

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## Comparison of contourlet transform and gray level co-occurrence matrix for analyzing cell-scattered patterns

Jun Zhang  
Gang Wang  
Yuanming Feng  
Yu Sa

**SPIE.**

Jun Zhang, Gang Wang, Yuanming Feng, Yu Sa, "Comparison of contourlet transform and gray level co-occurrence matrix for analyzing cell-scattered patterns," *J. Biomed. Opt.* **21**(8), 086013 (2016), doi: 10.1117/1.JBO.21.8.086013.

# Comparison of contourlet transform and gray level co-occurrence matrix for analyzing cell-scattered patterns

Jun Zhang, Gang Wang, Yuanming Feng, and Yu Sa\*

Tianjin University, Department of Biomedical Engineering, 92 Weijin Road, Tianjin 300072, China

**Abstract.** Distribution of scattered image patterns hinges on morphological and optical characteristics of cells. This paper applied a numerical method to simulate scattered images of real cell morphologies, which were reconstructed from confocal image stacks dyed by fluorescent stains. Two approaches, contourlet transform (CT) and gray level co-occurrence matrix (GLCM), were then used to analyze the simulated scattered images. The results showed that features extracted using GLCM contained more information than those extracted using CT. Higher classification accuracy could be achieved with a single GLCM parameter than CT and GLCM could achieve higher accuracy with fewer parameters than CT when using multiple parameters. Meanwhile, GLCM requires less computational cost. Thus, GLCM is more suitable and efficient than CT for the analysis of cell-scattered images. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.21.8.086013](https://doi.org/10.1117/1.JBO.21.8.086013)]

Keywords: contourlet transform; gray level co-occurrence matrix; scattered image; cell morphology; mitochondria.

Paper 160153RR received Mar. 10, 2016; accepted for publication Aug. 8, 2016; published online Aug. 23, 2016.

## 1 Introduction

Mitochondrion is one of the most important organelles in most cells; it supplies the energy used by cells and plays a significant role in many cell activities,<sup>1</sup> including cell differentiation and apoptosis. Cell-scattered light stimulated by coherent light can provide deep insight into a cell's morphology.<sup>2-4</sup> Previously, polarization diffraction imaging flow cytometry was developed, which could simultaneously obtain cross-polarized scattered images or a polarization-scattered image (p-SI) pair per cell carried by core flow and illuminated by a laser. It has been shown that the spatial patterns in p-SI images correlate with morphological features of biological cells,<sup>5-8</sup> but it is difficult to distinguish scattered images obtained from different cell morphologies. Therefore, an effective assay algorithm for analyzing the scattered images corresponding to the cell morphology is necessary.

This paper reports a simulation method based on an accurate process of scattered light simulation<sup>6,9-12</sup> and a reconstructed real cell morphology obtained in previous studies<sup>6,12-14</sup> to simulate the polarized cell-scattered images. The main characteristic of scattered images is pattern distribution, which can be treated as a type of frequency information. Contourlet transform (CT) can decompose an input image into several subimages that contain different scales and direction frequency information.<sup>15-18</sup> Gray level co-occurrence matrix (GLCM) analyzes the probability of pixel pairs, which is also frequency information.<sup>19,20</sup> Thus, these two algorithms were applied to analyze scattered images simulated from real cell morphologies with different mitochondria volumes and refractive indices (RIs). The results our comparative study showed that features extracted using GLCM contain more information than those extracted using CT.

Higher classification accuracy could be achieved with a single GLCM parameter than CT, and GLCM could achieve high accuracy with fewer parameters than CT when using multiple parameters. Meanwhile, GLCM requires less computational cost. GLCM is a more appropriate assay method than CT for cell-scattered images.

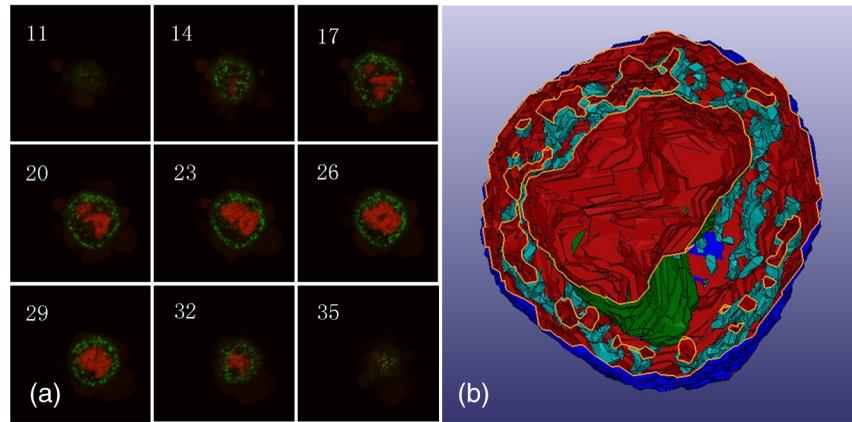
## 2 Materials and Methods

### 2.1 Reconstructed Cell Three-Dimensional Structure and Refractive Index Model

The reconstructed cell three-dimensional (3-D) structures were acquired using a laser scanning confocal microscope (LSM510, Zeiss, Germany). Cells were stained with two fluorescent dyes, Syto 61 and Mito Tracker Orange CMTMRos (S11343 and M-7510, Life Technologies), using the same protocol as described in previous papers.<sup>6,12</sup> Syto 61 marks the nucleic acids mainly distributed inside the nucleus, and Mito Tracker Orange CMTMRos mostly adheres to mitochondria. Stimulated by two laser beams with wavelengths of 633 and 532 nm, two 12-bit fluorescent intensity images of the nucleus and mitochondria were saved in the red and green channels, respectively, of an image stack with a resolution of 0.07  $\mu\text{m}$  in the  $X$ - and  $Y$ -directions and 0.5  $\mu\text{m}$  in the  $Z$ -direction as  $F_r(r)$  and  $F_g(r)$ , where  $r$  is the voxel position inside the cell. Approximately 60 slices were acquired for each cell at different focal planes with a step size of 0.5  $\mu\text{m}$  in one stack. Figure 1(a) shows some images at different focal planes of one normal human prostate epithelial cell, termed PCS (PCS440010, ATCC).

