Micron 80 (2016) 73–82

Contents lists available at ScienceDirect

Micron

journal homepage: www.elsevier.com/locate/micron

A method for quantitative analysis of clump thickness in cervical cytology slides

Yilun Fan*, Andrew P. Bradley

The University of Queensland, School of Information Technology and Electrical Engineering, St Lucia, QLD 4072, Australia

ARTICLE INFO

Article history:
Received 30 July 2015
Received in revised form 3 September 2015
Accepted 3 September 2015
Available online 28 September 2015

Keywords:
Virtual microscopy
Digital slide
Focal depths
Cytology

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 materials 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 μm in 88% and 97% of the fields of view respectively. Overall, cellular material can be over 30 μm thick and the distribution of cells is skewed towards the cover-slip (top of the slide). However, the median clump thickness is 10 μm and only 31% of clumps contain more than three nuclei. Therefore, by finding a focal map of the specimen the number of 1 μm 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 μm spaced focal planes to 18.6. This has the potential to reduce scan times and raw image data by over 25%.

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; Lu et al., 2015a; Lahrmann et al., 2013). 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 (Al-Janabi et al., 2012; El-Gabry et al., 2014; Ameisen et al., 2013). 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 (Wright et al., 2013; El-Gabry et al., 2014). Among them, the scan speed is the main bottleneck, because acquisition has to 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 (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 intraepithelial 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 (Wright et al., 2013; Evered and Dudding, 2011). Additionally, interpreting cellular objects requires the microscope to operate at high magnification (40×), preferably at resolutions close to

* Corresponding author.
E-mail address: y.fan3@uq.edu.au (Y. Fan).

http://dx.doi.org/10.1016/j.micron.2015.09.002
0968-4328/

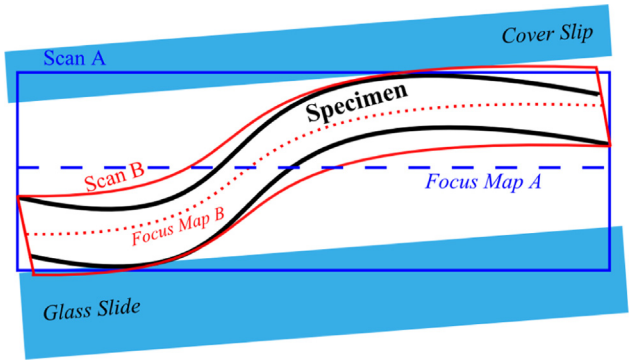

Fig. 1. An illustration of potential relative positions of a microscope slide, cover-slip and specimen. Two examples of two multiple focal-plane scans (Scan A and Scan B) are shown, viewed from the side and not in scale, their associated focus maps are shown as dash lines.

the diffraction limit (~0.2 μm/pixel), where the depth-of-field is reduced to around 1 μm (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 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.

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 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 over all possible tilt angles. Unfortunately, as Scan A in 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 (Fan et al., 2014). For example, the red dashed line in 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 (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., 2014). 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×), 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 1.5 μm interval achieved better diagnostic accuracy than those scanned with 5 focal planes at 1 μm interval (Evered and Dudding, 2011). A more recent work argues that scanning only three focal planes (with 1 μm 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 analyses 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 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 Thinprep 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 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) (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 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.