Realistic optical cell modeling and diffraction imaging simulation for study of optical and morphological parameters of nucleus

Jun Zhang,[1] Yuanming Feng,[1] Wehnhuan Jiang,[2] Jun Q. Lu,[2] Yu Sa,[1] Junhua Ding;[3] and Xin-Hua Hu[2,*]

[1]*Department of Biomedical Engineering, Tianjin University, Tianjin 300072, China.*
[2]*Department of Physics, East Carolina University, Greenville, NC 27858, USA*
[3]*Department of Computer Science, East Carolina University, Greenville, NC 27858, USA.*
[*]*hux@ecu.edu*

Abstract: Coherent light scattering presents complex spatial patterns that depend on morphological and molecular features of biological cells. We present a numerical approach to establish realistic optical cell models for generating virtual cells and accurate simulation of diffraction images that are comparable to measured data of prostate cells. With a contourlet transform algorithm, it has been shown that the simulated images and extracted parameters can be used to distinguish virtual cells of different nuclear volumes and refractive indices against the orientation variation. These results demonstrate significance of the new approach for development of rapid cell assay methods through diffraction imaging.

OCIS codes: (170.1530) Cell analysis; (110.1650) Coherence imaging; (050.1940) Diffraction.

References and links

1. D. Zink, A. H. Fischer, and J. A. Nickerson, "Nuclear structure in cancer cells," Nat. Rev. Cancer **4**, 677-687 (2004).
2. L. T. Perelman, V. Backman, M. Wallace, G. Zonios, R. Manoharan, A. Nusrat, S. Shields, M. Seiler, C. Lima, T. Hamano, I. Itzkan, J. Van Dam, J. M. Crawford, and M. S. Feld, "Observation of periodic fine structure in reflectance from biological tissue: a new technique for measuring nuclear size distribution," Phys. Rev. Lett. **80**, 627-630 (1998).
3. J. R. Mourant, M. Canpolat, C. Brocker, O. Esponda-Ramos, T. M. Johnson, A. Matanock, K. Stetter, and J. P. Freyer, "Light scattering from cells: the contribution of the nucleus and the effects of proliferative status," J. Biomed. Opt. **5**, 131-137 (2000).
4. D. Arifler, M. Guillaud, A. Carraro, A. Malpica, M. Follen, and R. Richards-Kortum, "Light scattering from normal and dysplastic cervical cells at different epithelial depths: finite-difference time-domain modeling with a perfectly matched layer boundary condition," J. Biomed. Opt. **8**, 484-494 (2003).
5. I. Itzkan, L. Qiu, H. Fang, M. M. Zaman, E. Vitkin, I. C. Ghiran, S. Salahuddin, M. Modell, C. Andersson, L. M. Kimerer, P. B. Cipolloni, K.-H. Lim, S. D. Freedman, I. Bigio, B. P. Sachs, E. B. Hanlon, and L. T. Perelman, "Confocal light absorption and scattering spectroscopic microscopy monitors organelles in live cells with no exogenous labels," P. Natl. Acad. Sci. USA **104**, 17255-17260 (2007).
6. O. C. Marina, C. K. Sanders, and J. R. Mourant, "Correlating light scattering with internal cellular structures," Biomed. Opt. Express **3**, 296-312 (2012).
7. T. Kim, R. Zhou, M. Mir, S. D. Babacan, P. S. Carney, L. L. Goddard, and G. Popescu, "White-light diffraction tomography of unlabelled live cells," Nat. Photon. **8**, 256-263 (2014).
8. Y. Sung, W. Choi, C. Fang-Yen, K. Badizadegan, R. R. Dasari, and M. S. Feld, "Optical diffraction tomography for high resolution live cell imaging," Opt. Express **17**, 266-277 (2009).
9. Y. Zhang, Y. Feng, C. R. Justus, W. Jiang, Z. Li, J. Q. Lu, R. S. Brock, M. K. McPeek, D. A. Weidner, L. V. Yang, and X. H. Hu, "Comparative study of 3D morphology and functions on genetically engineered mouse melanoma cells," Integr. Biol. **4**, 1428-1436 (2012).
10. R. S. Brock, X. H. Hu, D. A. Weidner, J. R. Mourant, and J. Q. Lu, "Effect of detailed cell structure on light scattering distribution: FDTD study of a B-cell with 3D structure constructed from confocal images," J. Quant. Spectrosc. Radiat. Transfer **102**, 25-36 (2006).
11. M. A. Yurkin, and A. G. Hoekstra, "The discrete dipole approximation: an overview and recent developments," J. Quant. Spectrosc. Radiat. Transfer **106**, 558-589 (2007).
12. K. M. Jacobs, L. V. Yang, J. Ding, A. E. Ekpenyong, R. Castellone, J. Q. Lu, and X. H. Hu, "Diffraction imaging of spheres and melanoma cells with a microscope objective," J. Biophotonics **2**, 521–527 (2009).