After acquisition, the confocal image stack was processed with in-house software. The cell was segmented into three regions, nucleus, mitochondria, and cytoplasm, which were denoted as  $\Omega_n$ ,  $\Omega_m$ , and  $\Omega_c$ , respectively.  $\Omega_n$  was further divided into three subregions denoted as  $\Omega_{n1}$ ,  $\Omega_{nm}$ , and  $\Omega_{nh}$  with

\*Address all correspondence to: Yu Sa, E-mail: [sayu@tju.edu.cn](mailto:sayu@tju.edu.cn)



**Fig. 1** PCS cell confocal images (a) typical confocal images of a PCS cell stack marked by the slice index at the upper left corner, (b) a section of the cell morphology reconstructed from (a); dark blue, green, and light blue represent the external surfaces of cell membrane, nucleus, and mitochondria, respectively.

relatively low, medium, and high RIs according to their pixel intensities, respectively; larger intensity means more nucleic acids, which should have a larger RI. To achieve accurate simulated diffraction images, cell 3-D structures should have the same resolution in all three directions. Thus, interpolation between confocal slices was performed. As Amsterdam discrete dipole approximation (ADDA) code treated scatterer as an array of voxel that has the same size in three directions, equal resolution in three dimensions is also necessary for accurate simulation. Because of the different resolution of confocal images between different directions,  $0.07 \mu\text{m}$  in  $X$  and  $Y$  and  $0.5 \mu\text{m}$  in  $Z$ , the 3-D cell structures were interpolated in  $Z$  direction to obtain the same resolution as in  $X$ - and  $Y$ -directions, and the precise cell structure used for ADDA code. Details of segmentation and interpolation have been described elsewhere.<sup>12,14</sup> The reconstructed 3-D cell structure is shown in Fig. 1(b) in which the nucleus, mitochondria, and cytoplasm can be clearly observed.

Because of the heterogeneity between and inside organelles, the cell RIs were modeled as the sum of a constant term, a mean RI value with a specific value in a specific region, and a randomly fluctuating term that indicates the heterogeneity inside organelle as follows:

$$n_{\alpha}(r) = n_{\alpha 0} + (n_{\alpha 0} - n_w) \cdot A_{\alpha} \cdot \text{RND}, \quad \forall r \in \Omega_{\alpha}, \quad (1)$$

where  $\alpha = c, m, nl, nm, \text{ or } nh$ , which indicates the region or sub-region,  $n_{\alpha 0}$  is the mean value of RI in the region of  $\Omega_{\alpha}$ ,  $n_w$  is the RI of water,  $A_{\alpha}$  is the fluctuation amplitude, and RND denotes random numbers uniformly distributed in  $[-1, 1]$ . To simulate the heterogeneity inside organelle and distinguish different organelles,  $A_{\alpha}$  was set to 10% in this paper, which gave every organelle an independent RI range. Combining reconstructed cell 3-D structures with RI, a cell model was obtained. By varying RI or modifying the cell structure, a series of cell models was derived from one confocal stack.

## 2.2 Simulation of p-SIs Using a Cell Model

To obtain the scattered light distribution of a cell model in a host medium, an open source code of a parallel ADDA algorithm for discrete dipole approximation was applied to simulate angularly resolved Mueller matrixes  $S_{ij}$  where  $i, j = 1, 2, 3, 4$ .<sup>11,21</sup> The cell model was divided into several voxels in three directions, and

each voxel was treated as a dipole. For example, the PCS cell mentioned above was treated as a  $116 \times 110 \times 86$  dipole array, and the number of dipoles per wavelength used for this cell was 3.8.  $S_{ij}$  were calculated from all fields scattered by these dipoles that were excited by an incident light. With the incident wavelength of 532 nm and RI of host media (water)  $n_w = 1.334$ , the polarizabilities could be calculated. ADDA code could be executed on a parallel computing cluster to calculate Mueller matrixes ( $S_{ij}$ ), which are functions of scattering polar angles  $(\theta, \phi)$ . Mie theory was used to validate ADDA code with a  $10\text{-}\mu\text{m}$  sphere, which was represented by a  $126 \times 126 \times 126$  dipole array, and the number of dipoles per wavelength was 10.08. The normalized  $S_{11}$  calculated by Mie theory and ADDA are shown in Fig. 2(a). The absolute error and relative error between ADDA and Mie theory  $S_{11}$  are shown in Fig. 2(b).

The configuration of the simulation system is shown in Fig. 3(a). First, the scattered light was projected onto an input plane  $P_{in}$ , which is an arbitrary plane defined between the scatterer and the objective to obtain an input p-SI denoted as  $I_{in-kl}(y, z)$  used for the simulation process.  $I_{in-kl}(y, z)$  was obtained using a linear combination of Mueller matrix elements, which carry the spatial distribution of scattered light for specific scattered polarization  $k = p$  or  $s$  and incident polarization  $l = p, s$ , or 45 deg. A total of three pairs of input p-SI were obtained with three types of incident polarizations. For example, the input p-SI  $I_{in-kl}(y, z)$  of scattered polarization  $p$  and incident polarization  $p$  was calculated by

$$I_{in-pp}(y, z) = S_{11} + S_{12} + S_{21} + S_{22}. \quad (2)$$

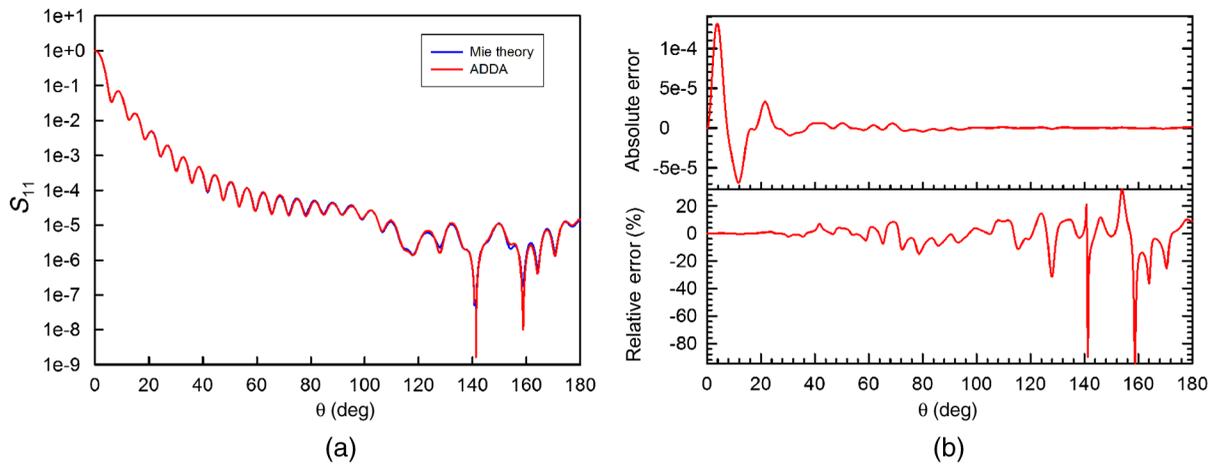
Other equations of different polarizations can be found in Refs. 8 and 22.

