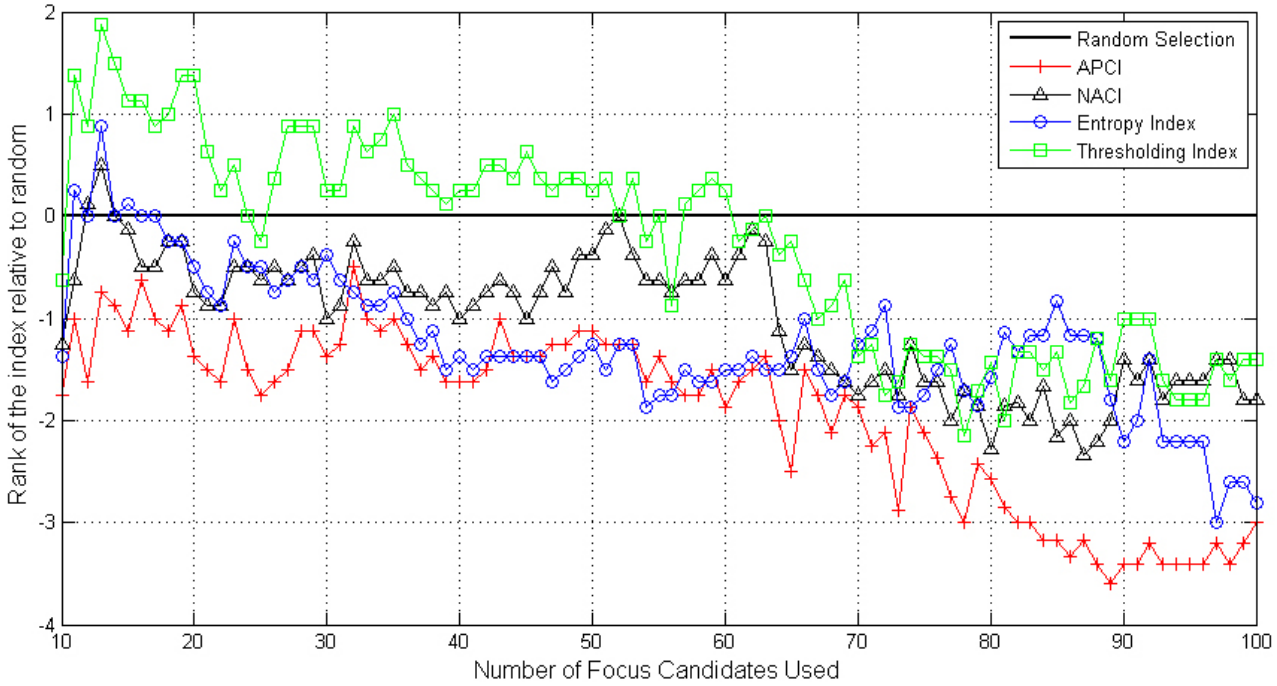


Figure 3.6: Average ranking of MSE from the focus map evaluation of eight slides. Note: ranking is relative to random selection (rank - 0) and so a negative rank means the MSE is better than that of random selection.

In order to summarise the focus map evaluation results, we rank the MSE of the focus maps generated from the same number of focus candidates for each slide against random selection as the baseline. As shown in Fig. 1- 8 (in Appendix), the MSE varies greatly between slides and metric and so averaging the MSE may bias the results between different slides/metrics. Therefore, we present a ranking that more directly compares the relative performance of selected focus candidates across each of the slides. The averaged ranks from the eight slides are shown in Fig. 3.6. Here, a negative rank means that the MSE of the focus map produced by a metric is smaller than that produced by random selection. Remarkably, only APCI achieves average rankings that are consistently better than random selection. Furthermore, according to the Sign Test (Dixon and Mood 1946), APCI is better than the three other metrics (NACI, EI and TI) at a very high level of significance ($p<0.005$). The other three metrics cross the random selection benchmark (rank 0) initially when the number of focus candidate being used is small (<15), but the rankings of all four metrics reduce as more focus candidates are used for the interpolation. In particular, TI requires more than 50 focus candidates before its performance is better than random selection.

Type	Artefacts	Sufficient No. Focus Candidates					Sufficient MSE
		Random	APCI	NACI	EI	TI	
FNA	No	52	45	**24**	26	53	12.9
FNA	Yes	-	**53**	68	86	-	19.0
FNA	Yes	-	**88**	117	123	191	5.4
FNA	Yes	113	**67**	126	96	105	95.6
Histology	No	38	56	**15**	66	66	2.0
Histology	No	30	**16**	47	31	45	3.4
Thin-prep	Yes	-	**18**	-	-	-	4.6
Thin-prep	Yes	-	**80**	-	125	-	24.6

Table 3.7: SNFC for each metric and slide. Note: a '-' means that the focus map did not achieve a sufficient MSE, the smallest number of sufficient focus candidates is shown in **bold**.

The ranking shown in Fig. 3.6, summarises the overall performance of the four metrics when the same number of focus candidates is used for focus map interpolation. Table 3.7 shows the SNFC and the corresponding sufficient MSE. Here, SNFC is broadly equivalent to settling or convergence time, where SNFC is defined to be the number of focus candidates required to achieve at least 50% of the globally minimum MSE observed from all four metrics, plus random selection, on each slide. Therefore, SNFC shows information of how quickly a metric converges to with 50% of the minimum MSE that can be expected on each slide.

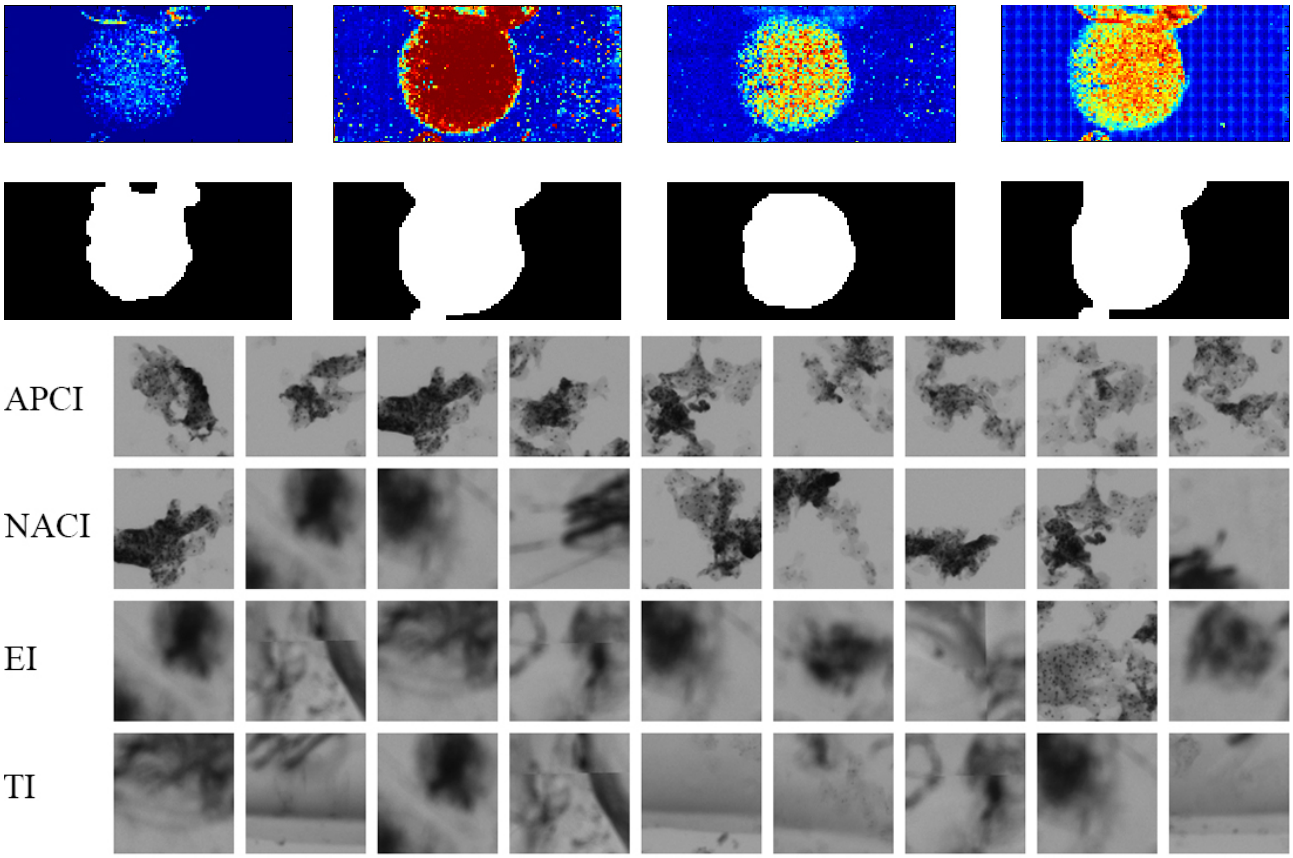
Table 3.7 shows that overall APCI achieves the lowest SNFC on six out of the eight slides, which means that on average the APCI focal map MSE converges most rapidly to the sufficient MSE. Moreover, among four focus-candidate selection metrics only APCI converge to the sufficient MSE on all eight slides. In contrast, random selection achieved the sufficient MSE on only or four out of the eight slides. The intensity based TI metric performed slightly better than random selection, but still failed to converge on three of the slides.

Viewed from the presence of slide artefacts, APCI again showed dominant performance in the presence of artefacts, with fastest convergence in all five artefact-affected slides. For the three slides without artefacts, NACI produced the minimum SNFC on two slides, while APCI produced one minimum. In terms of the artefact type, the two Thin-prep slides contain glue and dirt artefacts, and APCI showed a clear dominance. Particularly for the first Thin-prep slide in Table 3.7 (also illustrated in Fig. 3.7), the MSE curve of APCI is considerably below (smaller) all of the other metrics at all points (number of focus candidates) and the focus maps from all of the other metrics never reach an acceptable quality. The Random selection and TI were the worst performers in processing these two slides, both of them failed to converge. On the other hand, the FNA slides contain ink marker artefacts. Here APCI also showed superior effectiveness, producing the smallest SNFC for all three of these slides.

3.5 Discussion

3.5.1 Scan Map Segmentation

The results in Fig. 3.4 and 3.5 demonstrate that APCI, NACI and EI are all capable of correctly discriminating specimen from background. Despite the experiment evaluating a modest number of slides (40), these were chosen to be representative of three common specimen types containing commonly observed artefacts. Further, the scan maps obtained using APCI achieved the lowest average FPR, which implies that these scan maps contain fewer artefacts. Overall, the results show that the tile based delineation method is effective and that APCI is capable of distinguishing between (glue and ink related) artefacts and real specimen.

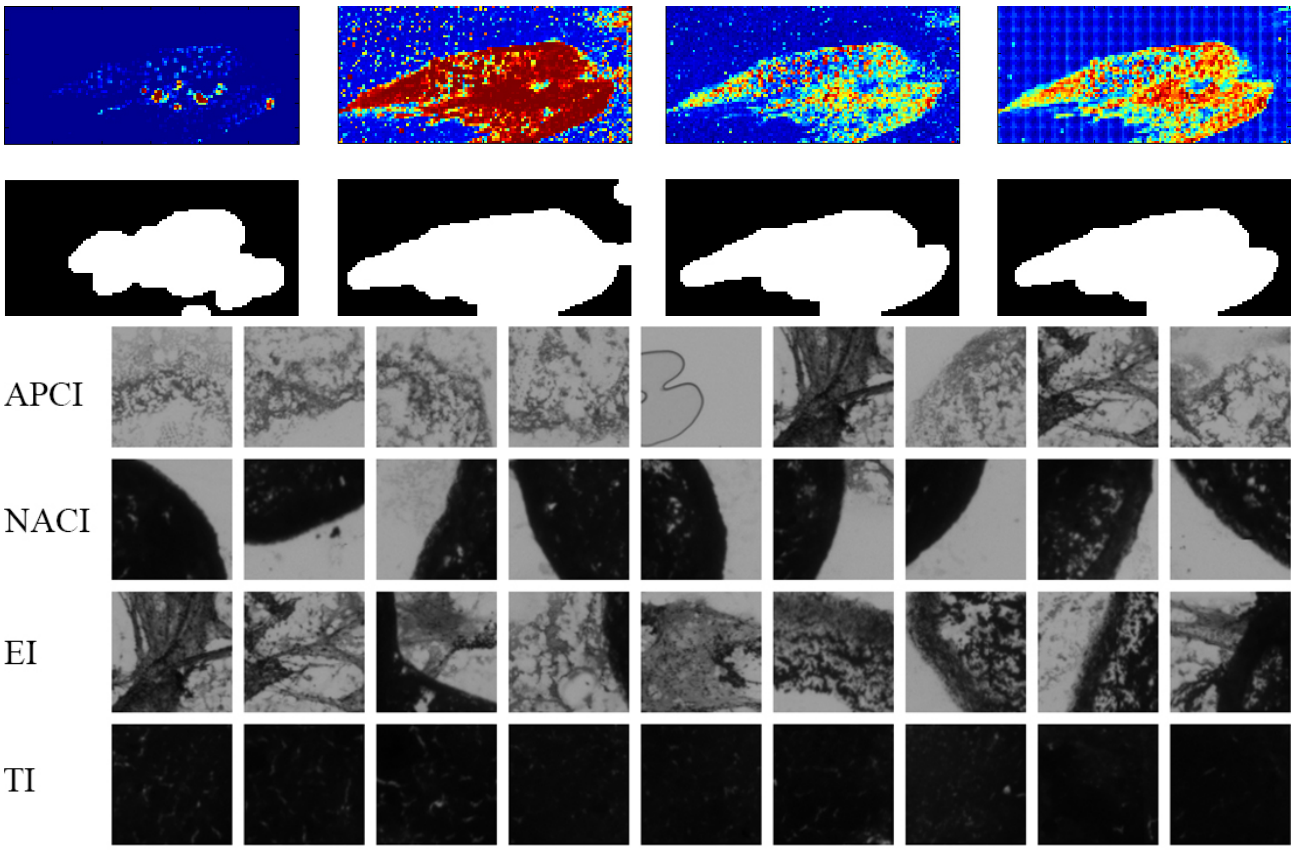


Ranks	1	2	3	4	5	6	7	8	9
APCI	0.823	0.818	0.818	0.817	0.815	0.806	0.801	0.794	0.792
NACI	0.818	**0.219**	**0.183**	**0.219**	0.758	0.634	0.710	0.815	**0.204**
EI	**0.219**	**0.351**	**0.190**	**0.256**	**0.183**	**0.217**	**0.270**	0.787	**0.239**
TI	**0.190**	**0.201**	**0.219**	**0.351**	**0.142**	**0.363**	**0.256**	**0.183**	**0.175**

Figure 3.7: Top: Tile matrix of the four metrics for a Thin-prep slide (top row) and the corresponding scan maps (bottom row). From left to right: TI, NACI, APCI and EI. Each pixel in the maps represents a single tile, red colour indicates high values (close to 1), while the dark blue indicate small values (close to 0). Middle and Bottom: The nine top ranked tiles by each of the four metrics, and in the table the values of these tiles evaluated by APCI. The APCI values of (blurred) outlier tiles are shown in **bold**.

To further interpret these results we first examine a Thin-prep slide that contains glue artefacts, the tile matrix and scan map of which are shown in Fig. 3.7. Compared with the low-resolution image of the slide shown in Fig. 3.3(left), it can be seen that the tiles contaminated by glue artefacts (mainly upper centre part of the slide) have a lower APCI value than those containing specimen (position of the specimen refers to the ground truth of the slide in Fig. 3.3). In contrast, the other three metrics, NACI, EI and TI all had some contaminated tiles highlighted, and hence these outlier tiles are included in the final scan map. To verify this observation, the nine top ranked tiles from each of the four metrics were found, as shown in Fig. 3.7. It is remarkable that the top ranked tiles by APCI all contain real specimen, while the top ranked tiles by TI all contain (out of focus) glue and dirt artefacts. By this measure EI is the worst performer as it assigned eight out of nine outlier tiles the highest index, while the NACI miss-identified four tiles. However, when all outlier tiles were evaluated by the APCI metrics, they all obtained small (<0.351 compared to 0.7-0.8) values. Clearly, APCI is robust to glue artefacts because these objects are on the surface of the cover-slip and so are out of focus in the low-resolution scan. The glue deposition on top of the cover-slip is at least 0.16mm (the thickness of the cover-slip) away from the specimen, which is larger than the focal depth of the objective lens (4x, around 100 µm) used for low-resolution image acquisition. Thus, the glue and other contaiminations at top and around it appear defocused, and hence produce a reduced phase diversity and a lower APCI value. On the other hand, specimen is correctly focused at image acquisition: the nuclei and cytoplasm of individual cells are clear and sharp, and so tiles containing these objects have a high APCI as shown in Fig. 3.7. Similarly, the objects in the specimen focal plane, but directly below the glue also appear defocused, because the transparent glue causes light to be diffracted and alters the pathway of the illumination. The light of objects under the glue will focus on different points and so the objects appear defocused and blurred, consequently, a smaller APCI value is obtained. However, the defocused dirt glue still have low opacities and these tiles give high TI values as seen in Fig. 3.7. The EI also assigned these tiles high values, because the light of the object will be unevenly distributed to more pixels when the object is defocused, and so the randomness of the pixel intensities increases. The NACI, on the other hand, is sensitive to objects with either strong edges or high intensity, which often yield tiles with high contrast.

We next examine an FNA slide with ink marker artefacts. The tile matrix, scan map and nine top ranked tiles by the four metrics are shown in Fig. 3.8, the low-resolution image and ground truth can be found in Fig. 3.3 (centre). Again, it can be seen that all of the top ranked tiles by the APCI metric contain specimen, while all tiles with high values of TI are fully occupied by the ink makers, which have a TI index of 1. Therefore, in the TI tile matrix, the tiles associated with markers are clearly visible and they have a strong influence on the scan map which has the largest FNR (34.62% compared with 4.67% (APCI), 1.22% (NACI) and 2.15% (EI)). The top ranked tiles by NACI contain edges from the ink markers, as these have the strongest contrast and so the highest correlation. The tiles assigned high values by APCI and EI all have significant edges, but EI favours edges with high contrast, which miss-identifies four tiles containing part of an ink marker (tiles ranked 3, 4, 7 and 8). It worth noting that some outlier tiles are also assigned high APCI values, i.e. the tile ranked 8 by NACI (0.637) and tiles ranked 3, 4, 7, 8 by EI (0.766, 0.785, 0.803 and 0.755 respectively), however



Ranks	1	2	3	4	5	6	7	8	9
APCI	0.847	0.847	0.847	0.846	0.843	0.842	0.841	0.840	0.840
NACI	**0.414**	**0.358**	**0.557**	**0.488**	**0.509**	**0.423**	**0.290**	**0.637**	**0.502**
EI	0.842	0.840	**0.766**	**0.785**	0.818	0.770	**0.803**	**0.755**	0.820
TI	**0.526**	**0.385**	**0.476**	**0.285**	**0.319**	**0.271**	**0.419**	**0.362**	**0.257**

Figure 3.8: Top: Tile matrix of the four metrics on an FNA slide (top row) and the corresponding scan maps (bottom row). From left to right: TI, NACI, APCI and EI. Each pixel in the map represents a tile, red colour indicates high values (close to 1), while the dark blue indicate small values (close to 0). Middle and Bottom: The nine top ranked tiles by each of the four metrics, and in the table the values of these tiles evaluated by APCI. The APCI values of (blurred) outlier tiles are shown in **bold**.

these are not among the top ranked tiles by APCI. Interestingly, the top nine ranked tiles by TI have APCI values not equal to zero. This is because the intensities in these tiles are not uniformly zero and thus contain small phase variations.

The segmentation results in Fig. 3.5 show that NACI is able to distinguish tiles containing both artefacts and specimens from background, but the above two examples illustrate that it cannot reliably distinguish specimen from artefact. We can also observe from the example tile matrices above, that NACI highlights tiles that contain dirt, glue, isolated cells and dust. This implies that the correlation due to either specimen or artefact is both much stronger than the background noise and residual illumination gradient. Figues 3.7 and 3.8 show that TI is not robust to glue and ink-marker artefacts, particularly in segmentation of cytology specimens. However, Fig. 3.4 shows that TI is much more effective at segmenting histology slides, producing results that are more competitive with other met-

rics. The first reason for this is that the histology specimen contain tissue samples that have higher and more uniform cell density than that of the cytology specimens. Secondly, the example slides contain only minor glue artefacts without the significant artefacts seen in Thin-prep and FNA slides. Therefore, TI is only suitable for tile-based delineation of Histology slides without any artefacts.

3.5.2 Focus Map Evaluation

The ranking results presented in Fig. 3.6 show that the MSE of focus maps produced using candidates selected via the four metrics eventually surpass that of random selection. This demonstrates that the selection of focus candidates based on information extracted from each individual tile is an effective approach. However, TI is clearly the least effective as it produced results worse than random selection even when a considerable numbers of focus candidate were used (around 50). As Fig. 3.7 and 3.8 illustrate, the tiles most highly ranked by TI are primarily outliers, which must subsequently be detected and handled via a robust interpolation scheme (as was done here (Hilsenstein 2005)). However, when the number of outliers is large, they dominate and so no longer appear to be outliers and so will have a significant effect on the focal depth estimated from the focal map. When this is the case even robust regression methods are likely to fail. The other three metrics exhibit performance that is superior to random selection once a reasonable number of focus candidates have been selected (around 15), but only APCI was significantly better than random selection over the complete range of focus candidates used (10 to 100).

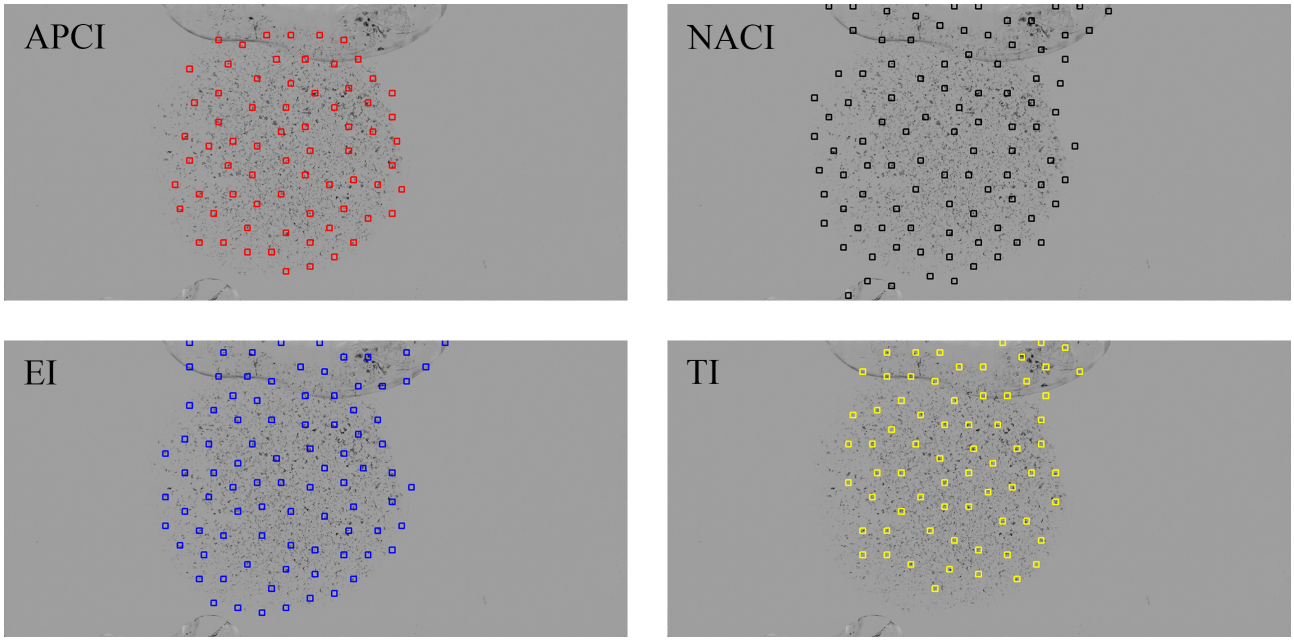


Figure 3.9: The APCI has the minimum number of focus candidates that locate on top of the glue artefacts. Notice: The minimum distance between two focus candidates was set to five tiles here for better illustration.

Table 3.7 and Fig. 3.6 indicate that not only are the focus maps constructed from the focus candidates selected by APCI consistently better (lower MSE) than the other three metrics (TI, NACI and EI), but that APCI requires a smaller number of focus candidates to construct an acceptable focus map.

The reasons for this are twofold: Firstly, the scan maps generated using APCI contain less artefacts. For instance, the APCI scan map of the Thin-prep example, shown in Fig. 3.7, contains the minimum number of artefacts and consequently far fewer "bad" focus candidates, as shown in Fig. 3.9. In contrast, the scan maps produced by NACI and EI cover the whole artefact area, and so contain more outlier focus candidates contaminated by glue artefacts. Secondly, the artefact affected tiles included in the scan maps, such as the ink markers in the FNA slide in Fig. 3.8, are given much lower rankings by APCI and so are far less likely to be selected as focus candidates. In fact, as Fig. 3.9 shows, the highest ranking tiles from APCI are almost exclusively "good" focus candidates that contain primarily specimen. Therefore, based on these results we can conclude that APCI is the best metric considered here for the selection of focus candidates due to its ability to rank in-focus specimen ahead of (out of focus) artefacts.

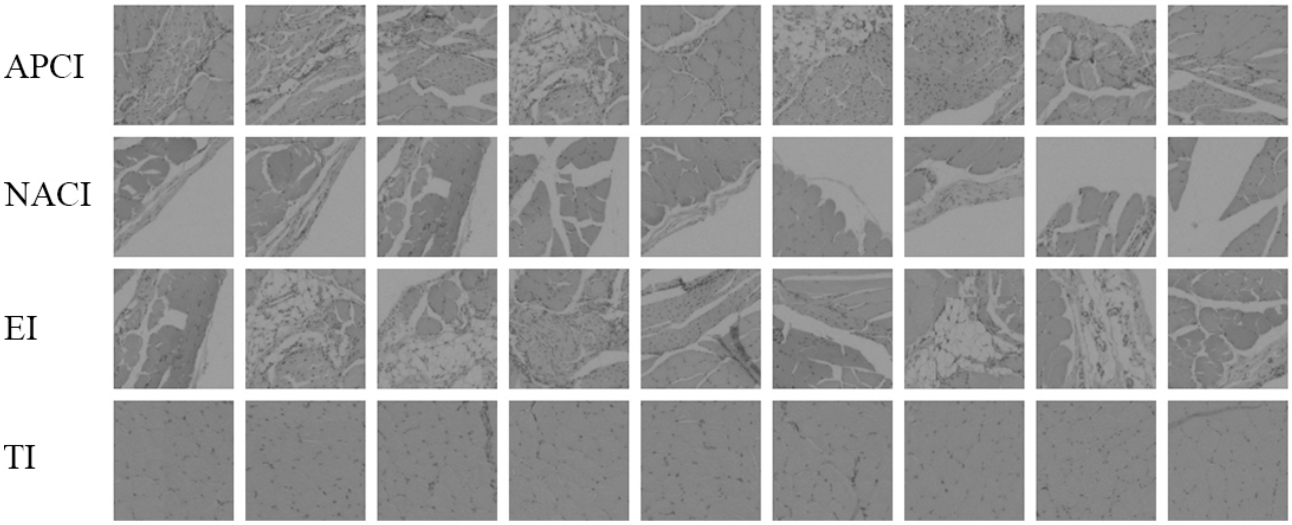


Figure 3.10: Nine top ranked tiles by four metrics from a Histology slide.

On processing histology specimens having few slide artefacts, all four metrics had similar performance in generation of scan maps. For example in Fig. 3.10 that shows the top ranked tiles from a histology slide, all tiles from all metrics contain a variable degree of specimen that are likely to give valid focal positions. As a result, the quality of focus candidates selected by all these metrics are samilar to each other, shown by the fact that all focus maps by four metrics converged to almost the same MSE, examples shown in Fig. 5 and Fig. 6. However, Fig. 3.10 indicates that the NACI metric tends to select tiles with high contrasts, i.e., on the edge of specimen and background, while APCI and EI select tiles containing almost entirely specimen. From Table 3.7 it can be seen that APCI produces the minimum SNFC on the majority of specimens. This means that APCI requires a smaller number of focus candidates to construct an accurate focus map. In practice, this means that APCI will minimise the amount of time spent on focusing on high-resolution FOVs, thus minimising scan times. Furthermore, Table 3.7 shows that on the two Thin-prep slides APCI is the only metric that reliably reaches the sufficient MSE criterion.

Obviously, the higher the resolution of the low-resolution image, the more information can be extracted for tile analysis. However, increasing the resolution, increases both the number of FOVs

acquired and the data to be processed. There is clearly a trade-off between the low-resolution scan times and scan map and focal map accuracy. Here we have chosen to compromise on a $4\times$ objective and demonstrated that accurate scan and focus maps can be produced. This not only saves considerable amount of high-resolution scan time, but also avoids wasting time on a re-scan if the high resolution images are found to be unsatisfactory. Particularly for APCI, which measures the phase diversity of the tile, the size of cellular material will be smaller at lower magnifications and so the accuracy of the method may be affected. Therefore, standalone macro camera with a unity (or lower) magnification may not be suitable as the depth of field may be larger than the depth of the specimen.

Future work will improve the threshold selection criteria for the modified Otsu's method. Moreover, the proposed modification of the Otsu's method assumes that there are only two classes (foreground and background), which could be further improved by using a Gaussian Mixture Model that explicitly models multiple classes consisting of different types of cellular material (such as cytoplasm and nucleus) and artefacts (such as ink markers) (Douglas 2009; Li et al. 2012). The method to date, has only been tested on conventionally stained (PAP and HE) bright-field microscopy images. However, in principle the technique would appear to be suitable for other types of specimens such as those acquired in fluorescence mode. Therefore, more types of specimens would be used to refine the proposed method.

3.6 Conclusions

This chapter presented a novel tile-based algorithm capable of both delineating a microscopic specimen and selecting a subset of FOVs that are good candidates from which to construct a high-resolution focus map. In particular, we investigated four tile evaluation metrics for their suitability for both detecting specimen and highlighting good focus candidates. The experimental results on a set of 40 real-world pathology specimens demonstrated that while the NACI, APCI and EI were superior to TI at specimen delineation, only APCI was capable of selecting good focus candidates while ignoring artefacts.

The proposed method provides the knowledge about lateral distribution of the specimen on the slide as well as a list of focus candidates. The focus profile map is next estimated and the number of focal planes to be scanned in the subsequent high resolution is decided, of course based on the knowledge of topology and thickness of the specimen, which will be investigated in the next chapter.

4

QUANTITATIVE ANALYSIS OF CELL AND CELL CLUMP THICKNESS IN CERVICAL CYTOLOGY SLIDES

This chapter presents the work to attain **Aim 1.2**, which is to quantitatively analyse the 3D distribution of the cytology specimen in order to determinate the thickness of the cytology specimens, and hence the optimal number of focal planes and the minimal scan time for scanning of these cytology specimens. As mentioned in chapter 1, there are inevitable overlaps and thick cell clumps on the liquid cytology preparation methods prepared slides, a method for the quantitative evaluation of the thickness of cell and cell clumps in a LBC cervical cytology specimen is therefore proposed. The method utilises an exhaustive high-resolution scan of the specimen, an over-complete wavelet transform for detecting the focal depth of objects and a level set segmentation that distinguishes the foreground cervical cell specimen from the background. The accuracy and effectiveness of the method is evaluated with Thin-prep prepared cervical cytology specimens. Using the method, the distribution and cell clump thickness, the topology of the specimen surface and the optimal number of focus candidates can be determined, all of which are important in order to optimise the high resolution acquisition of cervical cytology specimens.

Most of this chapter is included in a journal article that has been published on the *Micron*.

Related publication details

Fan, Y. and Bradley, A.P., 2016. A method for quantitative analysis of clump thickness in cervical cytology slides. Micron, 80, pp.73-82.

Manuscript revision history	
Submitted to *Micron*	30 July 2015
Interim decision	
Revision submitted	3 September 2015
Accepted	3 September 2015
Published/Available online	28 September 2015

4.1 Introduction

Cervical cancer screening is an important health issue among women world-wide that aims to detect pre-cancerous and cancerous processes of the cervix. The development of liquid based cytology (LBC) preparation methods and synthetic stains have provided clear and consistent cell preparations that are essential for effective, large scale cervical cancer screening (Bentz 2005; Lahrmann et al. 2013; Lu et al. 2015a). The introduction of WSI has been welcomed by pathologists because it enables remote consultation, quality assurance and importantly the potential for computer aided diagnosis for cervical cancer screening (Al-Janabi et al. 2012; Ameisen et al. 2013; El-Gabry et al. 2014). The diagnosis of cervical specimens by human experts has been considered a challenging task, because there are a huge number (10,000-20,000 in Pap smear slide) of cells that need to be examined and only a small fraction of them may be visually abnormal (Malm et al. 2013). With WSI, the specimens are first scanned as digital images and suspicious cells can be tirelessly and consistently detected by digital image processing algorithms (Zhao et al. 2004). However, WSI faces some significant issues limiting its adoption in clinical use, such as poor standardisation of image quality and resolution, large file size of the digitised slides and slow acquisition speeds, especially when multiple focal planes are required (El-Gabry et al. 2014; Wright et al. 2013). Among them, the scan speed is the main bottleneck, because acquisition has be fast enough for clinical use, but improved image quality and acquisition of multiple focal planes both result in significantly longer scan times.

