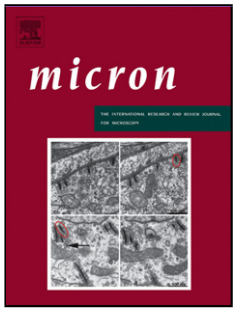
Accepted Manuscript



Title: A Method for Quantitative Analysis of Clump Thickness in Cervical Cytology Slides

Author: Yilun Fan Andrew P. Bradley

PII: S0968-4328(15)30039-1
DOI: http://dx.doi.org/doi:10.1016/j.micron.2015.09.002
Reference: JMIC 2238

To appear in: *Micron*

Received date: 30-7-2015
Revised date: 3-9-2015
Accepted date: 3-9-2015

Please cite this article as: Yilun Fan, Andrew P. Bradley, A Method for Quantitative Analysis of Clump Thickness in Cervical Cytology Slides, *Micron* (2015), http://dx.doi.org/10.1016/j.micron.2015.09.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

We developed an algorithm to estimate the thickness of cervical cytology specimens.

The proposed algorithm reached an accuracy of 1 micron at 90% of times.

The algorithm was used to quantitatively analysis of ten normal Thin-prep cervical cytology slides.

It was found that the distribution of cells is skewed towards the cover-slip (top of the slide).

It was also proved that considering the thickness of focal points produced focus maps in superior qualities compared to conventional ones.

Accepted Manuscript

*Manuscript
ACCEPTED MANUSCRIPT

A Method for Quantitative Analysis of Clump Thickness in Cervical Cytology Slides

Yilun Fan and Andrew P. Bradley

The University of Queensland, School of Information Technology and Electrical Engineering, St Lucia, QLD 4072, Australia.
Email: y.fan3@uq.edu.au

Abstract

Knowledge of the spatial distribution and thickness of cytology specimens is critical to the development of digital slide acquisition techniques that minimise both scan times and image file size. In this paper, we evaluate a novel method to achieve this goal utilising an exhaustive high-resolution scan, an over-complete wavelet transform across multi-focal planes and a clump segmentation of all cellular material on the slide. The method is demonstrated with a quantitative analysis of ten normal, but difficult to scan Pap stained, Thin-prep, cervical cytology slides. We show that with this method the top and bottom of the specimen can be estimated to an accuracy of 1 micron in 88% and 97% of the fields of view respectively. Overall, cellar material can be over 30 microns thick and the distribution of cells is skewed towards the cover-slip (top of the slide). However, the median clump thickness is 10 microns and only 31% of clumps contain more than three nuclei. Therefore, by finding a focal map of the specimen the number of 1 micron spaced focal planes that are required to be scanned to acquire 95% of the in-focus material can be reduced from 25.4 to 21.4 on average. In addition, we show that by considering the thickness of the specimen, an improved focal map can be produced which further reduces the required number of 1 micron spaced focal planes to 18.6. This has the potential to reduce scan times and raw image data by over 25%.

Keywords: virtual microscopy, digital slide, focal depths, cytology.

ACCEPTED MANUSCRIPT

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 [3, 16, 14]. The introduction of whole-slide-imaging (WSI) has been welcomed by pathologists because it enables remote consultation, quality assurance and importantly the potential for primary diagnosis for cervical cancer screening [1, 7, 2]. 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 [20]. With WSI, the specimens are first scanned as digital images and suspicious cells can be tirelessly and consistently detected by digital image processing algorithms [27]. 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 [26, 7]. 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.

The time-quality trade off is closely related to the nature of the complex three-dimensional distributions of cytology 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. Even with the advent of LBC preparation techniques cells often overlap to form thick cell clumps that span multiple focal planes [15]. In addition, important diagnostic cells are often found within these cell clumps, such as secretory cells from the endocervic and low-grade squamous intraepithelial lesions [15]. 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 [26, 8]. Additionally, interpreting cellular objects requires the microscope to operate at high magnification (40x), preferably at resolutions close to the diffraction limit ($\sim$0.2μm/pixel), where the depth-of-field is reduced to around one micron [5]. This means that even isolated cells can span multiple focal planes [7].

