Analysis of cellular objects through diffraction images acquired by flow cytometry

Jun Zhang,[1] Yuanming Feng,[2,1,3,5] Marina S. Moran,[3] Jun Q. Lu,[3] Li V. Yang,[4] Yu Sa,[1] Ning Zhang,[1] Lixue Dong,[4] and Xin-Hua Hu[3,1,*]

[1]*Department of Biomedical Engineering, Tianjin University, Tianjin 300072, China*
[2]*Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China*
[3]*Department of Physics, East Carolina University, Greenville, North Carolina 27858, U.S.A.*
[4]*Department of Oncology, Brody School of Medicine, East Carolina University, Greenville, North Carolina 27834, USA*
[5]*fengyu@ecu.edu*
[*]*hux@ecu.edu*

Abstract: It was found that the diffraction images acquired along the side scattering directions with objects in a cell sample contain pattern variations at both the global and local scales. We show here that the global pattern variation is associated with the categorical size and morphological heterogeneity of the imaged objects. An automated image processing method has been developed to separate the acquired diffraction images into three types of global patterns. Combined with previously developed method for quantifying local texture pattern variations, the new method allows fully automated analysis of diffraction images for rapid and label-free classification of cells according to their 3D morphology.

OCIS codes: (290.5850) Scattering, particles; (170.1530) Cell analysis; (110.1650) Coherence imaging.

References and links

1. A. K. Dunn, C. L. Smithpeter, A. J. Welch, and R. R. Richards-Kortum, "Finite-difference time-domain simulation of light scattering from single cells," J. Biomed. Opt. **2**(3), 262–266 (1997).
2. 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**(2), 131–137 (2000).
3. J. Q. Lu, P. Yang, and X. H. Hu, "SSimulations of light scattering from a biconcave red blood cell using the finite-difference time-domain method," J. Biomed. Opt. **10**, 024022 (2005).
4. 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. Transf. **102**(1), 25–36 (2006).
5. 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 850 nm," J. Biomed. Opt. **12**(3), 034032 (2007).
6. M. R. Melamed, T. Lindmo, and M. L. Mendelsohn, *Flow Cytometry and Sorting* (Wiley-Liss, 1990).
7. A. N. Shvalov, I. V. Surovtsev, A. V. Chernyshev, J. T. Soini, and V. P. Maltsev, "Particle classification from light scattering with the scanning flow cytometer," Cytometry **37**(3), 215–220 (1999).
8. D. I. Strokotov, A. E. Moskalensky, V. M. Nekrasov, and V. P. Maltsev, "Polarized light-scattering profile-advanced characterization of nonspherical particles with scanning flow cytometry," Cytometry A **79**(7), 570–579 (2011).
9. S. Holler, Y. Pan, R. K. Chang, J. R. Bottiger, S. C. Hill, and D. B. Hillis, "Two-dimensional angular optical scattering for the characterization of airborne microparticles," Opt. Lett. **23**(18), 1489–1491 (1998).
10. J. Neukammer, C. Gohlke, A. Höpe, T. Wessel, and H. Rinneberg, "Angular distribution of light scattered by single biological cells and oriented particle agglomerates," Appl. Opt. **42**(31), 6388–6397 (2003).
11. K. Singh, C. Liu, C. Capjack, W. Rozmus, and C. J. Backhouse, "Analysis of cellular structure by light scattering measurements in a new cytometer design based on a liquid-core waveguide," IEE proceedings **151**, 10–16 (2004).
12. X. T. Su, K. Singh, C. Capjack, J. Petrácek, C. Backhouse, and W. Rozmus, "Measurements of light scattering in an integrated microfluidic waveguide cytometer," J. Biomed. Opt. **13**(2), 024024 (2008).
13. X. Su, S. E. Kirkwood, M. Gupta, L. Marquez-Curtis, Y. Qiu, A. Janowska-Wieczorek, W. Rozmus, and Y. Y. Tsui, "Microscope-based label-free microfluidic cytometry," Opt. Express **19**(1), 387–398 (2011).