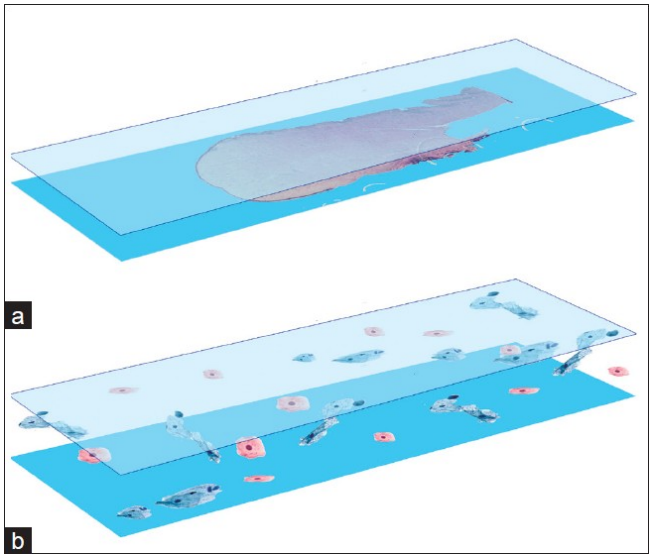


Figure 4.1: The spatial distributions of a typical histology specimens (a) and a typical cervical cytolocy specimen (b). Figure reproduced from (Lee et al. 2011) without permission.

For a digital slide scanners used to scan cytology specimens, the time-quality trade off is closely related to the nature of the complex 3D distributions of the specimens. Unlike histology specimens, which are prepared to be "continuous" and relatively flat, cytology cells are often sparsely and randomly distributed both spatially and between the glass slide and the cover-slip, as illustrated in Fig. 4.1. Even with the advent of LBC preparation techniques cells often overlap to form thick cell clumps that span multiple focal planes (Lee et al. 2011). In addition, important diagnostic cells are often found within these cell clumps, such as secretory cells from the endocervic and low-grade squamous intraep-

ithelial lesions (Lee et al. 2011). Therefore, it is no surprise to see that the acquisition of multiple focal planes achieves better diagnostic performances than acquiring only one or a few focal planes (Evered and Dudding 2011; Wright et al. 2013). Additionally, interpreting cellular objects requires the microscope to operate at high magnification (40×), preferably at resolutions close to the diffraction limit (∼0.2μm/pixel), where the DOF is reduced to around one micron (Castleman 1996). This means that even isolated cells can span multiple focal planes (El-Gabry et al. 2014). Given that scanning even a single focal plane at high resolution acquires a considerable amount of data (in order of GB) and can require significant acquisition times, the number of focal planes scanned should be minimised. In this way, the WSI system minimises not only the acquisition time, but also the storage space and subsequent image analysing time.

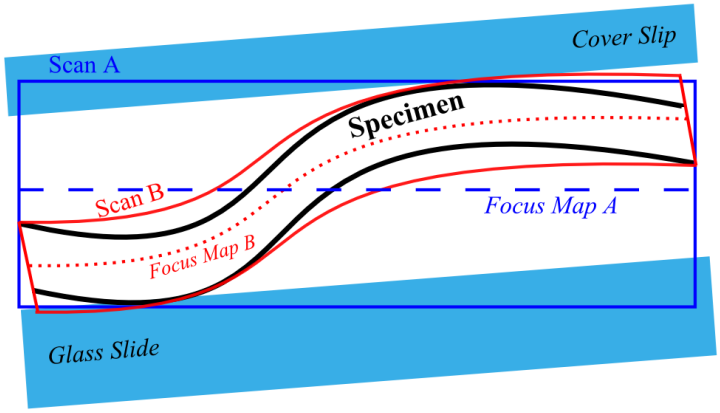


Figure 4.2: An illustration of potential relative positions of a microscope slide, cover-slip and specimen. Two examples of two multiple focal-plane scans (A in blue and B in red) are shown, viewed from the side and not in scale, their associated focus maps are shown as dash lines.

To achieve the minimum scan time, or to acquire a number of focal planes, the topology of the specimen and thickness of the specimen is required. To illustrate this, Fig. 4.2 shows relative the positions of a glass slide, cover-slip and a cytology specimen, the cells being deposited non-uniformly in the space between the cover-slip and glass slide. When viewed in high magnification the specimen is no longer a thin and flat mono layer. Rather, the specimen has finite thickness, which can also vary spatially (in Fig. 4.2 we assume the variation is the same at the top and bottom). In addition, the glass slide is unlikely to sit completely orthogonal to the optical axis and so the specimen appears to be tilted at a (small) angle in either spatial direction. This further increases the optical depth variation of the specimen.

The simplest scan strategy (Scan A in the Fig. 4.2) is to find the mean height of the specimen (shown as dashed blue line in the figure), say by averaging the height of a number of focal points, and then to scan an equal number of focal planes below and above that height. The number of focal planes should be equal to the thickness of the specimen plus the elevation of the specimen due to the tilt angle. It should be set high enough to be robust over all possible tilt angles. Unfortunately, as Scan A in the Fig. 4.2 shows much of image data acquired relates to out of focus or non-specimen areas. This approach is not only inefficient in scan time, but acquires unnecessary image data that then needs to be stored and potentially processed.

A better strategy is to attempt to estimate the topology of the specimen by fitting a plane through a number of focal points (Fan et al. 2014a). For example, the red dashed line in the Fig. 4.2 represents a focus map that better follows the topology of the specimen. The subsequent scan (B in Fig. 4.2) acquires a smaller number of focal planes, this is both a more time-efficient scan and reduces the amount of image data that needs to be stored (Fan et al. 2013). The quality of the focus map is determined by the number of focus points, because more focal points and data result in a better estimate (Fan et al. 2014a). If the specimen is assumed to have a consistent thickness, the minimum number of focal planes is then the same as thickness of the specimen and the effect of the tilt angle is removed. However, this approach may still be sub-optimal on specimens, such as those prepared with LBC, that have an inconsistent (spatially varying) thickness distribution.

Many previous works, mainly focused on qualitative analysis, have attempted to determine the optimal number of focal planes for scanning LBC cervical cytology specimens. For example, it was first shown that digital cervical cytology slides with seven focal planes achieved higher diagnostic accuracy than those of a single focal plane (both at $40\times$), but both were outperformed by conventional glass slides (Wright et al. 2013). Another study reported that cervical specimens scanned with 21 focal planes at a 1.5μm interval achieved better diagnostic accuracy than those scanned with 5 focal planes at 1μm interval (Evered and Dudding 2011). A more recent work argues that scanning only three focal planes (with 1μm interval) is able to achieve diagnostic performances close to that of a conventional glass slide (Donnelly et al. 2013). However, a majority of pathologist participated in this work reported that focusing over cell clusters were not as good as that in conventional microscope, and they did not prefer to use the virtual microscopy for the future diagnosis. These qualitative analysis failed to reach an agreement on the exact number of focal planes required to digitise the cervical cytology specimens because they do not know the exact thickness of these specimens.

In this chapter, a method to quantitatively analyse cervical cytology specimens is proposed, which estimates both the spatial location and thickness of every cell and cell clump. Specifically, the specimens are first exhaustively imaged **in 3D** at high resolution and multiple focal planes, every cell clump and nucleus is then segmented and an extended depth-of-field (EDF) algorithm, based on an over-complete wavelet transform, is utilised to determine the depths of each pixel of cell/clump. We purposely restrict our experimental slides to those with a normal diagnostic result so that we focus the study more on a demonstration of the usefulness of the method rather than a comparison of quantitative analysis between normal and abnormal slides. The usefulness of the method is demonstrated by finding the optimal number of focal planes required to acquire a "glass-faithful quality" digital version of these specimens. In this study, we quantitatively define the "glass-faithful quality" scan of cervical specimen to include the 95% of all ceullar objects in focus. In addition, we propose a novel method for focus map estimation that considers the thickness of the candidate focal points.

4.2 Methods

The proposed method is developed specifically for analysis of slides prepared in LBC as the cells are made to distribute in mono-layer. However, in principal the framework can be used to process other

types of cytology or histology specimens by selecting an appropriate imaging and cell segmentation strategy. The specimens are initially exhaustively imaged at high spatial resolution, as per Scan A in the Fig. 4.2, so as to image their full thickness with multiple focal planes. Each FOV, imaged at multiple focal planes, is then converted into a composite image with extended depth-of-field (EDF) (Bradley and Bamford 2004) prior to segmentation of cell clumps and nuclei (Lu et al. 2015a). The EDF algorithm also produces a depth estimate for all pixels in the image, which when combined with output of the cell segmentation results in a 3D map of all segmented objects (cells, clumps, nuclei). The main steps of the proposed method are illustrated in Fig. 4.3 and described in detail below.

4.2.1 Specimen Image Acquisition

The first step aims to acquire a complete set of FOV images from the specimen for later processing. To achieve high axial resolution, high magnification objectives are desired because the DOF is generally small at high magnification. Here, we use the conventional FOV scan method, which samples FOV images sequentially over the entire specimen (Fan et al. 2013), but other scan methods such as a line scan could also be used. The scan is equivalent to implementing a complete 3D scan of the specimen that aims to capture all cells at the greatest detail. Rather than building a focus map for faster scan speeds, we initially perform an exhaustive scan of the entire specimen, so that the effect of different focal maps and scan methods can be simulated.

The specimen image acquisition starts by generating the scan map of the slide and selecting one focus candiate near the centre of the specimen as initial focal point for the focusing algorithm. The scan map and focus candidate can be produced by the method proposed in the previous chapter. The selected focal point is then focused, which gives the rough height of the specimen and is used as the starting search height for the later focusing algorithm to minimize the search time. The entire specimen is then sampled FOV by FOV following a raster scan pattern with a small amount of overlap, the same as the FOV scan method as described in Chapter 2. Each FOV is focused and central focal plane is found as the best single focal plane (BSFP) to acquire the image of current field, i.e., the image that would be acquired if only a single focal plane of the specimen is to be sampled.

We again used normalized variance as focus metric and fast hill-climbing search as autofocus algorithm (Sun et al. 2004; Tello-Mijares et al. 2013). The algorithm searches a depth range with progressively smaller steps (here 20μm, 5μm and 1μm). To improve focusing accuracy, 15 focal steps were sampled in the last search interval and the one with the maximum normalised variance was selected as the BSFP . To avoid focusing on the top of the cover-slip, the search range was limited to be smaller than the height of the cover-slip from the specimen. Specifically, the focus algorithm searched between 100μm above and 100μm below the initial focus point previously found, given the height of the cover-slip was found to be around 150-170μm and the slide is approximately flat on the slide holder. When the slide is slanted, on the other hand, an approximate map associated with the center of the slide could be used to guide the acquisition.

Once the FOV was focused, a Z-stack of images (equal number above and below the BSFP) were acquired at an interval of 1μm, which is the same as the DOF of the imaging system. The height of

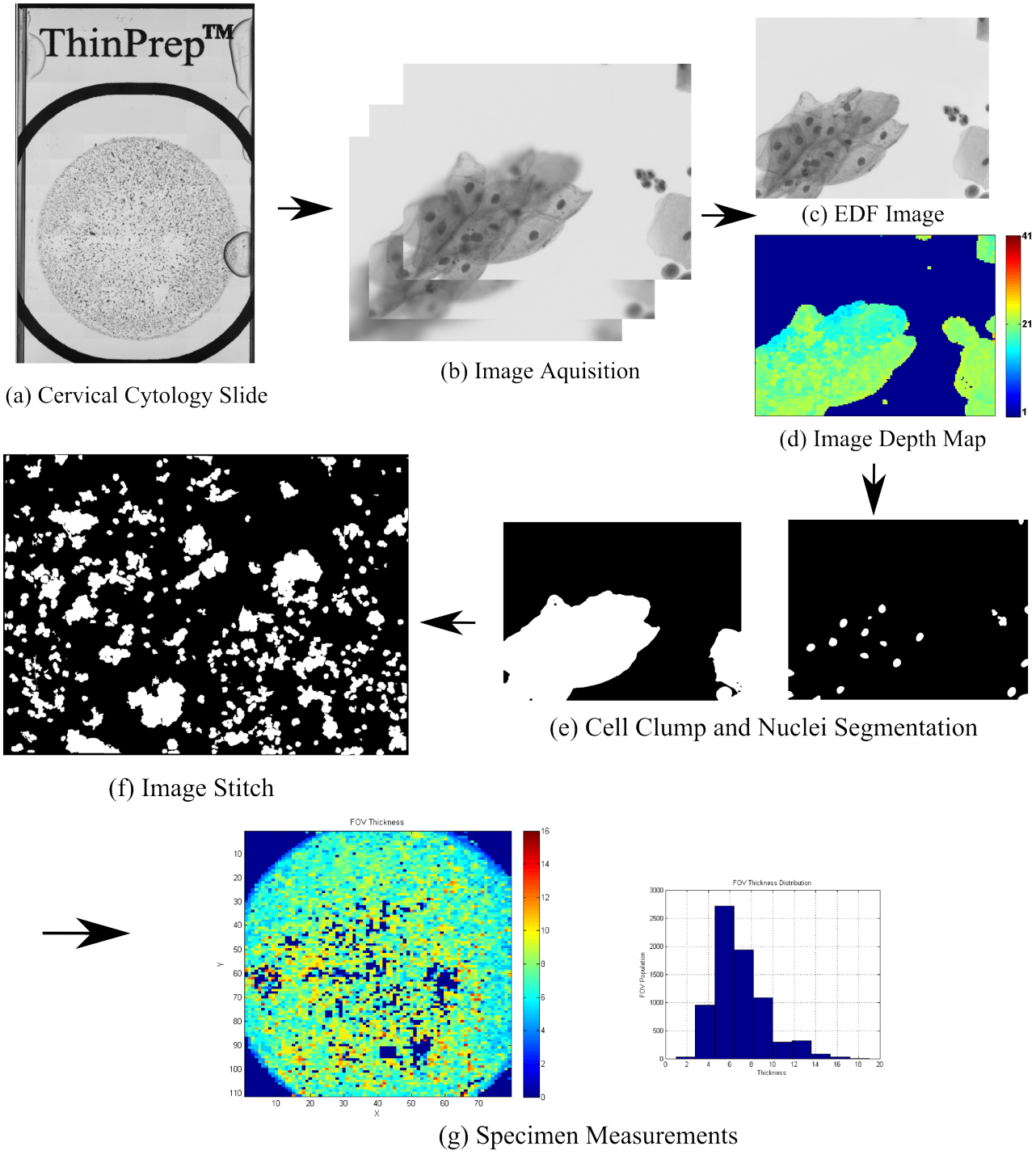


Figure 4.3: The cervical cytology slide (a) is first extensively scanned into 3D image stacks (b). Each 3D stack is processed to produce an extended depth-of-field image (c) and a depth map (d) showing which layer each pixel in the EDF image came from. The EDF image is then segmented for cells, clumps and nuclei (e). The segmented maps of FOVs are finally stitched together (f) for quantitative analysis of the entire specimen, such as the distribution of FOV thickness estimates (g).

the 3D stack was selected to be high enough to include all focal planes even from the thickest part of the specimen. In addition, an empty FOV threshold was calculated (in normalised variance) from an existing empty FOV images so that empty FOVs could be detected and acquisition terminated once a single empty FOV was acquired rather than wasting time on acquiring a Z-stack of empty FOVs. Sampling blank FOVs for image background correction was repeated every 30 minutes to cater for fluctuations in illumination experienced during these long scans (up to 60hrs per slide).

4.2.2 Extended Depth-of-Field Image and Depth Map Generation

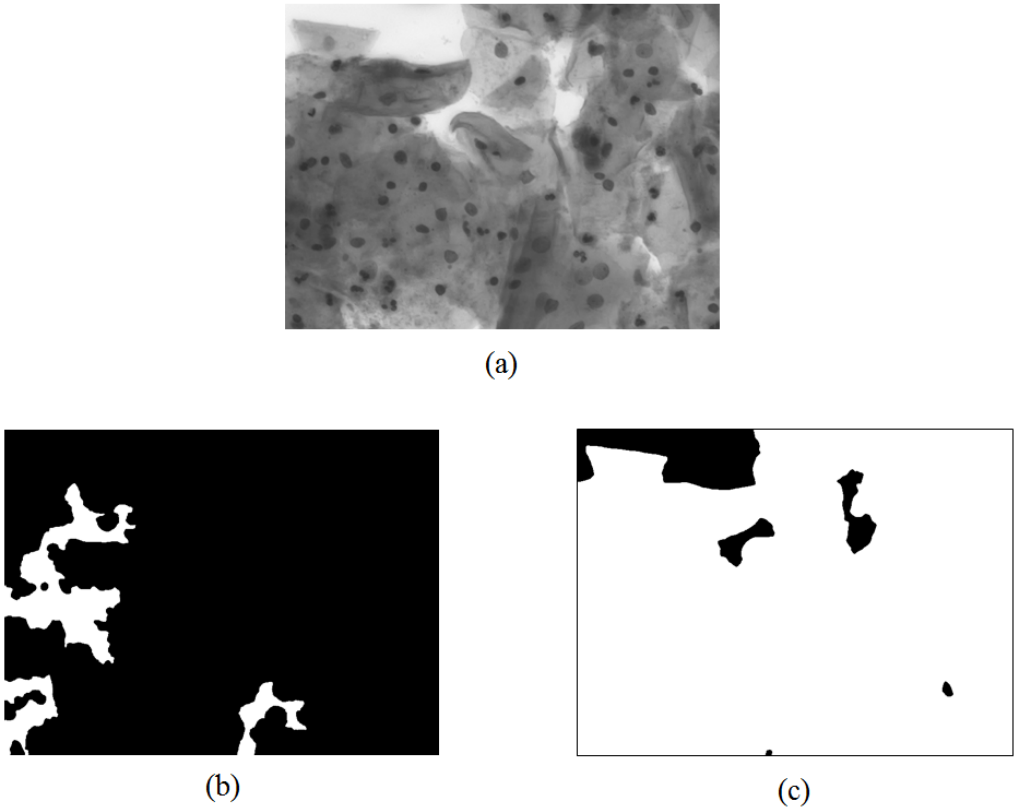
Each 3D image stack containing cells at multiple focal planes was then converted into a single EDF image with all objects in focus (c. in Fig. 4.3). The generation of an EDF image simplifies the cell segmentation process to a 2D space so that existing 2D segmentations algorithm could be applied, given that only the overall thickness of the cell and cell clump are interested rather than the exact 3D location of each cell. Many EDF algorithms have been developed to work in both the spatial and frequency domains such as in wavelets. Here, we used an algorithm, based on an over-sampled wavelet transform, previously developed for there analysis of cervical specimens (Bradley and Bamford 2004). The wavelet transform provides an effective method for detection of in-focus objects because these objects produce the large wavelet coefficients (Valdecasas et al. 2001). A seven level wavelet transform was performed, with lowest five levels being over-sampled to provide approximate shift invariance (Mallat 1989). The largest coefficients across the stack at each level were then selected within a 3x3 window and an inverse transform applied to return these largest wavelet coefficient to the image domain producing a composite image with all objects in-focus (Bradley and Bamford 2004).

While the largest coefficients are selected to produce the EDF image, the z position of these coefficients in the focal plane stack can also be used to estimate a depth map representing the focal positions of the selected in-focus objects. To avoid missing layers by spatial averaging, only the first three levels of the wavelet deposition were used to construct the depth map. As the first two levels had been down-sampled, the depth maps associated to these levels were up-sampled using the nearest neighbour interpolation. The lower bounds were first selected by finding the lowest layer. A median filter of size 8x8 was then used to smooth out any noise. To reduce edge effect, the boundaries of the image stacks were extended, using pixel reflection, by 10% of the original image size. The final depth map contains the height of each pixel with respected to the height of the stack. An example of depth map is shown in Fig. 4.3 (c), where the height of the stack was 41 layers. The EDF images and depth maps were produced concurrently with image acquisition in order to minimise overall processing time.

4.2.3 Cell Clumps and Nuclei Segmentation

The EDF images were next used for segmentation of cell clumps and nuclei by an previously proposed algorithm specially developed for the same task (Lu et al. 2015a). While the method is capable of segmenting individual overlapping cells, only the cell clumps and nuclei segmentation (the first two steps of the algorithm) are utilised here as this information is adequate to find the thickness of the specimen and the location and number of nuclei in each clump.

In outline, the clump segmentation first groups pixels that have similar grey values and spatially location using the quick shift algorithm (Vedaldi and Soatto 2008). The pixels in each group are then assigned to the mode value of all pixels in the group. An edge detector is used to find the the most prominent edges as rough boundaries of each cell clump. These boundaries are next refined by learning an unsupervised binary classification. The grey values of foreground and background pixels from this initial estimate are then fitted to a Gaussian mixture model, and the maximum likelihood estimate is used to re-classify each pixel as foreground or background. This refinement was repeated several (20) times, and a minimum area constraint imposed to remove small fragments that may not be cervical cells (such as the blood cells on the right side of Fig. 4.3(b)).



(a) (b) (c)

Figure 4.4: A FOV contained large proportion of cell clumps (a) was outlier when segemented with the complex method (b) but can be sucessfully segmented by the simple intensity thresholding method (c). The foreground is denoted as white in the binary map.

For FOV images that contained large proportion of cell clumps, the algorithm (Lu et al. 2015a) sometimes failed because it would classify cytoplasm as background as shown in Fig. 4.4.b. Therefore, these cases were detected by comparing the mean intensity of the segmented background to that of six previous FOVs and where appropriate this mean threshold was used to segment these images as shown in Fig.. 4.4.c.

Using the segmented cell clumps as the initial search area, candidate nuclei with stable connected components were detected by the maximally stable extremal regions algorithm (Matas et al. 2004). Nuclei that did not meet a minimum size constraint or had an eccentricity larger than 0.9 were treated as artefacts and not analysed.

4.2.4 Image Stitch and Specimen Measurements

The segmented FOV images and depth maps were next stitched together into larger images so that cell clumps spanning over multiple FOVs were as complete as possible. The FOV images can be aligned by various methods, such as image correlation and stage positional readings such as from the encoders fitted to the each axis of the stage. In stitching the depth maps, the heights of cells with respect to the height of the stack needs to be converted into absolute heights with respect to the stage origin. The binary (segmentation) map of cell clumps was next associated with the depth maps for extraction of various specimen measurements.

Connected regions with sufficient size were considered as foreground specimens. The depths of each pixel of the specimen were extracted from the depth map and the maximum differences between the heights of each cellular object were used to define the specimen thickness. The cell clumps without detected nuclei were excluded, while the number of nuclei per cell clumps was recorded. Additionally, the thickness of each FOV was also calculated by calculating the maximum height difference of all foreground material in each FOV.

4.3 Experiments

4.3.1 Specimen Preparation and Imaging

Cervical cytology slides classified within normal limits by cytopathologists were scanned to evaluate the usefulness of the proposed algorithm. Specifically, ten slides were selected to contain cells in both sparse and dense distributions that would be "difficult" to scan. The slides were PAP stained and made from a automated mono-layer slide preparation system (ThinPrep 2000 Processor).

A motorized bright-field microscope (Zeiss Axio Imager.M1) was used to acquire the specimen images, which was fitted with an infinity-corrected $40\times/0.75$ NA objective lens with a $1.2\times$ projection lens. The effective depth-of-field of the system was around 1 μm. A charge-coupled device (CCD) camera (Kodak KAI-2020-OM) with 1600x1200 pixels and 7.4μm square pixel size was used, with a binning factor of 2, resulting a pixel resolution of 0.3μm on all images. Notice the imaging system is not the same as that used in the previous chapter.

The XY stage of the microscope has a positional accuracy of $< 1\mu$m and is fitted with encoders in both axis with a resolution of 0.2μm. The Z stage has a better positional accuracy at 0.025μm. As the resolution of the XY encoders are higher than that of the images (0.2μ vs 0.3μm), the FOV images and depth maps were stitched based on the stage positional readings only, giving a misalignment error of around one pixel. A stack of 41 images (20 above and below the BSFP) were acquired at interval of 1μm for every non-empty FOV. The 41μm sampled depth range range was believed to be larger than the reported thickness of the specimens (15-20μm) (Evered and Dudding 2011). To exclude any outlier FOVs, the segmentation results and depths maps were manually examined for quality, especially where the specimen was thick.

4.3.2 Evaluation of Specimen Thickness Estimation

The key to estimating the local thickness of the specimen is to accurately identify the heights of the top and the bottom of the specimen in each FOV stack. The thickness is then found by calculating the height difference between the top and bottom layers containing in-focus material. Once the thickness of the specimen at each FOV is known and a scan interval less than the depth of field is used to acquire the image stack, the specimen is effectively fully sampled. That is, the virtual image stack includes all the focal plane information that a cytopathologist would have had access to in the original glass slide when viewed under the microscope. Therefore, it is reasonable to assume that they will be able to come to an equivalent diagnostic decision using a virtual slide acquired in this way.

To evaluate the accuracy of the proposed method, 238 (FOV) image stacks were randomly selected from the ten experimental slides. We then manually inspected individual stacks sequentially from the top slice to the bottom slice in order to find the slices that contain the top and bottom layers of the cervical cells. The heights of these slices were hence the groundtruth of the stack and used to compare with the height obtained from the proposed method.

4.3.3 Quantitative Specimen Measurements

As a demonstration of the usefulness of the proposed methodology, ten specimens are quantitatively analysed, with particular emphasis on their Z-dimensional thickness distribution. In particular, we measure properties of the scan, such as the number of FOVs and FOV layers that contained in-focus cells, the mean height and range of acquired FOVs. Properties of the scan once the tilt of the slide/specimen was removed by fitting a linear polynomial focal map, such as interquartile range of the specimen and the interquartile range of the top and bottom layer of the specimen to enable a comparison of topological complexity. Finally, we also present summary statistics of the final cell clump segmentation, such as the number of clumps, number of detected nuclei, the percentage of clumps with more than three nuclei and the median thickness of each clump.

As discussed in the introduction, the number of focal planes to acquire is minimised when the scan trajectory follows the centre of the specimen. With the positions of all cell clumps and FOV images known, the centre of the specimen can be estimated by fitting a polynomial surface through the BSFP. Here we used the `fit` function in MATLAB to produce a focal map with a predefined smoothness. The first experiment was to find best focus map by varying the degree of polynomial surface. Specifically, with all FOVs treated as focus points, the order of polynomial surface was increased from zero (constant Z) to fifth order (quintic). As the specimens are deposited inside a circular area, the same degree of polynomial were used for both lateral axes (XY). The zero order polynomial implies that the focus map is a plane parallel to the horizontal of the stage, as in Scan A of Fig. 4.2. A polynomial surface of order one effectively estimates the tilt of the slide. Higher order polynomial surfaces are then able to produce a focal map that estimates the varying deposition of cellular material over the slide.

With knowledge of the exact positions of all cells, the focus map can be evaluated quantitatively. In this way, the positions of the FOVs estimated during the specimen image acquisition (step 1 of the

proposed method) can be used to estimate the distance of all FOV images containing cells from the estimated focus map. Hence the total number of focal planes required can be estimated. Obviously, the smaller the number of focal planes required, the more representative the focus map is of the specimen and the faster the acquisition time will be. Therefore, we need a metric to indicate how much of the specimen is required to be scanned. Here, we propose that a slide is fully sampled and its digital slide is in "glass-faithful quality" if at least 95% of the cellular material in both spatial XY and lateral Z directions are imaged at, or below, the diffraction limit of the optics. As the experimental slides were imaged at z-interval equal to the DOF of the optics, this criterion translates to the acquisition of at least 95% of all FOV images containing in-focus cellular material. The 95% threshold, while arbitrary, is large enough to ensure that the vast majority of all in-focus cellular material is scanned. It is also low enough to avoid the time consuming and often unnecessary acquisition of FOV images of the outer areas of the specimen that are unlikely to contain diagnostic material. However, the exact value of this threshold requires further validation through a diagnostic trial on a large set of clinical specimens. This is future work, beyond the scope of the current paper.

4.3.4 Focus Map Estimation using Thickness

Conventionally, only the BSFP from each FOV, found by the focus algorithm, is used to estimate the focus map. In this study, we propose to consider multiple focal planes from each FOV when estimating the focus map. In this way, we explicitly consider the thickness of specimen and so thicker parts of the specimen have a greater contribution to the focus map. Specifically, if an FOV has a thickness of at least 5μm (estimated using the proposed method), then two focal planes on either side of the SBFP are included as focal points when estimating the focal map. For example, if a specimen is located between 151-155μm with a BSFP at 153μm, then all five focal depths from 151μm to 155μm are used to estimate the focus map rather than just BSFP at depth of 153μm. In this way, an FOV from a cell clump has more focal points contributing to the mean squared error fitting procedure of the focus map.

The data collected from the ten slides enables us to quantitatively evaluate this idea by constructing a focus map for each specimen and then comparing how many focal planes are required to scan the whole specimen. First a given number of FOVs are selected as focus points and then used to estimate the focus map (fit a polynomial surface) utilising the method proposed in section 4.3.3. Focus maps estimated in the conventional way that uses only the height of BSFP was used as the benchmark. In particular, the number of focal planes required for critical sampling of the specimen were compared as well as the number of focus points required such that the quality of the focal map map stays unchanged when more focus points are added. At this point we refer to the focus map as being settled. Based on our previous work in Chapter 3, the maximum number of focus points tested was 45, while the minimum number of that was 6, with an interval of 3. The best degree of polynomial surface found in the above experiment was used here.

As candidate focal points are selected stochastically, the experiments were repeated 100 times and the average and maximum number of focus planes calculated. Specifically, the selection of FOVs was

done so as to maintain a minimum spacing between focal points. This minimum spacing is directly related to the number of focal points required.

Given the deposition area of the specimen is nearly a perfect circle, we found this problem is almost identical to how many small circles (with the same size) can be fitted into a larger circle (a unit circle). If the circular specimen deposition area is treated as a large circle, the diameter of the small circle is hence the minimum distance between center of circle (positions of focus points) when certain number of these fully fill the large circle. The problems have been well studied such as in (Graham et al. 1998), (Lubachevsky and Graham 1997) and (Grosso et al. 2010), a collection of results from these works documented on (Herzlich 2015) was used in this study.

4.4 Results

4.4.1 Effectiveness of Specimen Thickness Detection

Absolute Error	Top	Bottom
0μm	173	163
1μm	58	47
2μm	6	11
3μm	0	7
4μm	0	3
5μm	1	1
>5μm	0	6

Table 4.2: The absolute error of detections of specimen top and bottom layer of 238 stacks.

In Table. 4.2, we present the difference between ground truth and the proposed algorithms on detecting the positions of top and bottom layers of each FOV that contain in-focus material. It can be seen from this data that the positions of the top layers was more accurately estimated than the depth of the bottom layers, shown by more FOVs with small errors. The specimen top layers of 173 (73%) image stacks had been accurately located, combined with another 58 (24%) stacks with acceptable error of just 1μm. On the other hand, slightly smaller number of image stacks had the same accuracy, with 163 (68%) for no errors and 47 (20%) for an error of 1μm. Furthermore, more numbers of stacks had larger errors (>1μm) for detecting the bottom layer. From observations, image stacks with large errors (>3μm) contained folded cell boundaries that span multiple layers.

4.4.2 Quantitative Measurements of Slides

Some general statistics of ten sample slides are shown in Table 4.3. Specifically, there are between 7406-8785 FOVs (or image stacks) acquired from each slide that contained cellular objects. Around 79,000-112,000 image slices were found to contain in-focus cells, which are required to be sampled in order to completely digitize these specimens. The mean heights of FOVs with respect to the stage origin which represents the center of the specimens (in Z) had a maximum difference of 93μm. The range of the heights of FOVs varied between 13-30μm across the ten slides, which was the distance

between the maximum and minimum height of the FOV. The distribution of the specimen can be better measured by the IQR, which reflects the height range of 50% of the middle of FOVs. The median IQR was 4μm, while Slide No.2, 4 and 5 had both high IQR and high ranges the height of FOVs. It is seen that slide No.2 had the highest range of the heights of FOVs and IQR, indicating the slide has the largest tilt angle. To eliminate the effect of glass tilt, the IQRs of height of FOVs were also calculated with a one degree polynomial surface (a linear estimation of slide tilt) as an reference. The IQRs were subsequently reduced (1.84-2.71μm). Interestingly, the IQR of heights of specimen top layers were much smaller (almost half) than that of bottom layers. The mean IQR was 1.95μm for specimen top layer and 4.32μm for specimen bottom layer.

