Chapter

Cytopathology Using High Resolution Digital Holographic Microscopy

Sarita Ahlawat, Purnima Sharma, Ankita Pandey, Durga Bisht, Aanisa Jan, Apoorv Pant, Ritika Malik, Sandeep R. Mathur, Kiran Agarwal, Smita Singh, Meeta Singh and Kedar Khare

Abstract

We summarize a study involving simultaneous imaging of cervical cells from Pap-smear samples using bright-field and quantitative phase microscopy. The optimization approach to phase reconstruction used in our study enables full diffraction limited performance from single-shot holograms and is thus suitable for reducing cost of a quantitative phase microscope system. Over 48000 cervical cells from patient samples obtained from three clinical sites have been imaged in this study. The clinical sites used different sample preparation methodologies and the subjects represented a range of age groups and geographical diversity. Visual examination of quantitative phase images of cervical cell nuclei show distinct morphological features that we believe have not appeared in the prior literature. A PCA based analysis of numerical parameters derived from the bright-field and quantitative phase images of the cervical cells shows good separation of superficial, intermediate and abnormal cells. The distribution of phase based parameters of normal cells is also shown to be highly overlapping among different patients from the same clinical site, patients across different clinical sites and for two age groups (below and above 30 years), thus suggesting robustness and possibility of standardization of quantitative phase as an imaging modality for cell classification in future clinical usage.

Keywords: cervical cell imaging, quantitative phase, cell classification and characterization, early cancer diagnosis

1. Introduction

Cervical cancer is the fourth most prevalent cancer among women worldwide. Human Papillomavirus (HPV) is known to be the main cause of cervical cancer [1]. It is well-known that cervical cancer has long latency period. Pre-malignant abnormalities in cervical cells can take up to a decade to progress to carcinoma. Early diagnosis of pre-cancerous cervical cells and treatment help in halting the progression of this fatal cancer [2]. The five year survival rate for patients suffering from cervical cancer has been documented to be over 60% [3]. In spite of the long latency

period for cervical cancer to develop, the mortality rate among cervical cancer patients is high in developing countries due to shortage of skilled clinicians and lack of effective screening tools [4, 5].

The cervix which is the innermost part of uterus is sub-divided between the endo-cervix and ecto-cervix regions. The endo-cervix is composed of glandular cells whereas the ecto-cervix is made up of squamous cells. Transformation zone is the place where these regions adjoin, and where most of the cervical cancer is known to originate [6]. Starting with the transformation zone the cells in the squamous region are typically classified as basal, para-basal, intermediate and superficial respectively. The majority of cells in a typical Pap-smear cell sample used for examination by clinicians are from the intermediate and superficial outer layers. The classification of precancerous cells as low grade and high grade is established through Bethesda system [7, 8] which is based on morphological changes in the cells (particularly the cell nuclei). Traditionally the detection of precancerous cervical cells is primarily performed using cytological screening. The widespread usage of liquid based cytology (LBC) in recent years has made the process of sample preparation and examination more uniform. However, screening methods based on visual inspection can suffer from both inter- and intra-observer variability. Machine learning based approaches have gained attention in this regard [9–13] and standard benchmark datasets of cervical cell images have been created [14, 15]. The goal of machine learning approaches is to make the process of cell classification at least semi-automated and to provide an assisting tool for cyto-pathologists.

The machine learning based studies have focused mainly on the 2D bright-field images of the cervical cells and their nuclei. Since cells are 3D objects, we believe that additional morphological information in the third (depth) dimension of cells, if available, can provide new information and help any image based cell classification. Digital Holographic Microscopy (DHM) is an interferometric imaging technology [16–18] which can fill this gap and provide quantitative phase information that may then be related to the depth dimension of the cells. When a coherent beam of light of wavelength λ is transmitted through a cell, the wave-front undergoes a phase change given by:

$$\phi(x,y) = \frac{2\pi}{\lambda}\int dz \;\; n(x,y,z). \tag{1}$$

Here n(x,y,z) represents the refractive index distribution within the cell relative to its surroundings. It is important to understand that phase provides new non-redundant information that cannot be derived from the usual 2D bright-field images. The phase change $\phi(x,y)$ cannot be measured directly by a 2D array sensor but may be recorded in the form of an interference fringe pattern. Here the coherent light source is first split into two beams, one of the beams passes through the cell sample and the other reference beam travels through free space before the two beams are recombined to record an interference pattern. As per Eq. (1), the phase function contains information on optical path length (product of refractive index and thickness) through the cell sample at location (x,y). While quantitative phase images have been shown to provide interesting new information about cancer cell morphology [19–22], clinically this modality is not yet popular and clinician are not trained to interpret quantitative phase images. We therefore follow a protocol where a focused bright-field image of a cervical cell is recorded along with phase image of the cell in the same focus plane. This way the clinicians can correlate with their traditional knowledge and treat phase images as an additional channel of morphological information. Recently we demonstrated such an approach for unsupervised organization of cervical cell images [23]. In the present imaging study

over a much larger sample size with samples collected from different clinical sites, we examine the structural changes in phase images of cervical cell nuclei and highlight their potential importance for cell classification. Even though 2D images are able to distinguish between the major stages of normal cells, the phase images allow one to observe the morphological changes as the cells evolve through these stages. Additionally we examine the structural characteristics of abnormal samples as identified by practicing cyto-pathologists.

Traditional DHM systems are based on single-shot off-axis interference configuration or the multi-shot phase shifting con-figuration. The single-shot off-axis systems are simpler and cheaper to build but the conventional Fourier filtering approach for phase reconstruction in these systems leads to sub-optimal phase resolution. The multi-shot phase shifting configurations offer full resolution but are hardware intensive and require stringent vibration isolation thus making them difficult to employ in clinical settings. In recent years our group has developed optimization based phase reconstruction algorithms [24–27] for single-shot DHM systems which offer the simplicity of hardware without compromising on resolution and quantitative phase accuracy. The full diffraction-limited resolution capability of our system allows us to treat the bright-field and phase images on par (with respect to their lateral resolution). The single-shot operation also reduces the cost of building a DHM system making it more accessible for wider deployment. Based on our imaging study we find that the phase images contain important morphological information associated with different classes of normal as well as abnormal cells. Further this phase information is seen to be robust across the samples from three clinical sites. Also the samples consisted of age group of 17–60 years of the subjects. The cell morphology captured as numerical parameters from the phase images can provide valuable additional information to clinicians over what they usually access with routine bright-field microscopy. Our results suggest that phase imaging can become an important clinical modality, and it should be possible to design phase-based software tools for clinicians to make better informed decisions with this new information. The Chapter is organized as follows. In Section 2 we explain the technique of digital holographic microscopy (DHM) and the nature of quantitative phase images along with our phase reconstruction methodology. Section 3 briefly describes the details of the cell samples used. The results are discussed in Section 4. In Section 4.1 we start by showing images of cervical cells in both bright-field and phase modalities to illustrate morphological changes in cervical cell nuclei. This is followed by PCA analysis of the cell data based on the morphological parameters derived from the cell images in Section 4.2. In Section 4.3 we describe our analysis to understand if the most important phase parameters for normal cells are consistent across different patients from same clinical site, across different clinical sites and between different age groups. Finally in Section 5 we provide concluding remarks.

2. Digital holographic microscopy (DHM)

DHM is an interferometric modality where the recorded image data $H(x,y)$ represents interference between the object wave $O(x,y)$ representing the light which has interacted with the cell sample and the reference wave $R(x,y)$ that has not interacted with the sample is described as:

$$H = |R|^2 + |O|^2 + R^* O + RO^* . \tag{2}$$

Here $*$ represents the complex conjugation of the corresponding wave-function. In the image plane holography case as in the present study, $O(x,y)$ represents the

resultant image field corresponding to the cell sample slide when observed through a 40x infinity corrected imaging system [23]. The interference is possible due to the use of a laser source which ensures that the object and reference waves remain temporally coherent at the detector plane and produce interference fringes with good contrast. Since our DHM system is also fitted with a white light LED illumination which allows recording of the cell sample in the usual bright-field mode for ease of interpretation by a clinician.