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

14. 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**(6), 067003 (2011).
15. K. M. Jacobs, J. Q. Lu, and X. H. Hu, "Development of a diffraction imaging flow cytometer," Opt. Lett. **34**(19), 2985–2987 (2009).
16. 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**(8-9), 521–527 (2009).
17. 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**(6), 1717–1726 (2011).
18. S. Yu, 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**(20), 22245–22251 (2012).
19. 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 (to be published).
20. Y. Feng, N. Zhang, K. M. Jacobs, W. Jiang, L. V. Yang, Z. Li, J. Zhang, J. Q. Lu, and X. H. Hu, "Classification of Jurkat T and Ramos B cells," Opt. Lett. submitted.
21. R. D. Castellone, N. R. Leffler, L. Dong, and L. V. Yang, "Inhibition of tumor cell migration and metastasis by the proton-sensing GPR4 receptor," Cancer Lett. **312**(2), 197–208 (2011).
22. M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, 1999).
23. C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn. **20**(3), 273–297 (1995).
24. C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," (2001), *http://www.csie.ntu.edu.tw/~cjlin/libsvm*.
25. A. G. Hoekstra, M. D. Grimminck, and P. M. A. Sloot, "Large Scale Simulations of Elastic Light Scattering by a Fast Discrete Dipole Approximation," Int. J. Mod. Phys. C **9**(01), 87–102 (1998).
26. M. A. Yurkin and A. G. Hoekstra, "The discrete dipole approximation: an overview and recent developments," J. Quant. Spectrosc. Radiat. Transf. **106**(1-3), 558–589 (2007).
27. M. A. Yurkin, A. G. Hoekstra, R. S. Brock, and J. Q. Lu, "Systematic comparison of the discrete dipole approximation and the finite difference time domain method for large dielectric scatterers," Opt. Express **15**(26), 17902–17911 (2007).
28. 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 (Camb) **4**(11), 1428–1436 (2012).
29. M. Moran, *Correlating the Morphological and Light Scattering Properties of Biological Cells*, Ph.D. Dissertation (East Carolina University, 2013).
30. N. Kanopoulos, N. Vasanthavada, and R. L. Baker, "Design of an image edge detection filter using the Sobel operator," IEEE J. Solid-State Circuits **23**(2), 358–367 (1988).
31. T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 881–892 (2002).

1. Introduction

It has become widely known that the elastically scattered light from single cells exhibits characteristic distributions in space if they are illuminated by a coherent beam. The features of the distribution have been shown theoretically and experimentally to correlate strongly with the morphology of the illuminated cell [1–5]. The forward and side scatters acquired with a conventional flow cytometer [6], however, are spatially integrated and provide very limited information on a cell's morphology even with a coherent laser beam. Consequently one needs to pursue spatially resolved strategies of light detection in search for morphology based assays of cells without fluorescent labeling. Over the last two decades, measurement of coherent light scatters under a flow condition have attracted numerous research efforts using angularly resolved [7, 8] or imaging [9–14] approaches. The studies of coherent scatters based on the imaging approach demonstrate clearly that the images present distinctive speckle patterns and the pattern parameters can be associated with the sizes and distribution of intracellular organelles such as mitochondria [13, 14]. But the images acquired in these studies were of low contrast and no clear roadmaps were provided on how to effectively and rapidly extract multiple image parameters from the acquired data for analysis and classification of cells.

Along the coherently imaging approach, we have developed a new flow cytometry method by combining system design improvement to acquire images of high contrast with automated image processing [15–19]. In this method, a jet-in-fluid design of flow chamber and a scheme of objective based off-focus imaging were established to significantly increase the image contrast by reducing light scattered off the interfaces outside the imaged cell but

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

within the field-of-view (FOV) of the cameras. In addition, we have developed image processing software based on the existing gray level co-occurrence matrix (GLCM) or Gabor transform algorithms to rapidly extract image parameters for automated classification of cells or characterization of particles [17, 18, 20]. We term this new method of imaging and analyzing cells on the flow cytometry platform as diffraction imaging flow cytometry (DIFC).

For DIFC measurement, a cell sample in culture medium is directly injected into a core fluid reservoir and pressurized by a syringe pump into the core fluid nozzle. When analyzing the acquired data, we found that the diffraction images could present a wide range of global and local patterns due to the coherent superposition of the light wavefields scattered from an imaged cellular object and other objects in or near the FOV. These patterns are highly sensitive to the morphology of the imaged objects. By microscopy one can easily observe that a cell sample contains not only viable cells of complex internal structures but also objects of different sizes and morphology such as various particles secreted by cells, damaged cells with fragmented structures or cellular debris, among others, and their aggregates. For cancer cell line samples of high viability and measured shortly after taken out of a cell incubator [21], most of the acquired diffraction images show "normal" global patterns with many speckles of moderate sizes and random shapes. Some acquired images, usually 20% or less, in these healthy samples present "abnormal" global patterns with small numbers of speckles of large sizes and random shapes or stripe-like patterns of increased symmetry. In our earlier studies, the diffraction images of abnormal global patterns were removed manually before the application of GLCM algorithm based on the recognition that these images are likely caused by very small sized objects or objects of highly symmetric morphology [22] instead of viable cells. We have shown the effectiveness of the GLCM algorithm in analyzing the local textures of the diffraction images of normal speckle patterns for cell classification [17, 20].