ACCEPTED MANUSCRIPT

Given that scanning even a single focal plane at high resolution acquires a considerable amount of data (in order of Gigabytes) and can require significant acquisition times, the number of focal planes scanned should be minimised. In this way, WSI systems can minimise the acquisition time, storage space and subsequent processing time of these images.

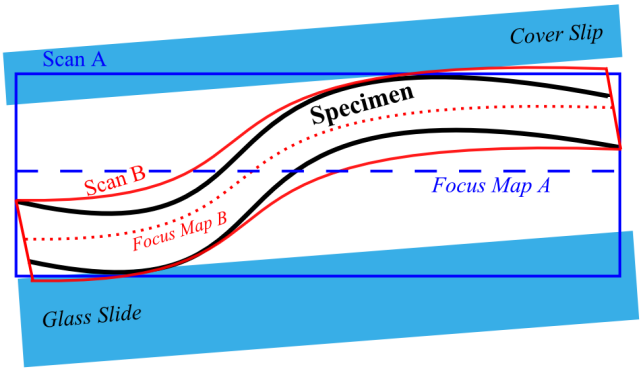


Figure 1: An illustration of potential relative positions of a microscope slide, cover-slip and specimen. Two examples of two multiple focal-plane scans (A in blue and B in red) are shown, viewed from the side and not in scale, their associated focus maps are shown as dash lines.

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. 1 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. 1 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. 1) 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

ACCEPTED MANUSCRIPT

over all possible tilt angles. Unfortunately, as Scan A in the Fig. 1 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 [10]. For example, the red dashed line in the Fig. 1 represents a focus map that better follows the topology of the specimen. The subsequent scan (B in Fig. 1) 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 [9]. 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 [10]. 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 40X), but both were outperformed by conventional glass slides [26]. 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 [8]. 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 [6]. 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 study, we propose a method to quantitatively analyse cervical cytology specimens, which estimates both the spatial location and thickness of every cell and cell clump. Specifically, the specimens are first exhaustively imaged **in three-dimensional (3D)** at high resolution and multiple focal

ACCEPTED MANUSCRIPT

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 height of each cell/clump. We purposely restrict our experimental slides to those with a normal diagnostic result so that we focus the paper 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" digital version of these specimens. In addition, we propose a novel method for focus map estimation that considers the thickness of the candidate focal points.

The paper is structured as follows, we first describe the details of the method for thickness analysis. Next, we outline the experiments that verify the effectiveness of the proposed method and demonstrate the method with quantitative analysis of ten Thin-prep cervical cancer slides. The spatial data from these slides are then used to develop and evaluate the novel method for focus map estimation.

2. Methods

The proposed method is developed specifically for analysis of LBC preparations. 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. 1, so as to image their full thickness with multiple focal planes. Each field-of-view (FOV), imaged at multiple focal planes, is then converted into a composite image with extended depth-of-field (EDF) [4] prior to segmentation of cell clumps and nuclei [16]. 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 3-dimensional map of all segmented objects (cells, clumps, nuclei). The main steps of the proposed method are illustrated in Fig. 2 and described in detail below.

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 depth-of-field is generally small at

