J. Biophotonics **10**, 634–644 (2017) / **DOI** 10.1002/jbio.201500338

Journal of BIOPHOTONICS

FULL ARTICLE

Development of a multispectral light-scatter sensor for bacterial colonies

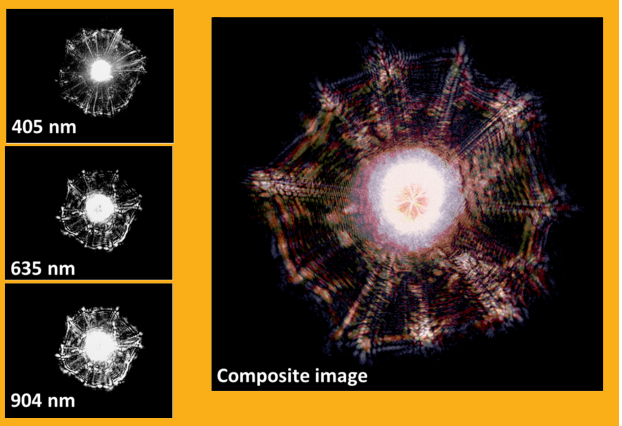
Huisung Kim[*, 1], *Bartek Rajwa*[2, 3], *Arun K. Bhunia*[4], *J. Paul Robinson*[2, 5], *and Euiwon Bae*[1]

[1] Applied Optics Laboratory, School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA
[2] Department of Basic Medical Sciences, Purdue University, West Lafayette, IN 47907, USA
[3] Bindley Bioscience Center, Purdue University, West Lafayette, IN 47907, USA
[4] Molecular Food Microbiology Laboratory, Department of Food Science, Purdue University, West Lafayette, IN 47907, USA
[5] Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907, USA

Received 20 December 2015, revised 16 May 2016, accepted 1 June 2016
Published online 14 July 2016

Key words: bacteria identification, diffraction, light scattering, multi wavelength, spectral measurement

We report a multispectral elastic-light-scatter instrument that can simultaneously detect three-wavelength scatter patterns and associated optical densities from individual bacterial colonies, overcoming the limits of the single-wavelength predecessor. Absorption measurements on liquid bacterial samples revealed that the spectroscopic information can indeed contribute to sample differentiability. New optical components, including a pellicle beam splitter and an optical cage system, were utilized for robust acquisition of multispectral images. Four different genera and seven shiga toxin producing *E. coli* serovars were analyzed; the acquired images showed differences in scattering characteristics among the tested organisms. In addition, colony-based spectral optical-density information was also collected. The optical model, which was developed using diffraction theory, correctly predicted wavelength-related differences in scatter patterns, and was matched with the experimental results. Scatter-pattern classification was performed using pseudo-Zernike (GPZ) polynomials/moments by combining the features collected at all three wavelengths and selecting the best features via a random-forest method. The data demonstrate that the selected features provide better classification rates than the same number of features from any single wavelength.



Three wavelength-merged scatter pattern from *E. coli*.

1. Introduction

Rapid identification and classification of microbial organisms are critical tasks in various areas such as bio-surveillance, biosecurity, clinical studies, and food safety. A series of recent foodborne disease outbreaks once again demonstrates the need for more reliable, accurate, and rapid analytical methods for detection and monitoring of pathogenic microorganisms such as *Escherichia coli*, *Listeria monocytogenes*, *Salmonella enterica*, *and Staphylococcus* spp. [1]. Compared to conventional detection methods,

* Corresponding author: e-mail: ebae@purdue.edu, Phone: +1-765-496-1125, Fax: +1-765-494-0539

label-free optical diagnostics not only delivers results in a fast, cost-effective, and less labor-intensive manner, but also provides accurate and non-destructive evaluation of the samples, allowing secondary confirmation with further verification. Owing to the wide range of the spectral region that is available for optical diagnostics, spectral imaging has been widely utilized in biomedical applications [2], and food quality and safety control [3–6]. In the area of applications related to food inspection, numerous applications of hyperspectral imaging to classify the quality of harvested vegetables [7], fruits [8], meats [9], and poultry [10] have been reported.