But for samples of primary cells or cultured cells measured about 8 hours or later after taking out of incubator, the abnormal images can count up to 70% of the acquired data. It is thus necessary to investigate the morphological variations underlying the abnormal diffraction images and develop a method to identify them. We found, however, that the GLCM analysis is not a reliable tool for this task. This can be attributed to the nature of the GLCM algorithm which quantifies the local textures in terms of "gray-level" or intensity variations among the neighboring pixels instead of the global patterns of the diffraction images that are recognizable by naked eyes within the FOV. Therefore, it is critical to develop an objective and practical method to separate those diffraction images of abnormal global patterns for subsequent classification of the images associated with the viable cells. This is particularly important to obtaining the training data to ensure the accuracy of automated classification based on algorithms such as the support vector machine (SVM) [23, 24]. For these reasons we have investigated various image processing algorithms for separating the images before the use of GLCM algorithm for quantifying local textures. In this report, we demonstrate that the abnormal diffraction images are most likely due to cellular fragments or debris and present an automated image processing method to rapidly identify them.

2. Methods

Figure 1(a) presents the schematic of an experimental DIFC system with more details given in [15, 16, 18]. The system contains a cw laser as the coherent light source with a beam of 532nm in wavelength, linear polarization and power variable from 10 to 160mW to illuminate single objects. The incident beam along the z-axis is focused on the core fluid carrying cells along the y-axis. The polarization direction of the incident beam can be adjusted with a half-wave plate to horizontal, vertical and 45° from the horizontal. The light scattered by the imaged object is collected using a 50x infinity-corrected objective (378-805-3, Mitutoyo) of NA = 0.55 followed by a polarizing beam splitter in an angular range of about 36° centered along the x-axis. Two tube lenses behind the beam splitter focus separately the outputs from the objective and beam splitter assembly onto two CCD cameras (LU075M, Lumenera) for acquisition of a s- and p-polarized image pair of 640x480 pixels and 12-bit pixel depth. The CCD cameras were aligned respectively to the focal planes of the two tube lenses using a

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

stationary target and a white light source. Afterwards the imaging unit consisting of the objective, optics and cameras was focused on the core fluid and the focal spot of the incident laser beam. At this location defined as $\Delta x = 0$, the imaging unit has a resolution of about 0.6μm and collects the scattered light over an angular range of about $\Delta\theta_s = \pm 18°$ in polar angle and $\Delta\phi_s = \pm 14°$ in azimuthal angle with respect to the incident beam direction or z-axis. The imaging unit was translated as whole to an off-focus position of $\Delta x = 100\mu m$ along the x-axis toward the flow chamber to acquire diffraction images with the white light source turn off. The off-focus imaging is necessary to obtain diffraction images that highly correlated to the morphology of the coherently illuminated cellular object as discussed in [16].

All diffraction image pairs investigated for this study were acquired from different suspension sample of the PC-3 cells (CRL-1435, ATCC) derived from an aggressive bone metastasis of a human prostatic adenocarcinoma. After data acquisition the images were normalized and converted into 8-bit format for later processing with the maximum, minimum and average pixel intensities stored. Visually the acquired diffraction images can be separated into three major types according to the global patterns as discussed earlier. For the convenience of discussion, we term each type as normal speckle, large speckle and stripe patterns as demonstrated by one example of the cross-polarized diffraction image pairs in Figs. 1(b)-1(d), respectively. In our previous numerical and experimental studies, we have shown that the images of the normal speckle patterns can be attributed to the coherent superposition of electromagnetic fields of light scattered from biological cells of sizes 5μm or larger with complex morphology [16, 17, 20]. In addition, we have also shown that the stripe patterns of high symmetry can arise from single or aggregated spheres or spherical particles of homogeneous structures [15, 18].