13. R. Pan, Y. Feng, Y. Sa, J. Q. Lu, K. M. Jacobs, and X. H. Hu, "Analysis of diffraction imaging in non-conjugate configurations," Opt. Express **22**, 31568–31574 (2014).
14. K. M. Jacobs, J. Q. Lu, and X. H. Hu, "Development of a diffraction imaging flow cytometer," Opt. Lett. **34**, 2985-2987 (2009).
15. K. Dong, K. M. Jacobs, Y. Sa, Y. Feng, J. Q. Lu, and X. H. Hu, "Study of cell classification with a diffraction imaging flow cytometer method," Proc. SPIE **7902**, 790215 (2011).
16. Y. Sa, J. Zhang, M. S. Moran, J. Q. Lu, Y. Feng, and X. H. Hu, "A novel method of diffraction imaging flow cytometry for sizing microspheres," Opt. Express **20**, 22245–22251 (2012).
17. Y. Sa, Y. Feng, K. M. Jacobs, J. Yang, R. Pan, I. Gkigkitzis, J. Q. Lu, and X. H. Hu, "Study of low speed flow cytometry for diffraction imaging with different chamber and nozzle designs," Cytometry A **83**, 1027-1033 (2013).
18. Y. Feng, N. Zhang, K. M. Jacobs, W. Jiang, L. V. Yang, Z. Li, J. Zhang, J. Q. Lu, and X. H. Hu, "Polarization imaging and classification of Jurkat T and Ramos B cells using a flow cytometer," Cytometry A **85**, 817-826 (2014).
19. H. Wang, Y. Feng, Y. Sa, Y. Ma, R. Pan, J. Q. Lu, and X. H. Hu, "Acquisition of cross-polarized diffraction images and study of blurring effect by one time-delay-integration camera," Appl. Opt. **54**, 5223-5228 (2015).
20. W. Jiang, J. Q. Lu, L. V. Yang, Y. Sa, Y. Feng, J. Ding, and X. H. Hu, "Comparison study of distinguishing cancerous and normal prostate epithelial cells by confocal and polarization diffraction imaging," J. Biomed. Opt. **12**, 071102 (2016).
21. M. N. Do, and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," Image Processing, IEEE Transactions on **14**, 2091-2106 (2005).
22. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," Data Min. Knowl. Disc. **2**, 121-167 (1998).
23. C. C. Chang, and C. J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Technol. **2**, 1-27 (2011).
24. L. F. Yousif, K. M. Stewart, and S. O. Kelley, "Targeting mitochondria with organelle-specific compounds: strategies and applications," ChemBioChem **10**, 1939-1950 (2009).
25. M. Unser, "Splines: a perfect fit for signal and image processing," IEEE Signal Process. Mag. **16**, 22-38 (1999).
26. R. Barer, "Refractometry and Interferometry of Living Cells," J. Opt. Soc. Am. **47**, 545-556 (1957).
27. R. Drezek, M. Guillaud, T. Collier, I. Boiko, A. Malpica, C. Macaulay, M. Follen, and R. Richards-Kortum, "Light scattering from cervical cells throughout neoplastic progression: influence of nuclear morphology, DNA content, and chromatin texture," J. Biomed. Opt. **8**, 7-16 (2003).
28. M. Friebel, and M. Meinke, "Model function to calculate the refractive index of native hemoglobin in the wavelength range of 250-1100 nm dependent on concentration," Applied optics **45**, 2838-2842 (2006).
29. O. Zhernovaya, O. Sydoruk, V. Tuchin, and A. Douplik, "The refractive index of human hemoglobin in the visible range," Phys. Med. Biol. **56**, 4013 (2011).
30. M. A. Yurkin, and A. G. Hoekstra, "The discrete-dipole-approximation code ADDA: capabilities and known limitations," J. Quant. Spectrosc. Radiat. Transfer **112**, 2234-2247 (2011).
31. E. M. Purcell, and C. R. Pennypacker, "Scattering and absorption of light by nonspherical dielectric grains," Astrophys. J. **186**, 705-714 (1973).
32. B. T. Draine, and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," J. Opt. Soc. Am. A **11**, 1491-1499 (1994).
33. K. Dong, Y. Feng, K. M. Jacobs, J. Q. Lu, R. S. Brock, L. V. Yang, F. E. Bertrand, M. A. Farwell, and X. H. Hu, "Label-free classification of cultured cells through diffraction imaging," Biomed. Opt. Express **2**, 1717-1726 (2011).
34. C. F. Bohren, and D. R. Huffman, "*Absorption and Scattering of Light by Small Particles*," (Wiley, New York, 1983), p. 65.
35. Y. Dong, and J. Ma, "Feature extraction through contourlet subband clustering for texture classification," Neurocomputing **116**, 157-164 (2013).
36. E. R. Dougherty, and R. A. Lotufo, *Hands-on Morphological Image Processing* (SPIE Optical Engineering Press, 2003).
37. S. Tanev, W. Sun, J. Pond, V. V. Tuchin, and V. P. Zharov, "Flow cytometry with gold nanoparticles and their clusters as scattering contrast agents: FDTD simulation of light–cell interaction," J. Biophotonics **2**, 505-520 (2009).
38. X. Su, Y. Qiu, L. Marquez-Curtis, M. Gupta, C. E. Capjack, W. Rozmus, A. Janowska-Wieczorek, and Y. Y. Tsui, "Label-free and noninvasive optical detection of the distribution of nanometer-size mitochondria in single cells," J. Biomed. Opt. **16**, 067003 (2011).
39. J. Q. Lu, P. Yang, and X. H. Hu, "Simulations of light scattering from a biconcave red blood cell using the FDTD method," J. Biomed. Opt. **10**, 024022 (2005).
40. H. Ding, J. Q. Lu, R. S. Brock, T. J. McConnell, J. F. Ojeda, K. M. Jacobs, and X. H. Hu, "Angle-resolved Mueller matrix study of light scattering by B-cells at three wavelengths of 442, 633 and 850nm," J. Biomed. Opt. **12**, 034032 (2007).
41. X. Li, A. Taflove, and V. Backman, "Recent progress in exact and reduced-order modeling of light-scattering properties of complex structures," IEEE J. Sel. Top. Quant. Electron. **11**, 759-765 (2005).
42. J. Zhang, Y. Feng, M. S. Moran, J. Q. Lu, L. V. Yang, Y. Sa, N. Zhang, L. Dong, and X. H. Hu, "Analysis of cellular objects through diffraction images acquired by flow cytometry," Opt. Express **21**, 24819–24828 (2013).
43. A. Wax, M. G. Giacomelli, T. E. Matthews, M. T. Rinehart, F. E. Robles, and Y. Zhu, "Optical Spectroscopy of Biological Cells," Adv. Opt. Photonics **4**, 322-378 (2012).

44. R. S. Brock, "*Modeling of Light Scattering by Biological Cells Using a Finite-Difference Time Domain Method*," (PhD Thesis, Department of Physics, East Carolina University, 2007).
45. M. Moran, "*Correlating the Morphological and Light Scattering Properties of Biological Cells*," (PhD Thesis, Department of Physics, East Carolina University, 2007).