The cited works employing spectral techniques rely on standard far-field imaging. However, cells and bacterial colonies are three-dimensional objects; optical interrogation of their whole volume can provide better classification accuracy. A label-free, non-destructive, automated detection technique called BARDOT (BActerial Rapid Detection using Optical scattering Technology) based on elastic-light-scatter (ELS) patterns of bacteria colonies was developed for rapid detection and classification of microbial organisms [11–13]. The applicability of the technology were reported for different organisms using a single-wavelength laser and a varying number of genera or species [14–23]. The merit of this technology is that the interrogation photons interact with the whole volume of a colony, thus collecting better phenotypic characteristics than reflective imaging. One inherent limitation of the current BARDOT instrument is that the classification performance suffers either when large numbers of species and strains are analyzed simultaneously or lower taxonomy of bacterial colonies were classified. Recently, we introduced scalar diffraction modeling of multispectral forward-scattering patterns for bacterial colonies [24] and verified the feasibility of the multispectral approach on bacterial phenotyping. Here we report a new design and validation of a multispectral forward-scatter phenotyping instrument called MS-BARDOT that combines multiple diode lasers and an optical density (OD) unit with the conventional BARDOT system. The benefit of this design is that both the spectral forward-scattering patterns and the OD of a bacteria colony can be measured. To maximize the classification efficiency, both a random-forest (RF) method for proper feature selection and pseudo-Zernike (GPZ) polynomials/moments are utilized for the multi-spectral information analysis. Four representative bacterial genera were measured and analyzed with the proposed instrument and algorithm.

2. Materials and methods

2.1 Design of MS-BARDOT

As shown in Figure 1, the MS-BARDOT system consists of three major components: a multispectral forward scatterometer, a sequence controller, and a two-dimensional lateral stages. These components were designed with three cage-type R45 : T55 pellicle beam splitters (Thorlabs Inc., Newton, NJ, USA). Two pellicle beam splitters were positioned above, and the third one was positioned below the sample.

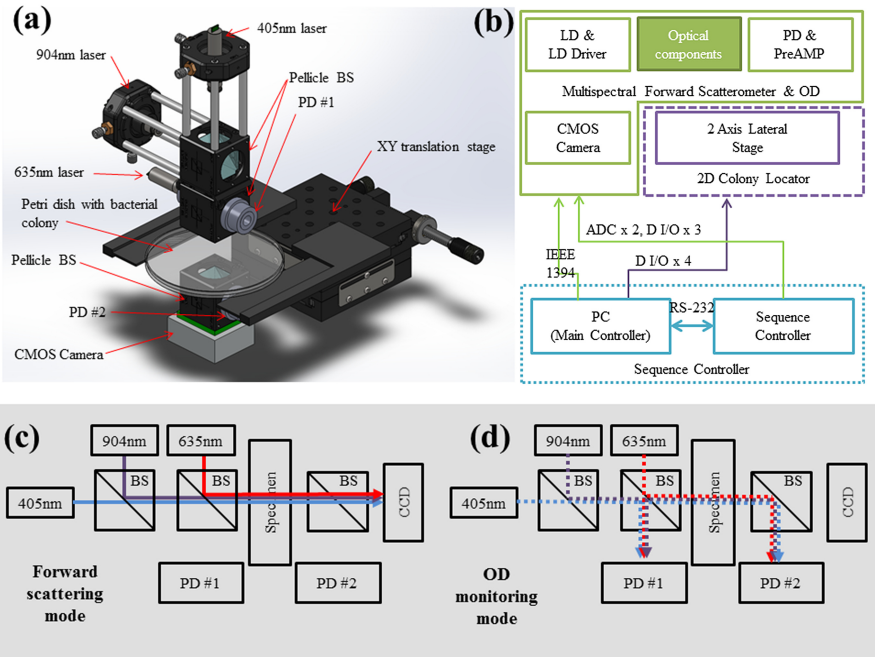


Figure 1 Schematic diagram of multispectral BARDOT with spectral OD measurement functionality. (**a**) Overall design of the proposed instrument. (**b**) Schematic block diagram. (**c**) Light path for forward-scattering mode, (**d**) Light path for OD-monitoring mode. Three different wavelength LDs and two PDs were integrated into the BARDOT system to simultaneously measure the spectral forward-scattering pattern and the OD of a bacterial colony on a semi-solid medium.