ACCEPTED MANUSCRIPT

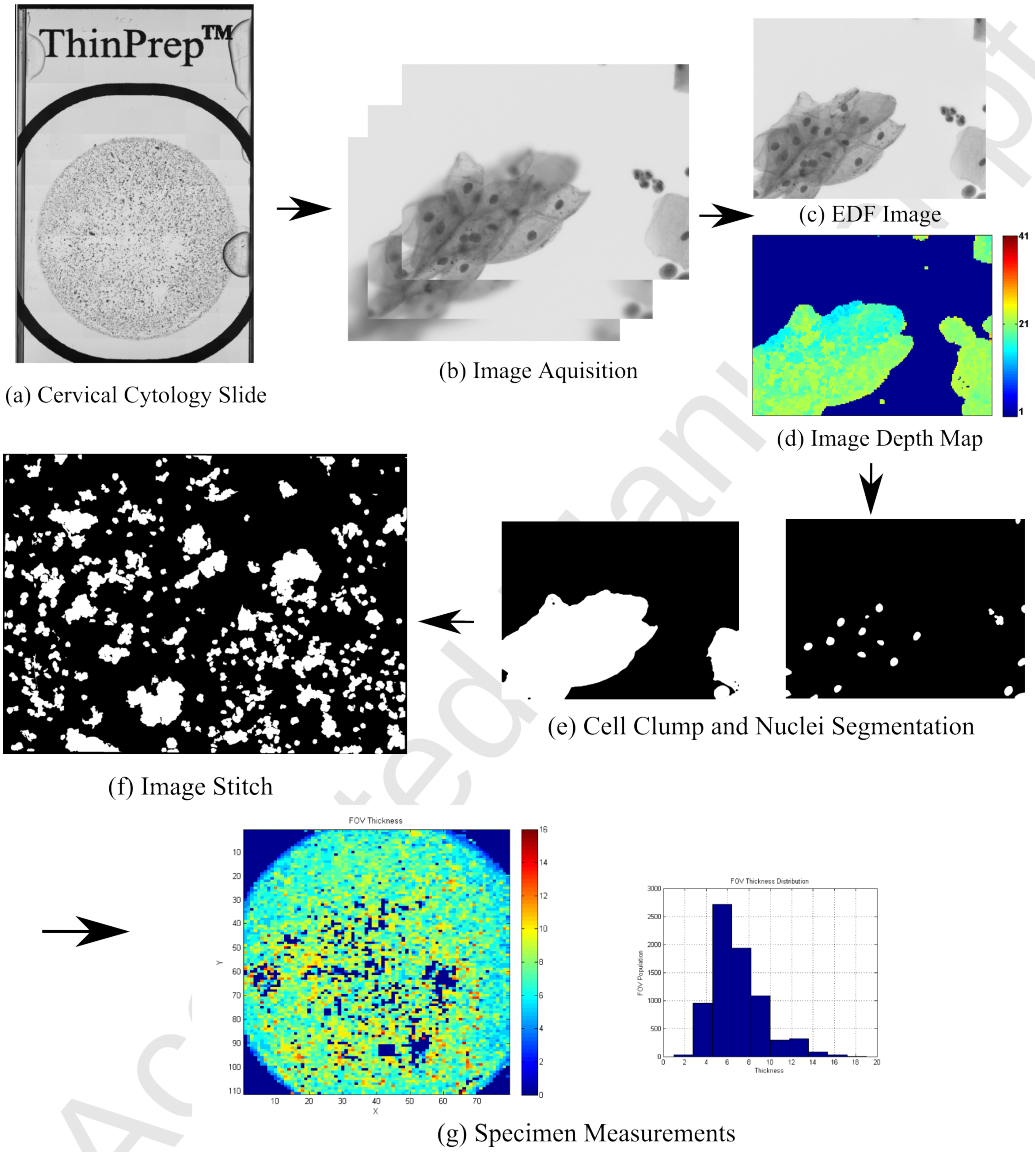


Figure 2: The cervical cytology slide (a) is first extensively scanned into 3D image stacks (b). Each 3D stack is processed to produce an extended depth-of-field image (c) and a depth map (d) showing which layer each pixel in the EDF image came from. The EDF image is then segmented for cells, clumps 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).

ACCEPTED MANUSCRIPT

high magnification. Here, we use the conventional FOV scan method, which samples FOV images sequentially over the entire specimen [9], but other scan methods such as a line scan could also be used. The scan is equivalent to implementing an 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 manually selecting the scan area of the specimen and a non-empty position near the centre of the specimen as the first focal point. 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 following FOVs. The entire specimen is then sampled FOV by FOV following a raster scan pattern with a small amount of overlap [9]. Each FOV is focused and central focal plane is found as the single best 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 used normalized variance as focus metric and fast hill-climbing search as autofocus algorithm [22, 23]. 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.

Once the FOV was focused, a z stack of images (equal number above and below the BSFP) were acquired at an interval of 1 μm, which is the same as the depth-of-field (DOF) of the imaging system. The height of 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 as the normalised variance of known empty FOV images so that empty FOVs could be detected and acquisition terminated once a single empty FOV was acquired. Image background correction was repeated every 30 minutes to cater for fluctuations in illumination experienced during these long scans (up to 60hrs per slide).

ACCEPTED MANUSCRIPT

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. 2). The generation of an EDF image simplifies the cell segmentation process to a two-dimensional space so that an existing 2D segmentation algorithm could be applied. 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 [4]. The wavelet transform provides an effective method for detection of in-focus objects because these objects produce the large wavelet coefficients [24]. A seven level wavelet transform was performed, with lowest five levels being over-sampled to provide approximate shift invariance [19]. 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 [4].