Slide Statistic	Slide No.										Average
	1	2	3	4	5	6	7	8	9	10	
Number of Valid FOVs	7563	7552	7947	7406	7825	8785	8242	7715	8136	8191	7936
Number of FOV Layers	93406	83707	100186	95081	105673	125014	101816	96969	112067	79504	99342
Relative Height of Specimen (mean height of FOVs) (μm)	28	25	82	42	0	3	49	41	40	93	-
Range of Height of FOVs (μm)	13	30	16	25	25	18	16	14	19	16	19
Interquartile Range of The Heights of FOVs (μm)	4	12	3	8	8	4	4	4	4	4	5.5
Interquartile Range of The Heights of FOVs Relative to Poly11 (μm)	1.84	1.84	2.71	1.91	2.32	2.52	2.57	2.45	2.60	2.44	2.32
Interquartile Range of Top Layer Relative to Poly11 (μm)	1.39	1.78	2.20	1.72	2.34	2.01	1.91	1.92	2.18	2.26	1.97
Interquartile Range of Bottom Layer Relative to Poly11 (μm)	2.65	4.83	5.41	4.34	4.50	3.60	4.49	2.97	4.30	4.64	4.17
Number of Nuclei	49873	75315	47919	64156	62749	56532	65527	50194	74138	68398	61480
Number of Clumps	12572	12927	12238	14179	8458	15186	9941	16796	16099	13514	13191
Number of clumps with less than or equal to 3 nuclei	8916 (71%)	8070 (62%)	9031 (74%)	10292 (73%)	4833 (57%)	11277 (74%)	6107 (61%)	13172 (78%)	11431 (71%)	8021 (59%)	9115 (69%)
Number of clumps with more than 3 nuclei	3656 (29%)	4857 (38%)	3207 (26%)	3881 (27%)	3625 (43%)	3909 (26%)	3834 (39%)	3624 (22%)	4668 (29%)	5493 (41%)	4075 (31%)
Median Clump thickness μm	10	9	9	10	11	12	11	9	11	9	10

Table 4.3: Statistics of ten cervical cytology slides prepared in Thin-prep.

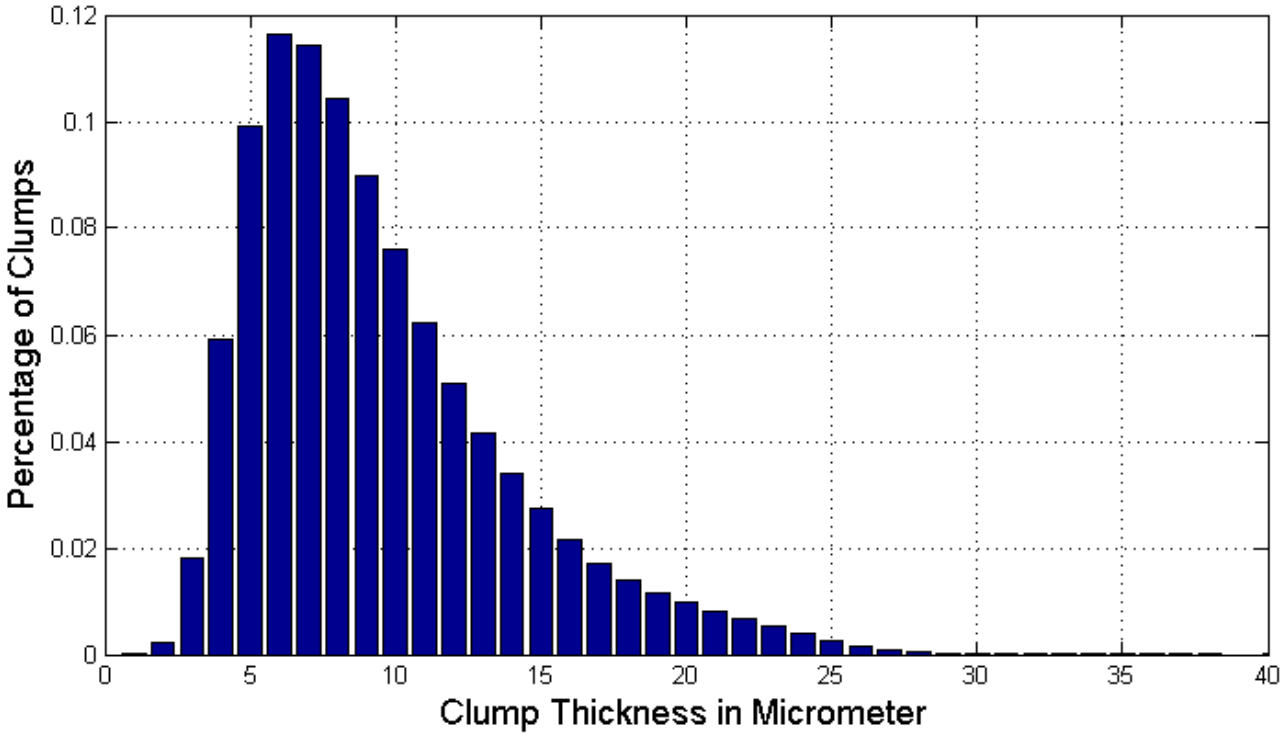


Figure 4.5: Histogram of Thickness of Cell Clumps from All Ten Slides.

From the perspective of the cells, there were between 8458 to 16796 cell clumps per slide, and the total number of cells counted by the number of nuclei were estimated at between 47919 to 75315 per slide. Around 22-43% of cell clumps have more than 3 nucleus, or consist of three cells. In addition, the distribution of thickness of all cell clumps in the ten slides is shown in Fig. 4.5. The IQR of the thickness of all cell clumps was 6μm spanning from 7μm to 13μm. Further more, the median thickness of cell clumps was 9μm, and 95% of cell clumps had thickness smaller than 19μm.

Polynomial Order	Slide No.										Average
	1	2	3	4	5	6	7	8	9	10	
0	21(19.0)	35(33.0)	25(23.0)	27(25.0)	29(27.0)	23(23.0)	24(22.0)	22(20.0)	25(23.0)	23(21.0)	25.4(23.6)
1	19(17.6)	23(21.0)	24(22.0)	21(18.3)	24(22.4)	23(20.6)	24(22.2)	20(18.2)	23(20.8)	20(18.2)	22.1(20.1)
2	19(17.5)	23(20.5)	23(21.2)	19(17.6)	24(21.9)	21(19.6)	22(20.6)	20(18.0)	23(21.0)	20(17.9)	21.4(19.6)
3	19(17.5)	22(20.1)	23(20.9)	19(17.6)	24(22.1)	21(19.6)	24(21.1)	20(18.1)	22(21.0)	20(17.9)	21.4(19.6)
4	19(17.6)	23(20.9)	23(21.1)	20(17.8)	25(22.5)	22(19.9)	23(20.9)	20(18.2)	23(21.1)	20(17.8)	21.8(19.8)
5	19(17.7)	23(20.5)	23(21.1)	20(17.8)	25(22.4)	21(19.7)	24(21.0)	20(18.3)	23(21.3)	20(17.8)	21.8(19.8)

Table 4.4: The number of focal planes (maximum of 100 simulations) required to sample 95% of the in-focus FOVs from the ten slides. The height range of these in-focus FOVs, in μm, is shown in brackets.

Focal Map	Slide No.										Average
	1	2	3	4	5	6	7	8	9	10	
BSFP Only	21(27)	24(15)	25(15)	21(24)	25(21)	22(45)	24(39)	21(18)	24(30)	21(9)	22.8(24.3)
FOV Thickness	17(21)	19(30)	19(33)	19(15)	21(15)	19(24)	19(33)	17(18)	19(21)	17(24)	18.6(23.4)

Table 4.5: The number of focal planes and settling number of focus points (in brackets) used to acquire 95% of the in focus material from the ten experiment slides in the worst case scenario (maximum number of focal planes required in 100 simulations).

The specimen complexity can also be illustrated by results shown in Table. 4.4, which shows the number of focal planes required to critically sample each specimen. Unsurprisingly, using the polynomial surface with zero degree as the focus map required the most number of focal planes to be scanned (Scan A in Fig. 4.2). The number of focal planes required dropped sharply if first order (planar) focus maps were used. Accounting primarily for the tilt of the slide. However, the number of focal planes required only decreased marginally when the order of the focal map was increased above third order. This is shown by slight reduction in the range of focal planes required to acquire 95% of the in-focus material and the insignificant reduction in the numbers of focal planes. Overall, the number of focal planes required varied between 19 and 23 over the ten slides. It is worth noting that the estimation of focus maps used all valid FOV and the height of FOV layers containing in-focus specimens as focus points.

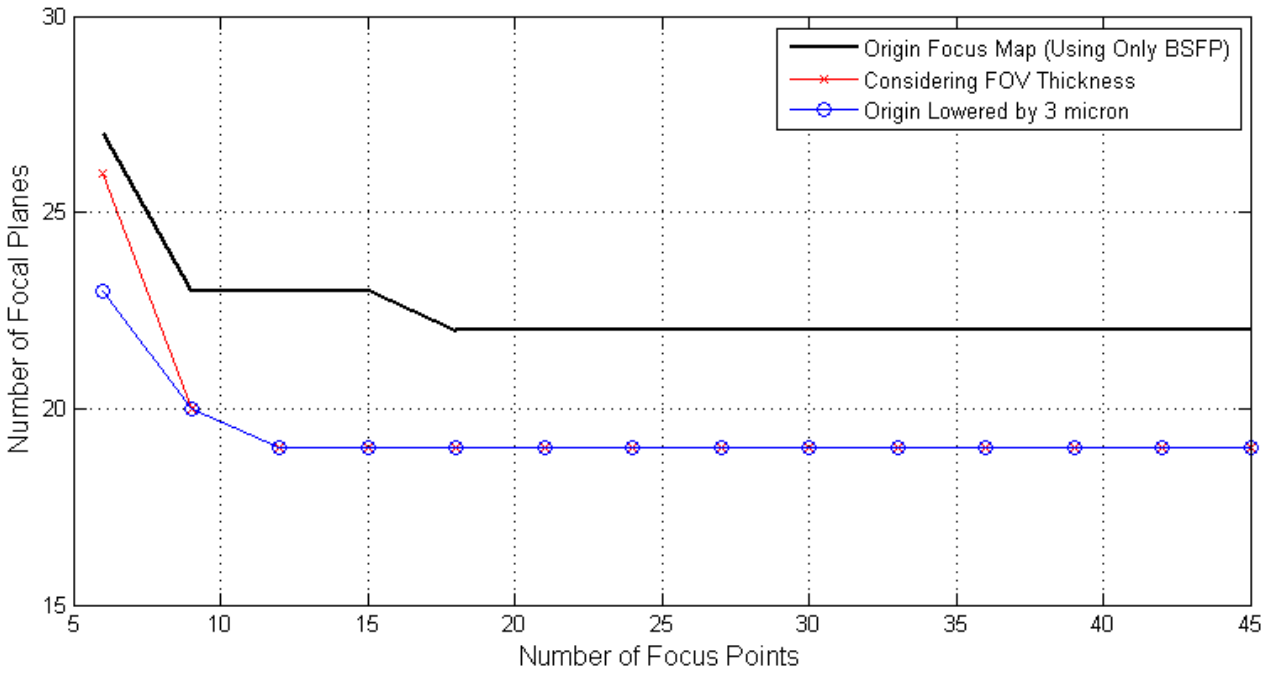


Figure 4.6: An example that shows the relationship between the number of focus points used for estimation of focus map and the mean number of focal planes required to acquire 95% of the in-focus material. Here we compare three focus map generating strategy: considering only the height of BSFP of the focus points (the origin focus map), considering the thickness of focus points and lowering the origin focus map by 3μm. The numbers of focal planes required are the maximum of 100 simulations.

4.4.3 Estimation of Focus Map Considering Specimen Thickness

The number of focal planes required for critical sampling of the ten experimental slides is further reduced by almost 20% from 22.8 to 18.6 when the thickness of the specimen inside the focus points is considered, as shown in Table 4. The relationship between the number of focal planes required for critical sampling of an example slide (NO.6) and the number of focus points used for focus map estimation using three different strategy is shown in Fig. 4.6. It can be seen that the required number of focal planes decreased quickly and remained the same after a small number of focus points (called

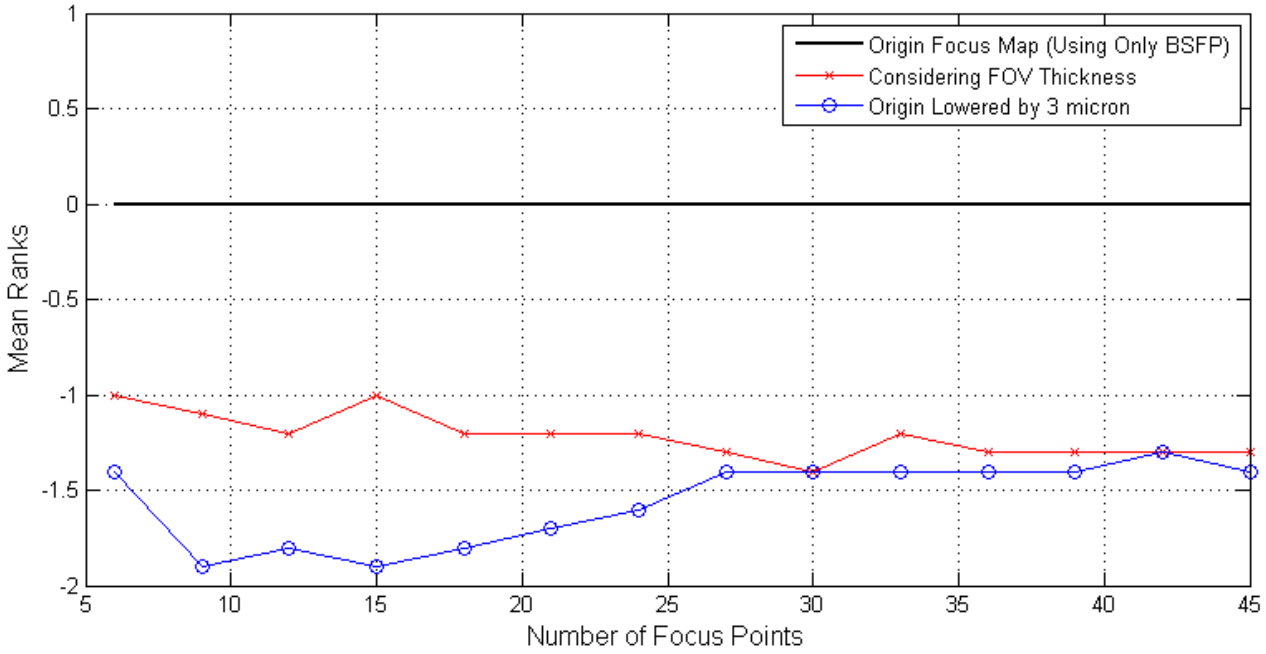


Figure 4.7: Ranks of the quality of estimated focus maps that consider only the height of BSFP of the focus points (the origin map), consider the thickness of focus points and are lowered from the origin focus map (by 3μm). The origin focus maps are used as the benchmark (with rank of zero) and the ranks are the mean of 100 simulations.

the settling number of focus points) being used. Specifically, the settling number of focus points for the slide was 18 if only using the height of BSFP for focus map estimation (the origin focus map), which was larger than that considering the FOV thickness (13). On the other hand, the numbers of focal planes required with the origin focus map were consistently larger than that considering the specimen thickness regardless how many focus points were used. Alternatively, the number of focal planes can also be reduced if the origin focus map were lowered, such as the one shown in the Fig. 4.6. The comparison of the three different strategy with all ten slides can be better illustrated by ranking them according to the required number of focal planes, shown in Fig. 4.7. With the focus map from considering only the height of BSFP as the benchmark, a negative rank means that the number of focal planes required by the method is smaller. Again, the focus maps constructed considering the thickness of the specimen achieved consistently superior quality regardless of the number of focus points used. And the lowered origin focus maps had even better performances.

4.5 Discussion

The evaluation results in Table. 4.2 showed that the proposed method is effective and accurate for estimating the thickness of the cervical cytology specimen thickness based on generated depth maps. The data obtained from the ten experiment slides hence enables a quantitative analysis of specimen, which can then be utilised to improve the focus map required for rapid whole slide digital acquisition. The result in Fig. 4.7 is a demonstration of the usage of this quantitative data that demonstrates that

generating a focal map based on specimen thickness was effective. Specifically, the evaluation results in Table. 4.2 show that the depths of the top and bottom layer of over 90% of the fields of view can be detected with of an error of less than 1μm (equivalent to the depth-of-field of the objective). It was also observed that image stacks with large errors contained poorly defined cell boundaries that spanned multiple focal planes and were difficult to distinguish even manually. In particular, as the EDF algorithm selects the maximum coefficient throughout the image stack, only one focal plane can be selected as the focal plane containing in-focus objects. In addition, these poorly defined layers did not contain cell nuclei, so they can be considered less critical for the subsequent diagnostic analysis of the acquired digital slide. Therefore, the estimation of the thickness of the cervical cytology specimens by the proposed method is sufficient for the later analysis.

The quantitative analysis of the ten specimens found that the median thickness of cell clumps was 10μm whilst the median thickness of the overall FOV was larger (12μm) as a whole FOV typically contains multiple clumps. From the perspective of the geometry of individual cells or acquiring fields of view, the specimens are considerably thicker than the DOF of the optics (which is around 1μm). This confirms the necessity of acquiring multiple focal planes during the digitisation of even "mono-layer" cytology specimens. More specifically, the simplistic flat scan map (as per Scan A in Fig. 4.2) required the largest number of focal planes (an average of 25.4). The inherent tilt angle of the slide has a significant effect on this, shown by a significant reduction in the number of focal planes required, as well as the IQRs of the height of FOVs, once a focal map that considers slide tilt is utilised. Not surprisingly, the number of focal planes required reduced as the complexity of the focus map was increased. This shows that increasing the order of the polynomial focal map allowed for a more complex surface that more closely matched the topology of the specimen. However, Table. 4.4 shows that focus maps above second order do not always show superior performance and are worse in some cases. This demonstrates that overall, on average, the surface complexity of the specimen surface is relatively smooth. Quantitatively, the average IQR of the height of FOVs (relative to first order polynomial surface) were around 2.3 μm. This shows that more than half of all cellular material is contained within a relatively narrow (focal) space, i.e., 4 μm of the single best focal plane. Further, the difference between the slide with highest IQR (2.57μm) and smallest IQR (1.84μm) was smaller than one micron (0.81μm). However, as Fig. 4.6 shows, significantly more (around 18) focal planes must be acquired to ensure that 95% of all in focus material is properly imaged. On the ten experimental specimens analysed here, the cubic polynomial surface was demonstrated to be sufficient to form a good focus map and is suitable as the basis for a rapid multi-focal plane scan.

The fact that cubic polynomial surfaces are sufficient for cytological scans does not necessary mean that the surface complexity of the specimens are in fact cubic. According to IQRs in Table. 4.3, the top of the specimen is relatively more flat (averaging 1.97μm compared to 2.32μm for the height of FOV), while the bottom of the specimen is less flat (averaging 4.17μm). This can most likely be explained by the placement of the cover-slip such that the specimen is compressed against the underlying glass slide. To further investigate this phenomena, the exact distributions of cell clumps and nuclei in different heights in the stack reflected in terms of areas were calculated. The distribution of the thickest slide (No.7) and the thinnest slide (No.10) are shown in Fig. 4.8. It can be seen that the

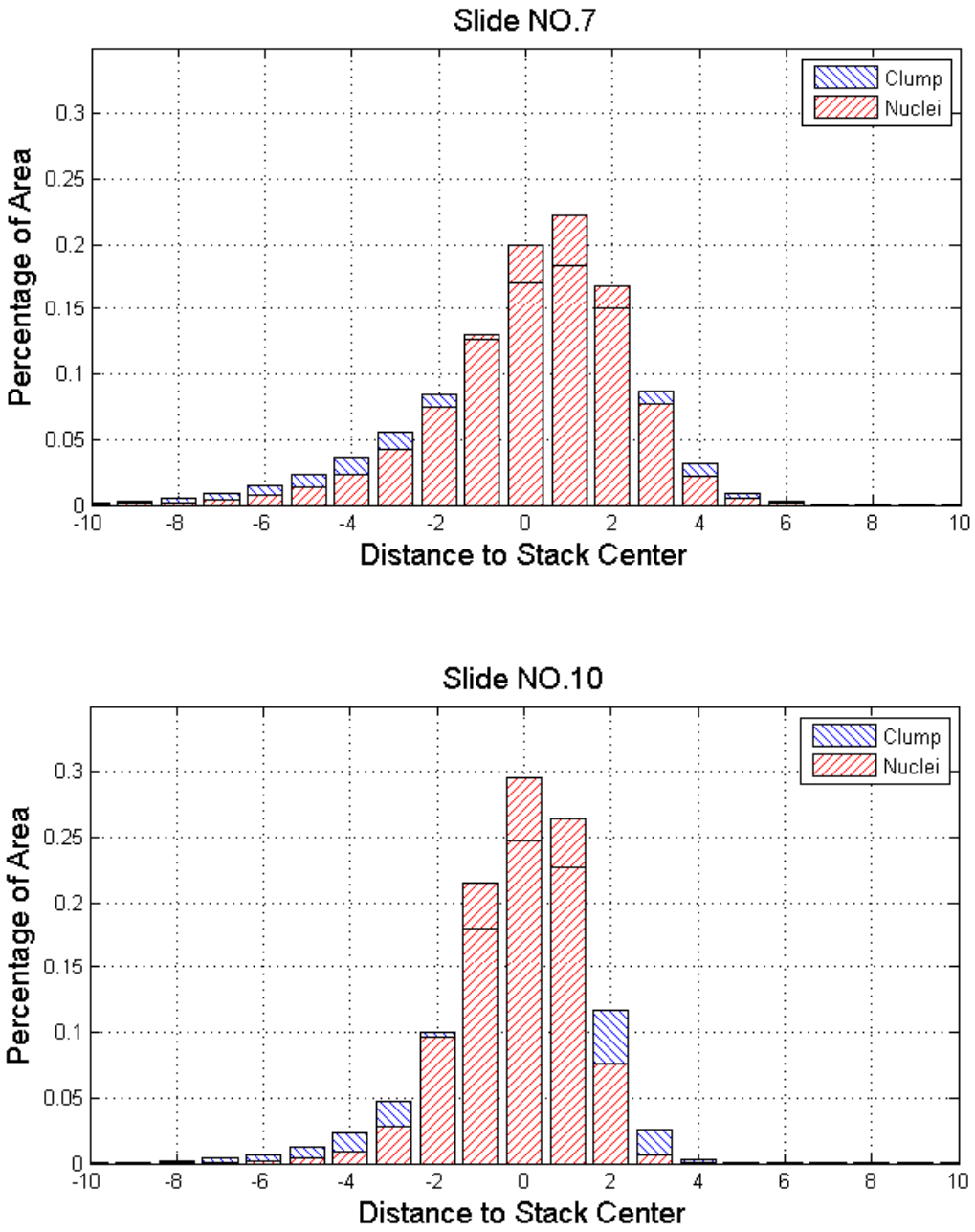


Figure 4.8: The height distribution of cell clumps and nuclei in the acquired image stack. The axis right is the direction of cover-slip.

overall distribution of the cells is skewed towards the cover-slip (which is on the right hand side of the graph). Specifically, there were almost no in-focus cells at layers acquired beyond 4μm from the central layer (the single best focal plane), shown as 0.13% of cell clumps for slide No.7 and 0.43% of cell clumps for slide No.10. In contrast, 5.02% and 9.62% of cell clumps from slide No.7 and No.10 respectively were located 4μm below the central layer. This is probably explained by the slide preparation protocol, particularly how the cover-slip interacts with the underlying sedimentation.

Similarly, the distribution of the cell clumps and nuclei across the entire slide showed the same trend, as illustrated in Fig. 4.9. The height of specimens were measured with respect to the "best" focus map (cubic surface) estimated using all FOVs as focus points. The overall shape of the distribution are

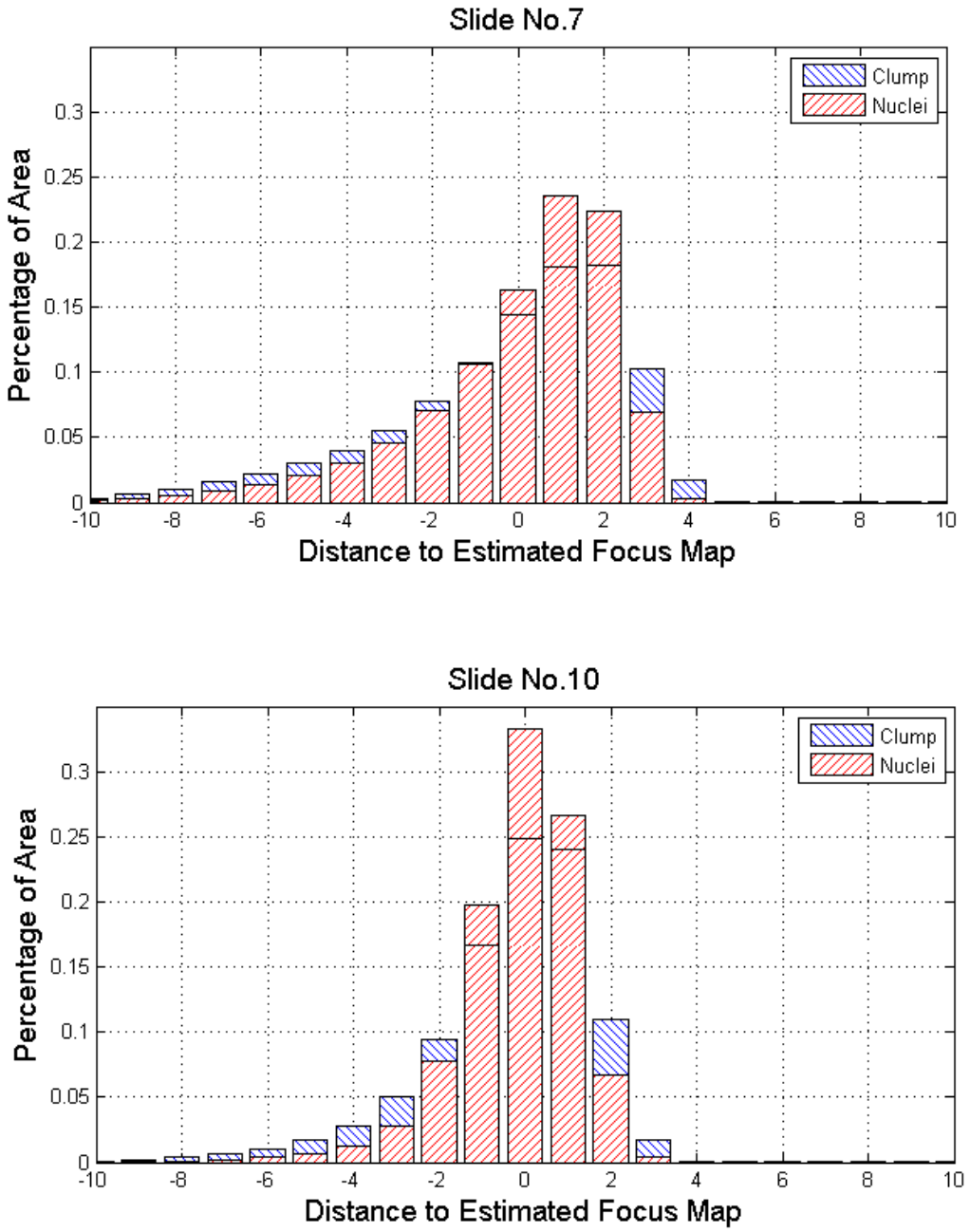


Figure 4.9: The height distribution of cell clumps and nuclei respect to the ground truth focus map in cubic polynomial. The axis right is the direction of cover-slip.

also skewed toward the cover-slip. Firstly, this confirms that the entire specimen top is flatter than the bottom of the specimen. Secondly, this similarity demonstrates that a knowledge of the distribution of the cells from a limited number of FOVs is capable of predicting cell distribution across the entire slide. This further justifies the effectiveness of considering the thickness of the specimen when estimating a focal map for cytological specimens.

It can also be seen in Fig. 4.8 and Fig. 4.9 that the nuclei were contained within a narrower distribution than that of the overall cell clumps. This can be simply explained by the fact that the nuclei are located at the centre of cell and will always be surrounded by other cellular material such a cytoplasm. Specifically, for the thinnest slide of the ten (No.10) scanned, 66% of cell clumps and 79% of nuclei

were located within a distance of 1μm of the focus map (in Fig. 4.9). Therefore, sampling only 3 layers at an interval of 1μm can capture sharp images of most of cells. This echoes the work by Bernd etl who claimed that it is possible to find and scan one "master" scan layer through the specimen where most of objects are in focus (Lahrmann et al. 2013). However, for slides thicker than slide No.10, significantly more than one layer is required to sample even just the nuclei.

The skewed distribution of the specimen also affects the estimation of the focus map for scans with multiple focal planes. The experimental results in Fig. 4.7 demonstrate the superior performance of estimation of focus maps when the thickness of focus points was taken into account in the fitting procedure. As shown in Fig. 4.8, the BSFP (layer 21) is not always located at the centre of the specimen in terms of its overall thickness. Instead, it is located at the "centre of mass" of the specimen (the centre of mass not being equal to the arithmetic mean due to the skewed distribution of the specimen between slide and coverslip). Therefore, the estimated focus map requires more focal planes to be scanned to adequately sample all focal planes of interest. On the other hand, considering the thickness of focus points (proposed in 4.4.3) eliminates the skewness distribution of specimens and is shown to be a better estimate of the true centre of the specimen. Alternatively, it is also possible to identify the amount of bias and correct the estimated focus maps. For instance, it was shown in Fig. 4.7 that focus maps lowered by 3μm from the origin focus maps (considering only the height of BSFP as the focus point) achieved even better performances as that considering the thickness of focus points.

The utility of the proposed method is not only limited to guiding the acquisition of cervical cytology specimens, it could also be used as a tool for quality assurance in slide preparation. Conventionally, the automatic slide preparation machines (e.g. ThinPrep 2000 Processor) do not have the ability to automatically evaluate the quality of the slides; this has to be performed by the cytotechnologists after the slide has been prepared. The proposed method enables the evaluation of the thickness of the slide so that the slide preparation machine can make another slide immediately if the cervical cell clumps are too thick. Alternatively, the method could also be used to quantitatively optimize the settings of the slide preparation machines (e.g. air pressures, spin speeds/times etc).

It is worth noting that all potential "outlier" focus points found during the estimation of focus map in this study were manually detected and excluded. In practice, dust and ink marker on top of the coverslip may result in outliers giving incorrect focal points that bias the focus map as evidence showed in last Chapter. However, specific methods can be developed to detect and remove these outliers such as that proposed in the last chapter and by the work in (Lahrmann et al. 2013) using machine learning techniques. Future directions for this work would include a more detailed analysis of the distribution of individual cervical cells from the development of fully 3D segmentation techniques, such as initial attempts outlined in (Lu et al. 2015b). In addition, the experiment for estimating the focus map, and in particular the selection of the 95% "glass-faithful quality" criterion should be extended to include a diagnostic comparison with the involvement of cytopathologists.

4.6 Conclusions

This chapter presented a novel method that enables the quantitative analysis of both the spatial distribution and thickness of cervical cytology specimens. The method relies on an over-complete wavelet transform to estimate the depth distributions of in-focus regions of the specimen and was accessed to be effective on over 90% of the acquired FOVs. The usability of the method was demonstrated by extensively analysing ten Pap stained Thin-prep slides. It was found that the majority of cells are located above the single best focus plane found using normalised variance as the focus metric. It was also demonstrated that knowledge of the thickness of the specimen is able to increase the quality of the estimated focus maps. Subsequently, it was shown that 95% of all in-focus cellular material can be imaged provided that a focal depth of 9μm on either side of the focal map is acquired, thus saving over 25% of raw image data and hence scan time.

When the proposed method is combined with the specimen delineation method proposed in the previous chapter, a focus profile map of the specimen including the optimal number of focal planes can be produced, having great potentials to reduce the overall scan time of cervical cytology specimen. The thickness distribution of the cells in the specimen is also useful in determining the optimal slide slant angle when the fast slanted scan method is utilised, which is to be investigated in the next chapter.

5 A METHOD TO CORRECT THE OPTICAL ABERRATION INDUCED BY SLIDE SLANT

This chapter presents the work to attain **Aim 2**, which is to investigate optical aberrations in the slanted scan method and to develop a method to correct these aberrations. For the fast slanted scan method to be accepted in clinics, digital slides produced by the slanted scan method must be in "glass-faithful quality", that is, containing multiple focal planes, at high resolution and are free of optical aberrations. However, the slide slant induces optical aberrations as illustrated in Fig. 5.4. To remove the effects of these aberrations, a two-stage method is proposed which would initially correct the bending artefacts of the optical aberration and then improve the contrast and resolution of the image in the second stage. This method is used in the generation of the digital slide after the scanning system has acquired the images via the slanted scan method.

Most of this chapter is included in a journal article that has been published on the journal *Micron*.

Related publication details

Fan, Y., and Bradley, A.P., 2016, April. A Two-stage Method to Correct the Optical Aberration Induced by Slide Slant. Micron, 87, p.18-32.