One-mW collimated beams from 405 nm, 635 nm, and 904 nm aser diode modules (Coherent Inc., Santa Clara, CA, Lasermate Group Inc., Walnut, CA, USA) were selected as light sources. Each module is attached to the port of the beam-splitter cage unit as shown in Figure 1(a). The selection of individual wavelengths was based on the spectral absorption of four different bacterial genera and the spectral availability of a commercial diode laser (Section 3.1). With a stacked beam-splitter structure, each multispectral forward-scattering pattern and OD from a single colony can be captured simultaneously. Consequently, all the multispectral forward-scattering patterns and ODs can be captured in less than 5 seconds per colony. For pattern capture, a monochromatic CMOS camera (PL-B741, Pixelink, Ottawa, ON, Canada) with 1280(H) × 1024(V) pixels and 6.7 μm pixel size was positioned under the Petri dish at a distance of 39 mm, measured from the bottom surface of the dish to the surface of the image sensor.

For OD measurement, a pellicle beam splitter was positioned between the Petri dish and the CMOS camera (Figure 1). Two Si photodiodes (PDs) (Thorlabs Inc.) with active wavelength from 400 nm to 900 nm were mounted at each port of the pellicle beam-splitter cages. The PD attached to the middle beam-splitter cage (PD #1, in Figure 1(c)) monitored the intensity of incident light, while the PD integrated with the bottom beam-splitter cage (PD #2, in Figure 1(c)) measured transmitted light from a sample. Using the light-intensity information, the OD of the sample was computed for each wavelength. Figure 1(c, d) depicts beam paths from each light source to the sensor for forward-scatter and OD measurement mode respectively.

2.2 Spectral intensity calibration

Since the pellicle beam splitters, PDs, and CMOS sensors have inherent optical response characteristics, the light intensity and sensor sensitivity were calibrated optically and electrically to maintain similar input intensities for each wavelength. As shown in Figure 2(a), the experimental reflectance and transmittance ratios for the R45:T55 pellicle beam splitter were R56.8 : T43.1, R44.3 : T55.7, and R40.9 : T59.0 for 405 nm, 635 nm, and 904 nm, respectively. For the detailed calibration routine, see Figure S1. Considering the quantum efficiency of each sensor and the attenuation ratio of the beam-splitter unit, a spectral intensity-compensation factor for each sensor was computed and applied to the system, as shown in Figure 2(b) (see supplementary section and Figure S2 for gain calibration of the actual scatter patterns).

2.3 Spectral OD measurement

E. coli O157 : H7 EDL933, *L. monocytogenes* F4244, *S.* Enteritidis PT21 and *S. aureus* ATCC 25923 were selected as model organisms. For the agar plate preparation, all cultures were grown in 5 ml brain heart