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. 2 (c), where the height of the stack is 41 layers.

The EDF images and depth maps were produced concurrently with image acquisition in order to minimise overall processing time.

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 [16]. 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

ACCEPTED MANUSCRIPT

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 [25]. 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. 2(b)). For FOV images that contained large clumps, the algorithm [16] sometimes failed because it would classify cytoplasm as background. 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.

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 [21]. Nuclei that did not meet a minimum size constraint or had an eccentricity larger than 0.9 were treated as artefacts and not analysed.

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 were 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

ACCEPTED MANUSCRIPT

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.

3. Experiments

3.1. Specimen Preparation and Imaging

As mentioned before, 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. We used an infinity-corrected $40\times/0.75$ NA objective lens with a $1.2\times$ projection lens. Consequently, the he 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, giving an effective pixel resolution of 0.3μm on all images.

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) [8]. To exclude any outlier FOVs, the segmentation results and depths maps were manually examined for quality, especially where the specimen was thick.

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

ACCEPTED MANUSCRIPT

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.

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 best single focal plane (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

ACCEPTED MANUSCRIPT

were used for both lateral axes (x and y). 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. 1. 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 "glass faithful" if at least 95% of the cellular material in both spatial (x,y) 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 depth-of-field 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.

3.4. Focus Map Estimation using Thickness

Conventionally, only the single best focal plane (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

ACCEPTED MANUSCRIPT

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 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 [10], 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 previous 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 [10]. This minimum spacing is directly related to the number of focal points required to . 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 [11], [18] and [12], a collection of results from these works documented on [13] was used in this study.

4. Results

4.1. Effectiveness of Specimen Thickness Detection

In Table. 1, we present the difference between ground truth and the proposed algorithms on detecting the positions of top and bottom layers of each

ACCEPTED MANUSCRIPT

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

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.

4.2. Quantitative Measurements of Slides

Some general statistics of ten sample slides are shown in Table 2. 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 93 μm. The range of the heights of FOVs varied between 13-30 μm 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 interquartile ranges (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

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

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.

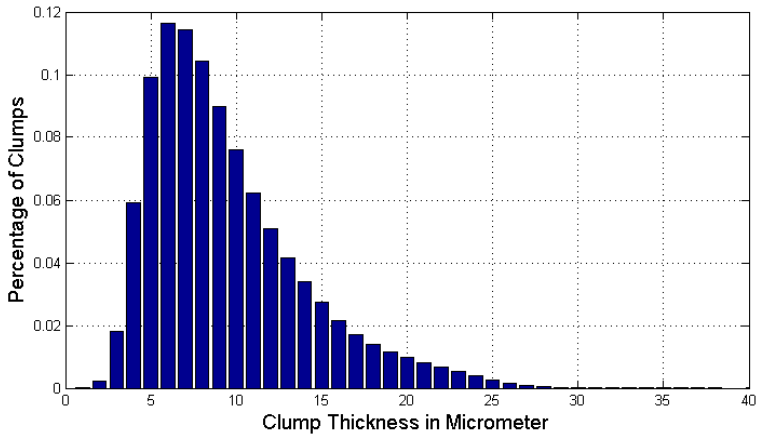


Figure 3: Histogram of Thickness of Cell Clumps from All Ten Slides.

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. 3. 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.

The specimen complexity can also be illustrated by results shown in Table. 3, 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. 1). 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

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 3: 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: 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).

Accepted Manuscript

ACCEPTED MANUSCRIPT

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.