A ray-tracing software program (Zemax-EE v2005, Zemax Development Corp.) was used to trace the rays of input p-SI  $I_{in-kl}(y, z)$  at the input plane  $P_{in}$  to the imaging unit at image plane  $P_{im}$  and obtain the p-SI  $I_{kl}(y, z)$  at  $P_{im}$ , which is the conjugate image of the object plane at defocus position  $x = 150 \mu\text{m}$ . Simulated p-SIs  $I_{kl}(y, z)$  are similar to the measured images obtained in our cell-scattered image measurement experiments with a microscope objective as indicated by their similar GLCM parameter distribution ranges (shown in Table 1). Some examples of simulated p-SI  $I_{in-kl}(y, z)$  and  $I_{kl}(y, z)$  and measured p-SI are shown in Figs. 3(b) and 3(c).

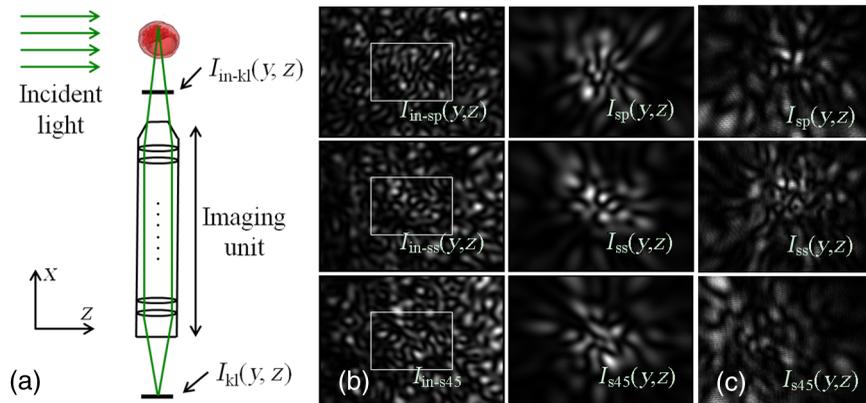
### 2.3 Analysis of Simulated p-SI with CT

CT is an algorithm that combines the merits of wavelet transform and directional filter banks. It can decompose the image into multiple subimages with different scales  $\gamma = 1, 2, 3, \dots, \gamma_{\max}$  and directions  $\delta = 1, 2, 3, \dots, \delta_{\max}$ , which enable simultaneous analysis of p-SI information at different scales and directions in the frequency domain. First,  $I_{kl}(y, z)$ , denoted as  $I_{kl,0-0}(y, z)$ , which can be treated as the low-pass image with overview information in scale 0, is processed by

the Laplacian pyramid algorithm.  $I_{kl}(y, z)$  is down-sampled by a weighted smoothing filter, so a low-pass image denoted as  $I_{kl,1-0}(y, z)$  is acquired, carrying the overview information of  $I_{kl,0-0}(y, z)$  in scale 1. The bandpass image  $I_{kl,1-d}(y, z)$  is then obtained by the difference of  $I_{kl,0-0}(y, z)$  and the up-sampled images of  $I_{kl,1-0}(y, z)$ . By repeating this procedure on low-pass images in each scale, a series of low-pass and bandpass images are acquired at different scales denoted as  $I_{kl,\gamma-0}(y, z)$  and  $I_{kl,\gamma-d}(y, z)$ , respectively, with overview and



**Fig. 2** Validation of ADDA by Mie theory: (a) Mie theory  $S_{11}$  with blue line and ADDA  $S_{11}$  with red line for  $10\text{-}\mu\text{m}$  sphere in common logarithm coordinates (b) absolute error and relative error between Mie theory  $S_{11}$  and ADDA  $S_{11}$  for  $10\text{-}\mu\text{m}$  sphere.



**Fig. 3** System setup and scattered images: (a) the configuration of the ADDA simulation system, (b) representative scattered images simulated from the cell in Fig. 1(b) by ADDA: the first column contains p-SIs at the input plane with the field of view indicated with the rectangle, and the second column contains p-SIs at the image plane; (c) three experimental cell-scattered images.

**Table 1** Distribution range of GLCM parameter for simulated and measured images.

	$SVA_{p45}$	$SEN_{p45}$	$MEA_{p45}$	$SVA_{s45}$	$SEN_{s45}$	$MEA_{s45}$
Simulated images	$46.24 \pm 8.09$	$4.80 \pm 0.18$	$23.12 \pm 4.04$	$47.50 \pm 12.19$	$4.56 \pm 0.38$	$23.74 \pm 6.09$
Measured images	$43.51 \pm 12.77$	$4.60 \pm 0.40$	$21.75 \pm 6.39$	$54.39 \pm 15.38$	$4.85 \pm 0.30$	$27.19 \pm 7.69$

detail information, where  $\gamma$  is the scale index. Second, bandpass images  $I_{kl,\gamma-d}(y, z)$  are processed by directional filter banks.  $I_{kl,\gamma-d}(y, z)$  is decomposed into several directional images denoted as  $I'_{kl,\gamma-\delta}(y, z)$  where  $\delta = 1, 2, 3, \dots, \delta_{\max}$ , which contain the detail information in specific directions. This report used  $\gamma_{\max} = 5$  and  $\delta_{\max} = 2^3$  for scales 1–4 and  $\delta_{\max} = 2^2$  for scale 5. In this study, 41 low-pass and bandpass subimages were obtained from one simulated p-SI, and 12 parameters were calculated from each subimage. These parameters were then tested by some stripe images with different periods or directions. The results showed that four of these 12 parameters varied significantly among the 41 subimages. Thus, these four parameters were chosen to characterize subimages, and the parameters energy  $E$ , contrast  $C$ , variance  $V$ , and fluctuation  $F$  were defined as

$$\begin{aligned} E_{kl,\gamma-\delta} &= \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} I'_{kl,\gamma-\delta}(y, z)^2 \\ C_{kl,\gamma-\delta} &= \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} \frac{\text{SLC}_{kl,\gamma-\delta}(y, z)}{4N_y N_z - 2(N_y + N_z)} \\ V_{kl,\gamma-\delta} &= \frac{1}{N_y N_z - 1} \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} [I'_{kl,\gamma-\delta}(y, z) - \text{AVE}_{kl,\gamma-\delta}]^2 \\ F_{kl,\gamma-\delta} &= \frac{\sqrt{V_{kl,\gamma-\delta}}}{\text{AVE}_{kl,\gamma-\delta}}, \end{aligned} \quad (3)$$