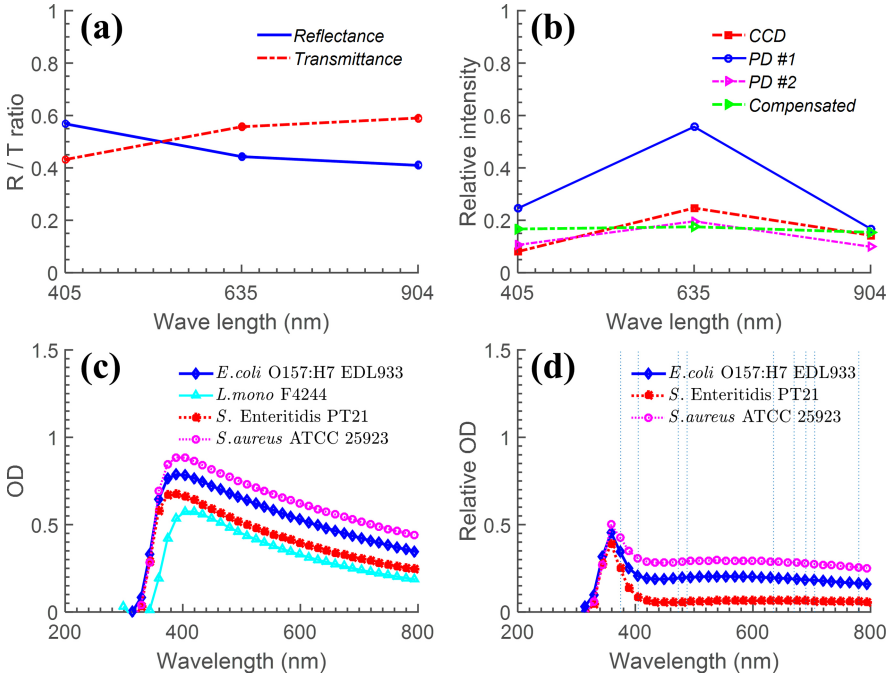


Figure 2 Wavelength-resolved light-intensity calibration and compensation. (**a**) Experimental result of reflectance and transmittance ratios for pellicle beam splitter for each wavelength. (**b**) Wavelength-resolved overall attenuation ratio by the integrated beam-splitter unit for each sensor and compensation result. Spectroscopic absorption results for four representative bacterial samples on liquid BHI stock. (**c**) ODs of interrogated genera from 300–800 nm. (**d**) Relative ODs of interrogated genera to that of *L. mono*, which had the lowest OD among the genera at a given wavelength. Solid lines represent the available laser lines as a module.

infusion (BHI, BD Diagnostics, Sparks, MD, USA) broth for 15 h at 37 °C and 130 rpm on an incubator shaker. The cultures were then serially diluted and surface plated on BHI agar plates (100 mm × 15 mm) to achieve bacterial counts of 50–100 CFU/plate. The plates were incubated at 37 °C until the colonies reached a diameter of 900–1100 μm. The colony diameters were measured using both a bright-field microscope equipped with a Leica DFC310 FX CCD camera, and Leica Application Suite V4.20 build 607 (all from Leica Microsystems, Bannockburn, IL, USA) using a 10× objective, and a BARDOT instrument. Colonies of *E. coli*, *Listeria*, *Salmonella*, and *S. aureus* were grown for 10.5 h, 22.5 h, 11.5 h, and 13.5 h, respectively, to obtain diameters of 1000 μm. The thickness of the agar for each plate was maintained at approximately 8 mm.

For the liquid-sample preparation, a pure colony of each genus was harvested and diluted in a single tube, and incubated for 12 h at 37 °C. Then, sample aliquots were transferred to disposable cuvettes, and each stock was serially diluted 3 times at a 1 : 10 ratio. ODs of the diluted samples were measured at 300–800 nm with a DU 800 spectrophotometer (Beckman Coulter Inc., Brea, CA, USA). Spectral absorption curves were recorded for 3 different replicates with 5 samples (a total of 15 data sets) for each bacterium, and average spectral response curves were calculated. For quantitative comparison, the area under the curve was calculated and used for normalization.

For the solid-sample experiments, 5 plates per organism were prepared for a single-day data set, and repeated on three different days in order to accommodate the biological variability. At least 20 random locations of bare BHI and 20 colonies were interrogated per plate. The mean value of the 20 data points was considered as the representative transmittance, and the results were computed as spectral OD. The spectral OD of BHI is defined as

$$\mathrm{OD}_{\mathrm{agar}}(\lambda) = -\log_{10}\left(\frac{I(\lambda)_{\mathrm{agar}}}{I(\lambda)_{\mathrm{input}}}\right) \tag{1}$$

The mean value of each agar OD was computed as 0.503 nm, 0.129 nm, 0.072 nm for 405 nm, 635 nm, and 904 nm, respectively. Since it is challenging to measure the actual OD of a bacterial colony without destroying the colony structure on semi-solid agar, we used an indirect method to obtain the colony ODs by subtracting the OD of the pure agar from that of colony and agar (see Figure S2(c)):

$$\mathrm{OD}_{\mathrm{colony}}(\lambda) = -\log_{10}\left(\frac{I(\lambda)_{\mathrm{agar+colony}}}{I(\lambda)_{\mathrm{input}}}\right) - \mathrm{OD}_{\mathrm{agar}}(\lambda) \tag{2}$$