Manuscript revision history	
Submitted to *Micron*	28 January 2016
Interim decision	10 March 2016
Revision submitted	1 May 2016
Accepted	1 May 2016
Published/Available online	7 May 2016

5.1 Introduction

As introduced in Chapter 1, whole slide imaging (WSI) based on bright-field light microscopy is becoming an important tool for research, education and clinical diagnosis (Al-Janabi et al. 2012; Ghaznavi et al. 2013). However, conventional digital slide acquisition methods for WSI are slow when

digitising thick specimens, such as cervical cytology specimens, as to acquire a (3D) multiple focal plane image the stage has to complete multiple passes over the specimen (El-Gabry et al. 2014; Fan and Bradley 2016a; Rojo et al. 2006). This is particularly problematic when imaging at high resolution (say 40×) where the depth of field of the optics is narrow (∼1 micron). Specifically, the data throughput rate of the popular FOV scanning approach is limited by the time consuming mechanical movements (stop-go-stop) of the microscope stage (Fan et al. 2013). Scan systems that utilise line scan cameras are capable of scanning faster than the FOV approach because the stage moves continuously and the line scan camera runs at a much higher frame rate (up to 70kHz). Recently, a novel slanted (specimen) scan method has been specifically designed to scan cytology specimens in 3D that has the potential to significantly improve data-throughput rates and hence scan speeds (Bamford and Mayer 2009; Fan et al. 2013). The slanted scan method purposely slants the microscope slide at a small angle (of 3-5 degrees) to the optical axis so that lines from multiple focal planes can be imaged simultaneously using a single area scan sensor or multiple line scan cameras. However, the slanted slide induces optical aberrations that resemble lens coma. The main artefacts include movement of objects in the image between different focal planes (due to bending of the optical point spread function) and a reduction in both image resolution and contrast compared to images acquired from the conventional flat scan (Arimoto and Murray 2004; Bamford and Mayer 2009). An example of this aberration is illustrated in Chapter 2 Fig. 2.5. Therefore, for the slanted scan approach to be practicable, and produce images of an equivalent quality to a conventional flat scan, methods need to be developed to ameliorate these artefacts.

Similar aberrations to those shown in Fig. 2.5 have been reported in fluorescence microscopy when using a water-immersion objective when the slide is slanted unintentionally (Arimoto and Murray 2004). Interestingly, the oil-immersion objective lens does not suffer from this type of aberration (Arimoto and Murray 2004) as the refractive index of oil is closely matched to that of the cover-slip. Specifically, when water is used as the immersion medium, the light refracts as it travels out from the cover-slip and into the immersion medium because the refractive index of the water (∼1.333) is smaller that that of the cover-slip (∼1.518). These refraction phenomena lead to optical path length differences and consequently cause wave-front aberrations in the final image. When oil is used as the immersion medium for high NA lenses, the immersion oil typically has the same refractive index as the cover-slip, and so the light does not refract when it enters the different medium. Consequently there is no optical path length differences and aberrations are not apparent. However, the vast majority of BF microscopes and WSI systems, in particular, use air as the immersion medium, which has an even lower refractive index than water (1.0) and so slide slant is particularly problematic. Not using an immersion medium does, however, significantly simplify the use of WSI systems, especially when lower magnification (20×) objectives provide sufficient image resolution.

Despite limited attention in the field of microscopy, generalisations of slide slant, such as tilted plane parallel plate (PPP), and the optical characteristics of light rays passing through them, have been widely investigated (Braat 1997), (Stallinga 2005) and (Sheppard 2013). In some high speed cameras, tilted PPP are intentionally included to split the light to different sensors (Zhong et al. 2015). In optical disk recording systems (such as CD-ROM and DVD) the optical disk can tilt unintentionally,

due to rotation, when it vibrates during operation. Here, the substrate of the disk acts as the PPP and limits the reading accuracy of the system (Goodwin 2007). These aberrations have been analysed by calculating the optical path differences of light rays passing through a tilted PPP as compared to an un-tilted PPP. It was found that the optical aberrations consist not only of coma distortion, but also spherical, astigmatism and other aberrations (Braat 1997). Among these aberrations, the lowest order piston, coma and spherical aberration dominate (Stallinga 2005). Piston aberration is a linear phase shift and hence does not affect the quality of the image. Given that most modern objectives for BF microscopy are specifically designed to compensate for the spherical aberration caused by the cover-slip, coma aberration dominates in the slanted scan approach and so is the primarily aberration that needs to be corrected. It was also found that the magnitude of these aberrations are proportional to the NA of the lens, tilt angle and the thickness of the PPP (in our case the cover-slip) (Braat 1997). This partially explains why the aberrations are noticeable in the high NA water-immersion objective lenses used in fluorescence microscopy and less apparent in the relatively low NA (air-immersion) objective lens commonly used in BF microscopes and WSI systems in particular.

To correct for the aberrations caused by tilted PPPs, various methods have been proposed, but these have mainly been concerned with the addition of corrective hardware, such as a PPP tilted in the opposite direction (Zhong et al. 2015) (Gerber and Mansuripur 1996). Here, we aim to develop a digital image processing approach to correct for these aberrations. This approach, simplifies the optical set-up and minimises any hardware modifications required to the microscope, so that the slanted scan methodology can be easily adopted and applied on existing BF microscopes.

In this chapter, we propose a novel two-stage deconvolution method designed specifically to correct the aberrations induced by slanted slides in high resolution BF microscopy. Specifically, we initially apply a phase deconvolution step to correct (only) the dominating coma aberration that is the primary cause the bending of the observed PSF. In the second stage we then apply a more conventional semi-blind deconvolution approach to correct the other aberrations and further improve image quality. In order to demonstrate the efficacy of this two-stage approach we compare it with a conventional semi-blind deconvolution method on both simulated and real data.

The remaining chapter is structured as follows. In Section 5.2, we briefly reviews the imaging model of BF microscopy and describe phase deconvolution. Section 5.3 describes the procedures for the improved estimation of the coma aberration and the novel two-stage method for correcting the aberrations due to the slanted scan. In Section 5.4, we describe our experimental set-ups with both simulated and real cytological image data.

5.2 Image Formation and Phase Deconvolution

5.2.1 Image Formation Model of Incoherent Imaging System

It is well known that the microscope image formation process can be modelled as is a convolution between the microscope optics (including illumination, glass slide, objective lens and imaging recording devices) and the physical specimen (object). And like most existing WSI systems, the slanted scan

system is based on a motorised BF microscope that has spatially incoherent (Kohler) Illumination. The 3D image formation process can be described as:

$$g(x, y, z) = o(x, y, z) \otimes h(x, y, z) + e \tag{5.1}$$

where the observed image g is the result of a convolution of the 3D specimen object o with the point-spread-function (PSF, the impulse response) of the system h, (x, y, z) are the spatial coordinates and e is acquisition (shot) noise. Within a BF microscope, h is the incoherent intensity PSF which is the squared magnitude of the complex-valued amplitude PSF, PSF_a, which is also known as the coherent PSF (Goodman 2005).

$$h(x, y, z) = |PSF_a(x, y, z)|^2 \tag{5.2}$$

The 3D psf_a can be found using the following equations:

$$PSF_a(x, y, z) = \iint P(k_x, k_y) \exp(2\pi i(k_x x + k_y y)) \exp(2\pi i k_z(k_x, k_z)z) dk_x dk_y \tag{5.3}$$

$$k_z(k_x, k_y) = \sqrt{(n/\lambda)^2 - (k_x^2 + k_y^2)} \tag{5.4}$$

where k_x and k_y are spatial frequency, $P(k_x, k_y)$ is the so-called pupil function in 2D, $k_z(k_x, k_y)$ is the defocus factor, n is the refractive index and λ is the illumination wavelength. The size of the pupil is defined as $\sqrt{k_x^2 + k_y^2} \leq NA/\lambda$, where NA is the numerical aperture as the objective lens.

As the convolution is in general implemented as a multiplication in the frequency domain for faster computational speed, the Fourier transform of the PSF is often utilised and is called optical transfer function (OTF). The 3D OTF of a bright field microscope is typically non-zero only within a toroid shape with dimensions dependent upon the sampling frequency and the NA of the objective lens (Hanser et al. 2004). In the Fourier domain the convolution operation in equation (5.1) becomes:

$$G(k_x, k_y, k_z) = O(k_x, k_y, k_z)H(k_x, k_y, k_z) \tag{5.5}$$

where, G, O and H are the Fourier transforms of g, o and h respectively.

From the above equations, it can be seen that the full response of the BF microscope can be described by the 2D pupil function. The 2D pupil function is complex valued with unit amplitude and a phase shift of zero when the system is ideal and free of aberrations. When the microscope comprises of aberrations, such as the coma aberration induced by a slanted specimen, then the phase shift of the system results in a deformation of the objects in the spatial domain. If the object is 3D, which it typically is, the phase shift in 3D is shown by a defocused phase shift within a defocused 2D pupil plane.

In optics, the phase and amplitude of the pupil function are normally described as Zernike polynomials (Hanser et al. 2004; Sheppard 2013). Specifically, coma aberration is the 8-th term in the Noll

index (Noll 1976). As the Zernike polynomials are orthogonal, the phase shifts generated for different coefficients can be combined to describe the pupil plane of any objective lens.

5.2.2 Phase Deconvolution

A deconvolution process can then be used to reverse (or invert) the image formation process in an attempt to recover the ideal image ($\hat{O}$) of the physical specimen (object), hence removing the effects of the optical aberrations. In microscopy, 3D deconvolution has been widely implemented for fluorescence microscopy, especially in wide-field microscopes, to recover the high-resolution 3D structure of specimens affected by out-of-focus blur (Sarder and Nehorai 2006). Here, the aim with a slanted specimen is not to remove all of the out-of-focus blur, but primarily to correct for the additional blur introduced by the slanted scan, e.g., the coma aberration, so that the resultant image is equivalent to an image scanned from a flat specimen. However, it may also be beneficial to remove some of the out-of-focus blur in order to improve contrast and resolution. By doing this, the end user of the digital slide, either a pathologists or a computer aided diagnostics (CAD) system, will be able to observe the same image quality and object features they are used to observing when making a diagnosis.

According to equation (5.5), the simplest way of recovering ($\hat{O}$) is to apply an inverse filter in the frequency domain. That is:

$$\hat{O}(k_x, k_y, k_z) = \frac{G(k_x, k_y, k_z)}{H(k_x, k_y, k_z)} \tag{5.6}$$

Unfortunately, the direct inverse filtering approach is problematic because a typical imaging system (H) is low-pass in nature and so the inverted system response ($1/H$) may be unstable, is high-pass and amplifies high frequency noise. To better understand this problem, the OTF (H) can be treated as the combination of two parts, namely the amplitude component, the modulation transfer function (MTF), and the phase component, called the phase transfer function (PTF). The MTF describes the contrast reduction effect of the imaging system at each spatial frequency, which approaches zero near the maximum passing frequency of the objective lens, as defined by the radius of the pupil. The noise amplification is a direct result of the inversion of the MTF, which amplifies high frequencies regardless of whether they are signal or noise. The PTF, on the other hand, describes the spatial shift of each frequency as they traverse along the optical path. Coma aberration, as an example, is primarily caused by a phase shift in the pupil plane, which eventually produces an optical aberration in the spatial (x, y, z) domain.

Although dedicated filters have been designed to reduce the effect of noise amplification of inverse filters, such as the Wiener filter and Tikhonov regularisation. These techniques specifically attenuate the noisy high frequencies, but are dependent upon an accurate estimate of the noise in the real imaging system, which is often difficult to obtain (Sarder and Nehorai 2006). An alternate approach, which seems particular suitable to removing coma aberrations, is to inverse filter only the phase component of the OTF, i.e., the PTF. By doing this, the amplitude of the observed image is completely preserved (effectively a unity MTF is assumed) and so noise is not amplified and the resolution is not degraded which is beneficial when processing high-resolution images. Importantly, an inverse filter of only the

phase delay is equivalent to introducing an equal but opposite phase shift to that which caused the coma distortion, i.e., the slanted slide. This approach is a digital equivalent of the correction method proposed in (Gerber and Mansuripur 1996), where a PPP tilted in the opposite direction is inserted into the optical path. We name this deconvolution with direct filtering using only the PTF *phase deconvolution*. The direct implementation of phase deconvolution is:

$$\hat{O}(k_x, k_y, k_z) = |G(k_x, k_y, k_z)| \exp(i(\phi_{G(k_x,k_y,k_z)} - \phi_{H(k_x,k_y,k_z)})) \tag{5.7}$$

To distinguish phase deconvolution with the conventional deconvolution using the complete OTF, we refer the conventional deconvolution as *intensity deconvolution* in this chapter.

5.3 The Proposed Methods

As previously described phase deconvolution, on its own, does not correct for aberrations described by the MTF of the optical system. Therefore, we propose a two-stage method to correct all of the aberrations introduced by slide slant. *Stage-1* aims to correct the distortion due to the dominating coma aberration by introducing an opposite phase shift using phase deconvolution. Only the coma aberration is corrected in the first step, but it is the largest aberration induced by the slide slant which causes the main visual distortion (the skewed PSF). As mentioned before, phase deconvolution does not amplify noise in the image, but neither does it improve the resolution of the image. Therefore, the output image from *stage-1* is further processed in *stage-2* utilising a conventional intensity deconvolution approach that removes the remaining aberrations (such as spherical distortion and astigmatism) and so improves image resolution and quality. Importantly, if it is assumed that different slides are slanted at the same angle, one can obtain an improved estimate of the coma aberration coefficient (Z_{coma}) from multiple specimens/slides, but this calibration step is only required for a particular slant angle, type of specimen and imaging system. In addition, as *stage-1* removes the most significant aberration (coma), the intensity deconvolution of *stage-2* can utilise a conventional PSF estimation technique by imaging a non-tilted (i.e., flat) specimen (PSF_{flat}).

5.3.1 Estimation of Coma Aberration

One complicating factor in BF microscopy is that 3D deconvolution is perhaps even more challenging than in fluorescence microscopy. This is because gaining an accurate estimate of the overall optical PSF is complicated by low signal-to-noise and the light reflected and refracted between the optics and the specimens (Tadrous 2010). Methods for estimating the PSF can be classified into three categories, namely analytical, experimental and computational (Markham and Conchello 1999). First, if the exact parameters of the imaging system parameters (e.g., NA, refractive index and the distance between optics and specimen) are known, the PSF of the system can be calculated using a suitable imaging model, such as the classical diffraction based Gibson and Lanni model (Gibson and Lanni 1991) (Kirshner et al. 2013). However, this method suffers when the system is used in non-design conditions,

such as when the slide is slanted, or when the compound lenses themselves contain intrinsic aberrations. In these cases, a better method is to estimate the PSF directly by imaging a small spherical bead having a size just below the resolution of the objective. The advantage of this approach is that the measured PSF contains both the intrinsic and extrinsic aberrations. However, the images acquired from sub-resolution beads have a low signal-to-noise ratio (SNR), which limits their applicability to high resolution deconvolution (Tadrous 2010). A further problem with this approach is that the PSF obtained from a sub-resolution sphere may differ from the effective PSF when imaging real specimens due to differences in the thickness and refractive index of the slide, cover-slip or specimen. However, recently a non iterative deconvolution method has been developed that estimates the PSF directly from real specimens (Tadrous 2010). The third and most general approach is to use blind deconvolution that estimates the PSF and the (deconvolved) image of the object simultaneously (Holmes and O'connor 2000; Kim and Naemura 2015).

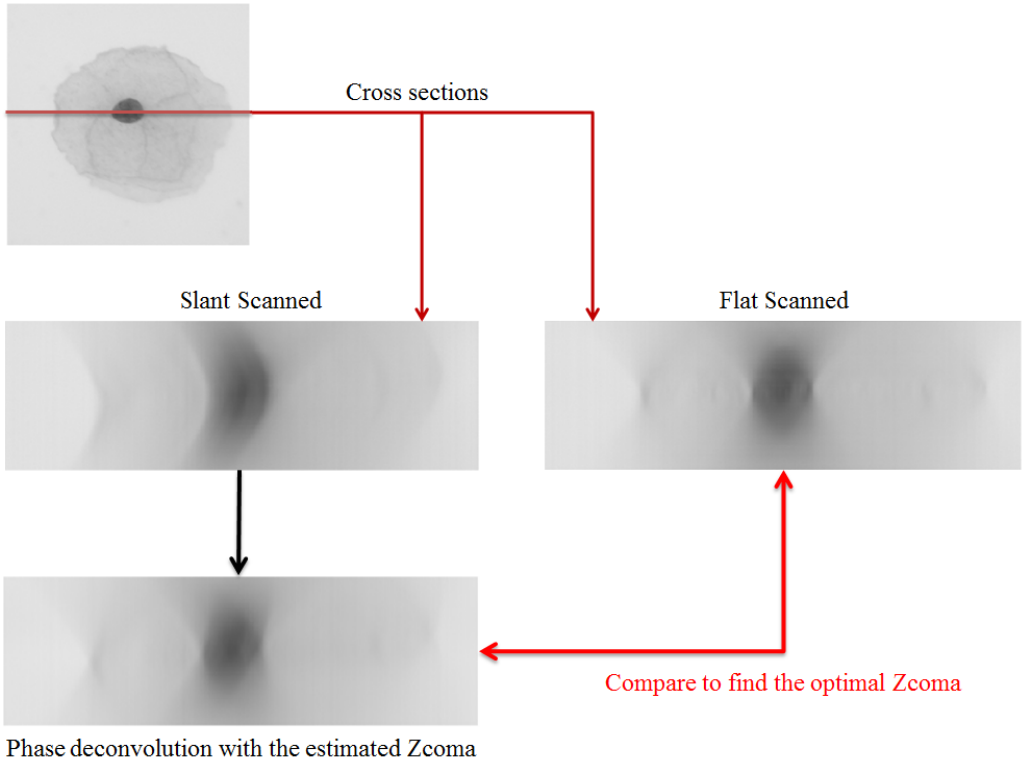


Figure 5.1: The slanted scan image stacks are deconvolved with the estimated OTF and compared with the flat scanned image stacks to find the optimal OTF that caused the amount of coma aberration induced by the slant of the slide.

The estimation of coma aberration in this study adapts the third approach, which is to use the phase deconvolution method and real specimens to find an optimal coma coefficient that minimises the degree of bending of the PSF. The principle of the calibration procedures is based on the assumption that when the correct amount of phase shift is introduced to the slanted scan images, the bending in the PSF will be minimised so that the shape of the PSF will then be the same as the PSF observed without coma aberration (i.e., specimen imaged flat). Specifically, the difference between the PSF estimated via the phase deconvolution of a slanted specimen and directly from a flat scanned image is minimised when the coma aberration is correctly estimated for that particular slant angle. This idea

is further illustrated in Fig. 5.1 and described in detail in the following.

To recover an accurate estimate of the coma coefficient, real specimens are used as the imaging target. Specifically, thin and free-lying (squamous) cervical cytology cells are selected. The main reason to select a thin object is because it is necessary for estimating PSF_{flat} utilised in the second stage intensity deconvolution (described later). To avoid artefacts due to the image edges when performing Fourier Transforms, the selected cells were all free-lying with no other cellular objects visible in the image edges.

The method for finding the optimal coma coefficient of a particular slant angle is described as follows:

(a) Image Acquisition and Pre-processing: The specimen cells are first imaged flat and then imaged slanted, utilising a slide holder specifically designed and constructed to slant the slides. For the acquisition of each cell, the in-focus depth of the nucleus is used as the midpoint of the acquired image stack and the same number of image slices are imaged above and below this central plane. The in-focus depth is found using standard auto-focusing method that greedily searches for a focal depth that maximises the normalised variance of the image (Sun et al. 2004). The distance between the slices (z-interval) is set to be the same as the lateral pixel size so that the final volumetric (3D) data is isotropic. The height of the stack is not strictly limited, but in the bottom and top slices the cell image is completely out-of-focus. When the specimen was slanted, the cell will naturally occupy a larger focal depth than when imaged flat, and so the image stacks for the slanted scan are higher in number. To reduce the background noise, multiple images (20-30) are acquired at each focal depth and averaged (Tadrous 2010). We also band-limited each acquired image by setting the frequency components outside the pupil plane to zero to further reduce the noise. In addition, the rotational angle between the camera and the microscope stage needs to be pre-calibrated and corrected after the image acquisition so that the image axis aligns with that of stage axis. This is an important procedure as the coma aberration is directional, in the same direction as the slant angle. Other standard pre-processing techniques applied include flat-field correction for uneven background illumination and intensity correction between slices due to illumination fluctuations (Holmes and O'connor 2000). To reduce computational burden, the image stacks are cropped into smaller images, around 512x512 pixels in size, with the individual cells located in the centre of the image and no other objects in the image. The intensity of the whole 3D stack is finally inverted by plusing the maximum intensity of the stack prior to deconvolution (Holmes and O'connor 2000).

(b) Phase Deconvolution: We first build the pupil plane containing only the estimated coma aberration:

$$P(k_x, k_y) = Aexp(i * Z_{coma} * W_{coma}(k_x, k_y)) \tag{5.8}$$

where Z_{coma} is the estimated coma aberration coefficient, A is the circular pupil with unit amplitude (assuming no aberration so cut-off at maximum resolving frequency of the objective), and the W_{coma} is the phase map of the lowest-order coma aberration as described in Zernike polynomials (Noll 1976). The pupil is used to construct an OTF using equation (5.2) for phase deconvolution (5.7). The phase deconvolution is performed in the frequency domain and finally transformed back in to spatial domain

in order to compare it with the flat image stack.

(c) Post-processing: As mentioned before, the specimen cells are slanted with the slide at an angle, the stacks from the above step have to be corrected so that the origin of the specimen is the same as the flat scan. Here, we use simple cubic resampling in spatial domain to rotate the orientation of the slanted slides. First, we resample along the Z (optical) axis to remove the tilt from the specimen. Then, we resample the specimen along the direction of the tilt, as the projection of the slanted specimen is shortened.

The resampled stacks are then aligned with the flat stack in 3D before the comparison. The in-focus slices of both the resampled slanted stack and the flat stack are first found by again finding the slice in the whole stack that has the maximum normalised variance (Sun et al. 2004). The translation between the two in-focus slices is then used to align these two stacks laterally (using normalised cross correlation (Lewis 1995)). As the two in-focus slices are pointing towards the centre of the same specimen, the axial positions of these two slices can be used to align the stack axially. Finally, the intensities of both the processed slanted scan stack and the flat scan stack are normalised.

The optimal coma coefficient for this particular slide slant angle is found when the root mean square error (RMSE) between the resampled stack and the flat stack is minimized. This RMSE function is convex and we apply a generalized pattern search algorithm to find the optimal coefficient value (Audet and Dennis Jr 2002). This calibration operation only need to be done once for one particular slide slant angle and the imaging set-up.

5.3.2 Two-stage Algorithm for Aberration Correction

The specimens are first imaged with the slide slanted and pre-processed using in same way as described in Section (a) above (5.3.1) except now only a single image is acquired at each focal depth (no image averaging is performed).

Stage-1: The first stage of the algorithm aims to remove the phase distortion due to the coma aberration using phase deconvolution. After the optimal coma aberration coefficient (Z_{coma}) for an particular angle is found, the coefficient is used to generate the OTF using the equations (5.2) and (5.8). The OTF is then used for phase deconvolution of the input image as per equation (5.7). Note, no frequency constraints are required because the magnitudes of the signal is not modified.

Stage-2: As previously mentioned, the phase deconvolution only corrects the phase distortion introduced by the coma aberration, but the magnitude of the signal is not changed, hence the image contrast may still be poorer than that of a flat scanned image. The second stage of the algorithm applies an intensity deconvolution algorithm to further improve the contrast of the image stack.

The 3D deconvolution technique is generally used in fluorescence microscopy in order to remove the out-of-focus blur and improve the contrast and resolution of the weak fluorescence signals. If the deconvolution is carefully implemented (a suitable method is used and an accurate PSF is estimated), the out-of-focus blur can be significantly removed resulting in superior image quality and higher resolution, that are comparable to images acquired on a confocal microscope (Sarder and Nehorai 2006; Soulez et al. 2012). However, our aim here is to improve the image contrast to the level where the

images are equivalent to those acquired from a BF microscope and a perfectly flat specimen. So, we decided to use iterative intensity deconvolution methods for the *stage-2*, so that we can control the amount of deconvolution applied and hence avoid the noise to be amplified to unacceptable levels. In other words, the extent of image contrast improvement is controlled by the number of iterations the deconvolution algorithm performs.

Conventional deconvolution methods require an accurate estimate of the PSF of the imaging system to produce high quality results. However, here in the proposed two-stage method the PSF is not directly accessible because the PSF after the initial phase deconvolution is an unknown intermediate state. Here, we use the PSF of the original microscope (PSF_{flat}) as the initial estimate of this intermediate PSF and utilize the Lucy-Richardson method to update it after each iteration of deconvolution. An estimation of the microscope PSF is used because it contains intrinsic aberrations of the objective lens, which is more accurate than a theoretically estimated one. In this study, we use the well-known Lucy-Richardson deconvolution method because the method was previously demonstrated to work well in bright field microscopy (Holmes et al. 1995). For comparison, the regularised deconvolution method based on total variation was also utilized (as discussed in the last section) but the overall effectiveness of the two-step method is initially evaluated with the Lucy-Richardson algorithm. In this way, the proposed two-stage method is directly compared with a well known and competitive one-stage algorithm that performs the deconvolution directly from the estimated PSF with slide slant.

The equations for updating the image and the PSF are:

$$f_{j+1}(x,y,z) = f_j(x,y,z)[h_j(-x,-y,-z) \otimes \frac{g(x,y,z)}{f_j(x,y,z) \otimes h_j(x,y,z)}] \tag{5.9}$$

$$h_{j+1}(x,y,z) = f_j(x,y,z)[f_j(-x,-y,-z) \otimes \frac{g(x,y,z)}{h_j(x,y,z) \otimes f_j(x,y,z)}] \tag{5.10}$$

where, $f(x,y,z)$ is the estimated 3D specimen image, $g(x,y,z)$ is the observed blurred image, h is the PSF and the j is the iteration number.

The PSF_{flat} is obtained from method proposed in (Tadrous 2010), in which the in-focus slice of a thin-flat object and its blurred 3D images are used to generate the PSF via a one step least-squares deconvolution. To achieve good quality images, the target specimen has to be thin and flat so that the central image can be equivalently used as the ideal in-focus image. To reduce noise, several such specimens are used and the generated PSFs are averaged. In this study we utilise squamous cervical cells, which because of their thickness does meet the assumption that the central image approximates the ideal in-focus image. Therefore, we create an extended-depth-of field (EDF) image to approximate the ideal in-focus image (Bradley and Bamford 2004). The final PSF_{flat} is used as the initial value of h, i.e., h_0 in the intensity deconvolution.

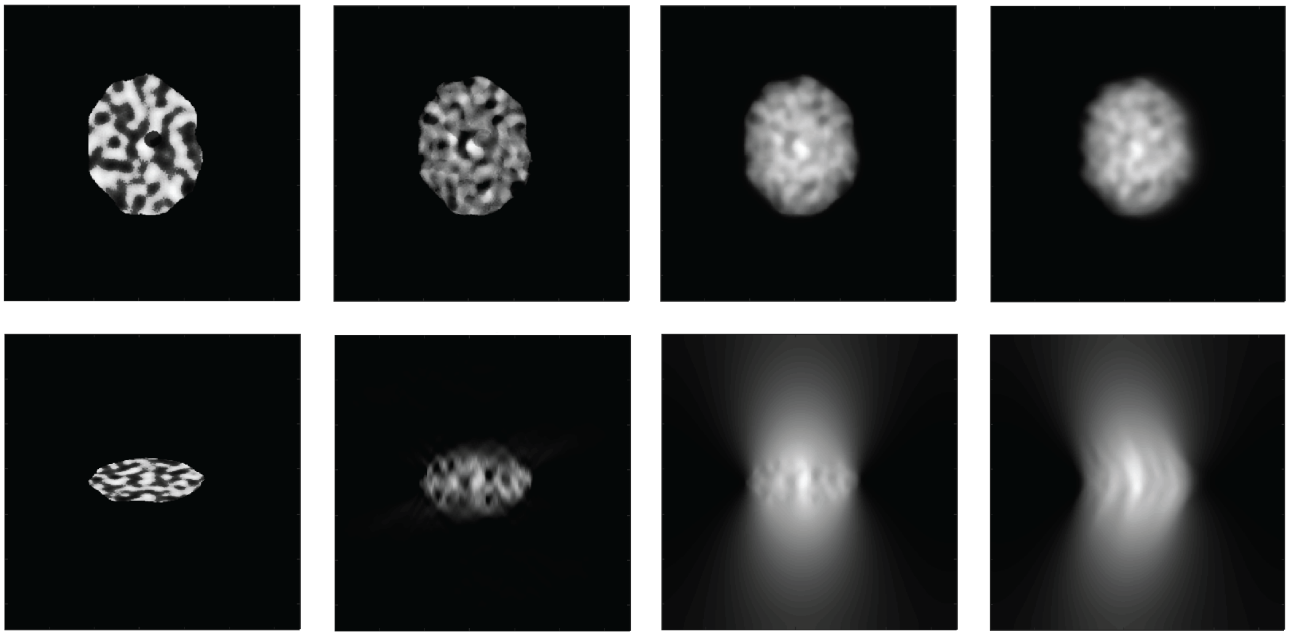
To enforce convergence, the PSF in equation (5.10) is constrained in the spatial domain based on the shape of a bright field PSF, and the image $f(x,y,z)$ is band-limited by the maximum passing

frequency of the imaging system at each iteration (Holmes and O'connor 2000) Due to the size of the 3D data, all the deconvolution operations were implemented in the frequency domain.

Post-processing: The post processing resamples the stack against the slant angle of the slide in order to correct the orientation of the specimen into the flat position. This resampling process is the same as described in (c) of Section 5.3.1). As the intensity deconvolution increases the contrast and dynamic range of the image, the intensity of the foreground specimen is scaled so that its maximum intensity is the same as that in flat scanned images, or alternatively to an pre-defined intensity level.

5.4 Experimental Methodology

5.4.1 Simulation Data



(a) (b) (c) (d)

Figure 5.2: (a) shows the central *XY* (top row) and *XZ* (bottom row) slice of the original synthetic nucleus. The diffraction limited nucleus image data (b) is then blurred by an ideal PSF (c) and a PSF with aberrations induced by slide slant (d). Assuming the slide slant is in the *XZ* plane and the optical axis is *Z*. All images are normalised to have a maximum intensity of 1.0.

To evaluate the performance of proposed 3D deconvolution method, a synthetic 3D nucleus (Hela Cell) generated for 3D deconvolution of fluorescence microscopy was used. The original dataset and generation parameters can be found at (Keuper et al. 2012; Svoboda et al. 2009). The dimension of the data stack is 328x328x328, and each voxel is assumed to be a square of size of 0.15μm. Next, assuming the data stack was imaged by a diffraction limited microscope, which has a sampling resolution of 0.15μm, a numerical aperture of 0.75 and illuminated by light with wavelength of 525nm (which is the same as our experimental microscope described later), we impose a constraint on the frequency domain of the 3D data so that the maximum passing frequency is below the diffraction limit of our

synthetic microscope. Cross-sections of the high resolution nucleus and the diffraction limited nucleus are shown in Fig. 5.2.

In practice, the 3D cell nuclei can not be directly imaged with a BF microscope because the out-of-focus blur can be seen across all slices in the stack. In contrast, the diffraction limited cell nucleus could be imaged with an confocal microscope or indirectly estimated by deconvolution of a wide-field fluorescence microscope (with equivalent resolution power, assuming the same objective is used in the BF microscope). So, the original data stack containing the diffraction limited cell nucleus ((b) in Fig. 5.2) was used as ground-truth ($S_{groundtruth}$) to evaluate the ability of the deconvolution methods to recover the 3D structure of the object. The ground-truth of a BF microscope, on the other hand, is generated by applying convolution between the 3D nucleus with an ideal PSF of the microscope, noted as S_{flat}, which is necessary to test the outputs on recovering the specimen to the point equivalent to a conventional flat scan.

PPP Tilt Angle	Primary Astigmatism	Primary Coma	Trefoil	Secondary Coma
Zernike Coe Order	6	8	10	16
3.2	0.0237	-0.2558	-0.0002	-0.0169
4.2	0.0409	-0.3360	-0.0005	-0.0221
5.3	0.0652	-0.4244	-0.0010	-0.0279

Table 5.2: The estimated aberration coefficients.

To simulate the effect of aberrations induced by the slide slant at various angles we used the formulas proposed in (Sheppard 2013) to calculate the estimated aberrations when the slide is slanted at three angles, namely, 3.2 degrees, 4.2 degrees and 5.3 degrees. These are same angle as the experimental slanted slide holders described later. Then, we constructed a PSF containing these aberrations and convoluted them with the $S_{groundtruth}$ to obtain a data set S_{slant} simulating the nucleus imaged in a slanted scan at each slant angle. The nucleus images that estimate the slide slant of 4.2 degrees are shown in Fig. 5.2 (d). In all simulations, the slide slant is assumed to be in the XZ direction, where Z is the optical axis and a negative coma corresponds to a slide slant in an anti-clockwise direction from the X axis. The analytical aberrations of all three angles are listed in Table 5.2 assuming the same imaging system as described above and a cover-slip thickness of 145μm. Equations (5.2), (5.3), (5.4), and (5.8) were used to generate the PSFs containing these aberrations.