Reconstruction of single-shot holograms is traditionally performed using the Fourier transform method. However, due to the low-pass filtering nature of this method image plane phase recoveries with full pixel resolution cannot be obtained using this approach. This poses a problem as the bright-field images available will then seem to have higher resolution even though both have been recorded using the same microscope objective. In order to have both bright-field and phase images with same diffraction-limited resolution, we reconstruct of the complex object wave $O(x,y)$ using a sparse optimization method that has been developed by our group in recent years. In particular, recovery of the complex image field $O(x,y)$ is posed as an optimization problem where we minimize a cost function of the form:

$$\begin{aligned} C(O,O^*) &= C_1 + C_2 \\ &= \left\| H - \left(|R|^2 + |O|^2 + R^* O + R O^* \right) \right\|^2 + \psi(O,O^*). \end{aligned} \tag{3}$$

Here $\| \ldots \|^2$ denotes the squared L2-norm of the quantity inside. The reference beam $R(x,y)$ is estimated by a separate calibration step involving recording of a straight line interference fringe pattern without any sample followed by accurate estimation of carrier frequency to fractional fringe accuracy [28]. The first term of the cost function represents the least square data fit and the second term $\psi(O,O^*)$ is a suitable image domain constraint. We use the modified Huber penalty function as a constraint and use an adaptive alternating minimization scheme explained in detail elsewhere [23, 26] for recovering the complex object function $O(x,y)$ in the image plane. The modified Huber penalty is defined as:

$$\psi(O,O^*) = \sum_{k=all\ \ pixels} \left[\sqrt{1 + \frac{|\nabla O_k|^2}{\delta^2}} - 1 \right]. \tag{4}$$

The tuning parameter δ is made proportional to the median of the gradient magnitudes of the image solution in a given iteration. The Huber penalty acts like the edge preserving Total Variation penalty at pixels where the gradient magnitude $|\nabla O|$ is much larger than δ and acts like the smoothing quadratic penalty for pixels where the gradient magnitude is small compared to δ. Further the adaptive optimization strategy makes sure that the change in the solution due to error minimizing step is balanced by that due to Huber minimization step in every iteration. We point out that the optimization problem above involves real valued data (hologram H) whose solution is complex valued. The steepest descent directions evaluated in the algorithm need to be evaluated using Wirtinger derivatives with respect to O^*. In particular we note that the Wirtinger derivatives for the two terms of the cost function in Eq. (3) is given by:

$$\nabla_{O^*} C_1 = -2\left[H - |R+O|^2 \right] \cdot (R+O), \tag{5}$$

and

$$\nabla_{O^*} C_2 = -\nabla \cdot \left[\frac{\nabla O}{\sqrt{1 + \frac{|\nabla O|^2}{\delta^2}}} \right]. \tag{6}$$

It is important to note that the optimization procedure operates fully in the image domain making it possible to employ it over a region of interest and thus allowing full resolution reconstruction in near real time. In our study, a 256×256 pixel ROI phase reconstruction requires few seconds (<25 iterations) in a MATLAB implementation on a desktop with 3.1 GHz processor. The data consistency error for the reconstructed solution is within 5% relative error. A user can therefore select a region of interest near a cell nucleus for object wave reconstruction. The resolution and noise advantage of this optimization procedure over traditional Fourier filtering approach has been shown in a series of publications [24–27, 29], as a result, we will not discuss this point here once again. However for completeness we summarize the advantages of the optimization method in comparison to the traditional methods for image plane hologram processing in **Table 1**.

The phase map $\phi(x,y)$ as in Eq. (1) is the argument of the recovered complex object field $O(x,y)$ and is given by arctangent of the ratio of imaginary and real parts:

$$\phi(x,y) = \arctan\left(\frac{Im[O(x,y)]}{Re\,[O(x,y)]}\right). \tag{7}$$

Since the arctangent function is defined only over the range $[-\pi, \pi]$ the phase map defined in Eq. (4) is wrapped. A 2D unwrapping procedure based on the transport of intensity equation (TIE) [30] has been employed in our work in order to associate physical meaning to the phase map in accordance to Eq. (1). The steps involved in imaging are summarized in supplementary (**Figure 1**). A Pap-smear sample is first imaged in both bright-field and holographic modalities using a dual mode digital holographic microscope (fabricated by Holmarc Opto-Mechatronics Pvt. Ltd., Kochi, India). The holographic (or interferometric) image is used further

Processing method	Single/Multi-shot	Resolution
Fourier filtering	Single-shot	Low resolution
Phase shifting	Multi-shot	Full diffraction-limited
Optimization	Single-shot	Full diffraction-limited

Table 1.
Summary of resolution performance of image plane digital holographic methods.

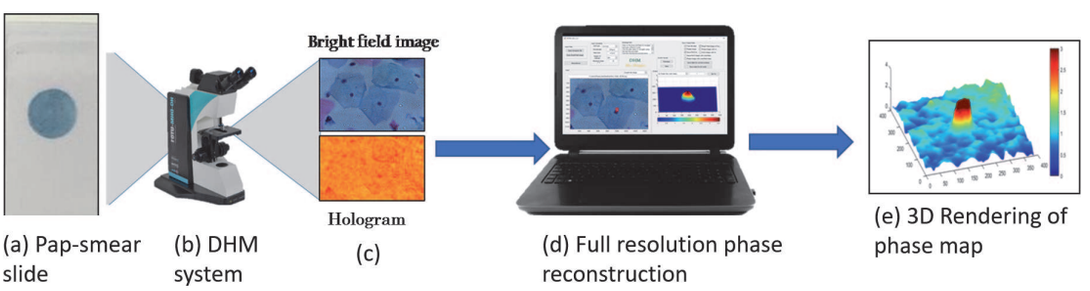

Figure 1.
Steps in imaging chain (a) Pap-smear slide, (b) dual-mode DHM system, (c) illustrative example of a bright-field image and a hologram recorded using the DHM system, (d) computer used for reading image from camera, phase reconstruction and computing morphological parameters from the bright-field and phase images, (e) illustrative example of phase map of a cell nucleus rendered as a surface plot.

Modality	Parameter	Parameter description[a]
Bright-field	Area	$\sum_{jk} M_{jk}$
	Nucleus perimeter	Number of boundary pixels of M
	Mean intensity of	R, G, B componentsAverage R, G, B values computed over mask M
	Variance of	R, G, B componentsVariance of R, G, B values computed over mask M.
	N/C Ratio	Ratio of areas (Nucleus)/(Cytoplasm) Labels assigned: low =1, mid =2, high = 3
Quantitative phase	Mean and maximum phase	$\overline{\phi}$ and maximum of ϕ computed over mask M
	Optical volume	(Area) $\times$ $(\overline{\phi})$ Computed over mask M.
	Variance of phase	Variance σ_{ϕ}^2 computed over mask M.
	Roughness at scales	1, 0.75, 0.5, 0.25 $\sum_{jk} \lvert(\nabla\phi)_{jk}\rvert$ computed over mask M.
	Moment of inertia	of nucleus $\sum_{jk}\left(\phi_{jk}\right) d_{jk}^2$ d_{jk} = distance between centroid and other pixels in mask M. Provides information about material distribution.
	Shift between geometric and phase centroid	$\lvert\vec{r}_{geom} - \vec{r}_{phase}\rvert$

[a] *M denotes a binary (0, 1) mask for individual cell nucleus, q denotes phase map. Both are defined over ROI of 256 × 256 pixels centered on cell nucleus.*

Table 2.
Morphological parameters evaluated for each cell nucleus imaged in this study. More details about these parameters are provided in ***Table 1*** *of ref. [23].*

for phase reconstruction as explained above. **Table 2** provides details about a number of morphological parameters derived from the cell images in the bright-field and quantitative phase modes. The morphological parameters were decided in consultation with practicing cyto-pathologists who participated in this study. We summarize them in **Table 2** for convenience of the reader. The N/C ratio which is the ratio of nucleus to cytoplasm areas has been included as list of three labels (low = 1, medium = 2, high = 3). This is because we found that a number of cells in the patient samples appeared in clusters and it was difficult to find boundaries of cytoplasm in simple automated manner in such cases.

3. Details of samples

We imaged a total of 48,006 cervical cells from 291 Pap-smear slides from three different hospitals in Delhi: AIIMS (All India Institute of Medical Sciences, New Delhi), LHMC (Lady Hardinge Medical College, New Delhi) and MAMC (Maulana Azad Medical College, New Delhi) (see **Table 3**). The samples were upto three years old (not from current patients) and stored in the repositories at the respective sites. They were collected by following the standard protocols for the Pap-smear examination. The patients varied in age from 16–70 years and came from varied geographical locations in India. The cell samples can be prepared conventionally or with Liquid-based cytology (LBC). For each method staining is performed with Pap stain for visualization with bright-field microscopy. For our study we have used

Clinical site	Preparation method	Number of patients normal / with condition [a]	Normal cells	Abnormal cells
AIIMS, New Delhi	Liquid based cytology	66 / 33	18801	532
	(ThinPrep)			
LHMC, New Delhi	Liquid based cytology	174 / 9	28007	214
	(SurePath)			
MAMC, New Delhi	Conventional	6/ 3	438	14
Totals		**246 / 45**	**47246**	**760**

[a] *As determined by clinicians.*

Table 3.
Details about cell samples used for imaging from the three clinical sites.

both types of samples. Two of the sites used liquid based cytology (LBC) slides prepared via ThinPrep and SurePath systems; while third site used conventional Pap-smears. In LBC method, samples are collected in liquid vials and the slide is prepared semi-automatically. The advantage of LBC is uniformity in sample preparation. On the other hand in conventional cytology the sample is applied directly to a slide for microscopic investigation. **Table 3** provides details about number of cells imaged. The classification of normal vs. abnormal cells in these samples was provided based on the bright-field images by practicing cyto-pathologists (S. R. M., M. S.,K. A., S.S.). The normal cells here include the superficial and intermediate cells while the abnormal cells include LSIL, HSIL, SCC, ASC-H and ASC-US type of cells [31].