For the serovar level classification, seven serovars of shiga toxin producing *E. coli* (*E. coli* O26, O45, O103, O111, O121, O145, and O157) with three stains per serovars were measured. The cultures were grown in BHI broth for 14 h at 37 °C in incubator shaker with 130 rpm. After the incubating, the cultures are serially diluted and surface plated on both BHI and SMAC agar plate (100 mm × 15 mm) to achieve a bacterial counts of 50–100 CFU/plate. The plates are incubated at 37 °C until the size of the colonies reached to diameter range of 700 ∼ 1100 μm. The diameters of the bacterial colonies are measured using a bright-field microscope (Leica Microsystems, Bannockburn, Illinois, USA) equipped with Spot software (Sterling Height, MI) using a 10× objective, and BARDOT. 10.5–11 h are taken for 1000 μm colony diameter on both of the nutrient agar plates. Utilizing the multispectral BARDOT, multispectral forward scattering patterns of the 50–60 colonies are measures for the each strain as a day duplication. Three different date duplications are collected for the experiment.

2.4 Modeling of spectral scattering

From our previous report [24], the multispectral forward-scattering pattern of a bacterial colony is modeled as

$$E_i(x_i, y_i, \lambda) = C \iint T(x_a, y_a, \lambda)\, \exp\left[i\Phi_{\mathrm{overall}}(x_a, y_a, \lambda)\right] \times \exp\left[-2\pi\, i(f_x(\lambda)\, x_a + f_y(\lambda)\, y_a)\right] \mathrm{d}x_a\, \mathrm{d}y_a \tag{3}$$

This model includes two major components: an amplitude modulator, $T(x_i, y_i, \lambda)$, and a phase modulator, Φ_{overall}, which consists of Φ_c, Φ_q, and Φ_r, defined as colony, quadratic, and Gaussian phase component, respectively. The maximum diffraction angle and the number of rings are also important parameters and are defined as

$$\theta(\lambda)/2_{\max} = \frac{1}{k}\left(\frac{d\ \Delta\Phi_{\mathrm{overall}}(\lambda)}{\mathrm{d}r}\right)_{\max} \tag{4}$$

$$N_{\mathrm{ring}}(\lambda) = \frac{\Delta\Phi_{\mathrm{overall}}(\lambda)_{\max}}{2\pi} \tag{5}$$

2.5 Spectral data analysis

The feature extraction/recognition of scatter patterns was performed using pseudo-Zernike (GPZ) polynomials/moments as described previously for a mono-

chromatic version of the method [12, 25]. The GPZ polynomials are formally defined as

$$k^{\alpha}_{p\lambda}(z) = z^{\frac{p+\lambda}{2}}(z^*)^{\frac{p-\lambda}{2}}\frac{(\alpha+1)_{p-|\lambda|}}{(p-|\lambda|)\,!_2} \times F_1\left(-p+|\lambda|,-p-|\lambda|-1;\ \alpha+1;\ 1-\frac{1}{(zz^*)^{1/2}}\right) \tag{6}$$

where * denotes the complex conjugate and $z = r\,e^{j\theta}$. The parameter α is user-tunable and scales the polynomial values. The repetition λ is set between 0 and p.