The first experiment with the simulation data aims to test the effectiveness of proposed calibration method and to explore the objective function. The three datasets estimating the slanted scan of the nucleus are used to recover the true coma aberration using the proposed method described in Section 5.3.1. Specifically, a list of coma coefficients spanning 0.12 above and below the true coma coefficient and spaced at 0.02 were tested. For example, if the true coma was -0.32 for a slant angle 4.2 degrees, the testing range was -0.32-0.12 to -0.32+0.12 at an interval of 0.12. With the simulation data, the intensity normalisation was not required to be implemented and the numerical metric for comparison was peak signal-to-noise ratio (PSNR), which is effectively RMSE in dB. For the phase deconvolution, the estimated OTFs contained only the coma aberration. Zero mean Gaussian noise of power 15dB was added to the S_{slant} stack in order to verify the effect of noise on the deconvolution

methods. The PSNR was calculated as:

$$PSNR = 20log_{10}\left(\frac{MAX_{S(x,y,z)}}{\sqrt{\frac{1}{mnd}\sum_{x=1}^{mnd}\left[S(x,y,z)-f(x,y,z)\right]^2}}\right) \tag{5.11}$$

where the $S(x,y,z)$ is the ground-truth and $f(x,y,z)$ is the processed 3D image.

The second experiment with the simulation data aims to evaluate the effectiveness of the proposed two-stage method in comparison with the conventional intensity deconvolution method. The conventional method is a single stage of the standard deconvolution performed with the PSF estimated from a slanted scan. For the proposed two-stage method, *stage-1* uses the OTF containing only the coma aberration, and *stage-2* uses the actual PSF used to generate S_{flat}. For the standard deconvolution, the PSF used was the actual PSF used to generate S_{slant} and containing all of the aberrations listed in Table. 5.2. As the actual PSFs were used, the full deconvolution only estimated the object at each iteration running only equation (5.9) and omitting equation (5.10). In order to judge the effectiveness of the method at recovering both the flat stack and the ground-truth, the output images were quantitatively compared by calculating the PSNR against the flat stack S_{flat} and the ground truth ($S_{groundtruth}$). Zero mean Gaussian noise of power 15dB was added to the S_{slant} datasets in order to verify the effect of noise on the methods.

5.4.2 Cytology Specimens

To verify the performance of the proposed algorithm on processing real specimens, several cervical cytology specimens were used in this study. The specimens were PAP stained and cover-slipped (# 1 coverslip 130-160 μm in thickness) using an automated mono-layer slide preparation system (ThinPrep 2000 Processor). All scanned slides were classified as normal by an experienced cytopathologist.

The specimens were imaged by a motorized BF microscope (Zeiss Imager M.1) with a Plan fluo 40$\times$ objective (NA = 0.75). The light from halogen lamp was band-pass filtered by a green light filter (525$\pm$25nm) to achieve the highest contrast on imaging PAP stained specimen. An Abbe condenser with the same numerical aperture of the objective was used to ensure the illumination was fully spatially incoherent. A charge-coupled device (CCD) camera (Kodak KAI-2020-OM sensor from SPOT Imaging) having a size of 1600x1200 pixels was used to acquire all specimen images. The pixel size of the imaging sensor is 7.4 μm, and a 1.2$\times$ relay lens was installed before the sensor to ensure the Rayleigh sampling criterion was met. The final image sampling resolution was at 0.15μm (giving an effective image resolution of $\approx$0.25μm) and all acquired images were in grayscale in this study.

Several slide holders, as shown in Fig. 5.3, were made to slant the slide at three different angles: 3, 4 and 5 degrees. The actual slant angle of each slide holder was measured before the acquisitions by calculating the angle between two cell nuclei. Specifically, let (x_1, y_1, z_1) and (x_2, y_2, z_2) be the positions of the stage when the two in-focus cell nuclei locating at the middle of the FOV, the angle between the line passing two cell with respect to the stage horizontal can be found by $\theta = tan^{-1}(z_1 - z_2, x_1 - x_2)$, assuming the slant direction is along the x axis and the y_1 and y_2 are approximately equal by selecting two cells along the same line y. If θ_{flat} is the angle of the line passing the two cells when

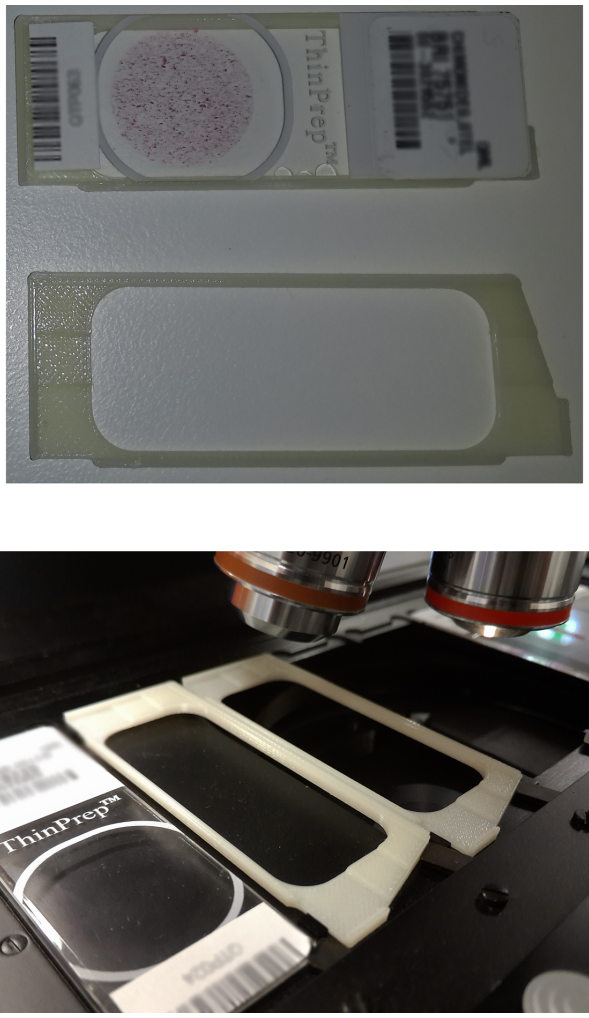


Figure 5.3: Top: Custom made slide holders for slanting the slide along the short side at a small angle. Bottom: The slide holders on the slide tray of the experimental microscope.

the slide is flat and θ_{slant} is the angle of the line passing the same two cells when the slide is slanted by one of the three slide slant holders, then the slide slant angle of the slide holder can be calculated from $\theta_{slant} - \theta_{flat}$. Three slides and three pairs of cells on each of the slide were used to calibrate each slant slide holder, and the measured average angles for the three slide holders were 3.2, 4.2 and 5.3 degrees respectively. Given the size of the imaging sensor, the range of focal depths can be seen on each FOV were 13μm, 15μm and 22μm respectively. According to the study in Chapter 4, the median thickness of this type of specimen is 9μm and 95% of cell clumps had thickness less than 19 μm, so these slide holders were sufficient to image most of thick cell clusters.

To calibrate the coma aberration coefficient at the three slant angles for the *stage-1* and the PSFs for *stage-2*, twenty free-lying and non-folded squamous cells from three LBC Thin-prep slides were selected and imaged. Specifically, 30 images with the same exposure time were acquired at each focal depth and averaged as an single image to reduce the image acquisition noise, and the FOV size of the image stacks containing free-lying cells were 512×512 pixels. The height of the flat stacks was 181 slices (spanning 27.15μm) ensuring the cell nuclei were sufficiently (out-of-focus) blurred at the top and bottom slices, and all the stacks were sampled with a Z-interval of 0.15 μm to ensure they were isotropic. The search ranges of the estimated Z_{coma} for each slide slant angle was bounded by the

analytical results presented in Table. 5.2. Those squamous cells were also used to obtain the PSF_{flat} and PSF_{slant}, where the slant PSFs (PSF_{slant}) were generated for each slide slant angle using the same method as that for producing the PSF_{flat}.

To evaluate the effectiveness of the proposed two-stage algorithm for aberration correction, images of the twenty squamous cells used above and an additional thick cell cluster were processed by the proposed method. As the phase deconvolution step is based on the pre-calibrated procedure described in Section 5.3.1 the images of these cells did not require averaging and so contain the native noise level of the acquisition system. The 3D data was also isotropic at spatial resolution of 0.15μm. The output images from the proposed method were qualitatively evaluated by visual examination.

5.4.3 Effect of Down-sampling the *Z*-axis

In 3D deconvolution of fluorescence stained specimens, the specimens are generally acquired in isotropic fashion to ensure their maximum continuity and hence the best 3D reconstruction results. However, in BF based slide scanners for WSI, it is extremely slow and hence not practical to acquire images at spacing close to optical resolution limit (0.25μm), because hundreds of images are required for thick specimens like cytology specimens. The distance between the slices are normally set to be approximately equal to the DOF of the optics, which is around 1 micron for the experimental microscope system, because within this range the cellular objects are in clear and sharp (Donnelly et al. 2013). The effects of increasing the distance between the 3D stack slices on the proposed correction method are investigated by down-sampling the stack along the *Z*-axis to be multiples of the DOF.

5.5 Results

5.5.1 Simulation Data

Fig. 5.4 shows the numerical results of the evaluation of the calibration method. Clearly, when accurate coma coefficient Z_{coma} were used, the corresponding PSNRs were the highest among all estimated Z_{coma} for all three slide slant angles. From the side images showed in the bottom of the Fig. 5.4, the blur of the nucleus after the phase deconvolution using the true Z_{coma} had nearly no bend and the side image (in the middle) was the closest to the S_{flat} showed in Fig. 5.2.

Fig. 5.5 shows the experiments with simulation data comparing the performance of two methods. When the output data stacks were compared with the flat stack, the outputs from the *stage-1* achieved the highest PSNR (iteration zero in left two graphs of Fig. 5.5), both with and without noise. Although the PSNRs then gradually decreased once *stage-2* was applied, the outputs from the both methods are closer to the ground-truth shown by an increasing PSNR (in the right column of Fig. 5.5). Importantly, the outputs from the proposed method achieved higher PSNR than from applying the conventional method at all iterations.

The above comparisons can be better illustrated by visualisation of output images as the deconvolution went on, shown in Fig. 5.6. It can be clearly seen that the bending of the PSF due to the coma aberration (as shown in Fig. 5.2) is almost entirely removed by the *stage-1* phase deconvolution. On

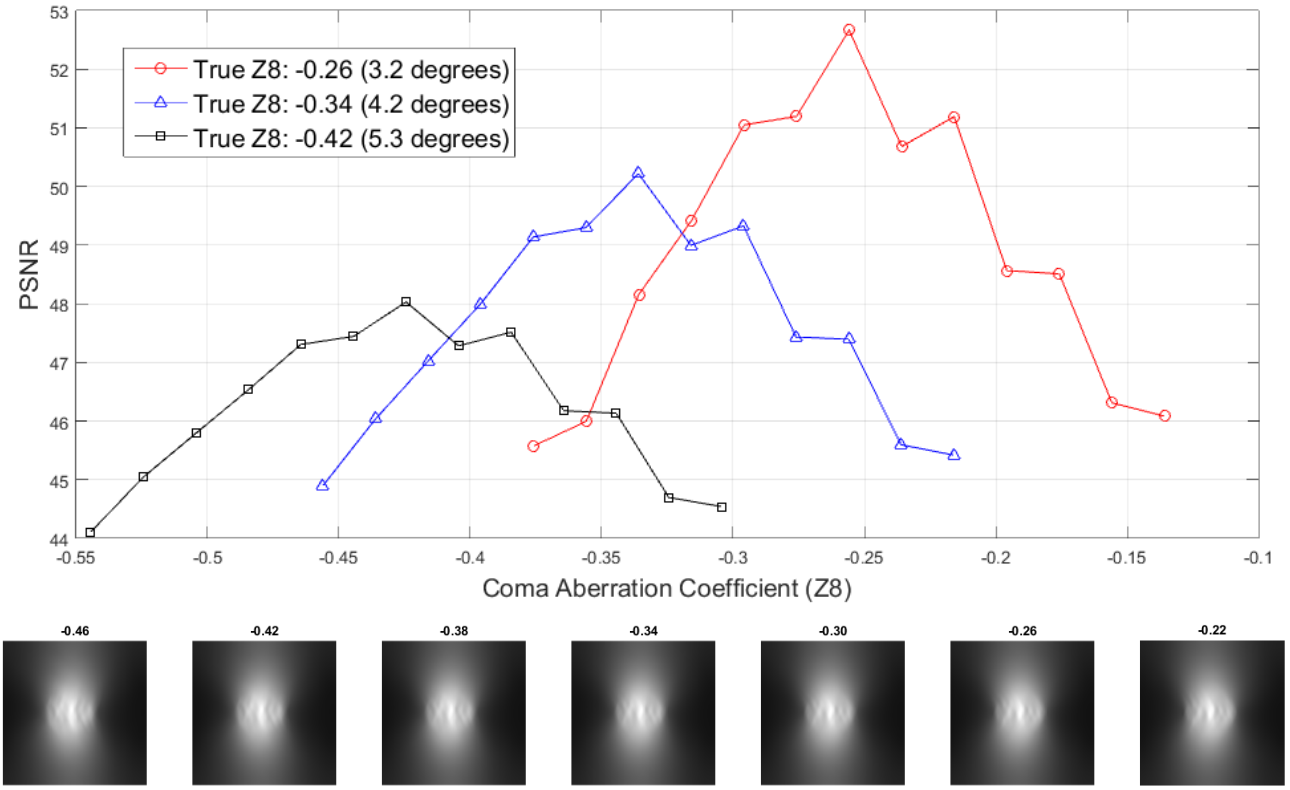


Figure 5.4: Top: The experimental results for calibrating the three estimating slide slant angles (3.2, 4.2 and 5.3 degrees). The PSNR is the highest when correct coma aberration coefficients Z_{com} were used. Bottom: The side image (XZ) of the slanted scanned nucleus (4.2 degree) after the phase deconvolution with various amount of coma aberrations (magnitude in Zernike coefficient labelled at the top of each image).

the other hand, the intensity deconvolution removes the blur and increases the contrast as expected, and the more iterations the closer the outputs to the ground-truth. However, the remaining blur is asymmetric along the Z-axis with the conventional method, while the blur is far more symmetric and close to that of flat scanned with the proposed method. In other words, the skew of the PSF due to the coma aberration was gradually removed using the conventional method, but the skew was completely removed with a single application of phase deconvolution.

5.5.2 Cytology Specimens

Fig. 5.7 shows the box plot of the calibrated coma coefficient Z_{coma} for the three slanted slide holders using the 20 cytology cells and the proposed method. The median coma aberrations Z_{coma} was -0.33 for 3.2 degree slide holder, -0.42 for 4.2 degree slide holder and -0.52 for 5.3 degree holder. The IQR for the three angles were 0.04 (3.2 degrees), 0.07 (4.2 degrees) and 0.08 (5.3 degrees) respectively.

Fig. 5.8 shows results of *stage-1* phase deconvolution of an example thin specimen. It can be clearly seen that the bending of the nucleus at all three angles have been effectively removed and the shape of the blurred nuclei were almost identical to that of free of coma aberration (imaged flat).

Fig. 5.9 shows the result of *stage-1* phase deconvolution on a thick cell cluster. From the side image (XZ) crossing the nucleus of two cells, the skewed blurs of nuclei were almost completely

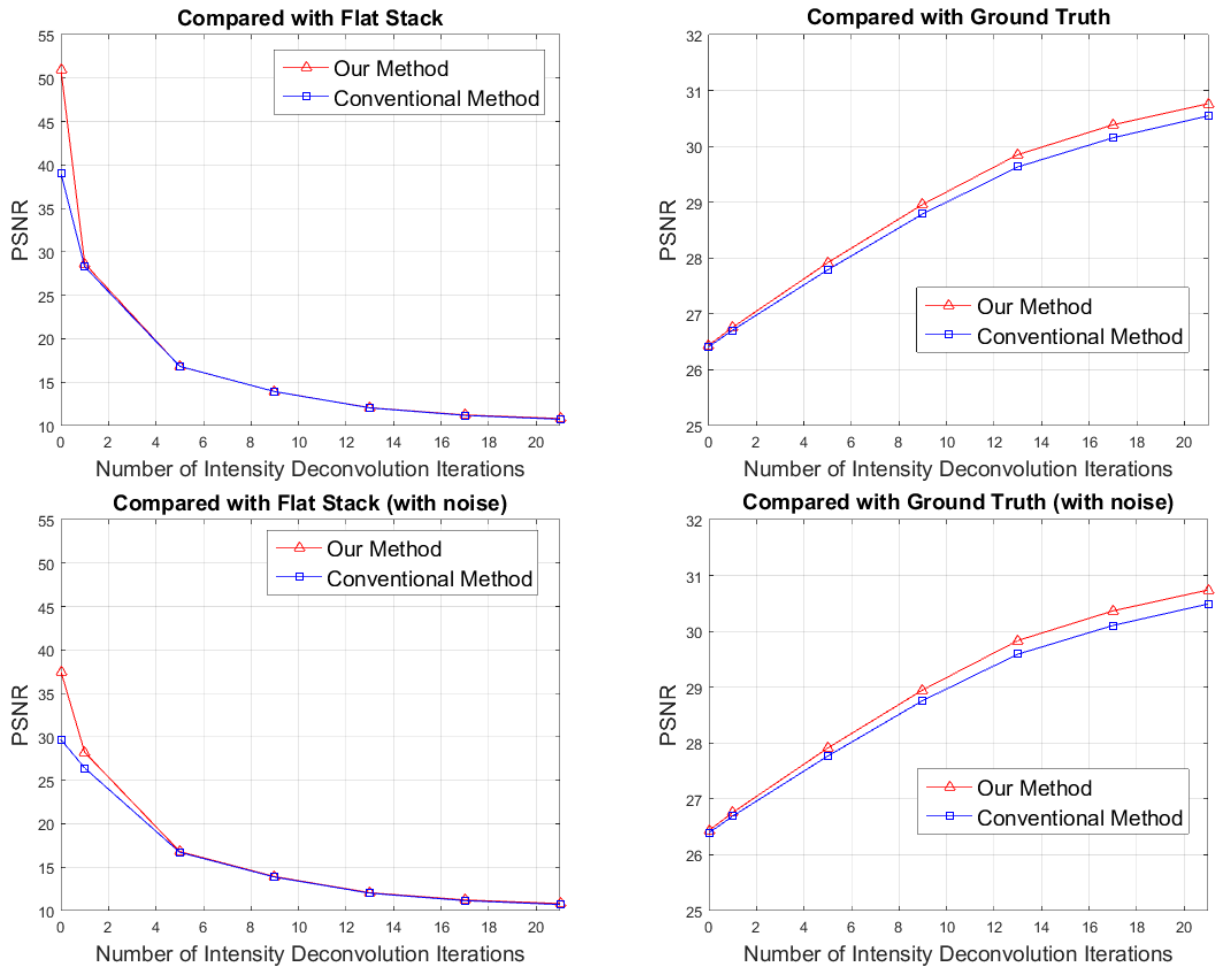


Figure 5.5: Both the outputs of *stage-1* (zero iteration) and *stage-2* (iterations 1 to 21) outperform the conventional method when compared with the flat stack (left-column) and ground-truth (right-column). The original stack simulated slide slant at 4.2 degrees. The evaluation using the original stack is shown in top row, and with addictive 15dB white Gaussian noise at bottom row. The zeroth iteration of the deconvolution means the implementation of the *stage-1* phase deconvolution only.

corrected for the cells at different depths (and hence different aberrations).

Fig. 5.10 shows output of intensity deconvolution of the same cell showed in Fig. 5.8. The input images were the cell imaged on a slanted slide holder of 3.2 degrees. It can be seen that the contrast of the cell after phase deconvolution was poor. However, it improved with further iterations of *stage-2* intensity deconvolution, as seen by richer details in cytoplasm texture. Interestingly, after a sufficient number of iterations (around 20), the contrast of the cytoplasm is even better than that scanned flat. In terms of the skewed PSF caused by the coma aberration, the skew was almost completely removed by the phase deconvolution while the bending is still visible after 20 iterations using the conventional deconvolution. It worth to noting that the images in the Fig. 5.10 were re-sampled to correct the orientation of the specimen, so the shape of the blur of the nucleus is closer to the flat scanned compared to that shown in Fig. 5.8.

To show the performance *stage-2* intensity deconvolution when improving the image contrast for different slide slant angles, the final output data stacks after 20 iterations of intensity deconvolution are shown in Fig. 5.11. The 3D stacks were processed into a single image in EDF, where the single

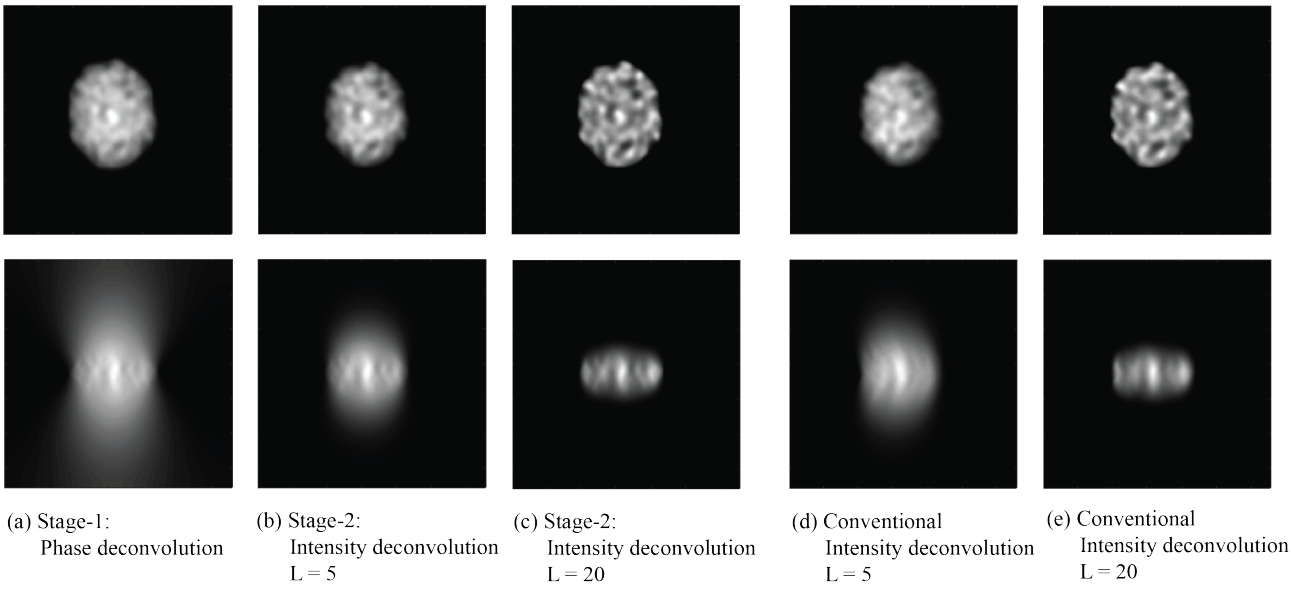


Figure 5.6: The middle (no. 164) *XY* (top row) and *XZ* (bottom row) slices of the outputs of the simulation data (blurred by the slant PSF) in Fig. 5.2 from the proposed two-stage method (a, b and c) and the conventional intensity deconvolution method (d and e). The actual PSFs were used to process the synthetic data. The data stacks was normalised and the dimension of these stacks were 328×328.

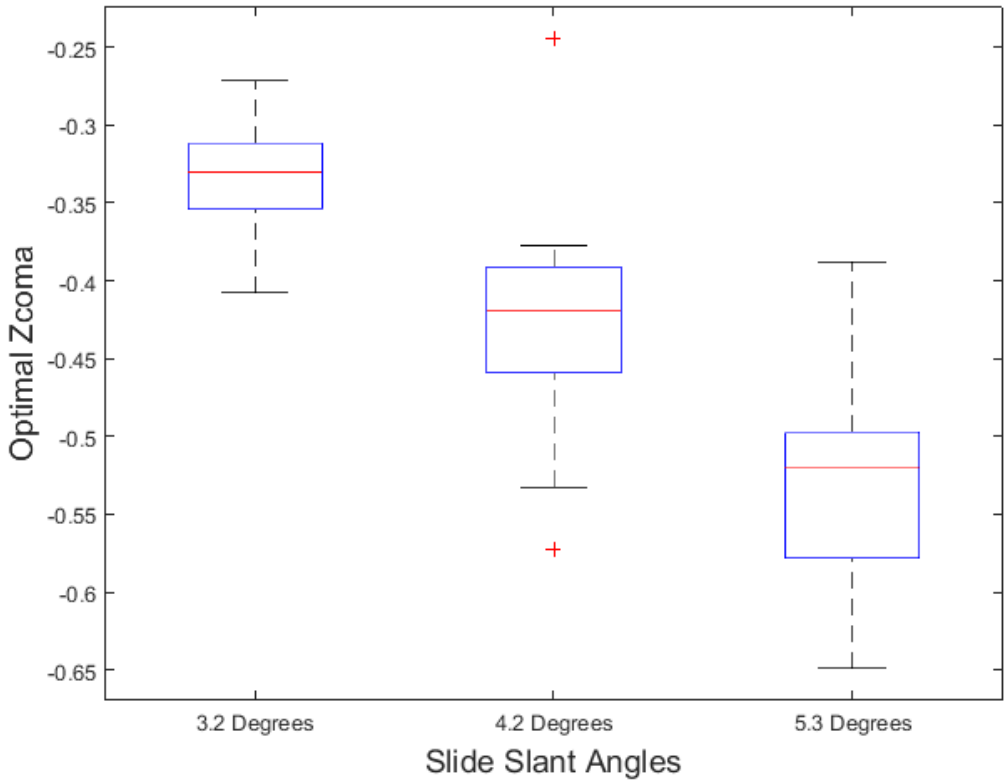


Figure 5.7: The box plot of recovered coma coefficients in Zernike (Z8) of 20 cells scanned at three different angles (3.2, 4.2 and 5.3 degrees).

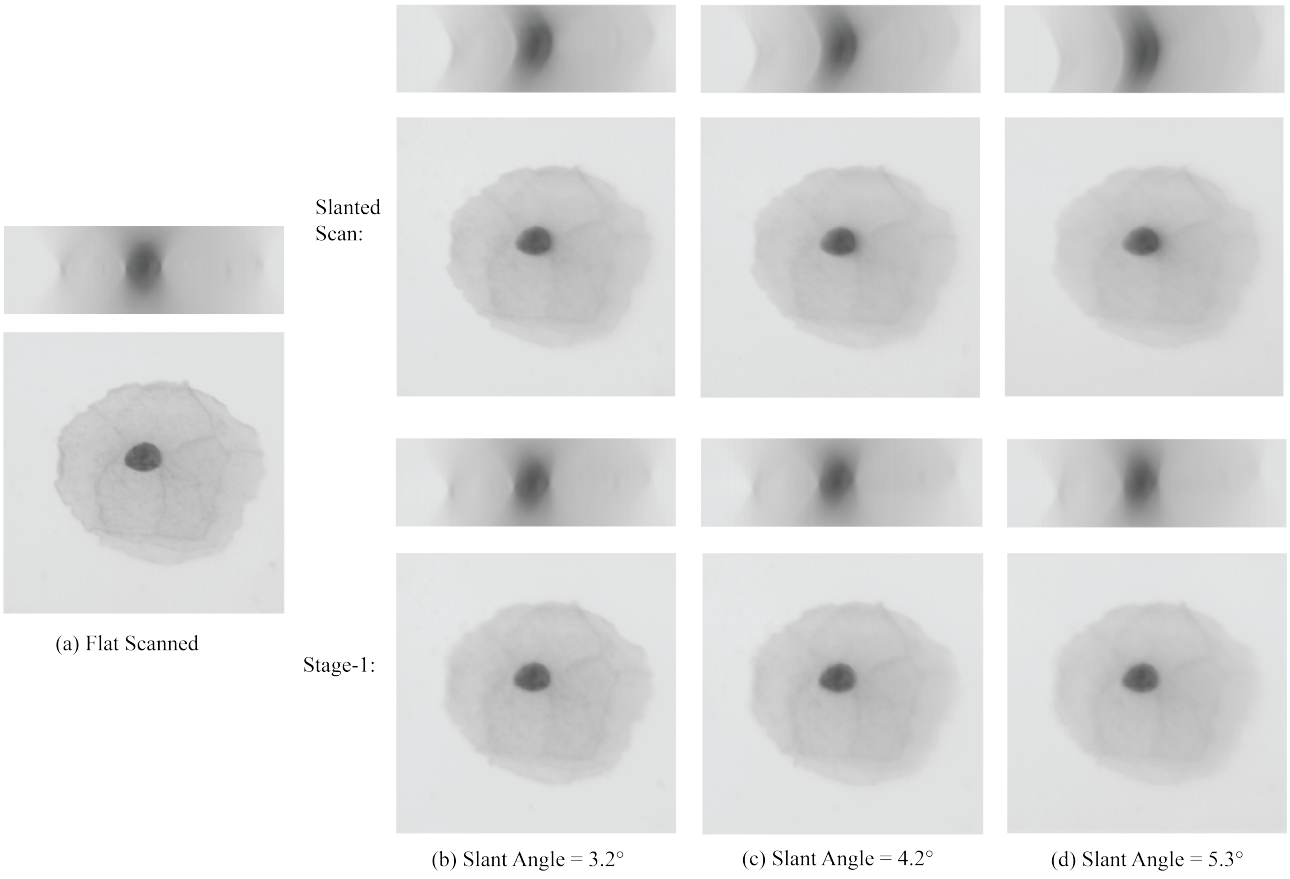


Figure 5.8: Phase deconvolution is able to effectively remove the bending of the PSF induced by slide slant at various angles. The first row of images show the *XZ* and *XY* slice of the image stack at different slide slant angles and the second row shows the the same slices after *Stage-1* phase deconvolution. The dimension of the *XZ* slices were 512×161 and *XY* slices were 512×512 with resolution of 0.15μm in all three dimensions.

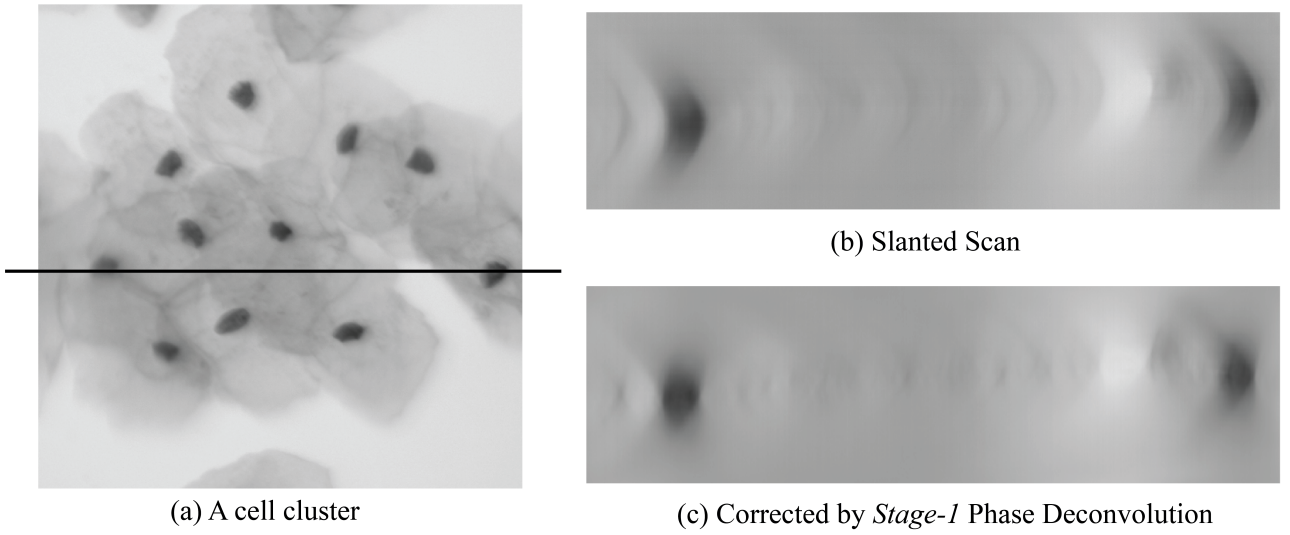


Figure 5.9: The *XY* image (a) and *XZ* side image (b) of a cervical cytology cell cluster acquired when the slide was slanted at 4.2 degrees, and the side image (marked by the dark line in (a)) after *Stage-1* phase deconvolution (c). The Z_{coma} used for phase deconvolution was -0.41. The dimension of the original image stack was $1024 \times 1024 \times 291$ with resolution of 0.15μm in all three dimensions.