4. Results

4.1 Illustrative bright-field and phase images of various cervical cell types

In this section we begin by providing sample images of cervical cell nuclei that were obtained using our dual-mode DHM system. While quantitative information obtained in terms of morphological parameters is certainly important, a large number of clinical sites worldwide typically use visual examination of cells using a bright-field microscope for cell classification. The importance of changes in nucleus structure in cancer diagnosis is already well-known [32]. With the illustrative examples in this section, we wish to qualitatively describe the morphological features observed in quantitative phase images of cervical cell nuclei in various stages. The simultaneous presentation of bright-field images (that pathologists can correlate to) and the quantitative phase images as shown here is important in our opinion from the perspective of clinical users. It is important to note that our single-shot full resolution phase reconstruction technique allows us to observe the quantitative phase images with the same resolution as the bright-field images. The examples shown here also aim to illustrate that the information contained in the quantitative phase images is different in nature from that in the bright-field images. Quantitative analysis of the images using morphological parameters as described in **Table 2** will be provided in the following sections. In **Figure 2** we show illustrative examples of normal cells in the intraepithelial squamous layer. A progression from intermediate to superficial stages is shown in **Figure 2(a)–(j)** respectively. As a cell progresses from intermediate to superficial stage the chromatin in the cell nucleus is known to condense. The superficial cells are in the outermost layer of ecto-cervix and have

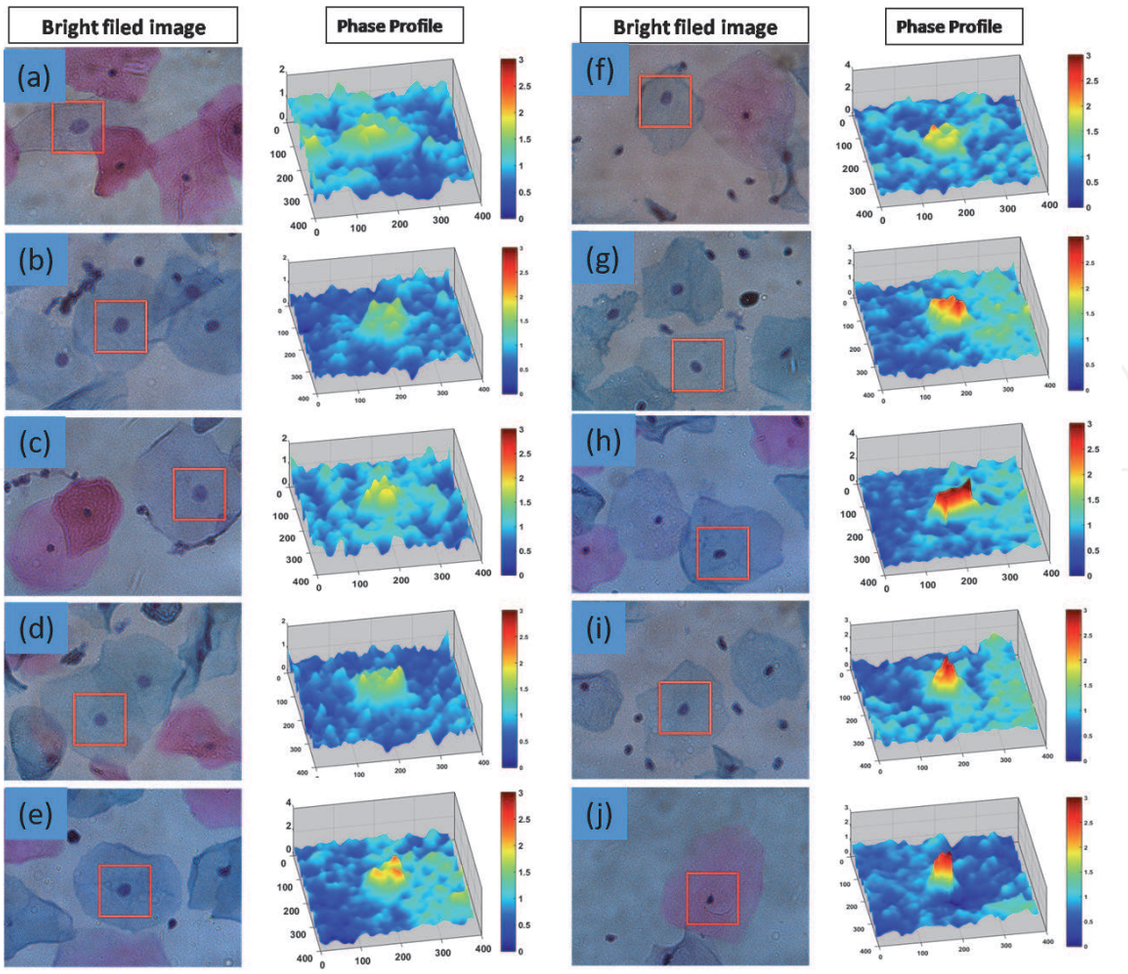

Figure 2.
Maturation of cells in intra-epithelial squamous layer from intermediate to superficial is shown from (a)-(j) in both bright-field and phase modes. The phase maps correspond to the ROIs marked in the bright-field images. As the nuclei progress to superficial stage, the nucleus area gets smaller and the phase profile is seen to get taller by approximately a factor of 2.

highly condensed pyknotic nucleus. While in the bright-field images the area of the nucleus is progressively decreasing and the color of the nucleus gets darker from intermediate to superficial stages, the accompanying phase map of the nucleus is seen to get taller by approximately a factor of 2. This change in the optical height profile cannot be inferred from the 2D bright-field images and the phase map is therefore seen to provide new morphological information. Further, from the phase profiles we also see that the evolution of the cells from intermediate to superficial stage happens via a continuous change. Next in **Figures 3–5** we examine the abnormal cell classes low grade squamous intraepithelial lesion (LSIL), high grade squamous intraepithelial lesion (HSIL) and squamous cell carcinoma (SCC) which progressively indicate higher grade abnormalities. The class LSIL consists of abnormal superficial and intermediate cells. Variable degrees of hyper-chromasia, nuclear size variation with coarsely granulated chromatin are identifiers of a typical LSIL cell. In the HSIL class, the degree of nuclear enlargement and hyper-chromasia is more than LSIL and the cells here are found in sheet-like aggregates. In both LSIL and HSIL cases the phase profiles of the cell nuclei are seen to have increased roughness or corrugations compared to the normal cells. In the SCC class which is considered to be a confirmed case of malignant cervical smear, the phase profile of the nucleus shows sharp narrow peaks with large phase values. It is once again important to note that the phase profile clearly provides new morphological information that is not readily available in the 2D bright-field images. Apart from the main classes above, the Bethesda system defines a class Atypical Squamous Cells

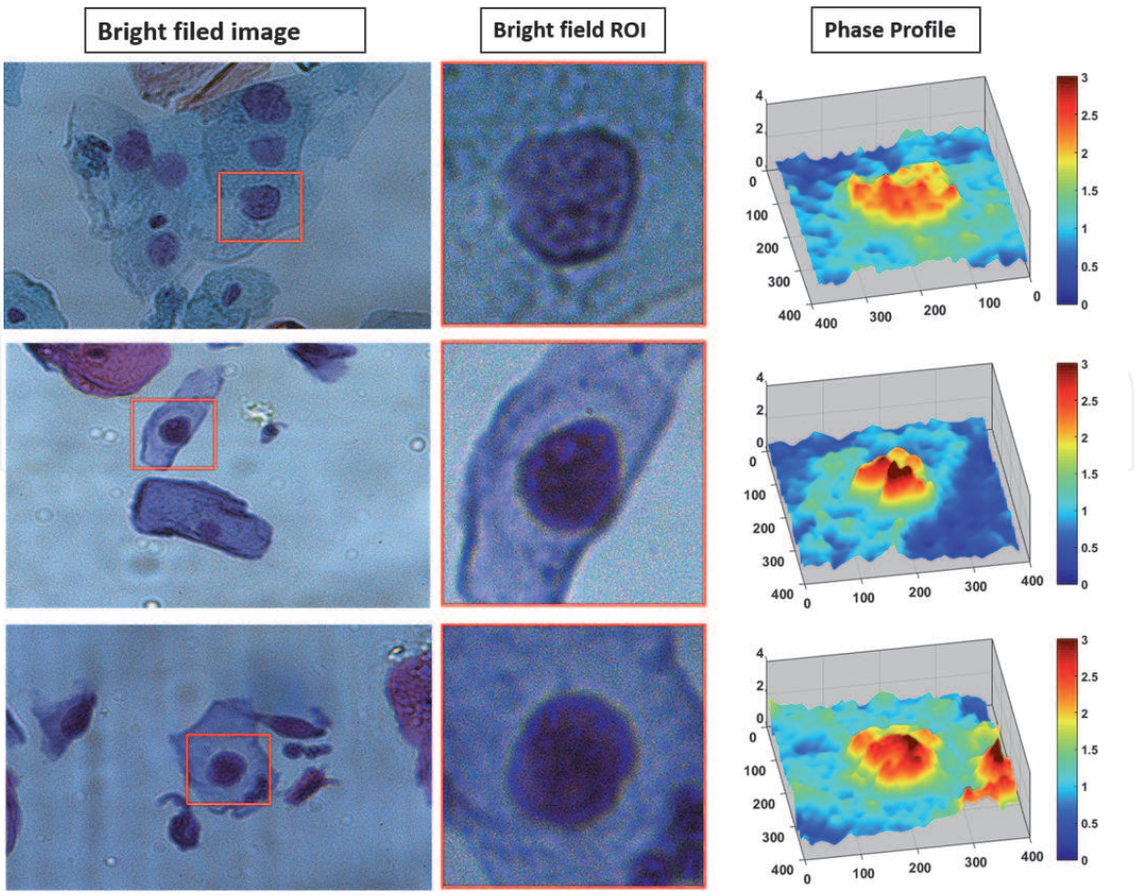

Figure 3.
Illustration of LSIL cells, the three columns show the bright-field image, selected ROI and the phase image of the ROI.