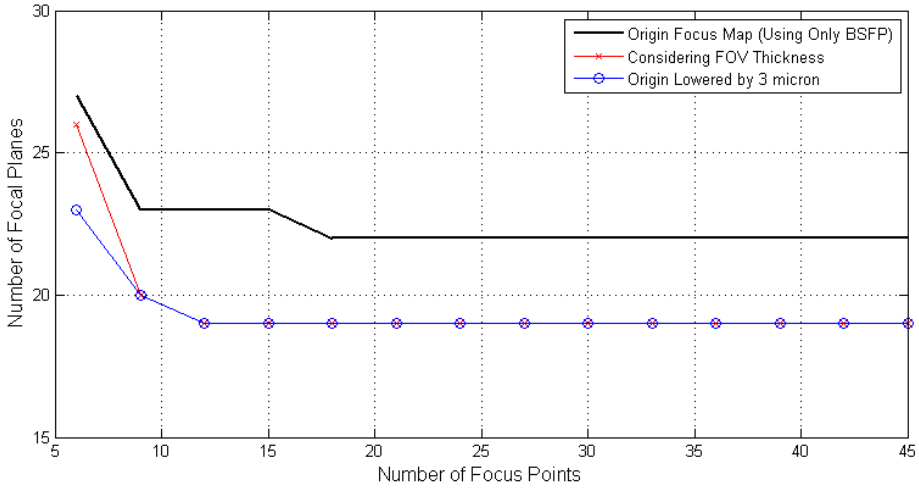


Figure 4: An example that shows the relationship between the number of focus points used for estimation of focus map and the mean number of focal planes required to acquire 95% of the in-focus material. Here we compare three focus map generating strategy: considering only the height of BSFP of the focus 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.

4.3. Estimation of Focus Map Considering Specimen Thickness

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. 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 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

ACCEPTED MANUSCRIPT

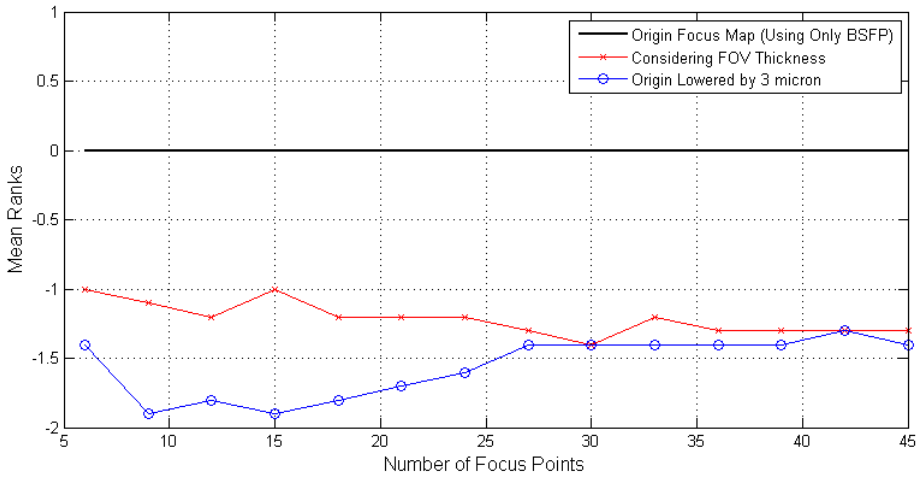


Figure 5: Ranks of the quality of estimated focus maps that consider only the height of best single focal plane (BSFP) of the focus points (the origin map), consider the thickness of focus points and are lowered from the origin focus map (by 3μm). The origin focus maps are used as the benchmark (with rank of zero) and the ranks are the mean of 100 simulations.

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. The comparison of the methods with all ten slides can be better illustrated by ranking them according to the required number of focal planes, shown in Fig. 5. 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.

5. Discussion

The evaluation results in Table. 1 show 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,

ACCEPTED MANUSCRIPT

which can then be utilised to improve the focus map required for rapid whole slide digital acquisition. The result in Fig. 5 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. 1 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 extended depth-of-field 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.

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 depth-of-field 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. 1) 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 3 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 interquartile range (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

ACCEPTED MANUSCRIPT

than one micron (0.81μm). However, as Figure 4 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.

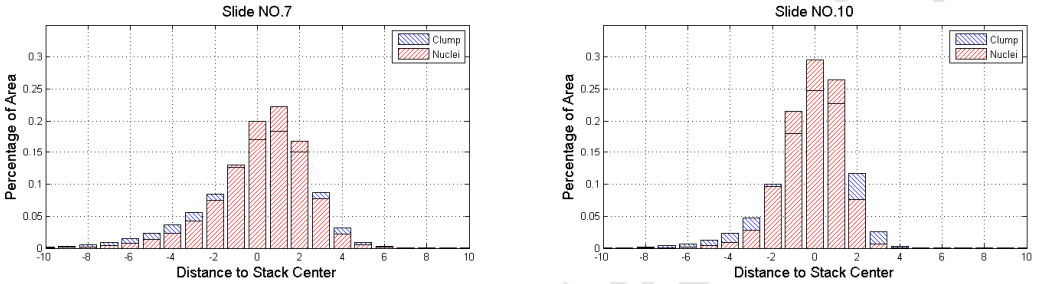


Figure 6: The height distribution of cell clumps and nuclei in the acquired image stack. The axis right is the direction of cover-slip.