1. Introduction

Nucleus is one of the largest organelles inside eukaryotic cells, provides the site for DNA and RNA synthesis, plays critical roles in cell development. Hence it serves as one of major targets for cell assay by morphology and is especially important for detection of abnormal conditions and cancer diagnosis [1]. Optical detection through coherent light scattering offers a much valued platform for its label-free nature and capacities to extract both morphology and molecular information. Characterization of nucleus by scattered light signals thus attracts active research efforts [2-7]. Determination of cellular and nuclear morphology is fundamentally a challenging inverse problem for their complex 3D structures. For example, structural reconstruction requires large amount of measured data per cell and often expensive computation that is too long for rapid assay [8, 9]. If one aims at only to distinguish cell types such as cancer from normal or apoptotic from viable cells, however, the goal may be achieved empirically with moderate amount of measured data per cell and powerful algorithms of pattern recognition. In either case, it is very useful to develop realistic optical cell models (OCMs) and accurate simulation tools for forward calculations of measured signals of scattered light. They can be employed, for example, to generate training data for algorithm development in search of the correlations between morphological features of cells and diffraction patterns of coherent light scatter.

In this report, we present a numerical approach based on previous studies for establishing realistic OCMs for generating virtual cells and accurate simulation of polarized diffraction image (p-DI) data [9-13]. The new approach takes the advantage of 3D cell morphology and molecular information acquired from the fluorescent confocal images to produce simulated p-DI data that are comparable to the measured ones acquired with a polarization diffraction imaging flow cytometry (p-DIFC) system [14-20]. To demonstrate the utility of the realistic OCMs, we have investigated the effects of nuclear morphology and refractive index (RI) on diffraction patterns against the orientation changes of OCMs derived from prostate cells. The simulated p-DI data were analyzed with a contourlet transform (CT) algorithm [21]. Classification of virtual cells using different OCMs by CT parameters has been performed with a support vector machine (SVM) algorithm [22, 23]. The results show that the CT parameters can serve as effective features for identifying nuclear effect in terms of volume and RI changes. The OCMs coupled with simulation tools yield a useful means for development and understanding of new single cell assay method.

2. Methods

2.1 Reconstruction of cell morphology and fluorescence distribution

An OCM yields the 3D distribution of RI or $n(\mathbf{r}, \lambda)$ with $\mathbf{r}$ as the voxel position of intracellular organelles and λ as wavelength of incident light. Development of an OCM requires both of morphological information of organelles important for concerned aspects of light-cell interaction and associated molecular information of polarizability. We obtained the morphology information by 3D reconstruction with a fluorescent image stack acquired by a laser scanning confocal microscope (LSM510, Zeiss). The cells were first double stained by fluorescent dyes of Syto 61 and MitoTracker Orange CMTMRos (S11343 and M-7510, Life Technologies) to visualize respectively the nucleus and mitochondria with details given elsewhere [9, 12]. In viable cells, Syto 61 binds to nucleic acids concentrated mostly inside the nucleus. The MitoTracker Orange accumulates preferably in mitochondria in response to the electric potential difference maintained across the inner mitochondrial membrane that is much larger than those across other membranes [24]. The two intracellular organelles of choice are not only critical for cell development and metabolism but also important in light

scattering due to their large contributions to the heterogeneity in $n(\mathbf{r}, \lambda)$. Two channels of the 12-bit image stack files were used to store the fluorescent intensity as $F_r(\mathbf{r})$ for Syto 61 in red channel and $F_g(\mathbf{r})$ for MitoTracker Orange in the green channel.

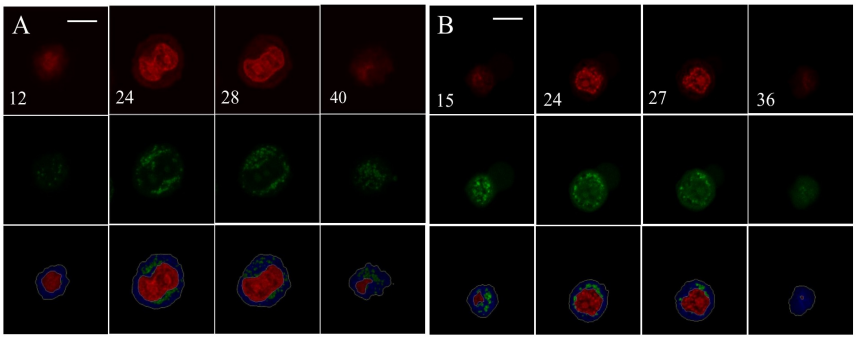

Fig. 1. Examples of the confocal image slices in red (F_r) and green (F_g) channels in top and middle rows of (A) one PC3 cancer cell; (B) one normal PCS cell. The bottom rows present segmented slices with nuclear region in red pixels of intensity F_r, mitochondria in green pixel of intensity F_g and cytoplasm in blue. Each column is labeled by the slice sequence number in the image stack and bar = 10 μm.

For each imaged cell, about 50 to 70 slices were acquired per stack by translating the cell through the focal plane of objective with a step size of 0.5μm. Examples of image slices are shown in Fig. 1 for one human prostate cancer PC-3 (CRL-1435, ATCC) cell and one of normal human prostate epithelial cells termed as PCS (PCS440010, ATCC). The details of cell maintenance, staining and confocal imaging have been described elsewhere [20]. The confocal image slices of a stack were imported into an in-house developed reconstruction software. Improved from the previous version [9], the software automates image processing to obtain segmented cell structure and heterogeneous distributions of the fluorescence intensity $F_r(\mathbf{r})$ and $F_g(\mathbf{r})$ for modeling $n(\mathbf{r}, \lambda)$. The process starts by automatic selection of the first and last image slices in the acquired stack containing F_r or F_g significantly above a background noise level, then aligns the selected image slices to correct the effect of cell motion during data acquisition. Pixels of each input image slice were separated into six exclusive region types: extracellular space, cytoplasm, mitochondria and nucleus with organelle's volume denoted as Ω_c, Ω_m and Ω_n, respectively. The nuclear region of Ω_n is further divided into three sub-regions of Ω_{nl}, Ω_{nm} and Ω_{nh} consisting of voxels of low, medium and high values of $F_r(\mathbf{r})$. Different segmentation algorithms have been combined, including histogram analysis, spatial gradient and water-mark techniques, to identify the boundaries of cytoplasm and mitochondria in the green channel and cytoplasm and nucleus in the red channel. Afterwards, multiple slices were added between input image slices by a linear B-spline interpolation scheme for determination of region type and values of F_r or F_g in separate color channels for pixels in the interpolated slices [25]. The interpolation produces a 3D structure with voxels of nearly equal sides of 0.07 μm length in all three directions using a 4x digital zoom. Finally, a 16-bit identifier is assigned to each voxel at $\mathbf{r}$ for labeling its region type and associated fluorescence intensity in output data files for establishing an OCM. In addition, 3D morphology parameters of the imaged cell can be determined from the output [9, 18, 20].