where  $k$  indicates the polarization of scattered light from cells;  $l$  indicates the polarization of incident light;  $N_y$  and  $N_z$  are the pixel numbers of CT image  $I_{kl,\gamma-d}(y, z)$  in  $y$ - and  $z$ -directions, respectively; and squared local contrast SLC and mean value AVE are defined as

$$\begin{aligned} \text{SLC}_{kl,\gamma-\delta}(y, z) &= \sum_{i=-1,1} \{ [I'_{kl,\gamma-\delta}(y + i, z) - I'_{kl,\gamma-\delta}(y, z)]^2 \\ &\quad + [I'_{kl,\gamma-\delta}(y, z + i) - I'_{kl,\gamma-\delta}(y, z)]^2 \} \\ \text{AVE}_{kl,\gamma-\delta} &= \frac{1}{N_y N_z} \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} I'_{kl,\gamma-\delta}(y, z). \end{aligned} \quad (4)$$

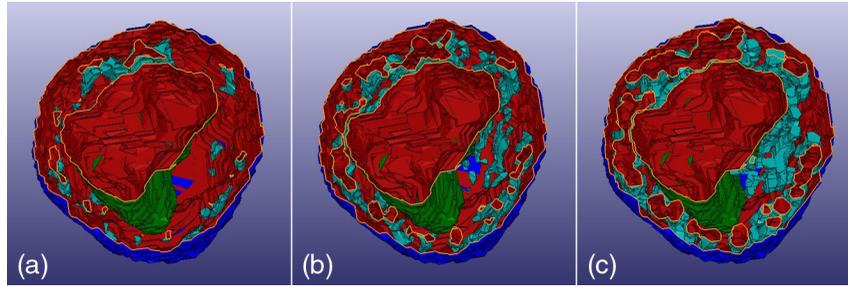
## 2.4 Analysis of Simulated p-SI with GLCM

GLCM is an image texture-analyzing algorithm based on the occurrence probability of gray level pairs. GLCM can be used to describe the orientation, amplitude, and period information of an image texture.<sup>19,20</sup> First, the GLCMs  $P(i, j, d, \psi)$  are calculated by the numbers of repetitive pixel pairs  $(i, j)$  in the p-SI  $I_{kl}(y, z)$ , which have an assigned distance  $d$  at an assigned angle  $\psi$  with intensities  $i$  and  $j$  where  $i, j = 1, 2, 3, \dots, G$ ,  $G$  is the maximum intensity,  $d$  can be any integer smaller than the image size, and  $\psi$  can be 0 deg, 45 deg, 90 deg, or 135 deg. Then, the normalized GLCMs  $p(i, j, d, \psi)$  are obtained by normalizing  $P(i, j, d, \psi)$  with the number of total pixel pairs. To study adjacent pixels, pixel pair distance  $d$  was set to 1, and four normalized GLCMs  $p(i, j, d, \psi)$  could be termed as  $p(i, j, \psi)$  with  $\psi = 0$  deg, 45 deg, 90 deg, and 135 deg. A total of 17 texture parameters (correlation COR, dissimilarity DIS, contrast CON, inverse difference moment IDM, entropy ENT, sum entropy SEN, difference entropy DEN, angular

second moment ASM, variance VAR, sum variance SVA, difference variance DVA, mean MEA, sum average SAV, cluster shade CLS, cluster prominence CLP,<sup>23</sup> maximum probability MAP, and minimum probability MIP) were extracted from each GLCM to characterize p-SI  $I_{kl}(y, z)$ , defined as follows:

$$\begin{aligned} \text{COR}_{kl,\psi} &= \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (ij) p(i, j, \psi) - \mu_x(\psi) \mu_y(\psi) \\ \text{DIS}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} |i - j| p(i, j, \psi) \\ \text{CON}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)^2 p(i, j, \psi) \\ \text{IDM}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i - j)^2} p(i, j, \psi) \\ \text{ENT}_{kl,\psi} &= - \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} p(i, j, \psi) \cdot \log[p(i, j, \psi)] \\ \text{SEN}_{kl,\psi} &= - \sum_{h=0}^{2G-2} p_{x+y}(h, \psi) \cdot \log[p_{x+y}(h, \psi)] \\ \text{DEN}_{kl,\psi} &= - \sum_{h=0}^{G-1} p_{x-y}(h, \psi) \cdot \log[p_{x-y}(h, \psi)] \\ \text{ASM}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [p(i, j, \psi)]^2 \\ \text{VAR}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [i - \mu(\psi)]^2 p(i, j, \psi) \\ \text{SVA}_{kl,\psi} &= \sum_{h=0}^{2G-2} (h - \text{SEN}_{kl,\psi})^2 p_{x+y}(h, \psi) \\ \text{DVA}_{kl,\psi} &= \frac{1}{G-1} \sum_{h=0}^{G-1} [p_{x-y}(h, \psi) - \bar{p}_{x-y}(\psi)]^2 \\ \text{MEA}_{kl,\psi} &= \sum_{i=0}^{G-1} i \sum_{j=0}^{G-1} p(i, j, \psi) = \mu_x(\psi) \\ \text{SAV}_{kl,\psi} &= \sum_{h=0}^{2G-2} h p_{x+y}(h, \psi) \\ \text{CLS}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [i + j - 2\mu(\psi)]^3 p(i, j, \psi) \\ \text{CLP}_{kl,\psi} &= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [i + j - 2\mu(\psi)]^4 p(i, j, \psi) \\ \text{MIP}_{kl,\psi} &= \min[p(i, j, \psi)] \\ \text{MAP}_{kl,\psi} &= \max[p(i, j, \psi)], \end{aligned} \quad (5)$$

where  $k$  indicates the polarization of scattered light from cells,  $l$  indicates the polarization of incident light,  $p(i, j, \psi)$  are the normalized GLCMs of p-SI  $I_{kl}(y, z)$ ,  $G$  is the maximum gray level of p-SI  $I_{kl}(y, z)$ ,  $H$  and  $W$  are the height and width of p-SI  $I_{kl}(y, z)$ , respectively, and



**Fig. 4** Reconstructed cell structures: (a) a section of the reconstructed cell morphology with eroded mitochondria by 3 pixels, (b) a section of the reconstructed cell morphology with original mitochondria, and (c) a section of the reconstructed cell morphology with dilated mitochondria by 3 pixels.