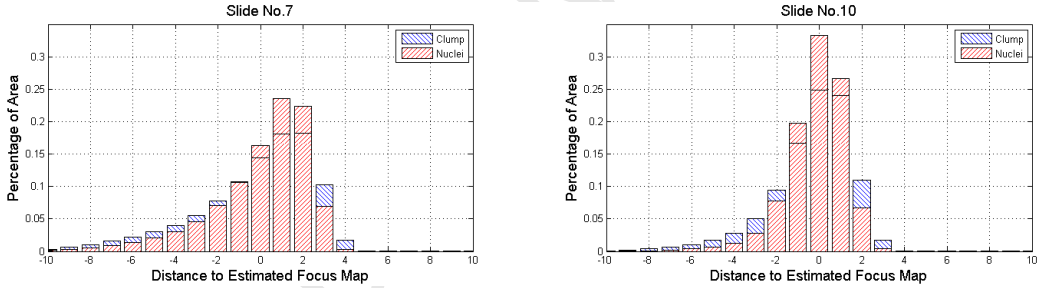


Figure 7: The height distribution of cell clumps and nuclei respect to the ground truth focus map in cubic polynomial. The axis right is the direction of cover-slip.

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. 2, 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)

ACCEPTED MANUSCRIPT

are shown in Fig. 6. It can be seen that the 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 specimen.

Similarly, the distribution of the cell clumps and nuclei across the entire slide showed the same trend, as illustrated in Fig. 7. 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 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. 6 and Fig. 7 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. 7). Therefore, sampling only 3 layers at an interval of 1μm is able to capture sharp images of most of cells. This echoes the work by Bernd *et al* who claimed that it is possible to find and scan one 'master' scan layer through the specimen where most of objects are in focus [14]. 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. 5 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. 6, 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

ACCEPTED MANUSCRIPT

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 4.3) 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. 5 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. suchin 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 cover-slip may result in outliers giving incorrect focal points that bias the focus map. However, these outliers can be detected and removed by some methods, such as by machine learning in [14], and an FOV evaluation method proposed in [10]. 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 [17]. In addition, the experiment for estimating the focus map, and in particular the selection of the 95% "glass faithful" criterion should be extended to include a diagnostic comparison with the involvement of cytopathologists.

ACCEPTED MANUSCRIPT

6. Conclusions

We have proposed a method to enable the quantitative analysis of both the spatial distribution and thickness of whole 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 demonstrated to be effective on over 90% of the acquired fields of view. The method was utilised to extensively analyse ten PAP stained Thin-prep slides. It was found that majority of cells are located above the single best focus plane found using normalised variance as the focus metric. It was subsequently demonstrated that knowledge of the thickness of the specimen is able to increase the quality of estimated focus maps. In this way, it was shown that 95% of all in-focus cellular material can be imaged provided a focal depth of 9μm on either side of the focal map is acquired.

7. Acknowledgements

Professor Bradley is supported by an Australian Research Council Future Fellowship (FT110100623).

Accepted Manuscript

ACCEPTED MANUSCRIPT

References

[1] Al-Janabi, S., Huisman, A., Van Diest, P., 2012. Digital pathology: current status and future perspectives. Histopathology 61 (1), 1–9.

[2] Ameisen, D., Deroulers, C., Perrier, V., Yunès, J., Bouhidel, F., Battistella, M., Legrès, L., Janin, A., Bertheau, P., 2013. Stack or trash? quality assessment of virtual slides. Diagnostic Pathology 8 (Suppl 1), S23.

[3] Bentz, J., 2005. Liquid-based cytology for cervical cancer screening.

[4] Bradley, A., 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). pp. 279–284.