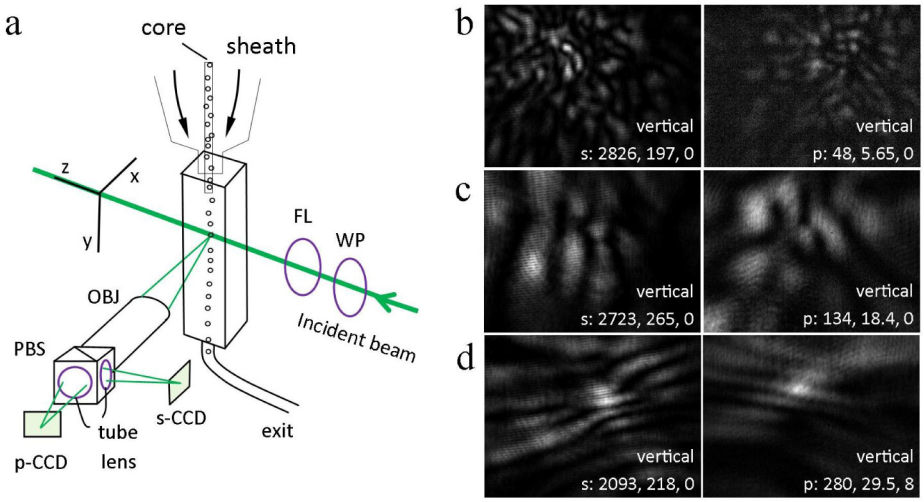

Fig. 1. (a) The schematic of the experimental DIFC system with the core and sheath fluids: WP = half-wave plate; FL = focusing lens; OBJ = infinity-corrected objective; PBS = polarizing beam splitter; CCD = camera recording either s- or p-polarized scattered light. (b) - (d): normalized cross-polarized diffraction image pairs labeled by polarization of the incident beam and scattered light, maximum, average and minimum pixel intensities of the 12-bit images.

To gain insights on the correlation between morphology and the features of diffraction images, we have performed numerical simulations of light scattering by different objects modeling intact cells, cellular fragments and aggregates of small solid particles. A parallel computing software based on the discrete dipole approximation (DDA) [25–27] has been adopted to model light scattering by cellular objects. For this purpose, we have developed a Gaussian random sphere method to obtain 3D and realistic cellular structures with nucleus and mitochondria based on the image stacks acquired with a confocal microscope [4, 28, 29]. Scatterers with full and partial cellular structures were simulated with the DDA software

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

using the coordinate system shown in Fig. 1(a). The angle-resolved 4x4 Mueller matrix (M_{ij}) with i or j = 1, 2, 3, 4 was first obtained and normalized by equating the full-sphere angular integral of the M_{11} element to 1 [3]. With the Mueller matrix representing a cellular object and a Jones vector representing a linearly polarized incident beam, we can express the ratios of the s- and p-polarized scattered light intensities with the incident beam intensity I_0 as

$$I_{p,45^\circ} / I_0 = (M_{11} + M_{13}) + (M_{21} + M_{23}), \quad (1)$$

$$I_{s,45^\circ} / I_0 = (M_{11} + M_{13}) - (M_{21} + M_{23}), \quad (2)$$

$$I_{p,ver} / I_0 = (M_{11} - M_{12}) + (M_{21} - M_{22}), \quad (3)$$

$$I_{s,ver} / I_0 = (M_{11} - M_{12}) - (M_{21} - M_{22}), \quad (4)$$

$$I_{p,hor} / I_0 = (M_{11} + M_{12}) + (M_{21} + M_{22}), \quad (5)$$

$$I_{s,hor} / I_0 = (M_{11} + M_{12}) - (M_{21} + M_{22}). \quad (6)$$

The s- and p-polarized diffraction images were simulated by projecting the angle-resolved elements M_{ij} onto an area in the Z-Y place corresponding to the CCD sensor centered on the x-axis. We have selected an angular range of $\Delta\theta_s = \pm 18°$ and $\Delta\phi_s = \pm 14°$ for obtaining the simulated diffraction images according to the experimental setup. In these calculations, we do not consider the imaging optics used in the experimental system shown in Fig. 1(a). Therefore, the simulated diffraction images should only be used as a guide to understand the correlation between object's morphology and cross-polarized diffraction image pairs instead of a tool for inverse solutions.

3. Results and discussion

Figure 2 shows examples of the simulated diffraction images by models of an intact cell, fragmented cellular structures and aggregates of different particles. For these simulations, we have assumed that the wavelength of the incident light beam is 532nm and all objects are immersed in a host medium of refractive index $n_h = 1.3340$. Through our numerical studies, we found that the heterogeneity of the nucleus and large number of mitochondria have to be considered to obtain diffraction images of normal speckle patterns similar to those in Figs. 1 and 3. For this purpose the refractive index inside the nucleus, n_n, was set to take the values of [1.3797, 1.3897, 1.3997, 1.4097, 1.4197] according to the fluorescence intensity of the nuclear dye (Syto-61, Life Technologies) while the index of mitochondria was set as $n_m = 1.4200$. The refractive indices of other cellular organelles are given by $n_{nu} = 1.4397$ for nucleoli, $n_{nm} = 1.4097$ for nuclear membrane and $n_c = 1.3675$ for cytoplasm of full cell structures.