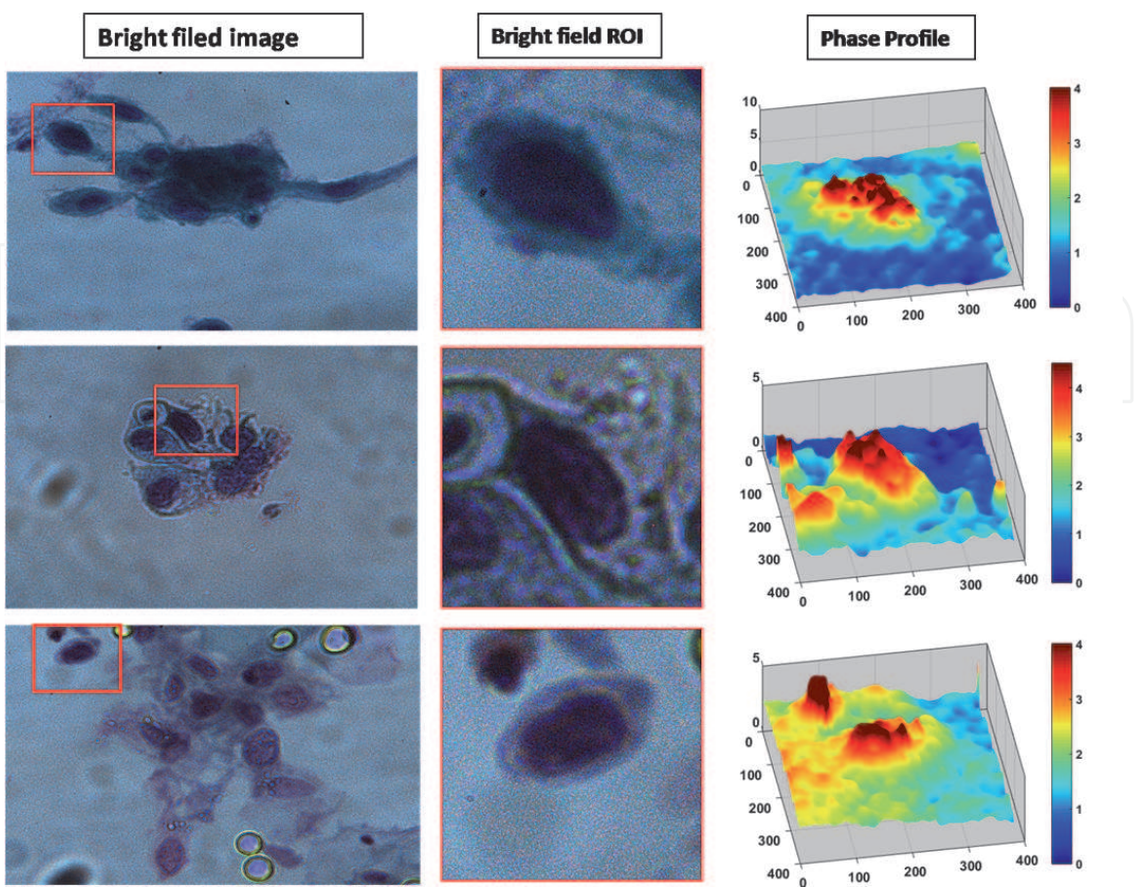

Figure 4.
Illustration of HSIL cells, the three columns show the bright-field image, selected ROI and the phase image of the ROI.

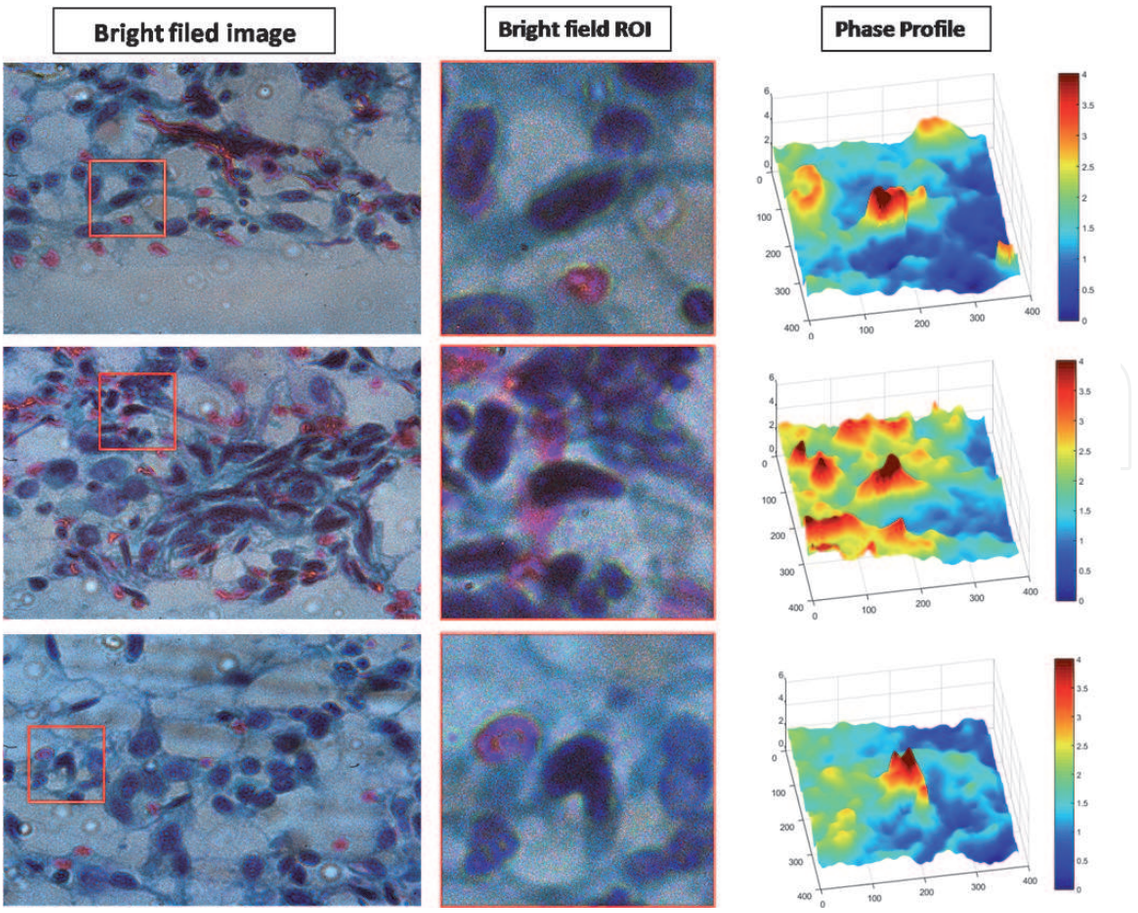

Figure 5.
Illustration of SCC cells, the three columns show the bright-field image, selected ROI and the phase image of the ROI.

ASC which include the samples that cannot be categorized as normal or abnormal typically. The nuclei in this class are typically larger in area, however, as seen in **Figure 6**, the average phase value in the nucleus is slightly lower compared to the LSIL, HSIL or SCC classes. In both ASC-US and ASC-H classes (**Figure 6**), the phase profile has undulations but the phase structure of ASC-H is flatter with lower average phase value in the nucleus as compared to that of the ASC-US cells. Further in some ASC nuclei, we observe local peaks in phase profile located near the nuclear boundary leading to a dip in the center. Finally we show rare abnormal cell cases in **Figure 7** that include inflamed, reactive, moon-type, virus-infected and koliocytotic classes. Just a few examples of these rare cells were present in our sample set. The phase profile for all these types appears corrugated with lower average phase values except for the virus infected cell type. While only a few representative images of each cell type have been shown simultaneously in bright-field and phase mode, we clearly observe that the phase profile offers distinctive morphological features that are not currently utilized in the clinical practice. This new information if incorporated in cyto-pathological examination, can be potentially valuable to clinicians.