2.2 Establishment of OCM and simulation of light scattering

Due to scarcity of reliable $n(\mathbf{r}, \lambda)$ data available in literature, we investigated various methods for assigning RI values to different intracellular organelles for establishing OCMs for this study by comparing the simulated p-DI data to measured ones. Two final choices are described here. The first method determines voxels' RI according to their region type based on the understanding that molecular polarizabilities differ mainly among organelles, which is given by the sum of a constant term and a randomly fluctuating term as follows

$$n_\alpha(\mathbf{r},\lambda) = n_{\alpha 0} + (n_{\alpha 0} - n_w)a_\alpha \cdot RND \quad \forall \mathbf{r} \in \Omega_\alpha, \tag{1}$$

where α (= c, m, nl, nm or nh) is an organelle or region type identifier, $n_{\alpha 0}$ is the mean value of n_α, n_w is the RI of water, a_α is the fluctuation amplitude and RND is random numbers uniformly distributed in [-1, 1]. This method utilizes organelle information with the fluctuation term to model heterogeneity in $n(\mathbf{r}, \lambda)$ but not the fluorescent information provided by $F_r(\mathbf{r})$ or $F_g(\mathbf{r})$. To incorporate fluorescent or molecular information, a second RI method was developed to derive OCMs from n_w with molecular contribution given by

$$n_\alpha(\mathbf{r},\lambda) = n_w + b_r F_r(\mathbf{r}) + b_g F_g(\mathbf{r}) \quad \forall \mathbf{r} \in \Omega_\alpha, \tag{2}$$

where α =c, m or n and b_r or b_g is respectively the specific RI increment by $F_r(\mathbf{r})$ or $F_g(\mathbf{r})$ of the dye targeting the mitochondrial biomolecules or nuclear acids. Note that $F_g(\mathbf{r})=0$ in Ω_n while $F_r(\mathbf{r})=0$ in Ω_m. The second method as expressed by Eq. (2) is based on a long held and reasonably validated view assuming a linear relation for a type of biomolecules between its density and specific contribution to RI beyond n_w [26-29]. Once RI values assigned to all voxels, an OCM was obtained and its morphology and/or RI can be modified to generate a series of OCMs derived from the same cell imaged by a confocal microscope.

To accurately simulate the distribution of coherent light scattered by single cells in a host medium of water, we have employed an open-source, parallel computing ADDA code of discreet-dipole-approximation (DDA) developed in C language by Yurkin et al. [11, 30]. The DDA model divides the scatterer into voxels of discrete dipoles and calculates the scattered wavefields from the dipoles in terms of the angularly resolved Mueller matrix $\{S_{ij}\}$ of 4x4 elements. The dipole voxels are excited by a given incident wavefields and their polarizabilities are determined by $n(\mathbf{r}, \lambda)$ [11, 31, 32]. Using the OCMs described above, we have executed the ADDA code on our parallel computing cluster to obtain $\{S_{ij}\}$ as functions of scattering angles of (θ_s, ϕ_s) from the incident light direction with λ = 532nm and $n_h = n_w$=1.334 μm for the RI of host medium.

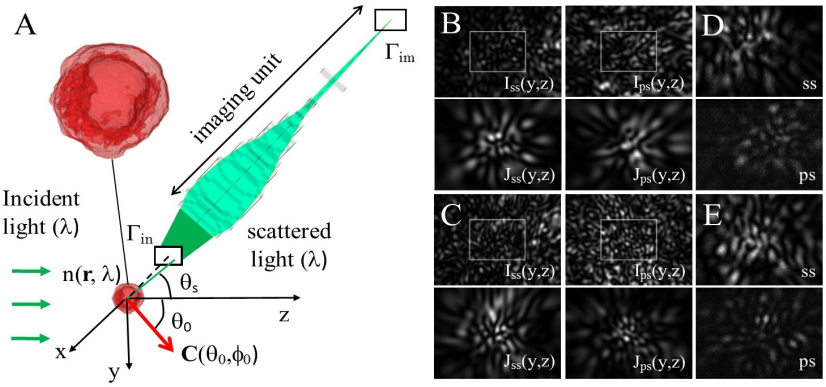

Fig. 2. (A) Configuration of simulation with a magnified view of an OCM of PC3 cell at orientation **C**, input plane Γ_{in} inside a water-filled flow chamber (not shown), the objective based imaging unit in air with shaded areas of two green tones indicating scattered light outside and inside the unit and imaging plane Γ_{im}; (B) simulated p-DI pairs using OCM of one PCS cell on Γ_{in} (top) and on Γ_{im} (bottom) with rectangles indicating the field of view by the simulated sensor on Γ_{im}; (C) same as (B) using OCM of the PC3 cell shown in (A); (D) measured p-DI pair of one PCS cell with the labels of image polarization followed by that of incident beam; (E) same as (D) for one PC3 cell.

The simulation configuration is plotted in Fig. 2(A) in which the orientation of an OCM is labelled as $\mathbf{C}(\theta_0, \phi_0)$ that is defined as the line connecting mass-centers of the cell and its nucleus. A linear combination of $S_{ij}(\theta_s, \phi_s)$ was first projected on an "input" plane Γ_{in} at x = – 0.15 mm inside the water-filled flow chamber to obtain a p-DI denoted as $I_{kl}(y, z)$ to be

measured by a microscope objective based imaging unit with (y, z) as the discrete pixel coordinates [12, 14, 18-20, 33]. The image $I_{kl}(y, z)$ can be expressed as a linear combination of Mueller matrix elements to represent the spatial distribution of the coherent light of a polarization k scattered by the cell excited by an incident beam of polarization l [34]. For example, it is straightforward to derive a p- or s- polarized p-DI on Γ_{in} for an incident beam polarized with l = s as

$$I_{ks} = e(S_{11} - S_{12} \mp S_{21} \pm S_{22}), \quad (3)$$

where e is a proportional constant, − is for k=s and + for k=p in the S_{21} term. Other expressions of p-DI on Γ_{in} can be found in [18, 19].