By comparing the simulated images, one can identify the two global pattern types of normal speckle and stripe as demonstrated respectively in Figs. 2(a) and 2(b). The diffraction images of normal speckle patterns can be obtained with a full and heterogeneous 3D cell structure built from the confocal image stacks acquired from cultured cells, double stained by nuclear and mitochondria fluorescent dyes [24], with equivalent diameters d ranging from 8 to 15μm (one example with d = 10μm is shown here). Those of stripe patterns, however, were calculated from cellular "shells" (only one example is shown). These shell structures were obtained by slightly displacing the centers of two full cell structures and removing the overlapped portion from one cell structure. The leftover shell from a full cell structure keeps only a small fraction of the original cell volume, set at 10% or 20%, with the emptied space filled with the host medium. Consequently, these shell structures possess high degree of symmetry in their structures and yield stripe patterns in the simulated diffraction images of high symmetry in comparison to those by the full cell structures. It is hence clear that the stripe patterns in diffraction images are likely caused by large sized objects with a nearly homogeneous internal structure in terms of the refractive index distribution. When the

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

illuminated objects reduce their sizes, the speckle sizes increase invariably as demonstrated by the images shown from Figs. 2(c)-2(f). Despite the above variations, we conclude from our numerical study that the global patterns of the simulated diffraction images by the side scattered do not depend sensitively on the size of the objects, and the transition from the normal to large speckle patterns is associated with the categorical shift in object sizes from the neighborhoods of 10μm to 2μm.

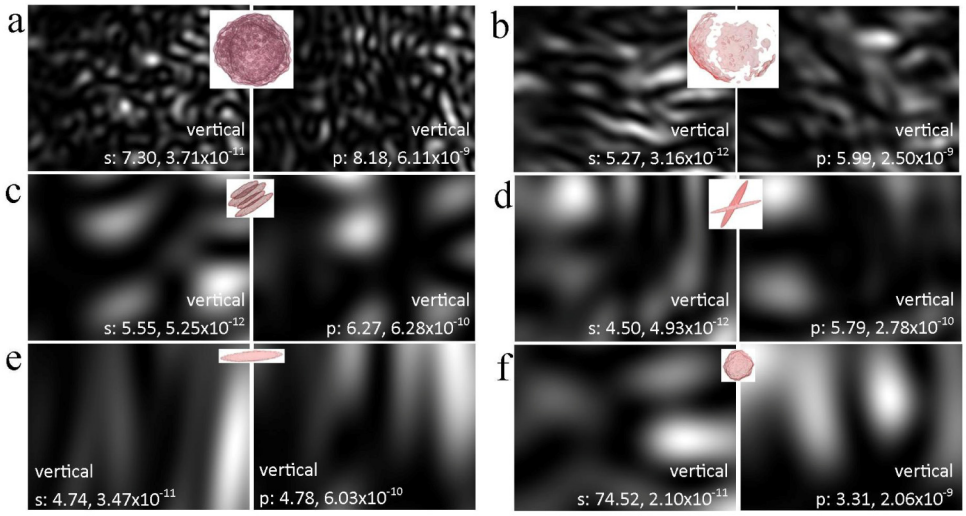

Fig. 2. The normalized cross-polarized diffraction images obtained through DDA simulations of light scattering by (a) a large cellular structure with nucleus and mitochondria and equivalent diameter d = 10μm; (b) a cellular "shell" structure with only 10% of volume of (a) after removing the overlapped region with another similar cell; (c) a stack of 4 parallel ellipsoids with long axial lengths d_1 = 2.7, 3.0, 3.9, 4.0μm and short axial lengths d_2 = 0.5μm; (d) angled stack of two identical ellipsoids of d_1 = 3.0, 6.0μm and d_2 = 0.5μm; (e) one ellipsoid of d_1 = 4.0μm and d_2 = 0.5μm; (f) a small homogeneous particles of cellular shape with d = 1.1μm. Each image is labeled by polarization of the incident and scattered light, ratio of maximum to averaged pixel intensity. The insets are cross-section views of the objects (not to scale).