$$\begin{aligned}
 p_x(i, \psi) &= \sum_{j=0}^{G-1} p(i, j, \psi) \\
 p_y(j, \psi) &= \sum_{i=0}^{G-1} p(i, j, \psi) \\
 p_{x+y}(h, \psi) &= \sum_{i=0}^{G-1} \sum_{\substack{j=0 \\ i+j=h}}^{G-1} p(i, j, \psi) \\
 p_{x-y}(h, \psi) &= \sum_{i=0}^{G-1} \sum_{\substack{j=0 \\ |i-j|=h}}^{G-1} p(i, j, \psi) \\
 \mu_x(\psi) &= \sum_{i=0}^{G-1} i p_x(i, \psi) \\
 \mu_y(\psi) &= \sum_{j=0}^{G-1} j p_y(j, \psi) \\
 \sigma_x(\psi)^2 &= \sum_{i=0}^{G-1} [p_x(i, \psi) - \mu_x(\psi)]^2 \\
 \sigma_y(\psi)^2 &= \sum_{j=0}^{G-1} [p_y(j, \psi) - \mu_y(\psi)]^2 \\
 \mu(\psi) &= \mu_x(\psi) = \mu_y(\psi). \tag{6}
 \end{aligned}$$

To combine the four-directional characteristic of p-SI  $I_{kl}(y, z)$ , the average of these 17 parameters over four directions was calculated and denoted as  $COR_{kl}$ ,  $DIS_{kl}$ ,  $CON_{kl}$ ,  $IDM_{kl}$ ,  $ENT_{kl}$ ,  $SEN_{kl}$ ,  $DEN_{kl}$ ,  $ASM_{kl}$ ,  $VAR_{kl}$ ,  $SVA_{kl}$ ,  $DVA_{kl}$ ,  $MEA_{kl}$ ,  $SAV_{kl}$ ,  $CLS_{kl}$ ,  $CLP_{kl}$ ,  $MAP_{kl}$ , and  $MIP_{kl}$ . Together with three other parameters, strength  $STR_{kl}$ , mass-z  $MSZ_{kl}$ , and mass-y  $MSY_{kl}$ , a total of 20 parameters were used to characterize one p-SI

$$\begin{aligned}
 STR_{kl} &= \sum_{y=1}^H \sum_{z=1}^W I_{kl}(y, z), \\
 MSZ_{kl} &= \frac{1}{STR_{kl}} \sum_{y=1}^H \sum_{z=1}^W I_{kl}(y, z) \times z, \\
 MSY_{kl} &= \frac{1}{STR_{kl}} \sum_{y=1}^H \sum_{z=1}^W I_{kl}(y, z) \times y. \tag{7}
 \end{aligned}$$

### 3 Results

#### 3.1 Simulated p-SIs

Erosion and dilation of three pixels of the mitochondria regions inside the cell for each slice, including the interpolated ones, generated two new virtual cell models, denoted as  $PCS_{-3}$  and  $PCS_3$ , whereas the original one was denoted as  $PCS_0$ ; these models are shown in Figs. 4(a), 4(b), and 4(c), respectively. The morphology parameters of  $PCS_{-3}$ ,  $PCS_0$ , and  $PCS_3$  are shown in Table 2. Furthermore, three virtual cell models of a human prostate cancer cell termed PC (CRL-1435, ATCC) were also derived, which are denoted as  $PC3_{-3}$ ,  $PC3_0$ , and  $PC3_3$ .

To eliminate the influence of cell orientation, 26 cell orientations, which could be treated as 26 virtual cells, were employed in the three virtual cell models described above by rotating the original one by specific Euler angles, which were uniformly distributed over the  $4\pi$  solid angle range. Six sets of virtual cells with different mitochondria volume ratios were used to execute ADDA simulation. The actual RI of mitochondria could not be precisely measured. Several RI models have been tested for simulated images and these simulated images were compared with experimental images by GLCM parameters. Then, two of these RI models  $N_a$  and  $N_b$  that yielded simulated images similar to experimental images were used in the simulation process to study the influence of mitochondria RI for diffraction images. Two sets of diffraction images were simulated with the same cell structure and the two RI models  $N_a$  and  $N_b$ . The RI mean values and standard deviations for cell organelles of RI models  $N_a$  and  $N_b$  used by ADDA are shown in Table 3.

For each cell model, each orientation and each RI, six p-SIs  $I_{kl}(y, z)$  were simulated with different polarizations. Figure 5

**Table 2** Morphology parameters of cell models.

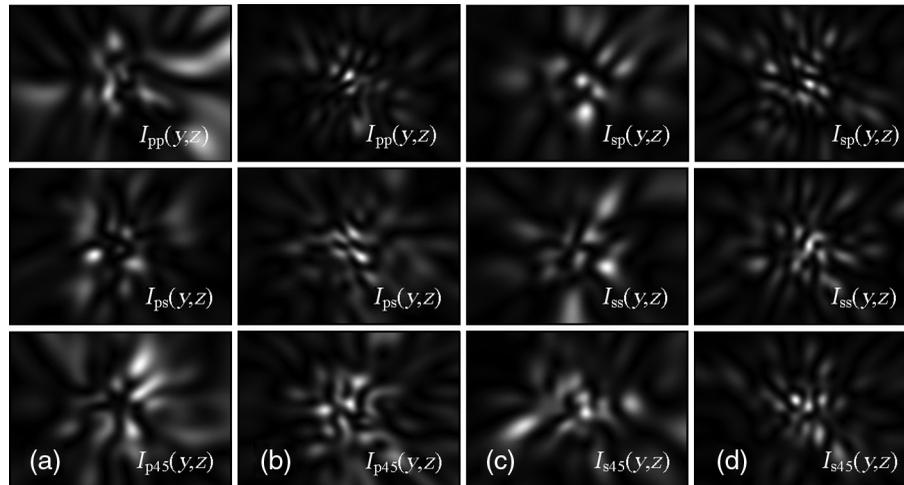
Cell model	$V_c^a$ ( $\mu m^3$ )	$V_m^a$ ( $\mu m^3$ )	$R_{mc}^a$ (%)
$PCS_{-3}$	1392	54	3.88
$PCS_0$		149	10.70
$PCS_3$		279	20.04

<sup>a</sup> $V_c$  is the volume of the cell,  $V_m$  is the volume of the mitochondria, and  $R_{mc}$  is the volume ratio of the mitochondria.