4.2 PCA analysis of quantitative parameters obtained from bright-field and phase images of cervical cell nuclei

A MATLAB based software was designed to compute a number of morphological parameters associated with cell nuclei that are listed in **Table 2** from each of the cell nuclei imaged in bright-field as well as quantitative phase mode. The cell nucleus measurement data was therefore consisted of an (N × p) matrix with N = 48,006

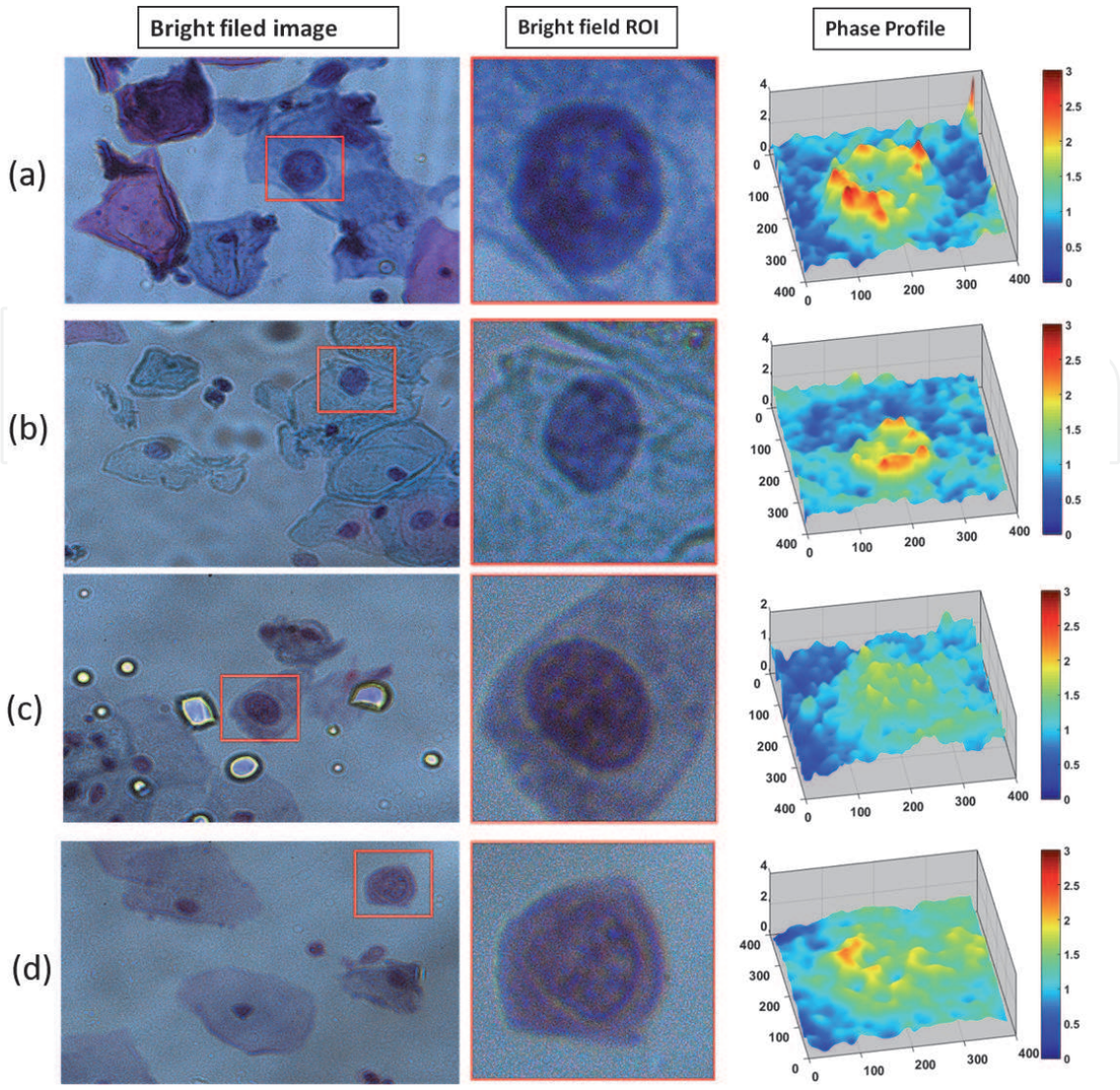

Figure 6.
Illustration of (a), (b): ASC-US cells and (c), (d): ASC-H cells, the three columns show the bright-field image, selected nucleus ROI and the phase image of the ROI.

and p = 20. Note that one parameter in the measurement set is the N/C ratio which was given three labels (low =1, mid =2, high =3) based on visual inspection of nuclei. This was done so that a number of cells of interest that appeared in clusters for which determining the cytoplasm boundary was difficult could be used in the analysis. All the other parameters were measured over the nucleus region in an automated fashion. For the present analysis, the cells were nominally labeled as superficial, intermediate and abnormal (including LSIL, HSIL, ASC-US, ASC-H, SCC) by practicing cyto-pathologists (S. R. M., M. S., K. A., S.S.). Since the number of abnormal cells was much smaller (1.6%) compared to the normal (superficial and intermediate) cells, a truncated data-set with 450 randomly selected cells from each of the three types (superficial, intermediate and abnormal) as per prior labelling was used to train the PCA. Denoting the truncated data matrix with 1350 rows and 20 columns (representing the measurements) with each column in standard form (zero mean and standard deviation 1) by A, the PCA solves the eigenvalue problem:

$$A^T A u_k = \mu_k u_k. \tag{8}$$

Here the superscript "T" stands for the transpose of the matrix A. The eigenvectors u_k are mutually orthogonal and are called as the principal vectors. All the

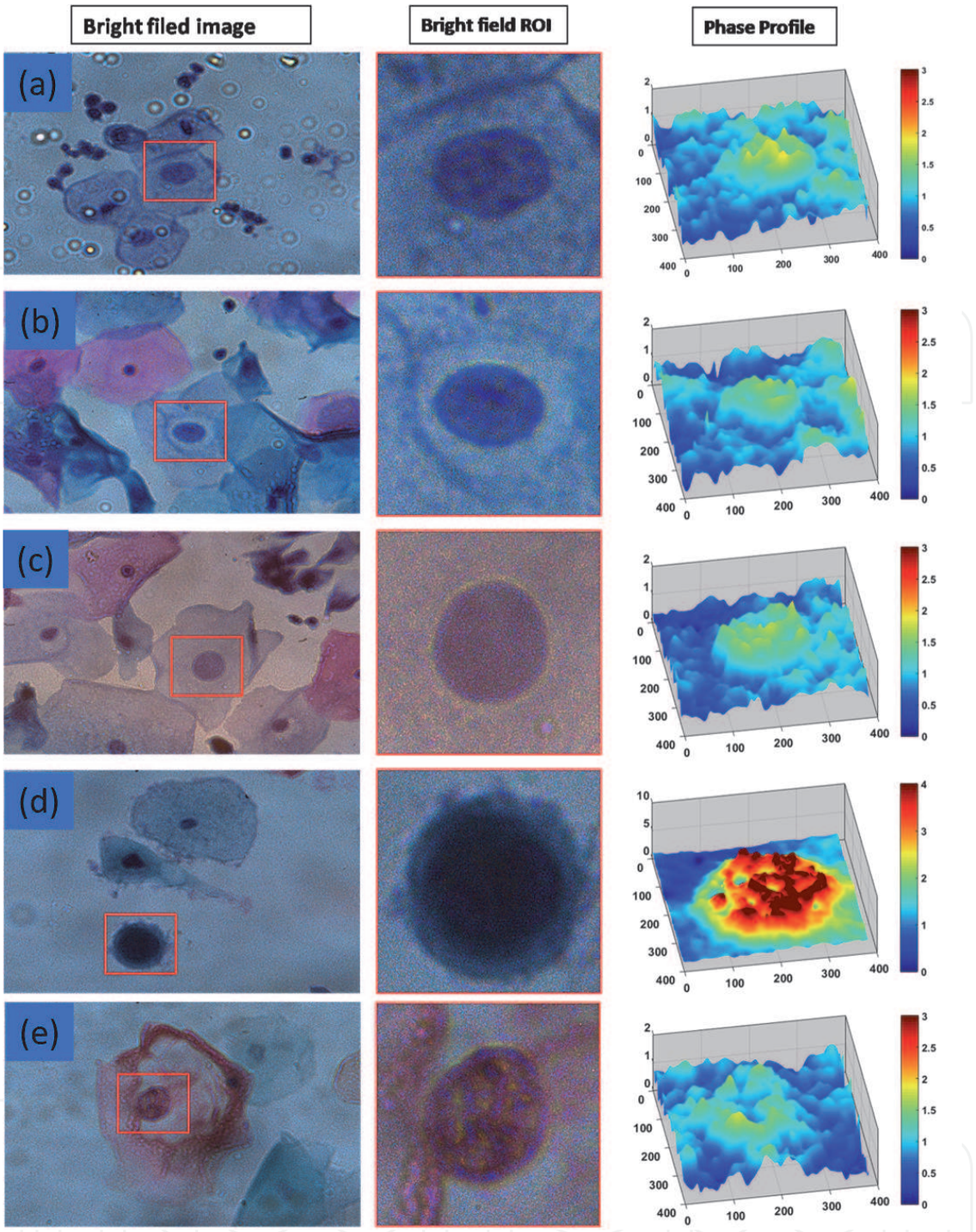

Figure 7.
Illustration of rare abnormal cell types: (a) inflamed, (b) reactive, (c) moon-type, (d) virus-infected and (e) koliocytotic; the three columns show the bright-field image, selected nucleus ROI and the phase image of the ROI.