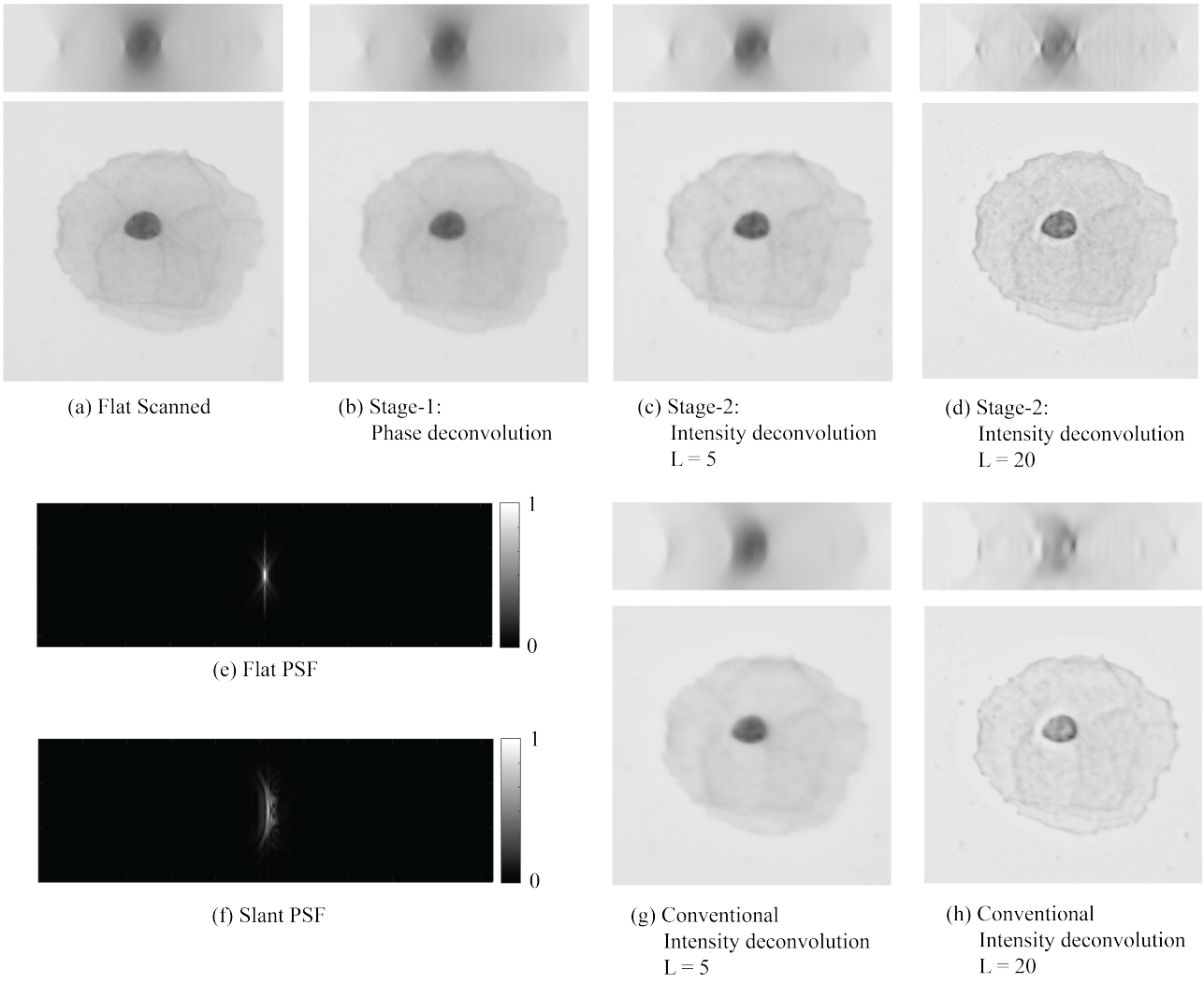


Figure 5.10: The middle slices *XY* (top row) and *XZ* (bottom row) of the same specimen as in Fig. 5.8 imaged flat are shown in (a). The central slices of the same cell processed by the proposed 2-stage method (d, c and d) and the conventional intensity deconvolution method (g and h). The PSFs used to for the intensity deconvolution are show in (e) and (f). To better visualise the PSF, the PSF stack was normalised and the square root of the values shown. The cell was imaged on the slide holder slanted by 3.2 degrees. The dimension of the stacks were $512\times512\times161$, and both the lateral and axial resolution were 0.15μm.

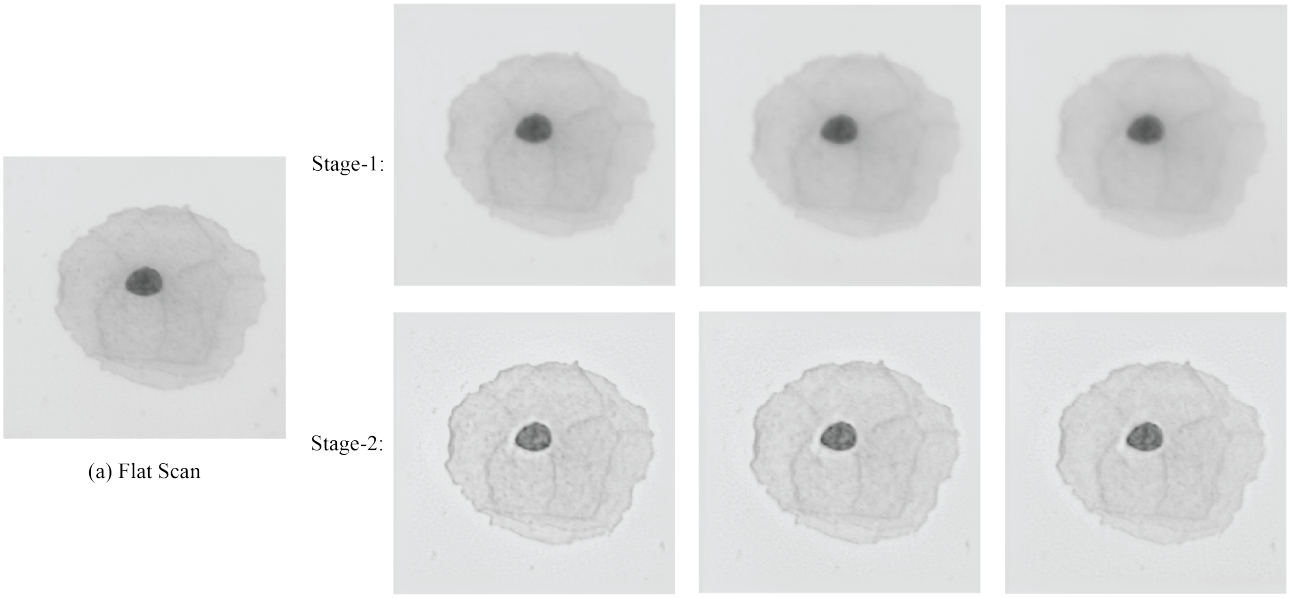


Figure 5.11: The slanted scan cell images after phase deconvolution of *stage-1* (top row) and 20 iterations of *stage-2* intensity deconvolution (bottom row). The cell is the same as showed in Fig. 5.8, but the images are generated as an extended depth-of-field (EDF) image so as to show all objects in focus objects in a single image. The size of the FOVs was 512×512 pixels.

image shows every detail of the cell in the stack. It can be seen that the contrast of the cytoplasm were improved in all cases and the cell edges and cytoplasm can be clearly distinguished even better than the flat scanned images.

5.5.3 Down-sampling the *Z*-axis

Fig. 5.12 shows the same slices of an experimental specimen with and without down-sampling in the *Z* (axial) direction after *stage-1* and *stage-2* (20 iterations) of the deconvolution. It is hard to directly visualise the differences between the slices at the same depths. Therefore, to highlight subtle differences, we set the image stack without down-sampling as the ground-truth and find the maximum differences compared to the down-sampled stack. The maximum differences for the slice pairs after the *stage-1* phase deconvolution were 2.8% (-4.8μm slice), 2.5% (mid slice) and 3.0% (+4.8μm slice). There were smaller errors for slices after the *stage-2* intensity deconvolution, namely 1.7% (-4.8μm slice), 2.5% (mid slice) and 2.0% (+4.8μm slice). If these images are finally saved as 8-bit images (with 256 grey levels), these errors are equivalent to only 4-8 gray levels, i.e., less than the 3 least significant bits. This is why they were hardly noticeable through human inspection. On the other hand, the processing time was dramatically reduced when the data was down-sampled along the *Z*-axis. The processing times for *stage-1* was 13.7 seconds for original data stack and only 1.7 seconds for stack down-sampled by factor of 8. The *stage-2* took longer, at 402.5 seconds and 46.7 seconds for original and down-sampled image stacks respectively. These times were recorded with a desktop computer having 16GB of memory and an Intel i7-3770 CPU on a programming environment of Matalab2015b.

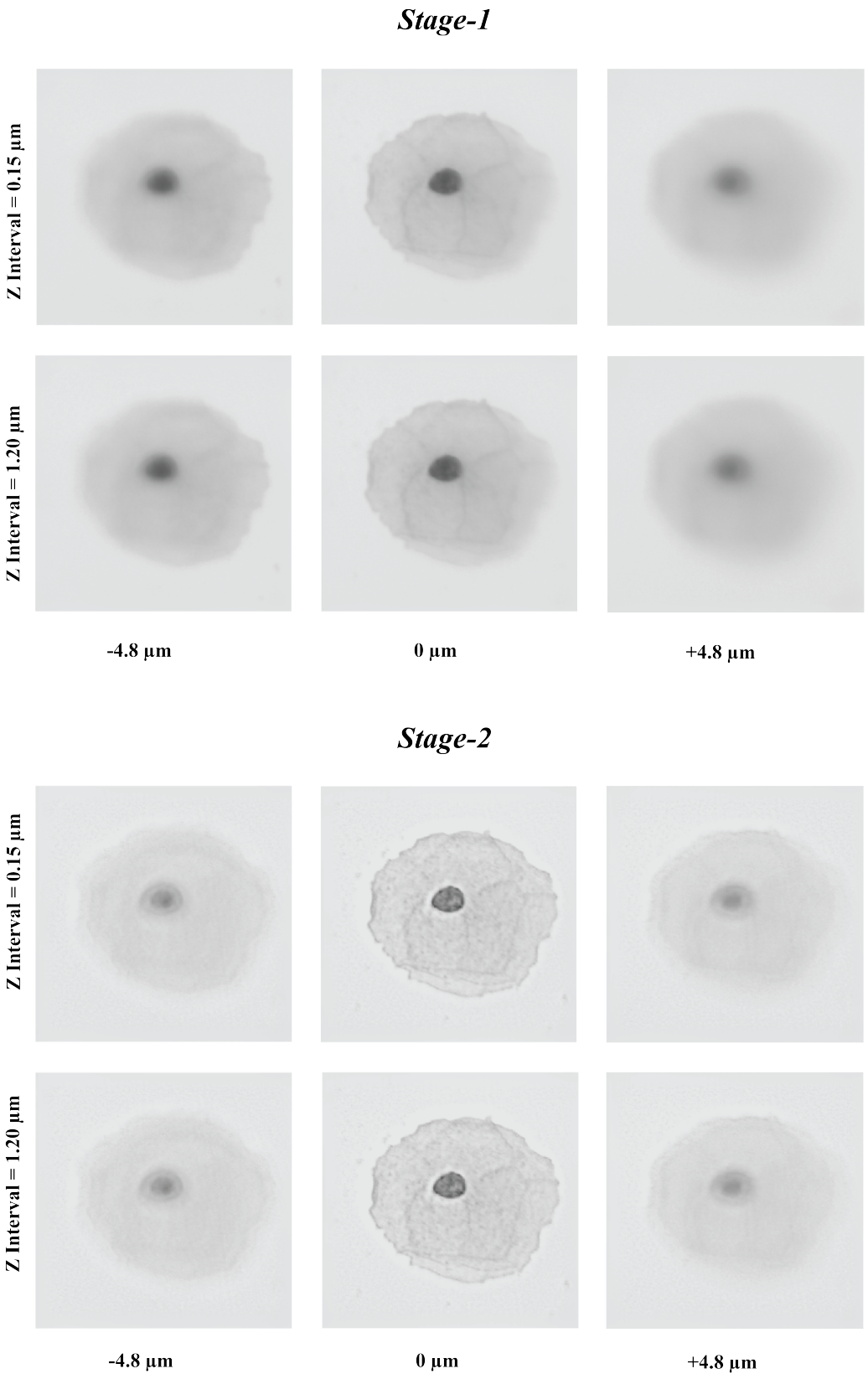


Figure 5.12: There are no noticeable artefacts after the stack was down-sampled by a factor of 8 (bottom row) compared with the original (isotropic) resolution (top row) with the proposed two-stage method. The axial distance between the slices were 0.15μm in the first row, and 1.2μm in the second row, i.e., approximately equal to the depth of field of the optics. Here, the 3D stacks were normalised to have the same dynamic range.

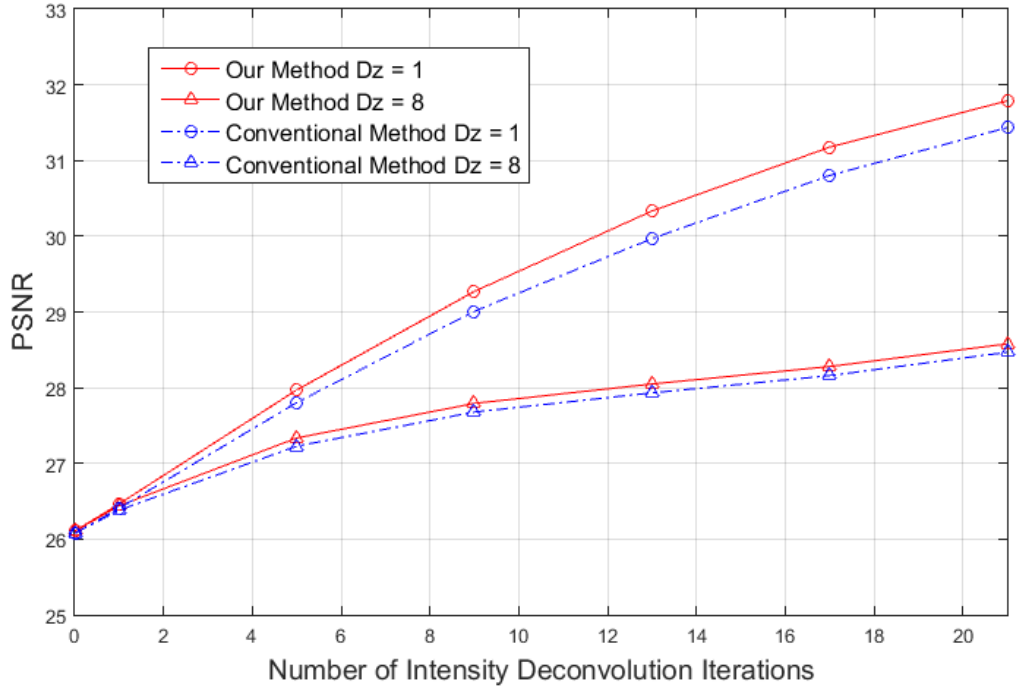


Figure 5.13: The effects of down-sampling on the both proposed method and conventional deconvolution method with the simulation data. Dz is the downsampling factor in the axial direction.

The effect of the downsampling on the deconvolution can be quantitatively evaluated using the simulation data as shown in Fig. 5.13. The downsampling factors were again one (no down-sampling) and 8 times down-sampling. The downsampling along the *Z-axis* had similar adverse effects on both the proposed two-stage method and the conventional method shown by decreased PSNR values at the same number of deconvolution iterations. However, the proposed method still outperformed the conventional method at all deconvolution iterations and down-sampling factors.

5.6 Discussion

The results with simulation data showed in Fig. 5.4 first demonstrated that introducing an opposite phase shift is capable of "correcting" the primary bending distortion due to the coma aberration. This also suggested the phase deconvolution could be used to search the actual coma coefficient (Z_{coma}) by minimising the differences between the flat and the corrected slanted PSF. The following experiments with cytology data presented in Fig. 5.8 and Fig. 5.9 then confirmed the hypothesis and demonstrated the effectiveness of the phase deconvolution and the proposed calibration method for correcting the coma aberration and for calibrating the coma coefficient due to the slide slant. Although the shape of the cell nucleus after the correction did not match exactly that scanned without slide slant, the movements of objects between different focal planes were greatly reduced. This was also caused by the changes of orientation of the cell due to the slide slant, where the nuclei (having a length of around 10μm or 68 pixels) were blurred at different focal depths.

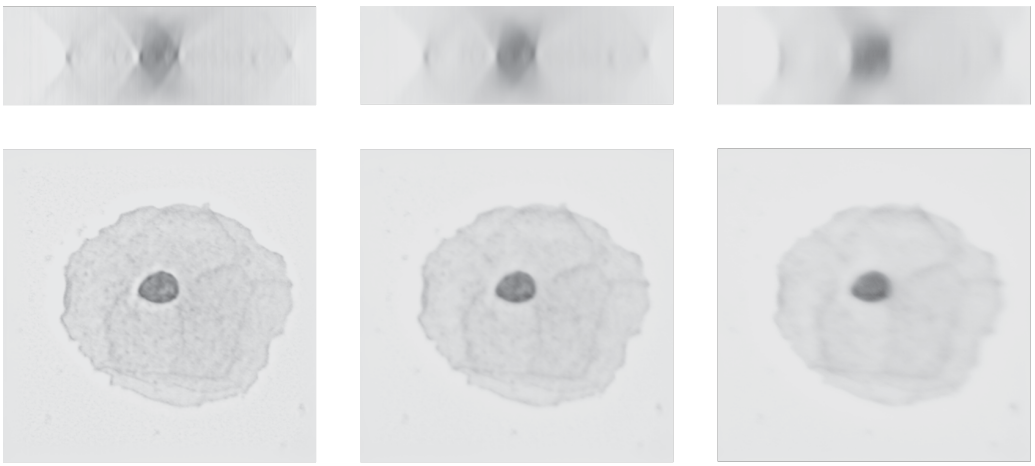
Compared to the analytical data shown in Table. 5.2, these recovered Z_{coma} coefficients were larger (in magnitude), which can be explained by the fact that the mediums on the two sides of the cover-

slip were different. The mediums are both air on the two sides of the PPP for deriving the proposed analytical functions in (Sheppard 2013), however with our experimental imaging system (and with most bright-field microscopes), the medium is air on one side of the cover-slip and is mounting medium on the other side. Thus, the conditions for the equations developed in (Sheppard 2013) no long hold. Even though, these analytical Z_{coma} coefficients are close to those found with the calibration method, so in practice they can be used to determine the search range for the optimal coefficient value. It is worth to noticing that the IQR of calibrated coma coefficients for the two smallest angles (3.2 and 4.2) were smaller than that for the slant angle of 5.3 degrees. This can be associated to the greater image contrast and resolution loss at larger angles, so the errors between the coma corrected stacks and the flat stacks were larger and hence it was more difficult to find the minima. The same phenomenon was indeed predicted by using the simulation data, shown in Fig. 5.4, where the phase deconvolution corrected images estimating the 5.3 degree slide slant angle had the smallest PSNR, indicating the poorest restored image quality.

The performance of the proposed two-sage method in improving the image resolution was first compared to the conventional one-stage method with the simulation data (Fig. 5.5). The proposed method outperformed the conventional method at almost every attempted iterations, because the dominating bending artefacts were corrected at the early stage (with the phase deconvolution). This was further demonstrated with the cytology specimens, where the reconstructed images were closer to that without slide slant. In contrast, the conventional method was inefficient not only in recovering the resolution but also in removing the bending artefacts, meaning a slower convergent speed.

Experimental results with both simulation data and real specimen data demonstrated that the effects of the aberrations are progressively corrected, making it a robust method for correcting asymmetry aberrations. Specifically, one can implement only the deterministic *stage-1* phase deconvolution to obtain images without bending PSF artefacts, while the image resolution is left unchanged requiring less computing power, which may be sufficient for small slide slant angles with limited image contrast losses. On the other hand, the iterative deconvolution method used in *stage-2* of the proposed method enables control of the trade-off between the final image quality and the expenditure of the computing power. In contrast, although the bending artefact is removed progressively when only the conventional deconvolution is used, and after the same number of iterations (20), the bending was still visible as shown in Fig. 5.10. Therefore, the proposed method is more suitable for correcting the asymmetry aberrations such as that induced by the slide slant.

While the effectiveness of the two-stage aberration correction method was successfully demonstrated with the standard Lucy-Richardson deconvolution algorithm, other deconvolution methods can also be utilised in *stage-2*. Therefore, as an example, we also applied the total variation (TV) regulator in the *stage-2* deconvolution (Kim and Naemura 2015; Soulez et al. 2012), with results shown in Fig. 5.14. It can be seen that the TV regulator has advantages of preventing and over emphasis of strong edges, such as nucleus cytoplasm boundary. However as a trade-off, it smoothed out details of cytoplasm texture, reducing some image contrast, which may not be preferred when the image resolution is critical. In addition, the TV regulator has a parameter that needs to be manually selected, depending on the amount of noises in acquired images and the image contrast.



(a) Proposed two-stage method

(b) Proposed two-stage method with total variation regulator

(a) Conventional one-stage method with total variation regulator

Figure 5.14: Strongest edges are smoothed with the use of total variation regulator, noticing the edges of nucleus in (a) and (b), but the regulator does not help correct the bending artefacts with the use of conventional one stage method (c).

The unnoticeable differences between resultant images with and without downsampling the Z-axis shown in Fig. 5.13 implied that increasing the distance between slices (up to the same as the DOF) in the proposed aberration correction method had limited effects . This can be explained by the fact that the axial resolving power of the BF microscopes is poor and not much focal information was lost with down-samplings. So, the Z interval between the adjunct focal planes equivalent to the DOF can be used to reduce the number of focal planes to be acquired while the proposed aberration correction method is still effective. In terms of performance, processing the down-sampled stack is obviously faster, but post-processing of a 512x512x21 stack containing only one cervical cell takes around 50 seconds, which is not acceptable given that some Thin-prep slides have more than 50,000 cells (as shown in last chapter). The optimal Z interval should also be decided by the requirements for data storage space and the need of end users (e.g. the pathologists) in terms of diagnostic performance (Donnelly et al. 2013).

The optimal slide slant angle depends on the final image quality, scan time, computing power, and storage space as well as the thickness of the specimen, illustrated in Fig. 5.15. Specifically, the greater the slide slant angle, the worse the aberrations are caused by the slide slant, and hence the poorer the image contrast and image resolution. Then, more deconvolution iterations are required in order to correct the adverse effects of the aberration. More available computing power means that the digital slide can be obtained more quickly. However, larger slide slant angles allow more focal planes to be acquired, assuming that the camera runs at a consistent frame rate, thus producing a better scan quality. Therefore, the optimal slide slant angle should be determined by taking into consideration the specifications of both the hardware and software available.

Future works for the research presented in this chapter would mainly include enhancements of the proposed aberration correction method. First, as mentioned before, regularised deconvolution methods should be investigated to further improve the image quality. In addition, faster deconvolution methods should be developed in order to reduce the processing time into acceptable levels. Moreover,

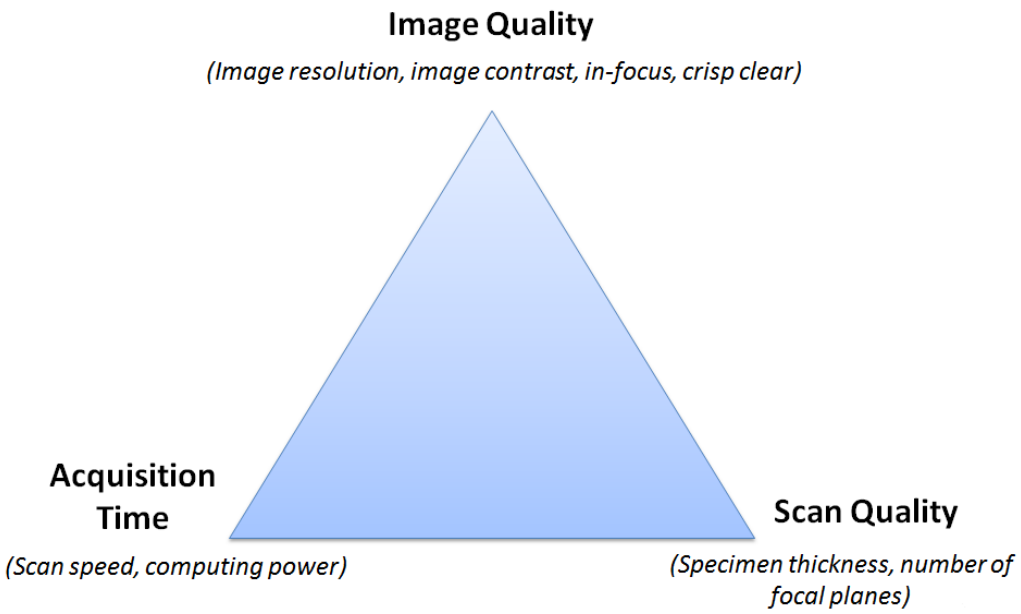


Figure 5.15: The optimal slide slant angle depends on the requirements for final image quality, acquisition time and scan quality.

given cervical cytology cells are sparsely distributed over the slide, deconvolution could only process FOVs or sub-images with an FOV containing cells in order to minimize the processing time. Furthermore, the calibration method should be further developed to estimate the other slide slant induced aberrations such as the spherical and astigmatic so that they can be at least partially corrected by the *stage-1* phase deconvolution step. This may also have the potential to improve image quality and reconstruction speed of *stage-2*. A further potential improvement to the calibration process could be based on phase retrieval which has been demonstrated to be capable of recovering microscope pupil function under coherent illumination Hanser et al. (2004); Zheng et al. (2013b).

5.7 Conclusions

This chapter presented a novel two-stage method for correcting the aberrations induced by slide slant in BF microscopes, which is required for the slanted scan method. Experiments with both simulated and real cervical cytology specimens demonstrated that phase deconvolution (*stage-1*) can effectively correct the main artefacts, the bending of the PSF, while the subsequent *stage-2* intensity deconvolution was able to improve the losses in image contrast and resolution. The resultant image quality of the cervical cytology specimen acquired using the proposed method was comparable to that scanned using the conventional (flat) method. Moreover, the proposed method was superior to the conventional method utilising only intensity deconvolution, especially in removing the bending artefact.

6 GENERAL DISCUSSION AND CONCLUSION

Digitalising the entire glass slides into digital images enabling whole slide imaging is increasing used in pathology, research and education thanks to the advances of the technology. In particular, the use of WSI can improve work-flow efficiency of pathology tests as the digital slide can be easily and efficiently stored, distributed and managed utilising modern computer technologies like digital databases and computer networks. More importantly, the digital slide enables computer aided diagnosis that has great potentials to further improve the productivity of pathology tests.

For WSI to be accepted in the clinics, the digital slide first must be "glass-faithful" and equivalent to the glass slide so that the same diagnosis as could be made using the original glass slide. Secondly, the digital slide acquisition systems for WSI must be fast enough in order to really improve productivity. However, the digitalisation of cytology specimens in "glass-faithful quality", such as cervical cytology specimens for cervical cytology screening, is slow with the conventional scan techniques and existing slide scanners due to the requirements to acquire specimens at multiple focal planes and at high spatial resolution.

This thesis presents a series of novel research in order to improve the scan efficiency for the acquisition of cervical cytology specimen in "glass-faithful quality". This following chapter summarizes the contributions, findings and implications of the work conducted in this thesis. The limitations of research in this thesis is then discussed as well as the directions that deserve further investigation.

6.1 Key Contributions and Findings

This thesis has primarily focused on two main research aims:

1. To develop methods for generating accurate scan maps and focus profile maps of the specimens as well as for determining the thickness of cytology specimens.

2. To investigate and correct the optical aberrations that occur due to the slide slant in the fast slanted scan method that is specifically designed to scan cytology specimens.

The above two aims were accomplished with the following key contributions:

- Chapter 2 proposed methods to quantitatively evaluate and compare the performance of three existing scan method, the FOV scan, the line scan and the slanted scan in terms of scan speed and hardware requirements.

- Chapter 3 developed a tile-based specimen delineation method that not only generates an accurate scan map of a specimen but also provides a ranked list of FOVs as focus candidates, required for estimating the focus profile map of the specimen. (Aim 1.1)

- Chapter 4 proposed a method for quantitative evaluation of the 3D distribution in LBC prepared cervical cytology specimens. (Aim 1.2)

- Chapter 5 developed a two-stage method that can effectively correct the optical aberrations induced by the slide slant in the slanted scan method. (Aim 2)

The main findings are summarised below:

- The performance analysis (Chapter 2) of three scan methods found that the stop-and-go scan strategy is limiting the scan speed of conventional FOV scan method. This can be overcome by moving the stage continuously and utilising line scan cameras as adopted by the line scan and slanted scan method, or using a pulsed illumination source. When the cameras are run at the same frame rates with constraints in the illumination intensity, the slanted scan system is capable of achieving larger data throughput rates than the line scan systems because multiple depths of the specimen can be acquired simultaneously when the specimen is slanted. So, the slanted scan is the most efficient method among the three analysed scan methods for digitising specimens in multiple focal planes.

- The proposed specimen delineation method (Chapter 3) proved an effective and robust achieving good segmentations of specimen deposition area as well selections of focus candidates from lower resolution specimen images. The experimental results with both histology and cytology slides indicated that when the APCI, NACI and TI were used as the tile evaluating metric, the proposed method was effective and robust to slide artefacts (e.g. cover-slip glue and ink marker). But when these metrics were used to select the focus candidates, only the APCI produced superior quality focus candidates, because the APCI is able to distinguish the out-of-focus artefacts by measuring phase diversity.

- The algorithm proposed in Chapter 4 demonstrated that it is feasible to accurately evaluate the 3D distribution of cell and cell clumps of cervical cytology specimens. In terms of accuracy, the experimental results showed that the proposed method achieved an accuracy of one micron more than 90% of the time. It was also demonstrated by ten experimental Thin-prep slides that the thickness of cell clumps, the size of cell clumps and the number of cells in the cell clumps of these slides could be extracted with the proposed method, and more importantly to be used to

optimise the scan parameters for minimising the overall scan time. Specifically, the cell clumps were found to be closely distributed towards the cover-slip, and this led to the development of a new method for generating the focus profile map that considers the thickness of the specimen. The method was shown to be effective in saving approximately 14% of scan time, and with the method average 19 focal planes and 23 focus candidates were sufficient for acquiring 95% of cellular objects in focus. Moreover, it was suggested that the focus profile maps with low-order (2nd or 3rd) polynomials are sufficient to direct the high resolution scan of ten experimental slides.

- An investigation of the optical aberrations occurring in images acquired when the slides was slanted (Chapter 5) found that the dominating effects of these aberration were the skew of the PSF and the loss of image contrast and resolution. The extent of these optical aberrations is related to the thickness of the cover-slip, the numerical aperture of the objective lens and the slide slant angle. To reduce the effects of these aberrations, a novel two-stage method was developed incorporating phase deconvolution as the initial step for removing the bending artefacts and conventional intensity deconvolution as the second step for improving the lost image contrast. The proposed phase deconvolution method successfully remove the PSF bending artefact and was also used to calibrate the coma coefficients for different slide slant angle. The experimental results indicated the effectiveness of the second step deconvolution method on improving the image contrast as shown by the recovery of fine cell clump textures. The proposed two-stage method was also compared with one step method that used only conventional intensity deconvolution, and it was shown that the proposed method was superior in removing the bending artefact because it was effectively mitigated in the stage-1 process the phase deconvolution. It was also found that the optimal slide slant angle is dependent upon the thickness of the specimen, the number of focal planes to be scanned, the available computing power, the processing time constraint as well as the final image quality. The experimental results also indicated that the scanning distance in axial direction (between image slices) can be as large as the DOF of the microscope without any noticeable image artefacts with the proposed correction method.

6.2 Implications of the Findings

The novel methods presented in this thesis provide a number of ways to improve the scan efficiency and scan quality of slide scanners for WSI for the acquisition of digital slides of cytology specimens, so these methods are anticipated to be used in both existing and future slide scanners.

The performance analysis of three scan methods first demonstrated the feasibility to estimate the scan times and hence quantitatively evaluate and compare scan methods for WSI. The findings with the proposed method indicated that the slanted scan method should be used among the three comparing methods for digitising the cervical cytology specimens at high spatial resolution and at multiple focal planes in order to achieve scan times less that ten minutes. Rather than slanting specimens (and slides),

an earlier method achieve multiple focal plane scanning by keeping the specimen flat but utilising a series of line scan cameras (9) positioned (via optical fibres) at different focal planes in the microscope image plane (where the camera is placed) (Price and Bravo-Zanoguera 2003). In principle, the systems using this method will have the same scan speed and scan times as that using the slanted scan, if the cameras are run at the same frame rates because they both do line scan and require only one pass of the stage because they acquire images at multiple focal planes simultaneously. But that system is much more complex and difficult to construct and would cost more, and the method is not readily applicable to existing microscopes. Therefore, in order to achieve a scan time of less than ten minutes in digitising cervical cytology specimens at multiple focal planes and at high spatial resolution, the slanted scan method is the most effective method.