**Table 3** RI models of cell models.

RI	$n_c^a$	$n_{nl}^a$	$n_{nm}^a$	$n_{nh}^a$	$n_m^a$
$N_a$	$1.3675 \pm 0.0019$	$1.4183 \pm 0.0049$	$1.4692 \pm 0.0078$	$1.5200 \pm 0.0107$	$1.4200 \pm 0.0050$
$N_b$					$1.5534 \pm 0.0127$

<sup>a</sup> $n_c$ ,  $n_{nl}$ ,  $n_{nm}$ ,  $n_{nh}$ , and  $n_m$  are the mean value and standard deviation of the RIs in regions  $\Omega_c$ ,  $\Omega_{nl}$ ,  $\Omega_{nm}$ ,  $\Omega_{nh}$ , and  $\Omega_m$ , respectively.



**Fig. 5** p-SI of different cell models: columns (a), (b), (c), and (d) were p-SIs of original cell morphology with RI model  $N_a$ , original cell morphology with RI model  $N_b$ , eroded mitochondria cell morphology with RI model  $N_b$ , and dilated mitochondria cell morphology with RI model  $N_b$ , respectively.

shows some simulated images with different cell structures, orientations, and RI models.

### 3.2 Classification Results for Simulated p-SIs Using CT and GLCM

For one p-SI pair with the same incident polarization and two scattered polarizations as described above, 82 subimages and four types of parameters were calculated for one subimage, so a total of 328 parameters were extracted to characterize one p-SI pair. Because too many parameters for just one p-SI pair consume a large amount of computational time, they were divided into four sets  $C$ ,  $E$ ,  $F$ , and  $V$  of 82 parameters to be analyzed separately. Simulated p-SIs with the same RI of  $N_b$ ; different cell models—PCS<sub>0</sub> versus PCS<sub>3</sub>, PCS<sub>-3</sub> versus PCS<sub>3</sub> and PCS<sub>3</sub> versus PCS<sub>3</sub>—and the same cell model PCS<sub>0</sub>; and different RIs— $N_a$  versus  $N_b$ —illuminated by a 532 nm wavelength laser with polarizations  $p$ ,  $s$ , or 45 deg were classified by an SVM algorithm<sup>24,25</sup> with four different kernel functions: linear (Lin), polynomial (Pol), sigmoid (Sig), and radial basis functions (Rbf). The sample size for every simulated p-SI set was 26. Given the small sample size of simulated diffraction images, all samples were used as training sets and no test process was performed. The goal of this study was to test the ability of CT and GLCM for characterizing cell diffraction images. Thus, the SVM was used only to achieve classification accuracy.

The brief classification results are shown in Table 4. All four sets of classification could achieve very high classification

accuracy  $A$ , greater than 98%, which was the ratio of the correctly classified p-SI number to the total p-SI number. This result shows that  $E$  parameters could result in better performance with a larger  $A$  for each set than the other parameters. The three highest classification accuracy levels indicated by larger  $A$  values with a single  $E$  parameter corresponding to the classification results in Table 4 with  $E$  parameters are shown in Fig. 6.

With a specific incident polarization  $k$ , a total of 40 GLCM features were applied to characterize one p-SI pair described above. The same four sets of p-SIs were classified by SVM with four of kernel functions. All four sets of classification results could also achieve very high classification accuracy  $A$ , greater than 98%, as shown in Table 5. The three highest classification accuracy levels indicated by larger  $A$  values with a single GLCM parameter corresponding to the results in Table 5 are shown in Fig. 6.

### 3.3 Classification Results for Experimental p-SIs Using CT and GLCM

Three sets of experimental PCS and PC3 cell images were also analyzed and classified as simulated images. The experimental cell image sample size is shown in Table 6. As with the simulated images, all samples are used as training sets.

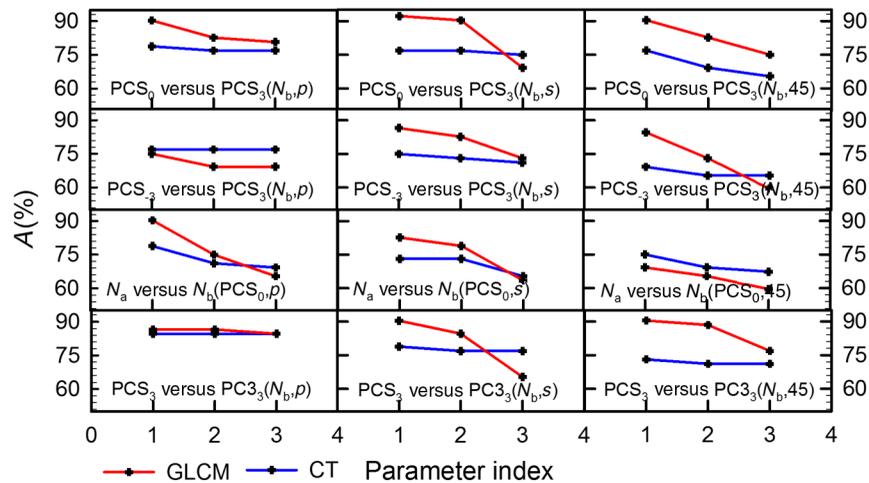
Tables 7 and 8 show the classification results of experimental cell images with CT and GLCM features. Both CT and GLCM algorithms can also achieve high classification accuracy with experimental cell images. As with the simulated results,

**Table 4** The SVM classification accuracy of different cell models or RIs with CT parameters.

Cell models	RI	$P^a$	$(A, N_p, \text{kernel})-p^b$	$(A, N_p, \text{kernel})-s^b$	$(A, N_p, \text{kernel})-45 \text{ deg}^b$
PCS <sub>0</sub> versus PCS <sub>3</sub>	$N_b$	$C$	100%, 50, Lin	96.15%, 23, Lin	100%, 8, Lin
		$E$	98.08%, 39, Lin	100%, 22, Lin	98.08%, 16, Lin
		$F$	88.46%, 22, Sig	86.54%, 28, Lin	84.62%, 7, Lin
		$V$	94.23%, 3, Lin	90.38%, 33, Sig	90.38%, 8, Lin
PCS <sub>-3</sub> versus PCS <sub>3</sub>	$N_b$	$C$	98.08%, 7, Lin	96.15%, 6, Sig	98.08%, 3, Lin
		$E$	98.08%, 63, Lin	96.15%, 13, Sig	94.23%, 20, Lin
		$F$	88.46%, 44, Rbf	88.46%, 38, Lin	78.85%, 14, Lin
		$V$	96.15%, 60, Pol	90.38%, 15, Rbf	88.46%, 24, Pol
PCS <sub>0</sub>	$N_a$ versus $N_b$	$C$	88.46%, 11, Lin	84.62%, 28, Lin	90.38%, 17, Lin
		$E$	98.08%, 20, Lin	94.23%, 77, Lin	96.15%, 43, Lin
		$F$	76.92%, 32, Lin	78.85%, 7, Lin	76.92%, 24, Lin
		$V$	88.46%, 31, Lin	82.69%, 14, Sig	80.77%, 8, Sig
PCS <sub>3</sub> versus PC3 <sub>3</sub>	$N_b$	$C$	100%, 3, Lin	100%, 5, Lin	100%, 4, Lin
		$E$	98.08%, 18, Lin	96.15%, 7, Lin	98.08%, 9, Rbf
		$F$	92.31%, 25, Lin	88.46%, 48, Lin	82.69%, 12, Lin
		$V$	100%, 14, Lin	94.23%, 17, Lin	96.15%, 9, Sig