[5] Castleman, K., 1996. Digital Image Processing, 2nd Edition. Prentice Hall.

[6] Donnelly, A., Mukherjee, M., Lyden, E., Bridge, J., Lele, S., Wright, N., McGaughey, M., Culberson, A., Horn, A., Wedel, W., et al., 2013. Optimal z-axis scanning parameters for gynecologic cytology specimens. Journal of pathology informatics 4.

[7] El-Gabry, E., Parwani, A., Pantanowitz, L., 2014. Whole-slide imaging: widening the scope of cytopathology. Diagnostic Histopathology 20 (12), 456–461.

[8] Evered, A., Dudding, N., 2011. Accuracy and perceptions of virtual microscopy compared with glass slide microscopy in cervical cytology. Cytopathology 22 (2), 82–87.

[9] Fan, Y., Gal, Y., Bradley, A., 2013. Performance analysis of three microscope slide scanning techniques. In: Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on. IEEE, pp. 1–6.

[10] Fan, Y., Gal, Y., Bradley, A., 2014. An algorithm for microscopic specimen delineation and focus candidate selection. Micron 66, 51–62.

Accepted Manuscript

ACCEPTED MANUSCRIPT

[11] Graham, R., Lubachevsky, B., Nurmela, K., Östergård, P., 1998. Dense packings of congruent circles in a circle. Discrete Mathematics 181 (1), 139–154.

[12] Grosso, A., Jamali, A., Locatelli, M., Schoen, F., 2010. Solving the problem of packing equal and unequal circles in a circular container. Journal of Global Optimization 47 (1), 63–81.

[13] Herzlich, 2015. The best known packings of equal circles in a circle (complete up to n=2600).
URL `http://hydra.nat.uni-magdeburg.de/packing/cci/`

[14] Lahrmann, B., Valous, N., Eisenmann, U., Wentzensen, N., Grabe, N., 2013. Semantic focusing allows fully automated single-layer slide scanning of cervical cytology slides. PloS one 8 (4), e61441.

[15] Lee, R., McClintock, D., Laver, N., Yagi, Y., 2011. Evaluation and optimization for liquid-based preparation cytology in whole slide imaging. Journal of pathology Informatics 2.

[16] Lu, Z., Carneiro, G., Bradley, A., April 2015. 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.

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

[18] Lubachevsky, B., Graham, R., 1997. Curved hexagonal packings of equal disks in a circle. Discrete & Computational Geometry 18 (2), 179–194.

[19] Mallat, S., 1989. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 11 (7), 674–693.

[20] Malm, P., Balakrishnan, B., Sujathan, V., Kumar, R., Bengtsson, E., 2013. Debris removal in pap-smear images. Computer methods and programs in biomedicine 111 (1), 128–138.

Accepted Manuscript

ACCEPTED MANUSCRIPT

[21] Matas, J., Chum, O., Urban, M., Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. Image and vision computing 22 (10), 761–767.

[22] Sun, Y., Duthaler, S., Nelson, B. J., 2004. Autofocusing in computer microscopy: selecting the optimal focus algorithm. Microscopy research and technique 65 (3), 139–149.

[23] Tello-Mijares, S., Flores, F., Bescos, J., Valdez, E., 2013. Efficient autofocus method for sequential automatic capturing of high-magnification microscopic images. Chinese Optics Letters 11 (12), 121102–121102.

[24] Valdecasas, A., Marshall, D., Becerra, J., Terrero, J., 2001. On the extended depth of focus algorithms for bright field microscopy. Micron 32 (6), 559–569.

[25] Vedaldi, A., Soatto, S., 2008. Quick shift and kernel methods for mode seeking. In: Computer Vision-ECCV 2008. Springer, pp. 705–718.

[26] Wright, A., Smith, D., Dhurandhar, B., Fairley, T., Scheiber-Pacht, M., Chakraborty, S., Gorman, B., Mody, D., Coffey, D., May 2013. Digital slide imaging in cervicovaginal cytology: a pilot study. Archives of pathology and laboratory medicine 137 (5), 618–624.

[27] Zhao, T., Wachman, E., Farkas, D., 2004. A novel scheme for abnormal cell detection in pap smear images. In: Biomedical Optics 2004. International Society for Optics and Photonics, pp. 151–162.

Accepted Manuscript