The development of the specimen delineation method proposed in Chapter 3 provided a novel frame-work for selecting focus candidates in the initial scan stage from a lower resolution image of the slide. As the method had been demonstrated to be effective and robust to both histology and cytology specimens when the slides are contaminated, the method should be used by slide scanners when slide artefacts, namely ink marker, excessive glue and dust, exist. The ink maker artefacts, in particular, are common when a pathology lab undergo transition from glass based into WSI based diagnosis, where early diagnosed slides are required to be converted into digital slides so that they can be digitally stored. Specifically, the APCI metric should be used when these artefacts are present because it was shown to be the most robust metric. As the APCI uses the depth clue to distinguish the foreground from background and artefacts, the lens used to acquire low resolution image of the slide should have suitable DOF as demonstrated in this thesis. Even when the slides do not have slides artefacts, the proposed method should be used on scanning cytology specimens, because the cytology cells are spatially distributed and the method is capable of rejecting blank FOVs as focus candidates. Moreover, if the focus candidates have fewer or no outliers (artefacts), fewer focus candidates would be required, thus resulting in a further saving in the overall scan time. The high quality focus candidates lay a good foundation for the subsequent estimation of the focus profile map as well as for the high resolution scan. In addition, the mechanical stage used to move the slide should have sufficient positional accuracy in order to allow these focus candidates found from the low resolution image to be revisited at high resolution. Therefore, the proposed method should be adopted by existing slide scanners that have sufficient positional accuracy.

In regards to the usage of the proposed cell clump 3D distribution evaluation method, an example has been demonstrated with LBC prepared cervical cytology specimens in Chapter 4. In details, the evaluations of ten experimental slides found that the cell clumps distributed more closely to the coverslip, and therefore, a new method for estimating the focus profile map, which considers the thickness of the focus candidates, was proposed and proved to be more effective than the conventional method. Consequently, this new focus profile map construction method should be adopted by the existing slide scanners to minimise the scanning time of cervical cytology specimens. Apart from this, the shape of the focus profile map, the optimal number of focus candidates, the optimal number of focal planes could also be quantitatively determined with the proposed method. In addition, similar studies could also be conducted on other types of specimens in order to study their 3D distribution, thickness profile

and hence optimise scan strategy. Furthermore, as already mentioned in the Chapter 4, the cell clump distribution evaluation method can be used to quantitatively study the slide preparation method, for example, to find out how long the cells should be centrifuged. Therefore, the proposed method can prove extremely useful for optimising the scan strategy of slide scanners for WSI.

The effectiveness of the proposed method for removing the optical aberration induced by the slide slant clearly indicates that those aberrations can be digitally corrected, and therefore, the slanted scan method could be applied to existing scan systems with minimal hardware modifications. Given the performance potential of the slanted scan method, this correction method ensures that the digital slides scanned by the slanted scan are free of optical aberrations, and hence the slanted scan method can be used for rapid acquisition of "glass-faithful" digital slides of thick cytology specimens. In addition, the relationships between the slide slant angle, the image quality and the computational power requirements found in this study would guide the development and usage of future slanted scan systems. Moreover, this work provided a unique tool for correcting the optical aberrations induced by the tilted plane parallel plate, which could be utilised in fluorescent microscopy and optical recording systems where similar aberrations had been reported.

6.3 Limitations of the Study

The limitations associated with this study are as follows:

- For the specimen delineation method proposed in Chapter 3, morphological filterings followed by a simple modification of the classic Ostu's thresholding method were used to distinguish the foreground FOV scan map from the background, which is not robust because the filtering sizes and the thresholding parameter were empirically selected despite the fact that the experimental slides were effectively segmented. method was demonstrated to be effective. The method should be further automated by utilising advanced classification methods such as Gaussian mixture model for both the specimen delineation and the selection of the focus candidates, probably with one or more proposed evaluating metric (image intensity, phase, and entropy) as features (Douglas 2009). In addition, the proposed method was only evaluated to be effective with a 4× objective lens used for low resolution image acquisition, more magnifications and objectives should be test the potential and suitability of the method.

- The experiment in Chapter 4 was conducted using only one type of LBC preparation method for the Pap test, the Thin-prep method, and another popular preparation method, SurePath, was not included. But the method can be readily applied to the SurePath prepared slides, because cervical cells are also prepared to deposit as a mono-layer on the slide, which can be processed with the selected segmentation method. A potential difference could be in the distribution density of cell clumps given that the deposition areas of SurePath slides are smaller (a diameter of 13mm for SurePath slides versus 20mm for Thin-prep slides). Nevertheless, the proposed method offers an effective stool to quantitatively evaluate such differences and to optimise their acquisitions.

- In experiments of both Chapter 3 and Chapter 4, a limited number of slides were used to evaluate the specimen delineation method and even fewer for the analysis of the 3D distribution of cells and the estimation of the focus profile maps, primarily due to the lack of high-throughput digital slide acquisition systems. Though the experimental samples were purposely selected as representative and the size of the data-set was sufficient to prove the effectiveness of the proposed methods, more data should be acquired and processed to comprehensively evaluate the method and study the specimens, e.g. cervical cytology specimens. Moreover, specimens and slides diagnosed as normal were used throughout experiments in order to focus the study on the effectiveness evaluation of proposed methods, abnormal slides should be specifically studied to investigate any potential impacts of these methods on the diagnostic performance.

- The images acquired for the development of aberration correction method in Chapter 5 were not acquired using the slanted scan method due to the lack of hardware. But, the slide slant induced aberrations were adequately sampled as the slides were slanted and specimens were imaged in isotropic, enabling the development and validation of the correction method. In this way, the future integration of the correction method into slanted scan systems only need limited evaluations, such as the effects of line scan artefacts.

- The effects of the optical aberrations due to the slide slant were not quantitatively studied such as the reduction in the resolution, because conventional resolution calibration targets (pre-measured lines and dots) do not have a cover-slip, which is the primary cause of the optical aberrations. In addition, the effectiveness of the correction method has been evaluated quantitatively with synthetic data and qualitatively by real data, but the method has not yet been assessed qualitative with real specimen.

6.4 Future Directions

The future work of this research mainly includes further enhancements of the aberration correction method for the slanted scan method. First, in order to quantitatively evaluate the optical aberrations due to the slide slant, the cover-slip could be added to the top of the calibration slide in the same way as the specimen slide is prepared. The effectiveness of the proposed method should also be quantitatively evaluated with real cytology specimen, which could be achieved by measuring the geometrical features of cervical cells such as the area and shape. Secondly, the proposed aberration correction method should be extended to process a full digital slide of the specimen and to work on the slanted scan systems. As processing a large FOV using a 3D deconvolution based correction method may require a large amount of computer memory, acquired images could be cropped into smaller FOVs and processed individually and then stitched together to form final digital slide. The effects of cells at FOV boundaries on the deconvolution should also be considered, which is necessary to create mon-

tage digital slides. Thirdly, it was found that existing deconvolution methods are too slow to process the entire slide within a acceptable time, so faster deconvolution methods as well as more effective computing frameworks like graphic processing units (GPUs) based programming could be selected to further reduce the processing time (Domanski et al. 2009). Alternative to developing faster deconvolution methods, hardware solutions such as employing an additional optic lens could also be developed to eliminate the need for the digital post-processing.

6.5 Conclusion

In summary, this thesis presented several novel methods developed for rapid and high quality acquisition of cervial cytology specimens for WSI. Specifically, a tile-based specimen delineation method and a novel tile evaluation metric (APCI) were demonstrated capable of not only generating a minimal scan map for the slide scanner but also a ranked list of FOVs required for interpolating a high quality focus profile map of the specimen. The proposed algorithm for thickness analysis of the specimen could be utilised to quantitatively investigate the thickness of cytology specimens for determining the optimal number of focal planes to be acquired and hence the minimal scan time. The high throughput slanted scan method has also been extensively investigated in this thesis, and for the first time, a correction method was proposed for effectively ameliorating the aberrations induced by the slide slant in this novel scan method. The combination of these proposed methods in this thesis have the potential to increase both overall scan efficiency and scan quality of whole slide scanners when acquiring "glass-faithful quality" images of cervical cytology specimens.

BIBLIOGRAPHY

3DHISTECH (2016). Pannoramic 250 flash iii. Available online at http://www.3dhistech.com/pannoramic_250_flash.

Al-Janabi, S., Huisman, A., and Van Diest, P. (2012). Digital pathology: current status and future perspectives. *Histopathology*, 61(1):1–9. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2559.2011.03814.x/full.

Altinay, D. and Bradley, A. P. (2011). An evaluation of multi-resolution microscope slide scanning algorithms. In *Digital Image Computing: Techniques and Applications, 2011. DICTA'11. Proceedings 2011*, pages 319 –324. IEEE. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6128702.

Altinay, D. and Bradley, A. P. (2013). Illumination effects in quantitative virtual microscopy. In *Computer Analysis of Images and Patterns*, pages 449–456. Springer. http://link.springer.com/chapter/10.1007/978-3-642-40246-3_56.

Altinay, D., Bradley, A. P., and Mehnert, A. (2010). On the estimation of extrinsic and intrinsic parameters of optical microscope calibration. In *Digital Image Computing: Techniques and Applications (DICTA), 2010 International Conference on*, pages 190–195. IEEE. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=5692563.

Ameisen, D., Deroulers, C., Perrier, V., Yunès, J., Bouhidel, F., Battistella, M., Legrès, L., Janin, A., and Bertheau, P. (2013). Stack or trash? quality assessment of virtual slides. *Diagnostic Pathology*, 8:S23. http://diagnosticpathology.biomedcentral.com/articles/10.1186/1746-1596-8-S1-S23.

Arimoto, R. and Murray, J. (2004). A common aberration with water-immersion objective lenses. *Journal of microscopy*, 216(1):49–51. http://onlinelibrary.wiley.com/doi/10.1111/j.0022-2720.2004.01383.x/full.

Audet, C. and Dennis Jr, J. E. (2002). Analysis of generalized pattern searches. *SIAM Journal on Optimization*, 13(3):889–903. http://epubs.siam.org/doi/abs/10.1137/S1052623400378742.

AustralianGovernment (2013). Review of the funding arrangements for pathology services. Available online at http://www.health.gov.au/.

Bamford, P. and Mayer, W. (2009). Method and apparatus and computer program product for collecting digital image data from microscope media-based specimens. US Patent 0295963.

BD (2016). Bd focalpoint slide profiler. Available online at https://www.bd.com/tripath/labs/fp_screening.asp.

Bengtsson, E. and Malm, P. (2014). Screening for cervical cancer using automated analysis of pap-smears. *Computational and mathematical methods in medicine*, 2014. http://www.hindawi.com/journals/cmmm/2014/842037/.

Bentz, J. (2005). Liquid-based cytology for cervical cancer screening. http://www.tandfonline.com/doi/abs/10.1586/14737159.5.6.857.

Boost, T. (2009). A comparison of screening times between the thinprep imager and conventional cytology. *Diagnostic cytopathology*, 37(9):661–664. http://onlinelibrary.wiley.com/doi/10.1002/dc.21069/abstract.

Braat, J. (1997). Analytical expressions for the wave-front aberration coefficients of a tilted plane-parallel plate. *Applied optics*, 36(32):8459–8467. https://www.osapublishing.org/ao/abstract.cfm?uri=ao-36-32-8459.

Bradley, A. and Bamford, P. (2004). A one-pass extended depth of field algorithm based on the over-complete discrete wavelet transform. In *Image and Vision Computing'04 New Zealand (IVCNZ'04)*, pages 279–284. http://espace.library.uq.edu.au/view/UQ:8934.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159. http://www.sciencedirect.com/science/article/pii/S0031320396001422.

Bradley, A. P. and Stentiford, F. W. (2003). Visual attention for region of interest coding in jpeg 2000. *Journal of Visual Communication and Image Representation*, 14(3):232–250. http://www.sciencedirect.com/science/article/pii/S1047320303000373.

Bradley, A. P., Wildermoth, M., and Mills, P. (2005). Virtual microscopy with extended depth of field. In *Digital Image Computing: Techniques and Applications, 2005. DICTA'05. Proceedings 2005*, pages 35–35. IEEE. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=1587637.

Castleman, K. R. (1996). *Digital Image Processing*. Prentice Hall, 2nd edition.

Cornish, T. C., Swapp, R. E., and Kaplan, K. J. (2012). Whole-slide imaging: routine pathologic diagnosis. *Advances in anatomic pathology*, 19(3):152–159.

Dixon, W. J. and Mood, A. M. (1946). The statistical sign test. *Journal of the American Statistical Association*, 41(236):557–566. http://www.tandfonline.com/doi/abs/10.1080/01621459.1946.10501898.

Domanski, L., Vallotton, P., and Wang, D. (2009). Two and three-dimensional image deconvolution on graphics hardware. In *Proceedings of the 18th World IMACS/MODSIM Congress, Cairns, Australia, July*, pages 13–17.

Donnelly, A. D., Mukherjee, M. S., Lyden, E. R., Bridge, J. A., Lele, S. M., Wright, N., McGaughey, M. F., Culberson, A. M., Horn, A. J., Wedel, W. R., et al. (2013). Optimal z-axis scanning parameters for gynecologic cytology specimens. *Journal of pathology informatics*, 4. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3908726/.

Douglas, R. (2009). Gaussian mixture models. *Encyclopedia of Biometrics*, pages 659–663. http://link.springer.com/content/pdf/10.1007/978-1-4899-7488-4_196.pdf.

Duby, J. M. and DiFurio, M. J. (2009). Implementation of the thinprep imaging system in a tertiary military medical center. *Cancer Cytopathology*, 117(4):264–270. http://onlinelibrary.wiley.com/doi/10.1002/cncy.20033/full.

El-Gabry, E., Parwani, A., and Pantanowitz, L. (2014). Whole-slide imaging: widening the scope of cytopathology. *Diagnostic Histopathology*, 20(12):456–461. http://www.sciencedirect.com/science/article/pii/S1756231714001753.

Evered, A. and Dudding, N. (2011). Accuracy and perceptions of virtual microscopy compared with glass slide microscopy in cervical cytology. *Cytopathology*, 22(2):82–87. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2303.2010.00758.x/full.

Fan, Y. and Bradley, A. P. (2016a). A method for quantitative analysis of clump thickness in cervical cytology slides. *Micron*, 80:73–82. http://www.sciencedirect.com/science/article/pii/S0968432815300391.

Fan, Y. and Bradley, A. P. (2016b). A two-stage method to correct the optical aberration induced by slide slant. *Micron*, 87:18–32. http://www.sciencedirect.com/science/article/pii/S0968432816300701.

Fan, Y., Gal, Y., and Bradley, A. P. (2013). Performance analysis of three microscope slide scanning techniques. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6691511.

Fan, Y., Gal, Y., and Bradley, A. P. (2014a). An algorithm for microscopic specimen delineation and focus candidate selection. *Micron*, 66:51–62. http://www.sciencedirect.com/science/article/pii/S0968432814001127.

Fan, Y., Gal, Y., and Bradley, A. P. (2014b). Microscopic specimen delineation using the auto phase correlation index. In *International Symposium on Biomedical Imaging (ISBI)*. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6868124.

Farahani, N., Parwani, A., and Pantanowitz, L. (2015). Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathol. Lab Med. Int.*, 7:23–33.

https://www.researchgate.net/profile/Liron_Pantanowitz/publication/279231201_Whole_slide_imaging_in_pathology_advantages_limitations_and_emerging_perspectives/links/55b0d5bc08ae32092e07336b.pdf.

for Research on Cancer, I. A. (2015). Globocan 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012. Available online at http://globocan.iarc.fr/pages/fact_sheets_population.aspx.

Gerber, R. E. and Mansuripur, M. (1996). Tilt correction in an optical disk system. *Applied optics*, 35(35):7000–7007. https://www.osapublishing.org/ao/abstract.cfm?uri=ao-35-35-7000.

Ghaznavi, F., Evans, A., Madabhushi, A., and Feldman, M. (2013). Digital imaging in pathology: Whole-slide imaging and beyond. *Annual Review of Pathology: Mechanisms of Disease*, 8(1):331–359. http://www.annualreviews.org/doi/abs/10.1146/annurev-pathol-011811-120902.

Gibson, S. F. and Lanni, F. (1991). Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *JOSA A*, 8(10):1601–1613. https://www.osapublishing.org/josaa/abstract.cfm?uri=josaa-8-10-1601.

Gilbertson, J. R., Ho, J., Anthony, L., Jukic, D. M., Yagi, Y., and Parwani, A. V. (2006). Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC clinical pathology*, 6(1):1. http://bmcclinpathol.biomedcentral.com/articles/10.1186/1472-6890-6-4.

Gonzalez, R. C. and Woods, R. E., editors (2008). *Digital image processing*. Harlow : Pearson/Prentice Hall.

Goodman, J. W. (2005). *Introduction to Fourier optics*. Roberts and Company Publishers.

Goodwin, P. C. (2007). Evaluating optical aberration using fluorescent microspheres: methods, analysis, and corrective actions. *Methods in cell biology*, 81:397–413. http://www.sciencedirect.com/science/article/pii/B9780124077614000154.

Grabe, N., Lahrmann, B., Pommerencke, T., von Knebel Doeberitz, M., Reuschenbach, M., and Wentzensen, N. (2010). A virtual microscopy system to scan, evaluate and archive biomarker enhanced cervical cytology slides. *Analytical Cellular Pathology*, 32(1-2):109–119. http://www.hindawi.com/journals/acp/2010/312048/abs/.

Graham, R., Lubachevsky, B., Nurmela, K., and Östergård, P. (1998). Dense packings of congruent circles in a circle. *Discrete Mathematics*, 181(1):139–154. http://www.sciencedirect.com/science/article/pii/S0012365X97000502.

Grosso, A., Jamali, A., Locatelli, M., and Schoen, F. (2010). Solving the problem of packing equal and unequal circles in a circular container. *Journal of Global Optimization*, 47(1):63–81. http://link.springer.com/article/10.1007/s10898-009-9458-3.

Halford, J., Batty, T., Boost, T., Duhig, J., Hall, J., Lee, C., and Walker, K. (2010). Comparison of the sensitivity of conventional cytology and the thinprep imaging system for 1,083 biopsy confirmed high-grade squamous lesions. *Diagnostic cytopathology*, 38(5):318–326. http://onlinelibrary.wiley.com/doi/10.1002/dc.21199/abstract.

Hamamatsu (2016a). Nanozoomer-xr. Available online at http://www.hamamatsu.com/jp/en/community/nanozoomer/Selection/NanoZoomer-XR.html.

Hamamatsu (2016b). Orca-flash4.0 v2 c11440-22cu. Available online at http://www.hamamatsu.com/us/en/product/category/5000/5005/C11440-22CU/index.html.

Hanser, B., Gustafsson, M., Agard, D., and Sedat, J. (2004). Phase-retrieved pupil functions in wide-field fluorescence microscopy. *Journal of microscopy*, 216(1):32–48. http://onlinelibrary.wiley.com/doi/10.1111/j.0022-2720.2004.01393.x/full.

Herzlich (2015). The best known packings of equal circles in a circle (complete up to n=2600). Available online at http://hydra.nat.uni-magdeburg.de/packing/cci/.

Hilsenstein, V. (2005). Robust autofocusing for automated microscopy imaging of fluorescently labelled bacteria. In *Digital Image Computing: Techniques and Applications, 2005. DICTA'05. Proceedings 2005*, page 15. IEEE. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=1587617.

Holmes, T. J., Bhattacharyya, S., Cooper, J. A., Hanzel, D., Krishnamurthi, V., Lin, W.-c., Roysam, B., Szarowski, D. H., and Turner, J. N. (1995). Light microscopic images reconstructed by maximum likelihood deconvolution. In *Handbook of biological confocal microscopy*, pages 389–402. Springer. http://link.springer.com/chapter/10.1007/978-1-4757-5348-6_24.

Holmes, T. J. and O'connor, N. (2000). Blind deconvolution of 3d transmitted light brightfield micrographs. *Journal of microscopy*, 200(2):114–127. http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2818.2000.00751.x/full.

Kaikai, G., Jun, L., Zichao, B., Xin, H., and Guoan, Z. (2015). Instantscope: a low-cost whole slide imaging system with instant focal plane detection. *Biomed. Opt. Express*, 6(9):3210–3216, DOI:10.1364/BOE.6.003210. http://www.osapublishing.org/boe/abstract.cfm?URI=boe-6-9-3210.

Kardos, T. F. (2004). The focalpoint system. *Cancer Cytopathology*, 102(6):334–339.

Kayser, K. et al. (2011). Quantification of virtual slides: Approaches to analysis of content-based image information. *Journal of pathology informatics*, 2(1):2. http://www.jpathinformatics.org/article.asp?issn=2153-3539;year=2011;volume=2;issue=1;spage=2;epage=2;aulast=Kayser.

Keuper, M., Temerinac-Ott, M., Padeken, J., Heun, P., Brox, T., Burkhardt, H., and Ronneberger, O. (2012). Blind deconvolution with psf regularization for wide-field microscopy. In *Biomedical*

Imaging (ISBI), 2012 9th IEEE International Symposium on, pages 1292–1295. IEEE. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6235799.

Kim, B. and Naemura, T. (2015). Blind depth-variant deconvolution of 3d data in wide-field fluorescence microscopy. *Scientific reports*, 5. http://www.nature.com/articles/srep09894.

Kirshner, H., Aguet, F., Sage, D., and Unser, M. (2013). 3-d psf fitting for fluorescence microscopy: implementation and localization application. *Journal of microscopy*, 249(1):13–25. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2818.2012.03675.x/full.

Kuglin, C. and Hines, D. (1975). The phase correlation image aligment method. *Proc.Int.Conf.on Cybernetics and Society*, pages 163–165.

Lahrmann, B., Valous, N. A., Eisenmann, U., Wentzensen, N., and Grabe, N. (2013). Semantic focusing allows fully automated single-layer slide scanning of cervical cytology slides. *PloS one*, 8(4):e61441. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061441.

Lee, R. E., McClintock, D. S., Laver, N. M., and Yagi, Y. (2011). Evaluation and optimization for liquid-based preparation cytology in whole slide imaging. *Journal of pathology Informatics*, 2. http://www.jpathinformatics.org/article.asp?issn=2153-3539;year=2011;volume=2;issue=1;spage=46;epage=46;aulast=Lee.

Leica (2016). Aperio at2 – high volume, digital whole slide scanning. Available online at http://www.leicabiosystems.com/digital-pathology/aperio-digital-pathology-slide-scanners/details/product/aperio-at2/.

Leong, F. and McGee, J. (2001). Automated complete slide digitization: a medium for simultaneous viewing by multiple pathologists. *The Journal of pathology*, 195(4):508–514. http://onlinelibrary.wiley.com/doi/10.1002/path.972/full.

Levi, A. W., Chhieng, D. C., Schofield, K., Kowalski, D., and Harigopal, M. (2012). Implementation of focalpoint gs location-guided imaging system. *Cancer cytopathology*, 120(2):126–133. http://onlinelibrary.wiley.com/doi/10.1002/cncy.20187/full.

Lewis, J. (1995). Fast normalized cross-correlation. In *Vision interface*, volume 10, pages 120–123. http://comupter-vision-summary.googlecode.com/svn/trunk/papers/FastNormalizedCrossCorrelation.pdf.

Li, K., Lu, Z., Liu, W., and Yin, J. (2012). Cytoplasm and nucleus segmentation in cervical smear images using radiating gvf snake. *Pattern Recognition*, 45(4):1255–1264. http://www.sciencedirect.com/science/article/pii/S0031320311003979.

Li, X., Liu, J., Xu, H., Gong, E., McNutt, M. A., Li, F., Anderson, V. M., and Gu, J. (2007). A feasibility study of virtual slides in surgical pathology in china. *Human pathology*, 38(12):1842–1848. http://www.sciencedirect.com/science/article/pii/S0046817707002419.

Lopez, X. M., D'Andrea, E., Barbot, P., Bridoux, A.-S., Rorive, S., Salmon, I., Debeir, O., and Decaestecker, C. (2013). An automated blur detection method for histological whole slide imaging. *PloS one*, 8(12):e82710. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0082710.

Lu, Z., Carneiro, G., and Bradley, A. P. (2015a). An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *Image Processing, IEEE Transactions on*, 24(4):1261–1272, DOI:10.1109/TIP.2015.2389619. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=7005499.

Lu, Z., Carneiro, G., Neeraj, D., and Bradley, A. (2015b). Segmentation of overlapping cervical cells from multi-layer cytology preparation volumes. In *DIEEE International Symposium on Biomedical Imaging (ISBI), New York, USA*.

Lubachevsky, B. and Graham, R. (1997). Curved hexagonal packings of equal disks in a circle. *Discrete & Computational Geometry*, 18(2):179–194. http://link.springer.com/article/10.1007/PL00009314.

Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):674–693. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=192463&tag=1.

Malm, P., Balakrishnan, B. N., Sujathan, V. K., Kumar, R., and Bengtsson, E. (2013). Debris removal in pap-smear images. *Computer methods and programs in biomedicine*, 111(1):128–138. http://www.sciencedirect.com/science/article/pii/S0169260713000606.

Mango, L. J. (1994). Computer-assisted cervical cancer screening using neural networks. *Cancer letters*, 77(2):155–162.

Markham, J. and Conchello, J.-A. (1999). Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur. *JOSA A*, 16(10):2377–2391. https://www.osapublishing.org/josaa/abstract.cfm?uri=josaa-16-10-2377.

Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10):761–767. http://www.sciencedirect.com/science/article/pii/S0262885604000435.

Mayerich, D., Abbott, L., and McCormick, B. (2008). Knife-edge scanning microscopy for imaging and reconstruction of three-dimensional anatomical structures of the mouse brain. *Journal of microscopy*, 231(1):134–143. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2818.2008.02024.x/full.

Network, E. (2016). Ethernet vs. camera link. Available online at http://www.edn.com/design/test-and-measurement/4381304/Ethernet-vs-Camera-Link.

Noll, R. J. (1976). Zernike polynomials and atmospheric turbulence. *JOsA*, 66(3):207–211. https://www.osapublishing.org/josa/abstract.cfm?uri=josa-66-3-207.

Nordin, B. (1989). The development of an automatic prescreener for the early detection of cervical cancer: Algorithms and implementation. *[Ph.D. thesis]*.

Nozaka, H., Miura, T., and Zheng, Z. (2013). Multi-layer virtual slide scanning system with multi-focus image fusion for cytopathology and image diagnosis. *IEICE TRANSACTIONS on Information and Systems*, 96(4):856–863. http://search.ieice.org/bin/summary.php?id=e96-d_4_856.

Olympus (2016). Vs120 virtual slide microscope. Available online at http://www.olympus-lifescience.com/en/microscopes/virtual/vs120/.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285–296):23–27.

Ou, X., Zheng, G., and Yang, C. (2014). Embedded pupil function recovery for fourier ptychographic microscopy. *Optics express*, 22(5):4960–4972. https://www.osapublishing.org/oe/abstract.cfm?uri=oe-22-5-4960.

Papillo, J. L., St John, T. L., and Leiman, G. (2008). Effectiveness of the thinprep imaging system: clinical experience in a low risk screening population. *Diagnostic cytopathology*, 36(3):155–160. http://onlinelibrary.wiley.com/doi/10.1002/dc.20779/abstract.

Paulsen, F. P., Eichhorn, M., and Bräuer, L. (2010). Virtual microscopy—the future of teaching histology in the medical curriculum? *Annals of Anatomy-Anatomischer Anzeiger*, 192(6):378–382. http://www.sciencedirect.com/science/article/pii/S0940960210001470.

Philips (2016). Intellisite ultra fast scanner. Available online at http://www.philips.com.au/healthcare/product/HCNOCTN442/intellisite-ultra-fast-scanner.

Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of tiled 3d microscopic image acquisitions. *Bioinformatics*, 25(11):1463–1465. https://bioinformatics.oxfordjournals.org/content/25/11/1463.full.

Price, J. and Bravo-Zanoguera, M. (2003). Automatic on-the-fly focusing for continuous image acquisition in high-resolution microscopy. US Patent 6,640,014. https://www.google.com.au/patents/US6640014.

Rankov, V., Locke, R. J., Edens, R. J., Barber, P. R., and Vojnovic, B. (2005). An algorithm for image stitching and blending. In *Biomedical Optics 2005*, pages 190–199. International Society for Optics and Photonics. http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=859158.

Rodríguez, A. C., Solomon, D., Herrero, R., Hildesheim, A., González, P., Wacholder, S., Porras, C., Jiménez, S., Schiffman, M., Group, C., et al. (2013). Impact of human papillomavirus vaccination on cervical cytology screening, colposcopy, and treatment. *American journal of epidemiology*, 178(5):752–760. http://aje.oxfordjournals.org/content/178/5/752.full.

Rojo, M. G., García, G. B., Mateos, C. P., García, J. G., and Vicente, M. C. (2006). Critical comparison of 31 commercially available digital slide systems in pathology. *International Journal of Surgical Pathology*, 14(4):285–305. http://ijs.sagepub.com/content/14/4/285.short.

Romer, D. J., Yearsley, K. H., and Ayers, L. W. (2003). Using a modified standard microscope to generate virtual slides. *The Anatomical Record Part B: The New Anatomist*, 272(1):91–97. http://onlinelibrary.wiley.com/doi/10.1002/ar.b.10017/full.

Sarder, P. and Nehorai, A. (2006). Deconvolution methods for 3-d fluorescence microscopy images. *Signal Processing Magazine, IEEE*, 23(3):32–45. http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=1628876.

Schiffman, M. (2007). Integration of human papillomavirus vaccination, cytology, and human papillomavirus testing. *Cancer cytopathology*, 111(3):145–153. http://onlinelibrary.wiley.com/doi/10.1002/cncr.22751/full.

Sheppard, C. J. (2013). Balanced diffraction aberrations, independent of the observation point: application to a tilted dielectric plate. *JOSA A*, 30(10):2150–2161. https://www.osapublishing.org/josaa/abstract.cfm?uri=josaa-30-10-2150.

Society, A. C. (2016). Survival rates for cervical cancer, by stage. Available online at http://www.cancer.org/cancer/cervicalcancer/detailedguide/cervical-cancer-survival.

Soulez, F., Denis, L., Tourneur, Y., and Thiébaut, E. (2012). Blind deconvolution of 3d data in wide field fluorescence microscopy. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1735–1738. IEEE.

Stallinga, S. (2005). Compact description of substrate-related aberrations in high numerical-aperture optical disk readout. *Applied optics*, 44(6):849–858. https://www.osapublishing.org/ao/abstract.cfm?uri=ao-44-6-849.

Sun, C., Beare, R., Hilsenstein, V., and Jackway, P. (2005). Mosaicing of microscope images. In *null*, page 50. IEEE. https://www.researchgate.net/profile/Changming_Sun/publication/4219272_Mosaicing_of_Microscope_Images/links/0912f5139a6fdb3892000000.pdf.

Sun, Y., Duthaler, S., and Nelson, B. J. (2004). Autofocusing in computer microscopy: selecting the optimal focus algorithm. *Microscopy research and technique*, 65(3):139–149. http://onlinelibrary.wiley.com/doi/10.1002/jemt.20118/abstract.

Svoboda, D., Kozubek, M., and Stejskal, S. (2009). Generation of digital phantoms of cell nuclei and simulation of image formation in 3d image cytometry. *Cytometry part A*, 75(6):494–509. http://onlinelibrary.wiley.com/doi/10.1002/cyto.a.20714/full.

Tadrous, P. (2010). A method of psf generation for 3d brightfield deconvolution. *Journal of microscopy*, 237(2):192–199. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2818.2009.03323.x/full.

Tello-Mijares, S., Flores, F., Bescos, J., and Valdez, E. (2013). Efficient autofocus method for sequential automatic capturing of high-magnification microscopic images. *Chinese Optics Letters*, 11(12):121102–121102. https://www.osapublishing.org/col/abstract.cfm?uri=col-11-12-121102.

Tips, O. T. (2016). Usb 2.0 vs. usb 3.0 vs. esata vs. thunderbolt vs. firewire vs. ethernet speed. Available online at http://www.online-tech-tips.com/computer-tips/usb-2-0-vs-usb-3-0-vs-esata-vs-thunderbolt-vs-firewire-vs-ethernet-speed/.

Tolles, W. E. and Bostrom, R. (1956). Automatic screening of cytological smears for cancer: the instrumentation. *Annals of the New York Academy of Sciences*, 63(6):1211–1218.

Turnkey, A. (2016). Line scan cameras. Available online at http://www.adept.net.au/cameras/lineScan.shtml.

Valdecasas, A., Marshall, D., Becerra, J., and Terrero, J. (2001). On the extended depth of focus algorithms for bright field microscopy. *Micron*, 32(6):559–569. http://www.sciencedirect.com/science/article/pii/S0968432800000615.

Vedaldi, A. and Soatto, S. (2008). Quick shift and kernel methods for mode seeking. In *Computer Vision-ECCV 2008*, pages 705–718. Springer. http://link.springer.com/chapter/10.1007/978-3-540-88693-8_52.

Weinstein, R. S., Descour, M. R., Liang, C., Barker, G., Scott, K. M., Richter, L., Krupinski, E. A., Bhattacharyya, A. K., Davis, J. R., Graham, A. R., et al. (2004). An array microscope for ultrarapid virtual slide processing and telepathology. design, fabrication, and validation study. *Human pathology*, 35(11):1303–1314. http://www.sciencedirect.com/science/article/pii/S0046817704004952.

Wright, A., Smith, D., Dhurandhar, B., Fairley, T., Scheiber-Pacht, M., Chakraborty, S., Gorman, B., Mody, D., and Coffey, D. (2013). Digital slide imaging in cervicovaginal cytology: a pilot study. *Archives of pathology and laboratory medicine*, 137(5):618–624. http://www.archivesofpathology.org/doi/abs/10.5858/arpa.2012-0430-OA.