To divide accurately the diffraction images into the three types of global patterns, we have developed an image processing method consists of four steps. In the first step, a normalized diffraction image I(z, y) was converted into a binary one B(z, y) using the average pixel intensity as the threshold. Afterwards, four Sobel operators given below for edge detection were applied to obtain the speckle borderlines or edges along four directions of horizontal (h), vertical (v), left diagonal (l) and right diagonal (r) respectively [30]

$$G_h = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}, G_h = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, G_l = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix}, G_r = \begin{pmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}. \quad (7)$$

By convoluting the four operators with B(z, y), one can derive a set of four directional edge images as E_a(z, y), where a = h, v, l, r, and a complete edge image E_T(z, y) by summing the 4 directional images. These images were made binary again by setting those pixels of maximum intensity for speckle borderlines to 1 against the background by all other pixels set to 0. Examples of the binary edge images are presented in Fig. 3(a). From the binary edge images the number of the pixels of 1 in each image can be quickly summed for a set of 5 borderline length parameters as [C_v, C_h, C_l, C_r] and C_T. Once these parameters were obtained we can separate those images of the stripe patterns from the other two types of speckle patterns. This is accomplished by comparing two C parameters in a pair of edge images, [C_v, C_h] or [C_l, C_r],

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

consisting of mutually perpendicular directions. If the following conditions are satisfied in one pair of the C parameters

$$C_1 < C_{th} \quad \text{and} \quad C_1 < 0.3C_2, \tag{8}$$

where C_1 is the lesser of the two C parameters in the pair and $C_{th} = 2500$ is a threshold, then the diffraction image is declared as the stripe type with the stripe directions along or approximately along the direction of the edge image with C_2.

In the second step of image processing, additional parameters are extracted from I(z, y) in frequency space for accurate grouping the rest of the diffraction images according to the speckle sizes in real space. A 2D fast Fourier transform (FFT) was used to map I(z, y) into the frequency space (u, v) as

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(z,y)\exp\{-2\pi i(uz+vy)\}dzdy, \tag{9}$$

from which a power spectrum image can be obtained as $P(u,v) = |F(u,v)|$. Analysis of the measured diffraction images led to the observation that the images of large speckle patterns as shown in Fig. 1(c) often contain bright spots of linear sizes of 150 or more pixels. Consequently, a frequency threshold was chosen as with Δ as the inter-pixel distance

$$f_{th} = \frac{1}{150(\Delta)} = 0.00667\left(\frac{1}{\Delta}\right). \tag{10}$$

With f_{th} we can derive a histogram $N(f)$ of high frequency pixels in $P(u, v)$ with $f = (u^2 + v^2)^{1/2}$ and N as the number of pixels with $f > f_{th}$ and $P(u, v) > 0.02 \cdot P(0, 0)$. The later ensures that the selected high frequency pixels have intensities above the noise background. The sum of $N(f)$ yields the number of pixels having high power and frequency, N_P, in the power spectrum image $P(u, v)$. Examples of edge images and $N(f)$ for each of the three pattern types of the diffraction images are demonstrated in Fig. 3.

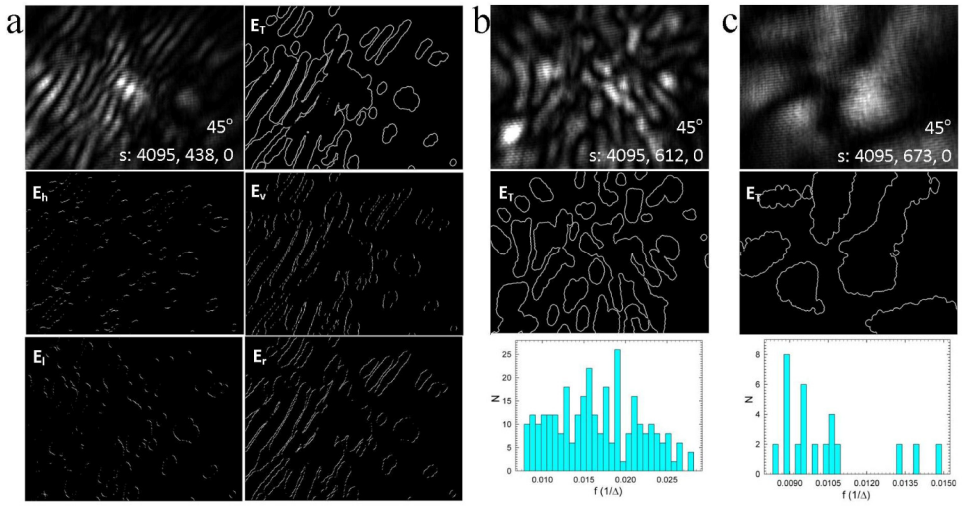

Fig. 3. Measured diffraction images I of different pattern types with their binary edge images E and frequency histogram N(f): (a) a stripe pattern and its five binary edge images of different C parameters as labeled; (b) a normal speckle pattern; (c) a large speckle pattern.