<sup>a</sup> $P$  indicates the CT parameters used for SVM  $C, E, F,$  or  $V$ .

<sup>b</sup>The classification accuracy  $A$  of incident polarization  $p, s,$  and  $45 \text{ deg}$  with parameter numbers  $N_p$  and the kernel function used in SVM.



**Fig. 6** The three highest classification accuracy levels indicated by larger  $A$  values with a single  $E$  parameter corresponding to the classification results in Table 4 with  $E$  parameters in blue lines and single GLCM parameter corresponding to Table 5 in red lines.

$E$  parameters resulted in better performance with a larger  $A$  than the other parameters. The three highest classification accuracy levels indicated by larger  $A$  values with a single  $E$  parameter corresponding to the classification results in Table 7 with  $E$  parameters and a single GLCM parameter corresponding to the results in Table 8 are shown in Fig. 7.

#### 4 Discussion

The actual variation of mitochondria volume change by erosion or dilation was obtained by replacing the mitochondria RIs of eroded regions with cytoplasm RIs or setting the dilated regions with mitochondria RIs instead of cytoplasm RIs. Thus, volume change essentially reflects local RI change. Visually, it is clear

**Table 5** The SVM classification accuracy of different cell models or RIs with GLCM parameters.

Cell model	RI	(A, N <sub>p</sub> , kernel)-p <sup>a</sup>	(A, N <sub>p</sub> , kernel)-s <sup>a</sup>	(A, N <sub>p</sub> , kernel)-45 deg <sup>a</sup>
PCS <sub>0</sub> versus PCS <sub>3</sub>	N <sub>b</sub>	98.08%, 2, Lin	100%, 32, Rbf	100%, 26, Lin
PCS <sub>-3</sub> versus PCS <sub>3</sub>	N <sub>b</sub>	96.15%, 34, Pol	96.15%, 19, Lin	98.08%, 25, Pol
PCS <sub>0</sub>	N <sub>a</sub> versus N <sub>b</sub>	98.08%, 25, Lin	94.23%, 2, Rbf	96.15%, 38, Pol
PCS <sub>3</sub> versus PC3 <sub>3</sub>	N <sub>b</sub>	98.08%, 30, Lin	98.08%, 2, Pol	96.15%, 3, Rbf

<sup>a</sup>The classification accuracy A of incident polarization p, s, and 45 deg with parameter numbers N<sub>p</sub> and the kernel function used in SVM.

that a change in mitochondria volume or RI greatly influences the p-SIs. The diffraction images have larger spots with larger RIs or more mitochondria as shown in Fig. 5. Therefore, changes in cell RIs can be studied by analyzing corresponding p-SIs. In this report, two algorithms, CT and GLCM, were applied to analyze the simulated p-SIs, which have different mitochondria volumes or different mitochondria RIs, and experimental cell images. The results showed that both CT and GLCM algorithms could precisely classify the same simulated and experimental p-SI sets with high accuracy, demonstrating that both parameters extracted by these two methods can be used to characterize and analyze p-SIs.

CT yielded more parameters than GLCM for a single p-SI and GLCM achieved larger classification accuracy A with a single parameter as compared with the CT method, indicating that the parameters extracted by GLCM contain more information and are more efficient to describe both the simulated and experimental p-SI as shown in Figs. 6 and 7. CT has disintegrated the original p-SI into 41 subimages, which generate only very few features of the original p-SI in one subimage. Therefore, even though CT extracts many subimages and parameters, this process did not significantly improve the performance of p-SI

analysis. A previous study also showed that the CT subimages carry more information at coarse scales  $\gamma \geq 3$  of p-SI patterns than the fine scales  $\gamma = 1$  or 2, and the former perform better than the latter in SVM classification. Meanwhile, a long period of time is needed for the calculation of these subimages and parameters.<sup>14</sup> CT consumed about 10 times computational cost than GLCM for both feature calculation and SVM classification. Therefore, the GLCM algorithm should be more suitable and efficient for analysis of cell p-SIs. GLCM coupled with cell-scattered images establishes a new rapid method that will benefit cell basic research including cell assay and cell structure studies.

Although diffraction images with different mitochondria RIs or volumes can be classified by the SVM algorithm, it is only suitable for binary situations. By using SVM multiple times and establishing a database with various types of cells, diffraction images provide opportunities for the development of fast and label-free cell assay methods in the future.

Since cell diffraction images mainly contain relatively low frequency information, and many subimages and parameters obtained by CT reflect relative high frequency information of the analyzed images, CT should be more appropriate for analyzing images with rich frequency information.

**Table 6** Experimental sample size of cell images used for SVM.

Cell model	Sample size		
	p	s	45 deg
PCS	602	407	390
PC3	731	733	794

**Table 8** The SVM classification accuracy of experimental PCS and PC3 cells with GLCM parameters.

Cell model	(A, N <sub>p</sub> , kernel)-p <sup>a</sup>	(A, N <sub>p</sub> , kernel)-s <sup>a</sup>	(A, N <sub>p</sub> , kernel)-45 deg <sup>a</sup>
PCS versus PC3	100%, 40, Lin	98.68%, 19, Lin	88.68%, 29, Pol

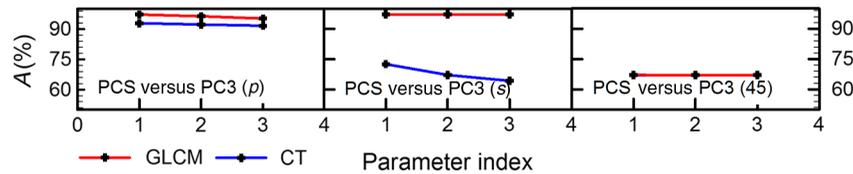
<sup>a</sup>The classification accuracy A of incident polarization p, s, and 45 deg with parameter numbers N<sub>p</sub> and the kernel function used in SVM.