2.3 Ray-tracing based diffraction image transform from Γ_{in} to Γ_{im}

Experimentally, the p-DIFC system employs an imaging unit aligned along the x-axis in Fig. 2(A) based on an infinity-corrected objective to acquire the coherent light scattered by a cell through Γ_{in} and record on its focal or image plane Γ_{im} with one or two camera sensors [18-20]. It has been previously shown that a unique relation exists between the p-DI pixels on Γ_{in} and Γ_{im} with the imaging unit translated off-focus towards the flow chamber ($\Delta x > 0$) [12, 13]. The off-focus positioning of the imaging unit in p-DIFC measurement allows optimization of magnification and image contrast with the same imaging unit. To make the simulated p-DI comparable to those measured by a p-DIFC system, we have validated a ray-tracing approach for modeling the diffraction imaging process using a commercial optical design software [13]. The approach imports $I_{kl}(y, z)$ into the ray-tracing software (Zemax-EE v2009, Zemax Development Corp.), which traces rays from each pixels of $I_{kl}(y, z)$ on Γ_{in} inside the flow chamber through the chamber wall of glass, air and imaging unit to each corresponding pixels on Γ_{im} as $J_{kl}(y, z)$ within the sensor area. Fig. 2(B) and Fig. 2(C) presents respectively the simulated p-DI pairs Γ_{in} and Γ_{im} derived from one PCS and one PC3 cell. The field of view on Γ_{im} and imaging unit parameters were set to the same values of 4.7mmx3.6mm for the camera sensor (Lm075, Lumenera) and Δx = 150μm for the off-focus distance and objective parameters used in our p-DIFC measurements [13, 20]. Compared to the two examples of measured p-DI pairs in Fig. 2(D) and Fig. 2(E), the simulated p-DI pairs of $J_{kl}(y, z)$ yield types of diffraction patterns similar to the measured data in terms of speckle size, number and distribution.

2.4 Contourlet transform (CT) analysis of simulated p-DI data

The CT algorithm has been developed as a 2D extension of the wavelet transform to extract smooth pattern contours from an input image at multiple scales of γ (=0, 1, …, γ_{max}) and directions of δ (=1, 2, …, δ_{max}) [21, 35]. CT first applies the Laplacian pyramid (LP) decomposing on the input image, designated by $\gamma = 0$ and $\delta = 0$ or 0-0 here, to generate a 2-fold down-sampled lowpass image of 1-0 by weighted pixel smoothing and a bandpass image as the difference between the input and up-sampled lowpass images. A directional filter banks (DFB) algorithm is then operated on the bandpass image to obtain δ_{max} directionally filtered images of 1-δ while LP is operated on the lowpass image for further decomposition to γ = 2, 3, …, γ_{max} followed by DFB filtering on subsequent bandpass images. In this study we used the CT algorithm with $\gamma_{max} = 5$ and $\delta_{max} = 2^2$ or 2^3 on each simulated p-DI of $J_{kl}(y, z)$ to obtain a total of 41 CT processed images of $J_{kl,\gamma-\delta}(y, z)$ with 9 for each pixel scale of γ=1 to 4 and 5 for γ=5. Each CT image $J_{kl,\gamma-\delta}(y, z)$ was first normalized and then characterized by 4 parameters of energy E, contrast C, variance V and fluctuation F defined as

$$E_{kl,\gamma-\delta} = \sum_y \sum_z J_{kl,\gamma-\delta}(y,z)^2, \quad (4)$$

$$C_{kl,\gamma-\delta} = \sum_y \sum_z \frac{SLC_{kl,\gamma-\delta}(y,z)}{4N_yN_z - 2(N_y + N_z)}, \quad (5)$$

$$V_{kl,\gamma-\delta} = \frac{\sum_y \sum_z [J_{kl,\gamma-\delta}(y,z) - MEA]^2}{N_y N_z - 1}, \tag{6}$$

$$F_{kl,\gamma-\delta} = \frac{\sqrt{V_{kl,\gamma-\delta}}}{MEA}, \tag{7}$$

with the squared local contrast SLC and mean value MEA given by

$$SLC_{kl,\gamma-\delta}(y,z) = \sum_{i=-1,1} \{[J_{kl,\gamma-\delta}(y+i,z) - J_{kl,\gamma-\delta}(y,z)]^2 + [J_{kl,\gamma-\delta}(y,z+i) - J_{kl,\gamma-\delta}(y,z)]^2\}, \tag{8}$$

$$MEA_{kl,\gamma-\delta} = \frac{\sum_y \sum_z J_{kl,\gamma-\delta}(y,z)}{N_y N_z}, \tag{9}$$

where N_y and N_z is the number of pixels in y and z directions respectively.

3. Results

3.1 Development of OCMs and populations for p-DI simulation

CT processing with $J_{kl}(y, z)$ as the input yields 164 parameters to characterize each simulated image at different pixel scales and directions. To investigate cell classification by these parameters, we set to examine if the morphology and RI changes made with paired OCMs can be recognized against variation in OCMs' orientations. OCMs were derived from the 3D structure of a PCS or PC3 cell and RI assignment by Eq. (1) or Eq. (2) using different parameters and designated as OCM(ID_{cell}, ID_{RI}). Tables 1 and 2 define the OCMs by selected morphology and RI parameters for ID_{cell} and ID_{RI}.

Table 1. Morphology parameters of OCMs[(1)]

ID_{cell} [(2)]	V_c	Vr_{nc}	Vr_{mc}	SVr_c	SVr_n	SVr_m
PCS_0	1392	27.5%	7.33%	0.626	1.07	6.85
PCS_{24}		5.75%			2.02	
$PC3_0$	2717	41.6%	4.78%	0.499	0.692	5.75
$PC3_{24}$		17.1%			1.02	

(1) V_c: cell volume in (μm^3); Vr_{nc}, Vr_{mc}: volume ratios of nucleus-to-cell and mitochondria-to-cell; SVr_c, SVr_n, SVr_m: surface-to-volume ratios of cell, nucleus and mitochondria in (μm^{-1}).
(2) The subscript indicates the number of eroded nuclear pixels on each image slice.