It is quite obvious that diffraction images with the normal speckle pattern tend to have C_T and N_P values larger than those of the images with large speckle patterns. Still we found the presence of noise due to spurious light and the large morphological variations of the imaged

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

objects can lead to significant fluctuations in the absolute values of C_T and N_P and their relative differences among samples of different cells. To achieve automated classification, one has to develop a calibration procedure to minimize the effect of these fluctuations. As the third step of image processing to be described below we have employed a k-means clustering technique [31] to rank and scale the extracted parameters of any given data set as the calibration procedure.

The k-means clustering technique allows separation of all dots into k groups according to their distances to k centers under appropriate conditions [31]. If a dot is closer to one center than the others, it is assigned into a group represented by that center. In this study we utilize two 2D scatter plots of $A_s(N_P, C_T)$ and $A_p(N_P, C_T)$ on the N_P-C_T plane as shown in Fig. 4(a) or 4(b) to represent a data set of diffraction image pairs measured with a cell sample. For comparison, we also present the simulated image pairs shown in Fig. 2 in the scatter plots of Fig. 4(a), which clearly indicates that the global patterns as characterized by the two parameters of N_P and C_T correlate strongly with the categorical sizes of the imaged objects.

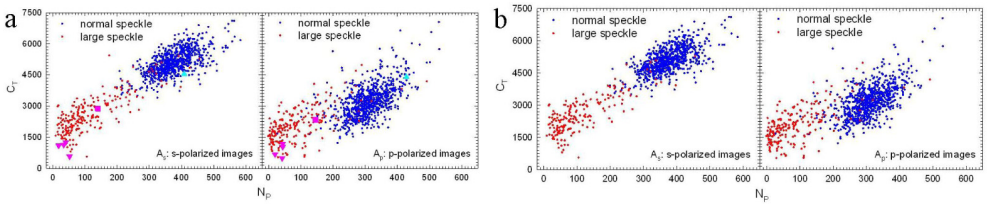

Fig. 4. Scatter plots of the diffraction image pairs in which the coloring of the dots represents the manual assignment by two of co-authors: (a) plots of 1000 image pairs acquired with a PC-3 cell sample and a 45° polarized incident laser beam and 6 simulated image pairs shown in Fig. 2 with the cyan-up-triangle symbol for 2(a), pink-square for 2(b) and pink-down-triangle for 2(c) - 2(f); (b) plots of 951 measured image pairs shown in (a) after removal of 35 image pairs of stripe patterns and 14 pairs of over- or under-exposure.

After removal of the image pairs of stripe patterns and over- or under-exposed pairs, the measured data set was re-plotted in Fig. 4(b) for subsequent analysis, which was also chosen as the reference set for calibrating other data sets as discussed below. To divide the image pairs into two pattern types with k = 2, two centers were initially assigned as (N_{Pai}, C_{Tai}) with a = s or p for each of the two plots. The initial centers were then updated by the averaged values of the N_P and C_T parameters in each plot and the process iterated until the two centers converged to the final values of (N_{Paf}, C_{Taf}). To reduce the effect of fluctuation in C_T and N_P, we modified the standard k-means clustering technique for the above iteration with two changes. First only those dots within a circle of radius R from a current center location were counted, with R determined empirically, to obtain the updated location. Secondly, the updating of the centers for the image type of large speckle patterns was limited to the region of small N_P and C_T based on the fact that these images cannot have large values. If the updating led to a new center for this group of dots outside the limited region, the initial center was used instead.