**Table 7** The SVM classification accuracy of experimental PCS and PC3 cells with CT parameters.

Cell type	P <sup>a</sup>	(A, N <sub>p</sub> , kernel)-P <sup>b</sup>	(A, N <sub>p</sub> , kernel)-s <sup>b</sup>	(A, N <sub>p</sub> , kernel)-45 deg <sup>b</sup>
PCS versus PC3	C	97.97%, 79, Lin	96.23%, 56, Lin	78.80%, 79, Lin
	E	98.50%, 78, Lin	97.81%, 69, Pol	86.15%, 82, Lin
	F	97.45%, 79, Lin	93.60%, 36, Lin	80.15%, 61, Lin
	V	97.75%, 75, Lin	95.09%, 45, Lin	78.12%, 54, Lin

<sup>a</sup>P indicates the CT parameters used for SVM C, E, F, or V.

<sup>b</sup>The classification accuracy A of incident polarization p, s, and 45 deg with parameter numbers N<sub>p</sub> and the kernel function used in SVM.



**Fig. 7** The three highest classification accuracy levels indicated by larger  $A$  values with a single  $E$  parameter corresponding to the classification results in Table 7 with  $E$  parameters in blue lines and single GLCM parameter corresponding to Table 8 in red lines.

## 5 Summary

This paper used virtual cell models reconstructed and interpolated from confocal image stacks stained by two fluorescence dyes to simulate polarization-scattered images with different mitochondria volumes modified by erosion and dilation or different mitochondria RIs. These simulated images were then analyzed with two algorithms, CT and GLCM. The results show that GLCM can carry more information than CT with the same number of parameters. Analysis of experimental images yields similar results as the simulated situation. GLCM is a more suitable and efficient method for cell-scattered image analysis than CT.

## Acknowledgments

The authors acknowledge financial support from the National Natural Science Foundation of China (Grant Nos. 81171342 and 81201148).

## References

- H. M. McBride, M. Neuspiel, and S. Wasiak, "Mitochondria: more than just a powerhouse," *Curr. Biol.* **16**(14), R551–R560 (2006).
- T. Kim et al., "White-light diffraction tomography of unlabelled live cells," *Nat. Photonics* **8**(3), 256–263 (2014).
- O. C. Marina, C. K. Sanders, and J. R. Mourant, "Correlating light scattering with internal cellular structures," *Biomed. Opt. Express* **3**(2), 296–312 (2012).
- J. R. Mourant et al., "Light scattering from cells: the contribution of the nucleus and the effects of proliferative status," *J. Biomed. Opt.* **5**(2), 131–137 (2000).
- K. M. Jacobs, J. Q. Lu, and X. H. Hu, "Development of a diffraction imaging flow cytometer," *Opt. Lett.* **34**(19), 2985–2987 (2009).
- K. M. Jacobs et al., "Diffraction imaging of spheres and melanoma cells with a microscope objective," *J. Biophotonics* **2**(8–9), 521–527 (2009).
- Y. Sa et al., "Study of low speed flow cytometry for diffraction imaging with different chamber and nozzle designs," *Cytometry* **83A**(11), 1027–1033 (2013).
- Y. M. Feng et al., "Polarization imaging and classification of Jurkat T and Ramos B cells using a flow cytometer," *Cytometry* **85A**(9), 817–826 (2014).
- R. S. Brock et al., "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**(1), 25–36 (2006).
- R. Pan et al., "Analysis of diffraction imaging in non-conjugate configurations," *Opt. Express* **22**(25), 31568–31574 (2014).
- M. A. Yurkin and A. G. Hoekstra, "The discrete dipole approximation: an overview and recent developments," *J. Quant. Spectrosc. Radiat. Transfer* **106**(1–3), 558–589 (2007).
- Y. Zhang et al., "Comparative study of 3D morphology and functions on genetically engineered mouse melanoma cells," *Integr. Biol.* **4**(11), 1428–1436 (2012).
- W. Jiang et al., "Comparison study of distinguishing cancerous and normal prostate epithelial cells by confocal and polarization diffraction imaging," *J. Biomed. Opt.* **21**(7), 071102 (2015).
- J. Zhang et al., "Realistic optical cell modeling and diffraction imaging simulation for study of optical and morphological parameters of nucleus," *Opt. Express* **24**(1), 366–377 (2016).
- M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Image Process.* **14**(12), 2091–2106 (2005).
- Y. Lu and M. N. Do, "Multidimensional directional filter banks and surfacelets," *IEEE Trans. Image Process.* **16**(4), 918–931 (2007).
- D. D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Trans. Image Process.* **15**(6), 1610–1620 (2006).
- L. Yang, B. L. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing* **72**(1–3), 203–211 (2008).
- R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE* **67**(5), 786–804 (1979).
- R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.* **SMC-3**(6), 610–621 (1973).
- M. A. Yurkin and A. G. Hoekstra, "The discrete-dipole-approximation code ADDA: capabilities and known limitations," *J. Quant. Spectrosc. Radiat. Transfer* **112**(13), 2234–2247 (2011).
- H. Wang et al., "Acquisition of cross-polarized diffraction images and study of blurring effect by one time-delay-integration camera," *Appl. Opt.* **54**(16), 5223–5228 (2015).
- L. K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Trans. Geoscience Remote Sens.* **37**(2), 780–795 (1999).
- C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discovery* **2**(2), 121–167 (1998).
- C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Syst. Technol.* **2**(3), 1–27 (2011).

**Jun Zhang** received his BS degree in biomedical engineering from Tianjin University in 2010 and he is a PhD candidate of biomedical engineering at Tianjin University. His current research interests include scattered image analysis, cell morphology, and image processing.

**Gang Wang** received his BS degree in biomedical engineering from Tianjin University in 2013 and he is a graduate student of biomedical engineering at Tianjin University. His current research interests include image processing and analysis.

**Yuanming Feng** is a professor of biomedical engineering at Tianjin University. His research interests include diffraction imaging flow cytometry and measurement technique of radiation-induced apoptosis.

**Yu Sa** received his BS and PhD degrees from Tianjin University and he is a lecturer in the Department of Biomedical Engineering at Tianjin University. His research interests include instrument development, diffraction imaging, and computational fluid dynamics modeling studies.