cell data corresponding to the 48006 cells was then projected on the PCA vectors. The plot of first two components of PCA for all the cells is shown in **Figure 8**. The color coding of black, blue, and red corresponds to cells that were labeled separately by cyto-pathologists as superficial, intermediate and abnormal (all classes) respectively based on the bright-field images of the nuclei. Typical bright-field and phase images of cells from the three different regions of the PCA plot are also shown for illustration. The PCA plot based on bright-field and phase information separates most of the cells in three different classes, despite some overlap in adjacent classes. In particular, it is interesting to observe that almost all the cells labeled as abnormal fall in the bottom right corner of the PCA plot. We further examine

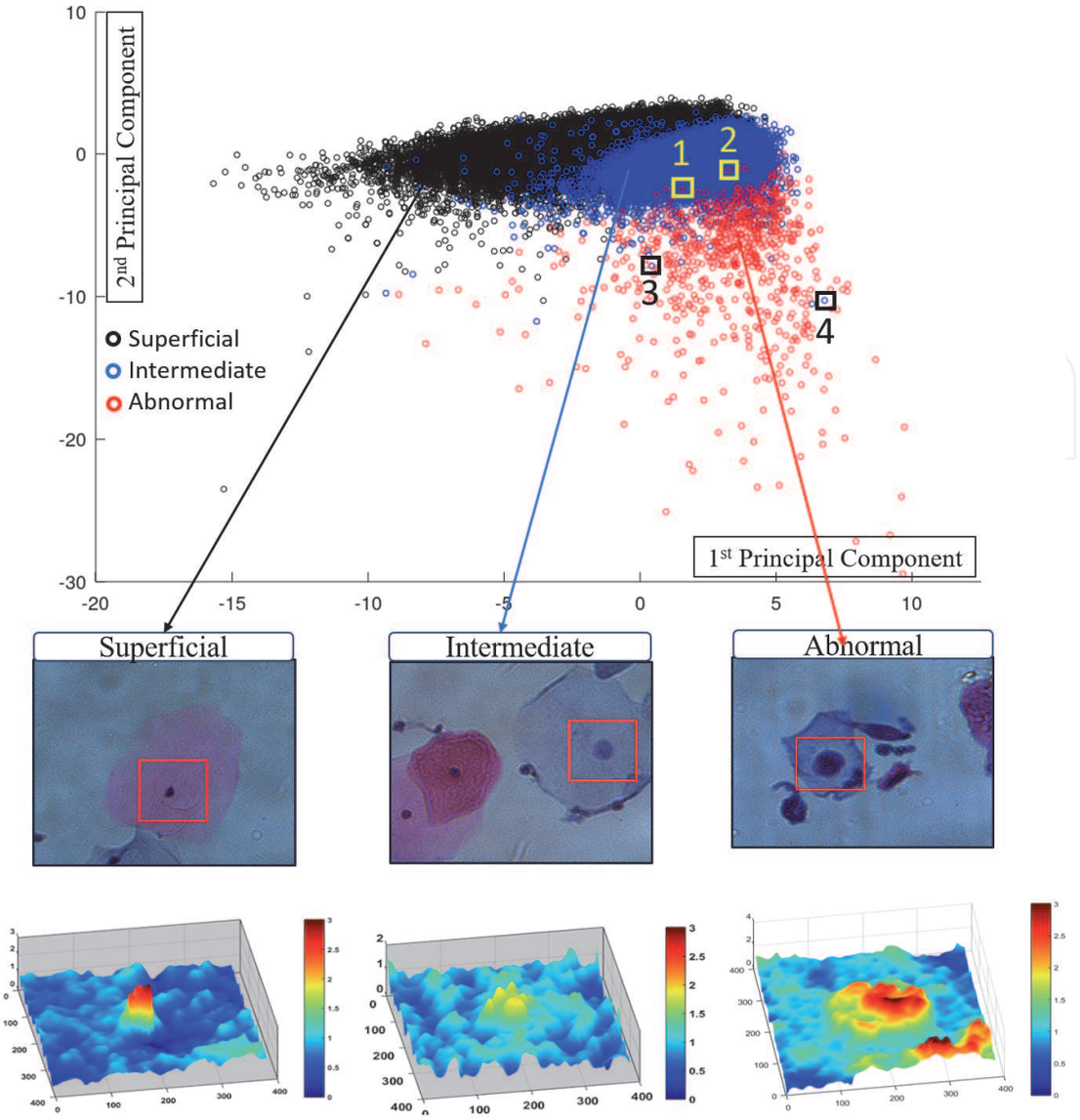

Figure 8.
Data corresponding to 48006 cells projected onto the first two PCA vectors. The color coding of black, blue, and red corresponds to cells that were labeled by cyto-pathologists as superficial, intermediate and abnormal (all classes) respectively.

four cells labeled as 1, 2, 3, 4 on the PCA plot that showed unexpected classification when phase parameters were used. Cells 1, 2 were labeled abnormal by the pathologist but were seen to be well within the intermediate region. Similarly the cells 3, 4 were labeled as intermediate but were observed to be well within the abnormal (red points) class. For a closer examination of this anomaly, we show bright-field and phase images of these cells in **Figure 9**. A re-examination of these cell images by pathologists suggested the following. Cell 1 is koliocytotic (abnormal) but it appears to have a dried up cytoplasm and leading to low phase values in the nucleus. The cell 2 is actually very similar to intermediate cells in general, but the pathologists labeled it as abnormal due to comparatively smaller sizes of other nuclei on the particular sample slide. The parameters associated with cell 3 are similar to abnormal cells but it is a rare example of enlarged intermediate cell. Finally cell 4 has folded cytoplasm leading to higher phase values although the cell may be considered to be of the intermediate class. Re-examination of these and other similar anomalies reveal that cervical cell classification has some aspects beyond simple numerical measurements performed on cell images (either in phase or bright-field modes) that need to be taken into account by any automated cell classification methodology. The issues like

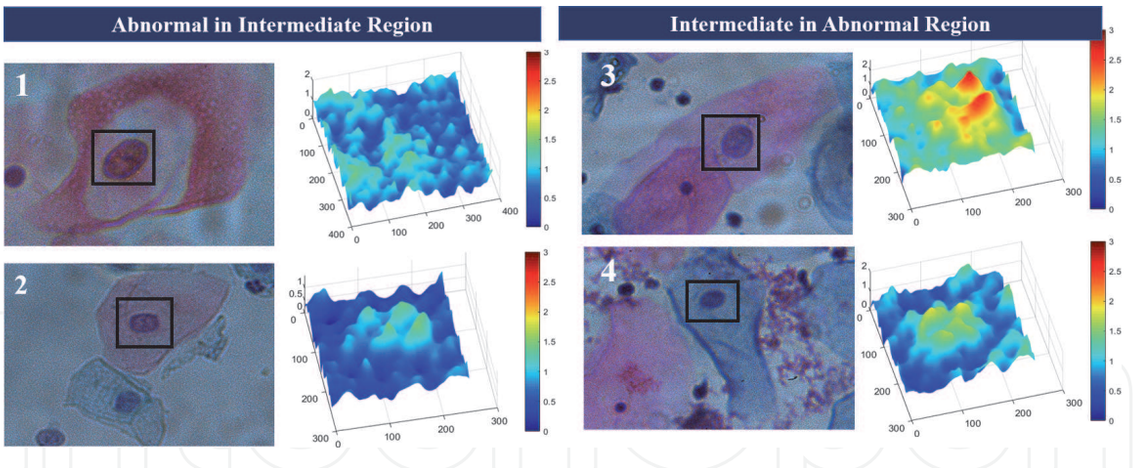

Figure 9.
Examples of intermediate and abnormal cells falling well within the abnormal and intermediate regions of the PCA plot in ***Figure 8****.*

folding of cell cytoplasm can for example be minimized with the LBC preparation methodology. PCA analysis was used here because the plot as in **Figure 8** can be generated essentially in an unsupervised manner, however, it is certainly not the best classification methodology available today. In future we hope to test the possibility of cell classification using more advanced machine learning ideas applied to this data.