Table 2. RI parameters of OCMs (mean±std)

ID_{RI}	Eq.	n_c	n_{nl}	n_{nm}	n_{nh}	n_m
1a	(1)	1.367± 0.0019	1.398±0.0037	1.429±0.0055	1.460±0.0027	1.487± 0.0088
1b			1.438±0.0060	1.509 0.0101	1.580±0.0142	
		n_c		n_n		n_c
2a	(2)	1.409±0.0495		1.429± 0.0362		1.487±0.0786
2b		1.437±0.0574		1.509 ± 0.0667		

The nuclear volume were changed to the given structure of either PCS or PC3 cell through erosion of 24 nuclear pixels next to the nucleus-cytoplasm border in each image slice of a stack, which converts that part of nucleus into cytoplasm with unchanged mitochondria [36]. To make comparable RI distributions determined by two Eqs. (1) and (2), we set the parameters of b_r and b_g in Eq. (2) to obtain same mean values of RI for regions of Ω_{nm}, Ω_n and Ω_m. Figure 3 shows additional examples of simulated images of $J_{ss}(y, z)$ on Γ_{im} of different

OCMs derived from the same PCS and PC3 cells in Fig. 2 at different orientations of **C**. The simulated p-DI exhibit pattern detail changes as results of nuclear volume change, choice of RI equations and parameters and orientation of the OCM relative to the incident light direction or z-axis. The changes, however, are subtle and very difficult to identify visually to distinguish, say, the nuclear changes from the orientation changes. One has to resort to quantitative image texture analysis and machine learning algorithms to investigate further.

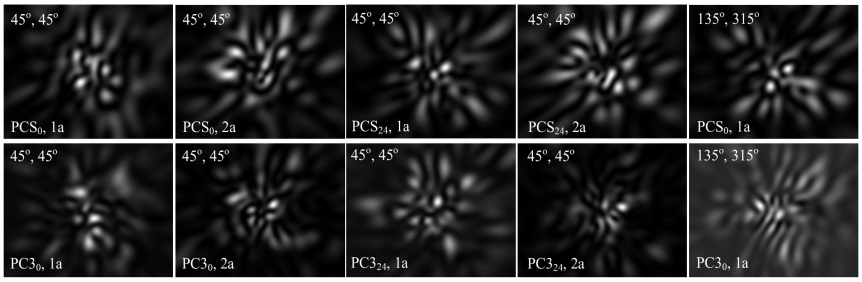

Fig. 3. Normalized p-DI images of $J_{ss}(y, z)$ of 640x480 pixels derived from different OCMs and orientations of $\mathbf{C}(\theta_0, \phi_0)$. The upper and lower labels are values of (θ_0, ϕ_0) and (ID_{cell}, ID_{RI}) for OCM.

To differentiate the nuclear effect from the orientation changes, we built groups of two "cell populations" with each population produced by one OCM of different orientations marked by $\mathbf{C}(\theta_0, \phi_0)$ as virtual cells. Two types of **C** variation were applied for cell population construction. Each small angle population consists of 25 virtual cells of the same OCM that include 1 with **C** along the flow direction of y-axis and 24 with **C** uniformly distributed at 3^o from the y-axis. Each large angle population contains 26 virtual cells with **C** uniformly distributed over the 4π solid angle range. Each group for classification study was composed of two populations with paired OCMs of either nuclear volume or RI changed for a specific polarization direction, p, s or 45^o, of the incident beam.

An SVM algorithm was employed with 4 different kernels of linear, polynomial, sigmoid and radial basis functions to investigate the classification of simulated p-DI data representing cell populations [20, 23]. As a powerful machine learning algorithm, SVM defines a feature space by transforming the CT parameters of the training image dataset with a chosen kernel function and search for an optimized model to distinguish the two populations in a group by multiple CT parameters of the training data. Because of the limited number of virtual cells in each population, all cells were used as the training dataset.

3.2 Effect of nuclear volume and RI changes on classification of virtual cell populations

We first employed OCM(PCS_0, 2a) and OCM(PCS_{24}, 2a) to build two small angle populations to examine the effect of nuclear volume on cell classification. Each virtual cell in a population produced one p-DI pair of $J_{pl}(y, z)$ and $J_{sl}(y, z)$ obtained through ADDA simulations and subsequent ray tracing transform for an incident light of l polarization. The CT was performed on the p-DI data to output CT images with selected examples shown in Fig. 4. For images obtained by DFB filtering on the fine pixel scales of $\gamma = 1$ and $\gamma = 2$ (not shown), they exhibit pixel distributions of small variance in intensity which can be best characterized by the energy or E parameters. For coarse pixel scales of $\gamma = 3$ to $\gamma = 5$, pixel patterns start to appear among different DFB filtered images of $\delta \neq 0$ which can be characterized by the C, F and V parameters.

To quantify the patterns of the input image of p-DI at different scales and orientations, we plot the four CT image parameters in six sets as indicated along the horizontal axis in Fig. 5. The lowpass image parameters are arranged as the first set from 1-0 to 5-0 followed by the other five sets of bandpass image parameters of different δ values (1 to 8 or 1 to 4 for $\gamma = 5$). Among the CT image parameters, the C (contrast) and V (variance) increase significantly

from the scale of γ = 1 to γ = 5 in either the lowpass set or among the bandpass sets, which is consistent with the pattern changes exhibited by the CT images presented in Fig. 4. The variations of the four CT image parameters at different pixel scales and orientations provide the basis for using them to classify virtual cells represented by the simulated p-DI data.

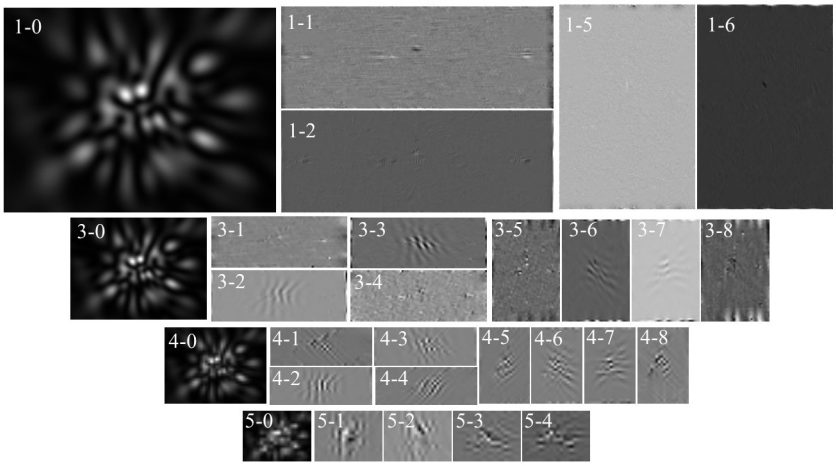

Fig. 4. Normalized CT images $J_{ss,\gamma-\delta}(y, z)$ generated from input image $J_{ss}(y, z)$ of 640x480 pixels shown in Fig. 2(B) with each image labelled by values of γ–δ. The lowpass image of 1-0 is of 320x240 pixels, 3-0 of 80x60 pixels, 4-0 of 40x30 and all 5-δ are of 20x15 pixels. The images of γ=3, 4 and 5 are scaled up 2-, 3- and 4-fold relative to the γ=1 images for clear viewing.