The k-means clustering technique makes it possible to select automatically one polarized image group with tighter clustering along the right diagonal direction among the two groups of s- or p-polarized images as the higher ranked data for subsequent classification based on a SVM algorithm [23, 24]. This is due to the fact that the higher ranked group consists of images often with higher signal-to-noise ratios among the two and the extracted parameters of (N_{P1}, C_{T1}) yield better results of SVM classification [20]. For the data set shown in Fig. 4(a) or 4(b), the group of s-polarized images received the higher rank with 1 = s followed by 2 = p. After correct ranking, the SVM analysis can be performed using a classification vector of four parameters $(N_{P1}, C_{T1}, N_{P2}, C_{T2})$, for each imaged object. To further eliminate the need for training the SVM algorithm repeatedly for different data sets, we developed a method to scale a new data set according to the reference set shown in Fig. 4(b). Specifically, the four

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

parameters extracted from a cross-polarized diffraction image pair in the new data set are scaled as follows

$$N_{Pi}' = \frac{N_{Pif}\big|_{ref}}{N_{Pif}} N_{Pi} \quad \text{and} \quad C_{Ti}' = \frac{C_{Tif}\big|_{ref}}{C_{Tif}} C_{Ti}, \tag{11}$$

where i = 1 or 2, N_{Pif} and C_{Tif} are the averaged parameter values obtained by the k-means clustering analysis on the new data set, $N_{Pif}\big|_{ref}$ and $C_{Tif}\big|_{ref}$ are the corresponding values of the reference data set.

In the final step of image processing, a SVM based approach was applied for automated classification of the imaged objects into the two types of speckle patterns after removal of those with the stripe patterns using the calibrated parameters of (N_{P1}', C_{T1}', N_{P2}', C_{T2}'). The SVM classifier was implemented by the LIBSVM library [24] with details described elsewhere [17, 20]. The SVM classifier was initially trained with 620 of the reference data set of 951 image pairs shown in Fig. 4(b) which were manually assigned into the two types of speckle patterns by two of the co-authors. Four different kernel functions were investigated for achieving accurate classification and it was found that the best results can be obtained with a linear kernel function. Through a five-fold training scheme [20] the criteria of classification were determined as the specific ranges of the calibrated parameter in (N_{P1}', C_{T1}', N_{P2}', C_{T2}'). To evaluate the SVM classification performance against visual examination, we employed a parameter of classification accuracy A given as

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \tag{12}$$

where TP is the number of correctly identified image pairs of normal speckle patterns, TN the number of correctly identified pairs of large speckle patterns, FP and FN respectively the number of incorrectly identified pairs of normal and large speckle patterns. The averaged values of A were found to be 97.1% and 97.6% for the training and test data groups of the reference data presented in Fig. 4(b). The image processing method trained with the reference data set has also been applied to another data set acquired from the same cell sample with horizontally polarized incident beam against visual examination. This data set contains 925 image pairs and A was found to be 97.3% against visual examination.

It should be noticed that in conventional flow cytometry the angle-integrated forward scatter signal is often used as the indicator of the volume of the measured objects while the side scatter signal can only indicate the degree of internal heterogeneity. From the simulated images shown in Fig. 2, it can be seen that the angle-integrated side scatter signal is equivalent to the average pixel intensity which indeed cannot be used as a useful indicator of the object's volume or size. The results of the diffraction image analysis presented in this report, however, demonstrate clearly that the global patterns of these images provide reliable information about the size of the imaged object and morphology as well. As the last note, the software of image processing method described above has been developed on the Matlab platform (Version 7.1, MathWorks). The computation time was about 0.79s per image pair on a standard personal computer with 2GHz CPU and the calculation of C values is the most time consuming portion at 44%. We are currently optimizing the algorithms and implementing a GPU execution code to significantly increase the speed of image processing.

4. Summary

Through numerous measurement of cultured and primary cell samples with the DIFC method, we have observed two classes of pattern variations in the diffraction images measured along the side scattering directions. One is associated with the global patterns within the FOV and the other with local textures among neighboring pixels. In this report, we show theoretically that the variation of global patterns associates primarily to the categorical differences in the

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013

cellular objects and, to a lesser degree, to the internal heterogeneity as well. An image processing method has been developed and presented for accurate and automated classification of measured diffraction images into three types of normal speckle, large speckle and stripe according to their global patterns. The method yields a critically important tool which allows automated and objective removal of the "abnormal" diffraction images of large speckle and stripe patterns due to most likely the fragmented cellular objects and debris before the detailed analysis of the images of normal speckle patterns [17, 20]. The latter is most likely associates with viable and intact cells and detailed analysis of the local textures by previously developed GLCM algorithm, among others, can enable accurate classification of cells according to their morphology.

Acknowledgment

We thank Dr. K.M. Jacobs and Mr. W. Jiang for data acquisition. Y. F. acknowledges the grant supports from NSFC (#81171342 and #81041107) and X.H. Hu acknowledges the grant support from the Golfers Against Cancer (2012-13-GAC).

#196723 - $15.00 USD Received 29 Aug 2013; revised 26 Sep 2013; accepted 26 Sep 2013; published 9 Oct 2013