We further performed a leave-one-out analysis of the PCA for the cell data to determine which of the 20 measurements influenced the PCA scores the most [33]. If the PCA eigenvectors are arranged as columns of a matrix U the scores Z for the data corresponding to the principal components may be expressed as:

$$Z = AU. \tag{9}$$

The plot in **Figure 8** thus corresponds to the first two columns of the score matrix Z. For the leave-one-out analysis, the PCA was performed at a time with only 19 parameters by leaving one of the measured parameters one by one. The data matrix, the eigenvector matrix and the score matrix corresponding to the case where j-th measurement ($j = 1, 2, 3, \ldots, 20$) is left out may be denoted by $A^{(-j)}$, $U^{(-j)}$ and $Z^{(-j)}$ respectively. The importance of the j-th measurement is judged by the Procrustese distance D_j between the first $M = 2$ columns of the score matrices Z and $Z^{(-j)}$. A specific parameter will be judged to influence the PCA the most if its corresponding Procustes distance D_j is higher. The top five morphological parameters in order of importance are shown in **Table 4**. The relative Procrustes distances

Morphological parameter	Modality	Relative Procrustes Distance D_j
Moment of inertia	Phase	1.0
Optical volume	Phase	0.85
N/C Ratio	Bright-field	0.83
Perimeter of nucleus	Bright-field	0.79
Mean phase of nucleus	Phase	0.76

Table 4.
Relative importance of numerical parameters using leave-one-out analysis applied to PCA.

are calculated by dividing all the distances D_j with $(j = 1, 2, 3, \ldots, 20)$ by the maximum among them. We find that among the top five parameters that influenced the PCA scores the most, three were derived from the phase images while two were derived from the bright-field images. It may be noted from **Table 4** that two phase based parameters (moment of inertia and optical volume) influence the PCA more than the commonly used N/C ratio criterion. We therefore believe that quantitative phase may prove to be an important future imaging modality in addition to the commonly used bright-field microscopy for cervical cell classification.

4.3 Consistency of phase parameters

Quantitative phase is not a standard clinical methodology for cell classification, however, as we showed in Section 4.2, quantitative phase may become an important modality to consider for future clinical use. It is therefore important to understand if the phase parameters for cervical nuclei are consistent across different subjects from same clinical site, age group of subjects or clinical sites with different sample preparation methodologies. Since our leave-one-out PCA analysis suggested that optical volume and moment of inertia are the most important phase parameters as explained in the previous section, we have plotted a few hundred randomly selected normal cells (superficial and intermediate) with respect to these phase parameters in **Figure 10**. In **Figure 10(a)** we show the plot for 200 normal cells each for five different patients from a single clinical site. **Figure 10(b)** shows the same plot for 200 cells each from three different clinical sites with different sample preparation protocols. In **Figure 10(c)** we show the plot once again for 500 normal cells for two different age groups (below and above 30 years). From these plots we observe that the normal cells from different categories as above show highly overlapping distributions for the most important phase parameters. We believe that this observation

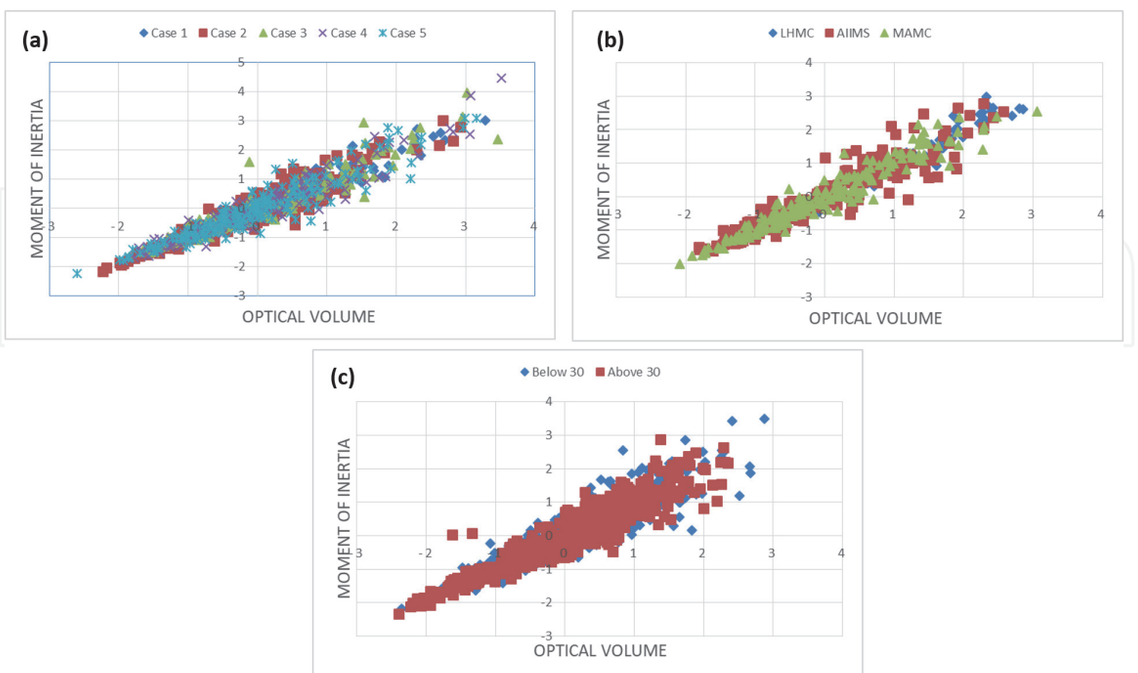

Figure 10.
Verification of consistency of the two most important phase parameters (moment of inertia and optical volume) decided based on the leave-one-out analysis; (a) plot of 200 cells each for 5 patients from the same clinical site (AIIMS), (b) plot of 200 cells each from three clinical sites with different sample preparation protocols, (c) plot of 500 cells each for patients below and above 30 years of age. The numerical values of moment of inertia and optical volume are normalized to standard form (zero mean and standard deviation one).

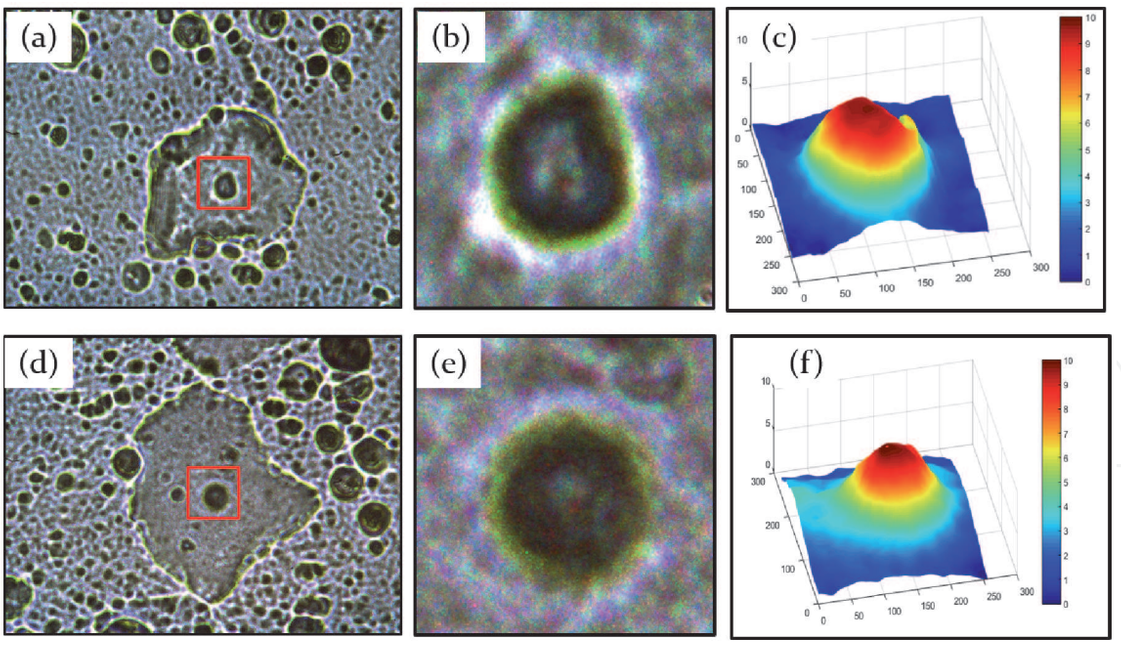

Figure 11.
Illustration of bright-field and phase imaging of unstained cervical cells: (a), (d): Bright-field images of unstained cells, (b), (e): Nucleus ROI selected from the bright-field image, (c), (f): Phase map of the unstained nuclei.

is very important for standardization and usage of quantitative phase imaging methodology in future clinical practice.

4.4 Observations on quantitative phase imaging of unstained cervical cells

In this section we briefly describe an interesting possibility of quantitative phase imaging of unstained cervical cell samples with two typical images of normal cells as shown in **Figure 11**. For unstained cervical cell samples, the Pap smear was prepared using the conventional method and cells were fixed with ethyl alcohol. While the staining protocols used in Pap-smear sample is a gold standard for diagnosis by cyto-pathologists, we note here that compared to stained cells, the phase signal observed from nuclei of unstained unprocessed cell samples is almost three times higher in magnitude. While interpretation of images of the unstained cells and their phase may require one to go through a learning process, the possibility of using unstained cell samples for diagnostic practice may offer an attractive alternative as the cell sample preparations will not involve any wet-lab processing and recurring costs associated with reagents.