Wright, P., Marshall, J., and Desai, M. (2010). Comparison of surepath® and thinprep® liquid-based cervical cytology using positive predictive value, atypical predictive value and total predictive value as performance indicators. *Cytopathology*, 21(6):374–378. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2303.2010.00772.x/full.

Yang, F., Deng, Z., and Fan, Q. (2013). A method for fast automated microscope image stitching. *Micron*, 48:17–25. http://www.sciencedirect.com/science/article/pii/S0968432813000231.

Yu, J.-H., Chen, Y.-C., and Shiou, F.-J. (2011). Development of a biomedical virtual microscopy slide scanner using a line scan charge-coupled device (ccd) camera. *Scientific Research and Essays*, 6(10):2101–2118. http://www.academicjournals.org/journal/SRE/article-full-text-pdf/F88C3D022843.

Zeiss (2015). Objective ec plan-neofluar 40x/0.75 m27. Available online at https://www.micro-shop.zeiss.com/?l=en&p=us&f=o&a=v&m=a&id=420360-9900-000.

Zeiss (2016). Slide scanner axio scan.z1. Available online at http://www.zeiss.com/microscopy/en_de/products/imaging-systems/axio-scan-z1.html.

Zhao, T., Wachman, E., and Farkas, D. (2004). A novel scheme for abnormal cell detection in pap smear images. In *Biomedical Optics 2004*, pages 151–162. International Society for Optics and Photonics. http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=838880.

Zheng, G. (2016a). personal communication.

Zheng, G. (2016b). Fourier ptychography. Available online at https://sites.google.com/site/gazheng/Fourier-Ptychography.

Zheng, G., Horstmeyer, R., and Yang, C. (2013a). Wide-field, high-resolution fourier ptychographic microscopy. *Nature photonics*, 7(9):739–745. http://www.nature.com/nphoton/journal/v7/n9/abs/nphoton.2013.187.html.

Zheng, G., Ou, X., Horstmeyer, R., and Yang, C. (2013b). Characterization of spatially varying aberrations for wide field-of-view microscopy. *Optics express*, 21(13):15131–15143. https://www.osapublishing.org/oe/abstract.cfm?uri=oe-21-13-15131.

Zhong, X., Wang, T., Yan, J., and Jin, G. (2015). Aberrations correction of tilted plane-parallel plate in convergent rays. *Applied Optics*, 54(7):1758–1764. https://www.osapublishing.org/ao/abstract.cfm?uri=ao-54-7-1758.

APPENDIX

The low resolution images of the slides used for evluation of quality of focus candidates in Chapter 3

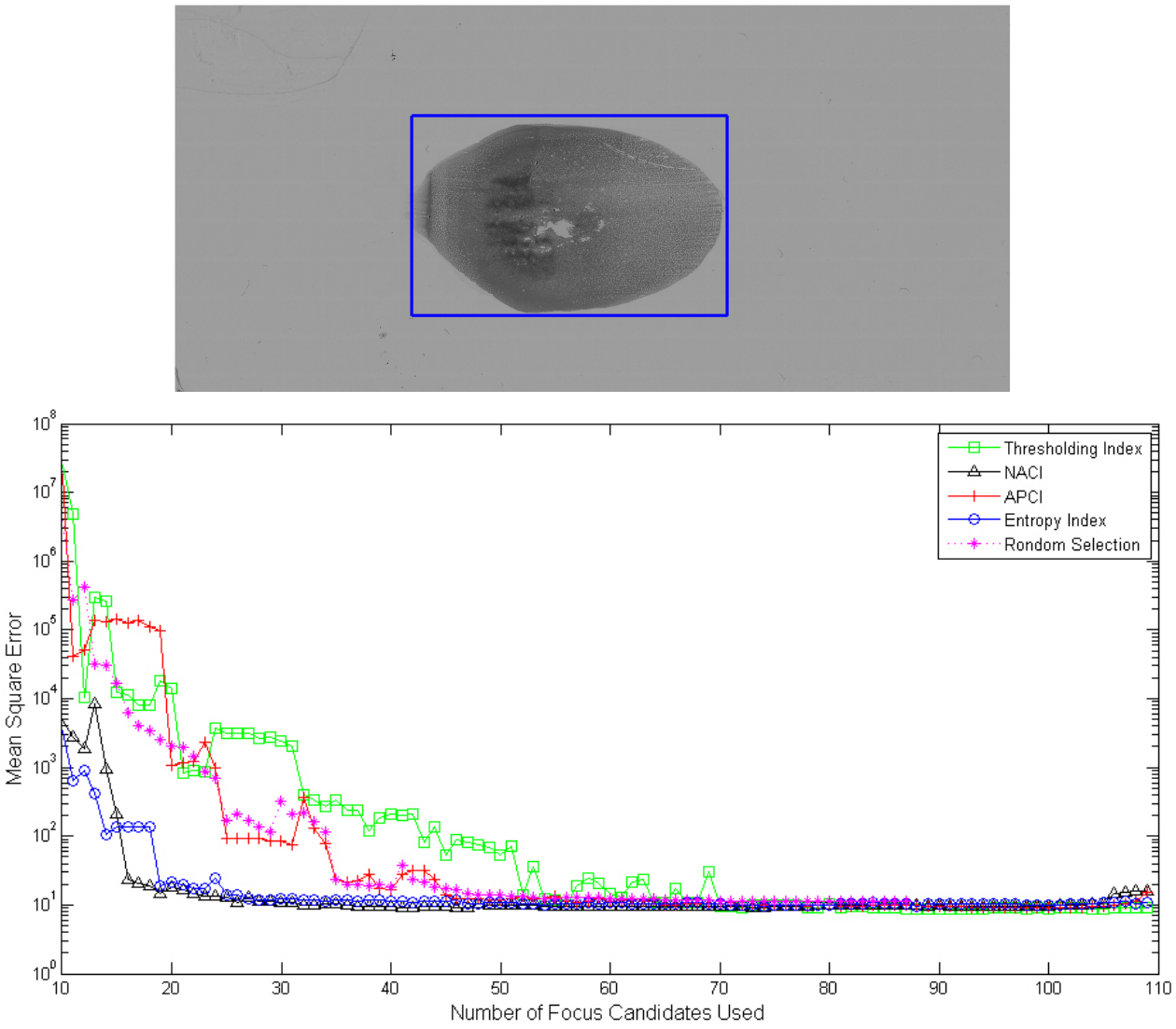


Figure 1: Top: low-resolution of a sample slide (Row 1 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

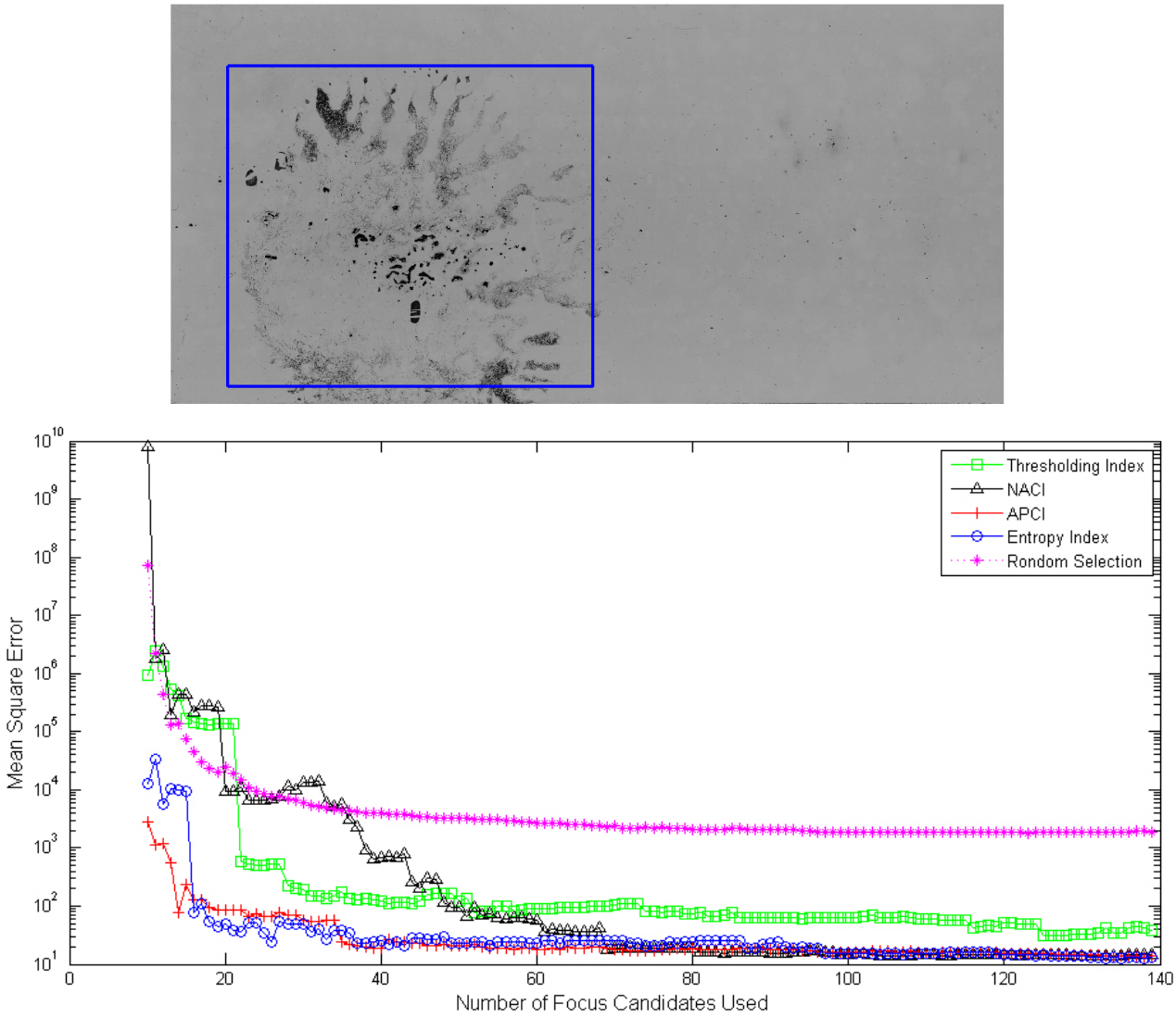


Figure 2: Top: low-resolution of a sample slide (Row 2 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

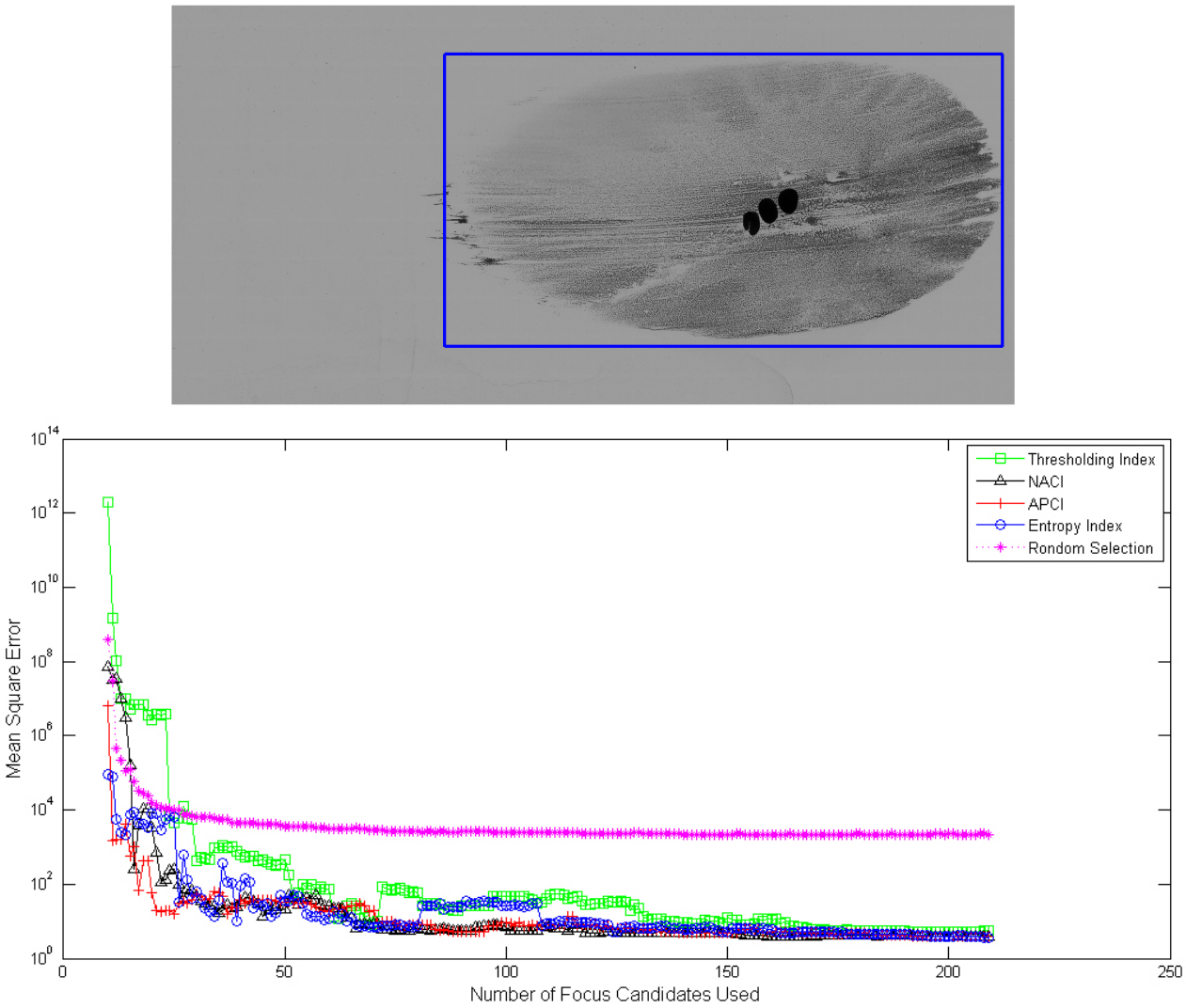


Figure 3: Top: low-resolution of a sample slide (Row 3 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

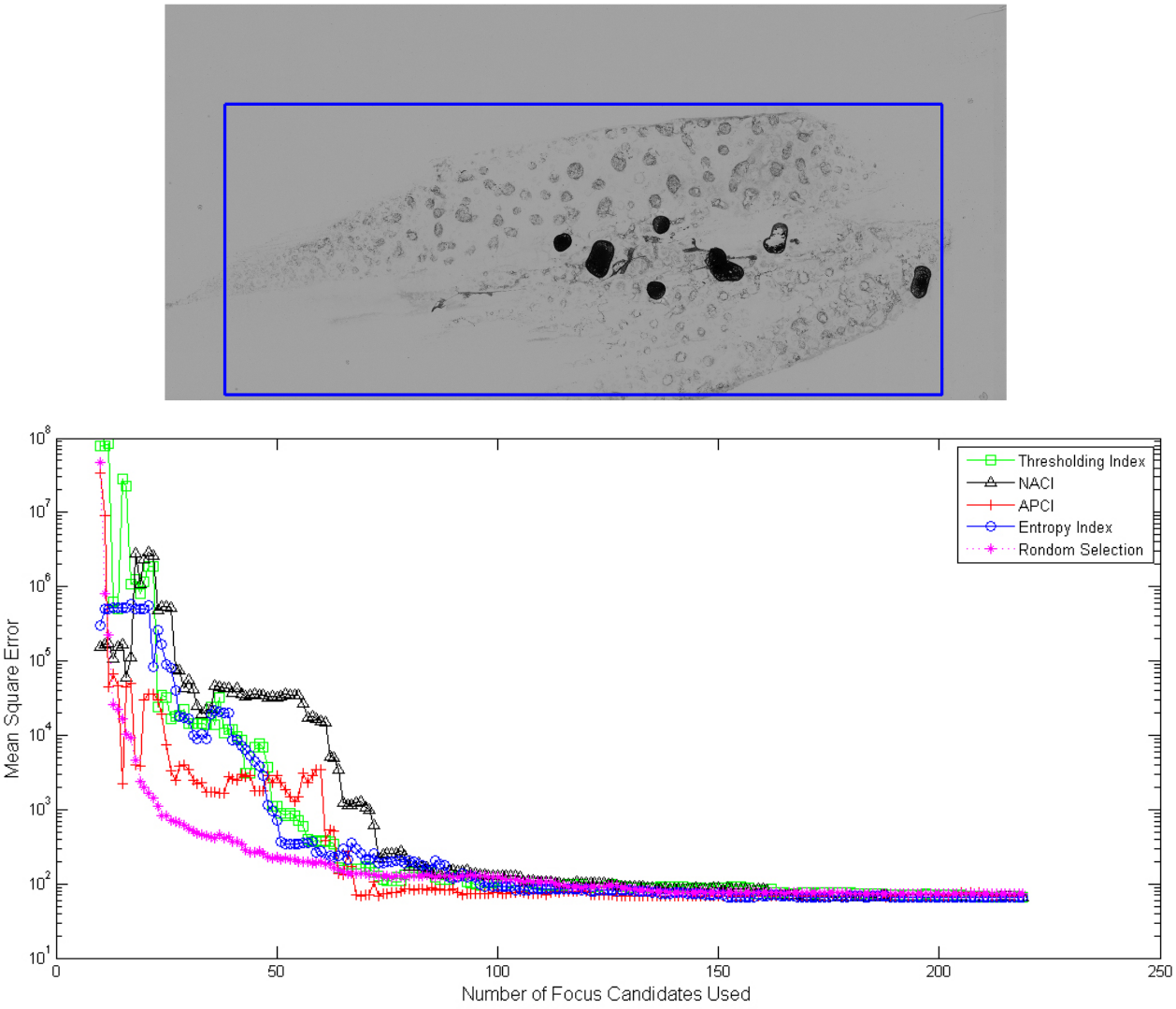


Figure 4: Top: low-resolution of a sample slide (Row 4 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

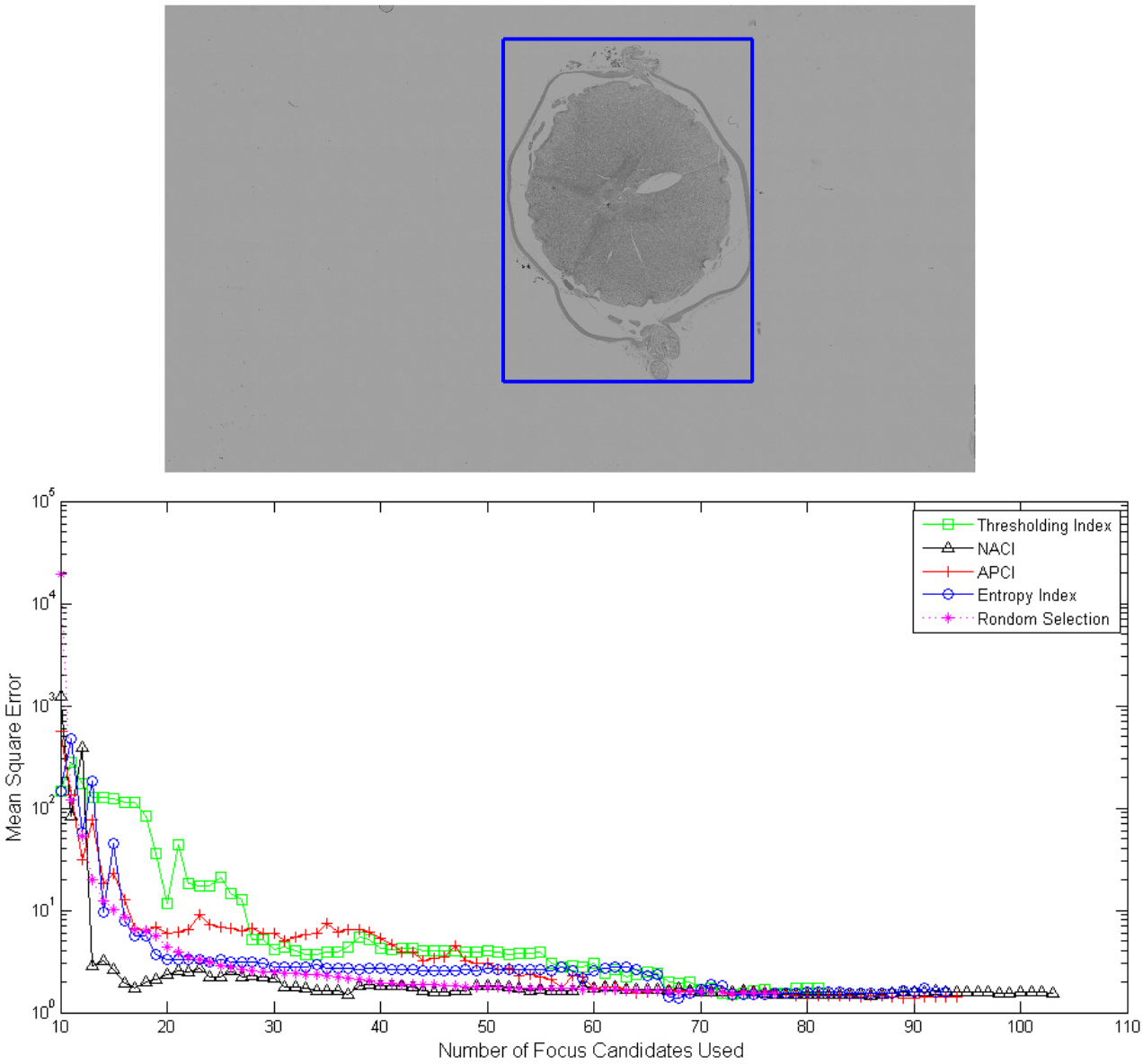


Figure 5: Top: low-resolution of a sample slide (Row 5 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

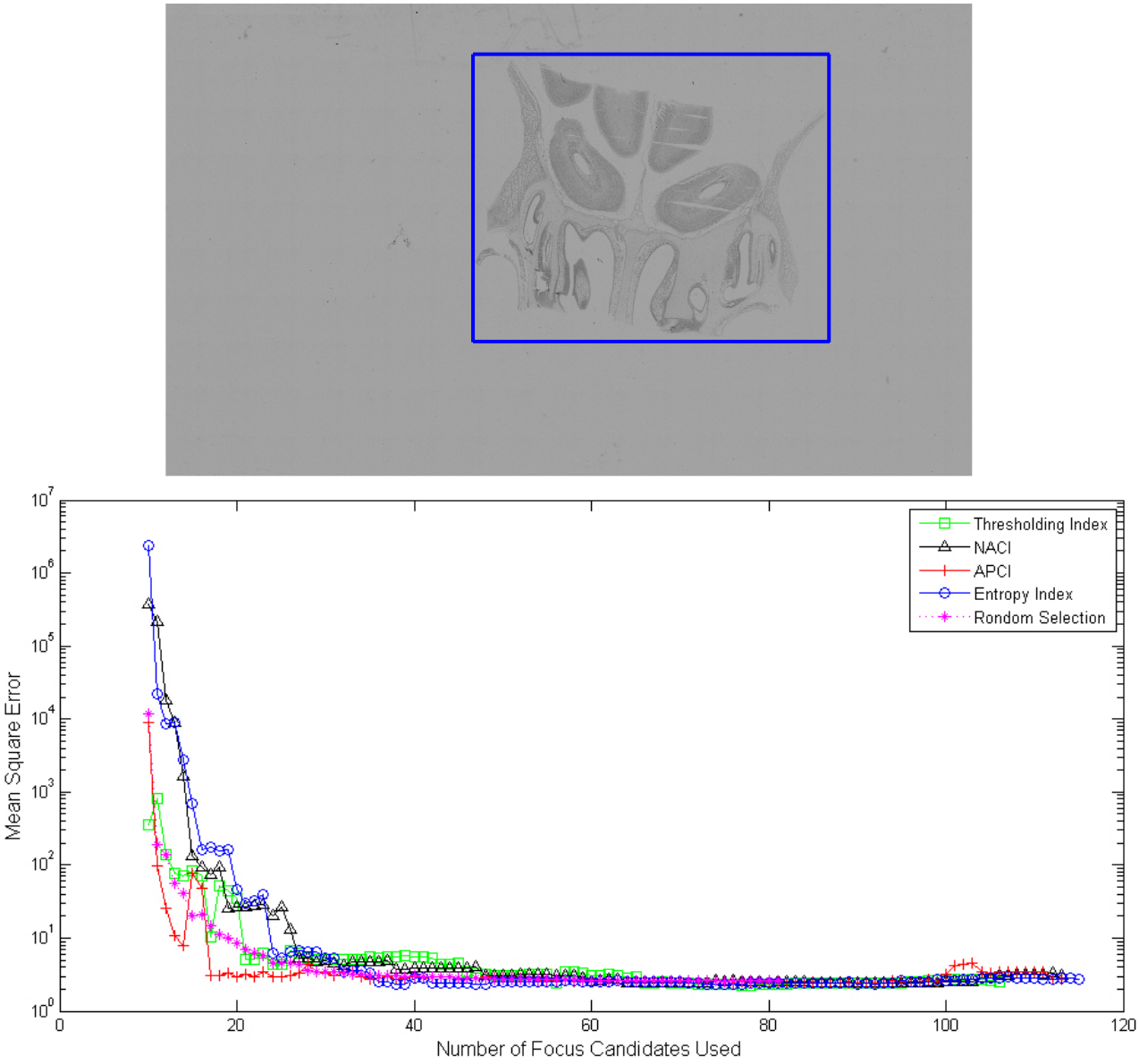


Figure 6: Top: low-resolution of a sample slide (Row 6 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

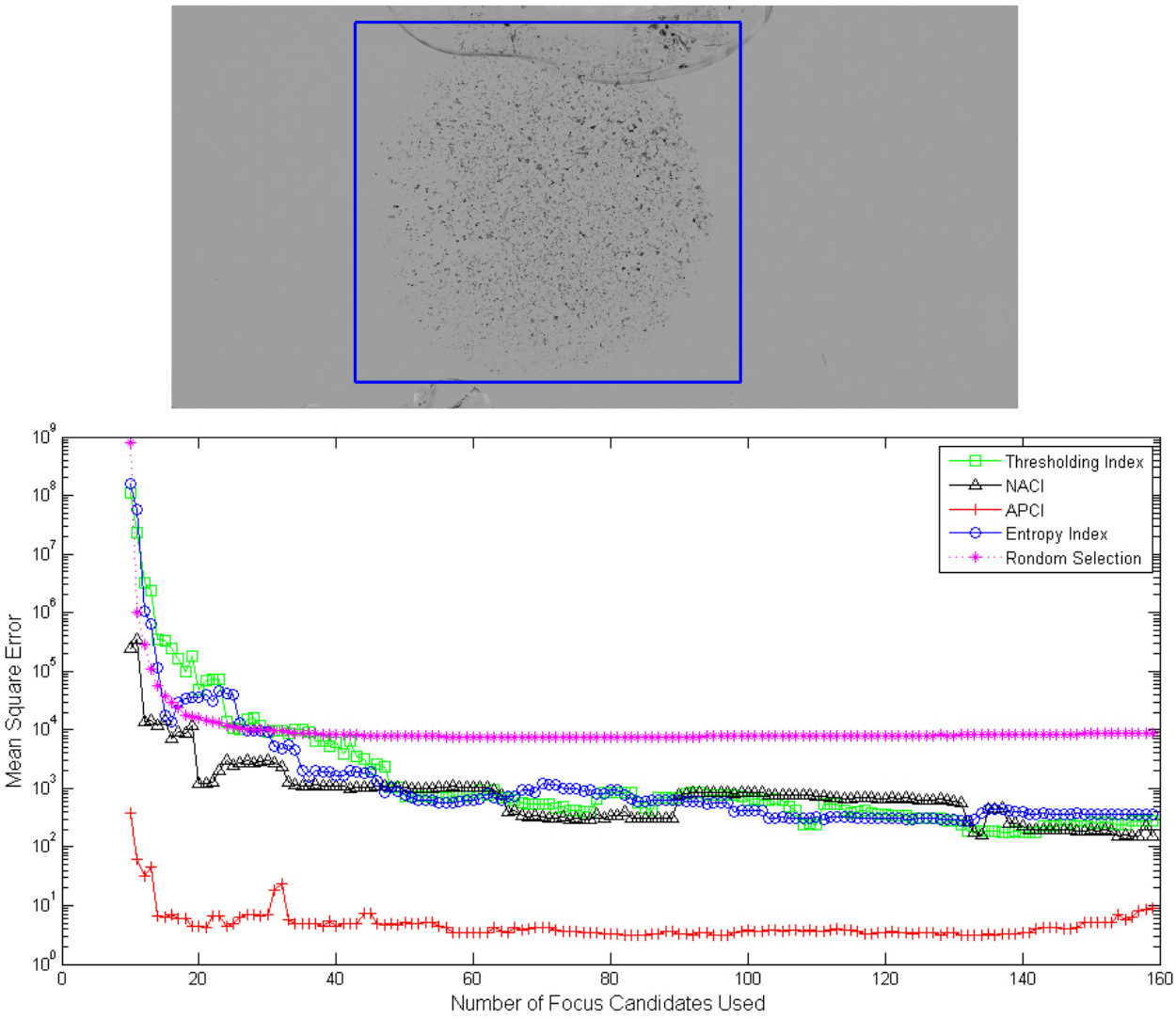


Figure 7: Top: low-resolution of a sample slide (Row 7 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.

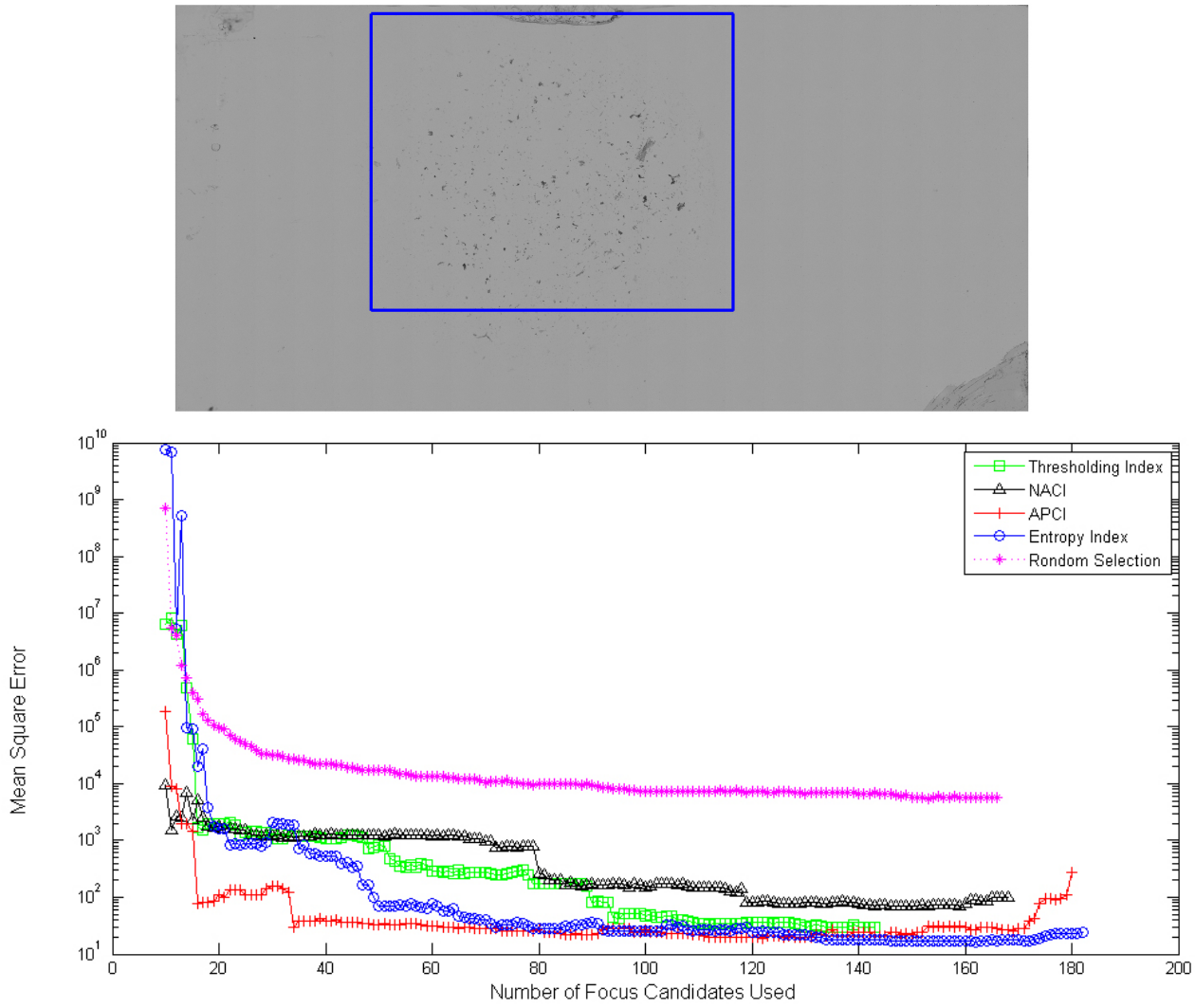


Figure 8: Top: low-resolution of a sample slide (Row 8 in Table. 3.7) and the blue box shows the part of the slide used for focus map evaluation. Bottom; The mean square error of the interpolated focus maps compared with the ground truth as the number of the sampled focus candidate increase.