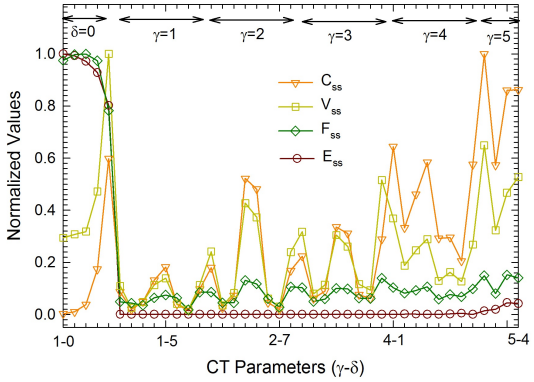

Fig. 5. The four CT image parameters of $J_{ss,\gamma-\delta}(y, z)$ shown in Fig. 4 plotted in groups with the first representing those extracted from lowpass images (δ=0) in the order of γ=1 to γ=5 followed by groups from bandpass images on the same scale with γ from 1 to 5 in the order of δ=1 to δ=8 or 4. The lines are for visual guide.

With two small angle populations derived from OCM(PCS_0, 2a) and OCM(PCS_{24}, 2a), we applied the SVM algorithm to examine the performance of CT parameters on cell classification by nuclear volume change against the variation of **C**. Because of large number of CT parameters, we performed SVM classification for each one of C, V, F and E parameters separately using a previously developed software [20] that calls the LIBSVM modules [23] to rank single parameters' performance. The simulated p-DI data of virtual cells in each population were treated as the training dataset and divided into 5 parts with 4 used for training and 1 for test. A classification accuracy A was defined to quantify performance for

each rotation of the part for test, which is given by the number ratio of p-DIs correctly identified for its OCM, as true-positive or true-negative, to the total number of p-DIs. The averaged value of A, A_{av}, was calculated after 5 rotations of the test part to rank single CT parameters with the top one having maximum A_{av} value followed by others with decreasing A_{av}. Afterwards, multiple parameters were used for classification training by combining the top N parameters, which together with the kernel function defines an SVM model for classification. An optimized SVM model is obtained when A_{av} reaches the maximum value which is typically larger than the maximum A_{av} of single parameters and the corresponding value of N is denoted as N_m.

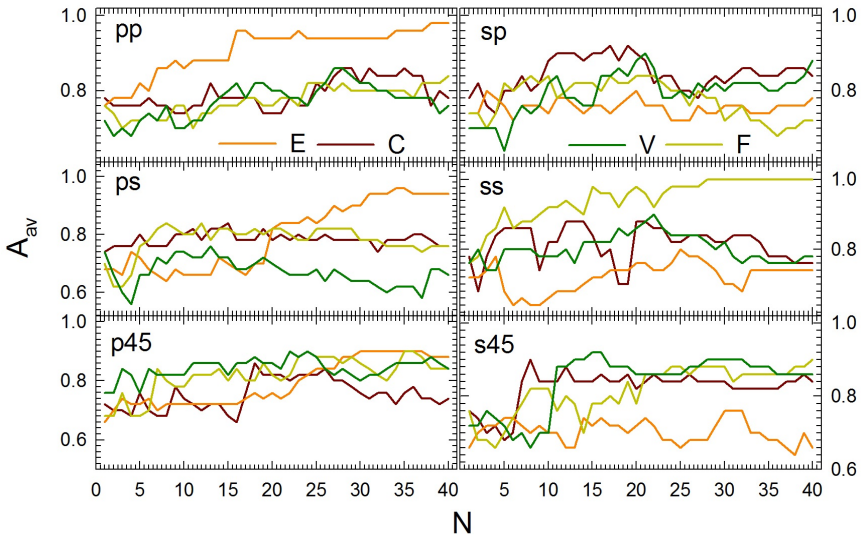

Fig. 6. The averaged classification accuracy A_{av} of SVM models versus the number of CT parameters N used to for the models with linear kernel function using four CT parameters extracted from $J_{kl}(y, z)$ of two populations derived from OCM(PCS_0, 2a) and OCM(PCS_{24}, 2a). The upper-left labels are polarization symbols of scattered light k (p or s) followed by that of incident beam l (p, s or 45°).

As representative data, Fig. 6 plots A_{av} of SVM models with $1 \le N \le 41$ and the linear kernel function that provided the larger maximum values of A_{av} than those obtained with other three kernel functions on classifying the two small angle populations of virtual cells. We found that the performance of single CT parameters is not very good with maximum A_{av} less than 78% for the top ranked ones. By using multiple parameters, the performance can improve significantly with $N = N_m$ parameters. In the case of classification of the virtual cells by the input images of $J_{ss}(y, z)$, A_{av} can reach 100% for $N_m \ge 28$ while in other cases maximum values of A_{av} ranges from 92% to 98%.

We have built 7 groups of two cell populations with each group derived from paired OCMs of either nuclear volume change or RI change and performed SVM classification on each group with one of the 4 kernel functions and different combination of the scattered light and incident light polarizations. In all cases, the optimized SVM models were obtained with the linear kernel function that yield highest values of A_{av}. The classification results with the linear kernel function are compiled in Table 3 to provide a brief summary of the classification results. Table 3 shows clearly that the paired OCMs with difference in either nuclear volume or RI of the nucleus can be identified from the simulated p-DI data of $J_{kl}(y, z)$ among virtual cells of clustered around y-axis or random orientations with accuracy at or above 92%. We note that the ability to distinguish two OCMs can be achieved with either one simulated p-DI or paired p-DI and in the latter case the number of CT parameter used in an optimized SVM model can be significantly reduced despite that fact that A_{av} may suffer. Considering the fact

that the two OCMs in each group have either identical morphology or very similar random distributions of RI, the simulation results in Table 3 suggest strongly that the p-DI data have the capacity for high-performance cell classification. They further corroborate our previous experimental results with the p-DIFC method for distinguishing cell types of highly similar morphology or lineage, which were achieved with little orientation control of cells' carried by a core fluid moving at a low speed of about 5 mm/s [18, 20].