5. Conclusions

In conclusion we have reported an image based study of cervical cells at various stages using bright-field as well as quantitative phase microscopy. Over 48000 cells have been imaged individually. The phase images of the cell nuclei were reconstructed using an optimization approach that provided same resolution as the bright-field images. This image data-set may be valuable for future application development using advance machine learning methods. The visual inspection of images shows interesting features in the phase images as the cells evolve from intermediate to superficial stages with distinct features associated with abnormal cells. This finding based on visual inspection is confirmed in the PCA analysis of the

morphological parameters of cells derived from both the bright-field and phase images of cell nuclei. A leave-one-out analysis applied to the PCA scores suggests that apart from the N/C ratio that has been used for identifying abnormal cells for decades, the other two parameters that influence the PCA the most are optical volume and moment of inertia of nucleus - both of which are derived from phase images. A consistency study suggests that the phase parameters associated with normal cells show highly overlapping distributions for multiple patients from same clinical site, for three clinical sites with different sample preparation protocols and for patients in two age groups. The consistency of phase parameter distributions for these cases further suggest that phase is a robust modality that can certainly be used in a standardized manner in clinical practice. We believe that quantitative phase may become an important imaging modality in addition to the bright-field imaging that is solely used in the current clinical practice. While this study has been performed for cervical cells we believe that our conclusions regarding importance of quantitative phase may possibly have a wider applicability.

Acknowledgements

The authors acknowledge financial support from Department of Science and Technology India (Award: IDP/MED/34/2016) and Biotechnology Industry Research Assistance Council India (Awards: FT/12/02(110)/15/0237 and BT/SIBRI45/733/17). The unstained sample shown in **Figure 11** of this chapter was shared by Dr. R. Sankaranarayanan (RTI International India). Incubation support from FITT-IIT Delhi is gratefully acknowledged.

Conflict of interest

The authors declare no conflict of interest.

Author details

Sarita Ahlawat[1], Purnima Sharma[1], Ankita Pandey[1], Durga Bisht[1], Aanisa Jan[1], Apoorv Pant[2], Ritika Malik[2], Sandeep R. Mathur[3], Kiran Agarwal[4], Smita Singh[4], Meeta Singh[5] and Kedar Khare[2*]

1 Phase Laboratories Pvt Ltd, Technology Business Incubator Unit, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India

2 Department of Physics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India

3 Department of Pathology, All India Institute of Medical Sciences, New Delhi, India

4 Department of Pathology, Lady Hardinge Medical College, New Delhi, India

5 Department of Pathology, Maulana Azad Medical College, New Delhi, India

*Address all correspondence to: kedark@physics.iitd.ac.in

References

[1] Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. The Lancet 2007; 370: 890–907.

[2] Andrae B, Andersson TML, Lambert PC, Kemetli L, Silfverdal L, Strander B, Ryd W, Dillner J, Tornberg S, Sparen P. Screening and cervical cancer cure: population based cohort study. BMJ 2012; 344:e900 doi: 10.1136/bmj.e900

[3] Benard VB, Watson M, Saraiya M, Harewood R, Townsend JS, Stroup AM, Weir HK, Allemani C. Cervical cancer survival in the United States by race and stage (2001–2009): Findings from the CONCORD-2 study. Cancer 2017; 123: 5119–5137.

[4] Cronje H. Screening for cervical cancer in developing countries. International Journal of Gynecology & Obstetrics 2004; 84: 101–108.

[5] Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, Hingmire S, Malvi SG, Thorat R, Kothari A, Chinoy R, Kelkar R, Kane S, Desai S, Keskar VR, Rajeshwarkar R, Panse N, Dinshaw KA. HPV screening for cervical cancer in rural India. New England Journal of Medicine 2009; 360: 1385–1394.

[6] Small Jr W, Bacon MA, Bajaj A, Chuang LT, Fisher BJ, Harkenrider MM, Jhingran A, Kitchener H, Mileshkin LR, Viswanathan AN, Geffney DK. Cervical cancer: a global health crisis. Cancer 2017; 123(13): 2404–2412.

[7] Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, Raab S, Sherman M, Wilbur D, Wright Jr. T, Young N. The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA 2002; 287: 2114–2119.

[8] Nayar R, Wilbur DC (ed.s). The Bethesda system for reporting cervical cytology: definitions, criteria, and explanatory notes. Springer International Publishing 2015.

[9] Marinakis Y, Dounias G, Jantzen J. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. Computers in Biology and Medicine 2009; 39: 69–78.

[10] GençTav A, Aksoy S, ÖNder S. Unsupervised segmentation and classification of cervical cell images. Pattern Recognition 2012; 45: 4151–4168.

[11] Chankong T, Theera-Umpon N, Auephanwiriyakul S. Automatic cervical cell segmentation and classification in Pap smears. Computer Methods and Programs in Biomedicine 2014; 113: 539–556.

[12] Song Y, Zhang L, Chen S, Ni D, Lei B, Wang T. Accurate Segmentation of Cervical Cytoplasm and Nuclei Based on Multiscale Convolutional Network and Graph Partitioning. IEEE Trans on Biomedical Engineering 2015; 62:2421–2433.

[13] Zhang L, Le Lu, Nogues I, Summers RM, Liu S, Yao J. DeepPap: Deep Convolutional Networks for Cervical Cell Classification. IEEE Journal of Biomedical and Health Informatics 2017; 21: 1633–1643.

[14] Jantzen J, Norup J, Dounias G, Bjerregaard B. Pap-smear Benchmark Data For Pattern Classification. In: Proc. NiSIS (Nature Inspired Smart Information Systems); 2005: 1–9.

[15] Plissiti ME, Nikou C. Cervical cell classification based exclusively on nucleus features. In: Lecture Notes on

Computer Science Springer. 2012; 7325: 483–490.

[16] Kim MK. Principles and techniques of digital holographic microscopy. SPIE Reviews 2010; 1: 018005.

[17] Park Y, Depeursinge C, Popescu G. Quantitative phase imaging in biomedicine. Nature Photonics 2018; 12: 578–589.

[18] Kemper B, Illy E. Digital Holographic Microscopy. PhotonicsViews 2020; 17(1): 32–35.

[19] Mihailescu M, Paun I, Scarlat E, Grigorescu I, Nedelcu O, Radu R. Digital holographic microscopy for phase images of cervical cells 3D structure. In: Conference on Lasers and Electro-optics/Pacific Rim, Optical Society of America 2015: paper 27P69.

[20] Benzerdjeb N, Garbar C, Camparo P, Sevestre H. Digital holographic microscopy as screening tool for cervical cancer preliminary study. Cancer Cytopathology 2016; 124: 573–580.

[21] Lam VK, Nguyen TC, Chung BM, Nehmetallah G, Raub CB. Quantitative assessment of cancer cell morphology and motility using telecentric digital holographic microscopy and machine learning. Cytometry Part A 2018; 93: 334–345.

[22] Lam VK, Nguyen T, Phan T, Chung BM, Nehmetallah G, Raub CB. Machine Learning with Optical Phase Signatures for Phenotypic Profiling of Cell Lines. Cytometry Part A 2019; 95: 757–768.

[23] Mangal J, Monga R, Mathur SR, Ahlawat S, Khare K, Unsupervised organization of cervical cells using bright-field and single-shot digital holographic microscopy. Journal of Biophotonics 2019; 12: e201800409.

[24] Khare K, Samsheerali PT, Joseph J. Single shot high resolution digital holography. Opt. Express 2013; 21: 2581–2591.

[25] Singh M, Khare K. Single-shot interferogram analysis for accurate reconstruction of step phase objects. J. Opt. Soc. Am. A 2017; 34: 349–355.

[26] Singh M, Khare K. Single-shot full resolution region-of-interest (ROI) reconstruction in image plane digital holographic microscopy. J. Mod. Opt. 2018; 65: 1127–1134.

[27] Rajora S, Butola M, Khare K. Mean gradient descent: an optimization approach for single-shot interferogram analysis. J. Opt. Soc. Am. A 2019; 36: D7–D13.

[28] Singh M and Khare K. Accurate efficient carrier estimation for single shot digital holographic imaging. Opt. Lett. 2016; 41:4871–4874.

[29] Singh M, Khare K, Jha AK, Prabhakar S, Singh RP. Accurate multipixel phase measurement with classical-light interferometry. Phys. Rev. A 2015; 91:021801.

[30] Pandey N, Ghosh A, Khare K. Two-dimensional phase unwrapping using the transport of intensity equation. Applied Optics 2016; 55: 2418–2425.

[31] Gray W, McKee G. Diagnostic Cytopathology 2 ed. Churchill Livingstone, Philadelphia. 2002.

[32] Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. Nature Reviews Cancer 2004; 4: 677–687.

[33] Krzanowski W. Cross-validation in principal component analysis. Biometrics 1987: 575–584.