Table 3. Values of A_{av} for SVM classification of two populations in 7 groups[(1)]

ID_{cell}	ID_{RI}	P/N_m [(2)]	angle	p45	s45	pp	sp	ps	ss
PCS_0 vs PCS_{24}	1a	V/41	small	94%	96%	96%	92%	90%	98%
PCS_0 vs PCS_{24}	2a	F/28	small	92%	92%	98%	92%	96%	100%
PCS_0 vs PCS_{24}	2a	F/20	large	90%	92%	90%	92%	90%	94%
$PC3_0$ vs $PC3_{24}$	2a	V/13	small	88%	90%	88%	92%	86%	88%
PCS_0	1a vs 1b	E/14	small	98%	96%	94%	96%	94%	94%
PCS_0	2a vs 2b	C/8	small	94%	90%	92%	98%	94%	92%
PCS_0 vs PCS_{24}	2a	F/7	small	94%		96%		94%	

(1) The A_{av} values in percentage on 7 groups were obtained with the linear kernel function. The values in top 6 rows were obtained with one simulated p-DI of $J_{kl}(y, z)$ and those on the bottom row were obtained with the paired p-DI of $J_{pl}(y, z)$ and $J_{sl}(y, z)$ as input data.

(2) The CT image parameter P and the values of N_m correspond to the maximum A_{av} values in red fonts in that group on the same row.

4. Discussion

Rapid assay of single biological cells through light scattering is a challenging problem due to the complex cell structure. Because of the difficulty in selecting and manipulating cells during measurement, accurate simulation of the light scattering process can play a critical role that is not only valuable but also irreplaceable in methodology development. This is especially true in the case of diffraction imaging to record spatial distribution of coherent light scattered by single cells with a flow cytometer. Previous studies by different groups, including ours, have developed simplified OCMs of constant or varying RI within a cell and/or its individual organelles built by spheres or spheroids [4, 37, 38], mathematical surfaces [39], reconstruction from confocal image stack data for nucleus only [10, 33, 40] or Gaussian random sphere models for nucleus and mitochondria [41, 42]. These models may be sufficiently accurate for investigation of certain aspects of scattered light distributions such as angularly (1D) or spectroscopically resolved measurements [43]. For diffraction images that exhibit complex texture patterns with very low degrees of symmetry, however, our investigations with simplified OCMs have shown that they are not capable of producing diffraction patterns similar to the measured data [44, 45]. The results in Fig. 2 present clear evidences that the similarity between simulated and measured p-DI data can be achieved by incorporating realistic morphological and molecular information into OCMs followed by accurate modeling of the imaging unit from Γ_{in} to Γ_{im}. The approach presented here allows development of realistic OCMs by using fluorescent confocal image data with the ability to account for the RI heterogeneity from measured fluorescence data with only two adjustable parameters in the case of Eq. (2) instead of artificial RI assignment for each regions of intracellular organelles. Furthermore, the accurate simulation of coherent light scattering through the imaging unit provided the opportunity to evaluate various OCMs. The simulated p-DI data, partly shown in Fig. 2 and 3, with different choices of parameters in Eqs. (1) and (2) proved clearly that variation of RI values among the voxel is essential to obtain the types of diffraction patterns comparable to the measured p-DI data.

In this study, we have employed the CT algorithm for processing of the simulated p-DI data and extraction of CT image parameters to represent virtual cells in populations derived from paired OCMs for classification. The summary data of Table 3 confirm clearly the observed fact that the orientation of OCMs or virtual cells affect little on distinguishing the nuclear effects by volume or RI changes. It should be noted that that the virtual cells of

different nuclear volumes were obtained from the same OCM with RI of the eroded nuclear voxels replaced by those of cytoplasm voxels instead of vacuole. Consequently, it is quite remarkable for p-DI data being able to detect the small changes of RI values. A closer look at the CT image parameters used to form optimized SVM models provides an interesting view of the p-DI features supporting such ability. For example, in the case of SVM classification of virtual cells derived from OCM(PCS_0, 2a) and OCM(PCS_{24}, 2a), A_{av} can reach 100% using only one p-DI image of $J_{ss}(y, z)$ with 28 parameters as shown in Fig. 6. Among the top 10 of the 28 parameters, 7 were extracted from the CT images of $\gamma = 4$ and 5 while 3 from images of $\gamma = 1$ and 2. Similar distribution of the top ranked CT image parameters can be seen in other cases of high values of A_{av}. These results confirm the observation of the CT images shown in Fig. 4 that the coarse pixel scales of $\gamma \geq 3$ for CT images carry more information on the patterns of the input p-DI data than the fine scales of $\gamma = 1$ or 2. Despite the useful insights provided by the CT algorithm, we would like to point out that the method needs to be improved since the numbers of CT parameters for optimized SVM models are quite large, leading to high cost of computing for image processing and training. With the realistic OCMs and accurate simulation tools described here, one can build large amount simulated p-DI data that are relevant to experimental studies. By combining with powerful data mining tools such as deep learning algorithms, the new approach provides opportunities to correlate biological cells' morphological and molecular features with the features of diffraction patterns for future development of rapid and label-free cell assay methods.

5. Summary

We have developed OCMs established from confocal image stacks of stained cells which can be used to generate virtual cells for accurate simulation of p-DI data. Classification of virtual cell populations of paired OCMs have been perform to demonstrate the utility of new approach of p-DI simulation. The results show that nuclear effect in terms of the morphology or RI changes can be recognized against variation of an OCM's orientation.

Acknowledgments

The authors wish to thank Dr. L.V. Yang for providing helps on cell maintenance and imaging, Dr. K.M. Jacobs on diffraction imaging measurements and Mr. E. King on maintaining the Biomedical Laser Laboratory's parallel computing cluster. Y. Feng acknowledges support by the NSFC (grants 81171342 and 81201148) and X.H. Hu acknowledges grant support from Golfers Against Cancer (2012-13-GAC).