The polynomial is defined in polar coordinates as

$$k^{\alpha}_{p\lambda}(r,\theta) = k^{\alpha}_{p\lambda}(r\,e^{j\theta}) = R^{\theta}_{p\lambda}(r)\,e^{j\lambda\theta}$$

where the real-value radial polynomial $R^{\alpha}_{p\lambda}(r)$ is given by

$$R^{\alpha}_{p\lambda}(r) = \frac{(p+|\lambda|+1)!}{(\alpha+1)_{p+|\lambda|+1}}\sum_{s=0}^{p-|\lambda|} \times \frac{(-1)^s(\alpha+1)_{2p+1-s}}{s!(p-|\lambda|-s)\,!(p+|\lambda|+1-s)\,!}\,r^{p-s} \tag{7}$$

The radial polynomial $R^{\alpha}_{p\lambda}(r)$ is computed using the recurrence relation

$$R^{\alpha}_{p\lambda}(r) = (M_1 r + M_2)\,R^{\alpha}_{p-1,\lambda}(r) + M_3 R^{\alpha}_{p-2,\lambda}(r)\,, \tag{8}$$

with the following:

$$M_1 = \frac{(2p+1+\alpha)\,(2p+\alpha)}{(p+\lambda+1+\alpha)\,(p-\lambda)}$$

$$M_2 = \frac{(p+\lambda+1)\,(\alpha+2p)}{p+\lambda+\alpha+1} + M_1\frac{(p+\lambda)\,(p-\lambda-1)}{2p-1+\alpha}$$

$$M_3 = \frac{(p+\lambda)\,(p+\lambda+1)\,(2p-2+\alpha)\,(2p-1+\alpha)}{2(p+\lambda+\alpha+1)\,(p+\lambda+\alpha)} + M_2\frac{(p+\lambda)\,(2p-2+\alpha)}{p+\lambda+\alpha} - M_1\frac{(p+\lambda)\,(p+\lambda-1)\,(p-\lambda-2)}{2(p+\lambda+\alpha)}$$

and

$$R^{\alpha}_{\lambda\lambda}(r) = r^{\lambda}$$

$$R^{\alpha}_{\lambda+1,\lambda}(r) = [(\alpha+3+2\lambda)\,r - 2(\lambda+1)]\,R^{\alpha}_{\lambda\lambda}(r)$$

3. Experiments and results

3.1 Spectral absorption from liquid samples

Figure 2(c) shows spectral-absorption measurements of the four genera in liquid BHI stock. All the genera showed peak OD values near 400 nm, and ODs gradually decreased as the wavelength increased. *S. aureus* had the highest OD value and *L. mono* had the lowest among the tested genera. Figure 2(d) shows the relative OD values of the interrogated bacteria with respect to *L. mono*. The vertical dashed lines represent the commercially available wavelengths of the laser diode (LD). The 405 nm, 635 nm, and 904 nm laser lines were selected on the basis of line separation that maximized OD differences among the interrogated bacteria. This result confirms the observation of other researchers postulating that the OD information can be used as a simple classification method for bacteria at the genus level [26].

3.2 Spectral OD of bacteria colonies on semi-solid BHI agar

Figure 3(a) shows the measurement points. The spectral absorption from pure agar areas displayed characteristics similar to those of the liquid samples (Figure 3(b)). At longer wavelengths, the OD of the BHI agar decreased. Meanwhile, the net OD from bacterial colonies showed less reduction as the wavelength increased. Furthermore, for *S. aureus* the net OD for 635 nm showed a peak value of 0.38, while that of *S.* Enteritidis was 0.22 (Figure 3(c)). The ODs at 904 nm showed the minimum ODs for all genera. In contrast to the liquid-sample result where *L. mono* showed the lowest OD, *S.* Enteritidis produced the lowest ODs at 635 nm and 904 nm by Eqs. (1), (2). Although single-wavelength OD values provide limited differentiability among genera, ODs determined at three wavelengths can be utilized as the first-step classification method.

To better estimate the difference, spectral OD differences were calculated (Figure 3(d)). The X-axis displays three combinations of OD difference (#1 = $OD_{405} - OD_{635}$, #2 = $OD_{405} - OD_{904}$, and #3 = $OD_{635} - OD_{904}$). Using this method, difference among the genera was visually enhanced, and it was easier to recognize their spectral OD variations. For instance, *S. aureus* had negative OD difference at #1, while the other genera showed 0 to 0.07 OD difference, which means that each genus had similar ODs for 405 nm and 635 nm. For #2, OD differences between 405 nm and 904 nm, *L. mono* had the highest while *S. aureus* had the smallest OD difference,

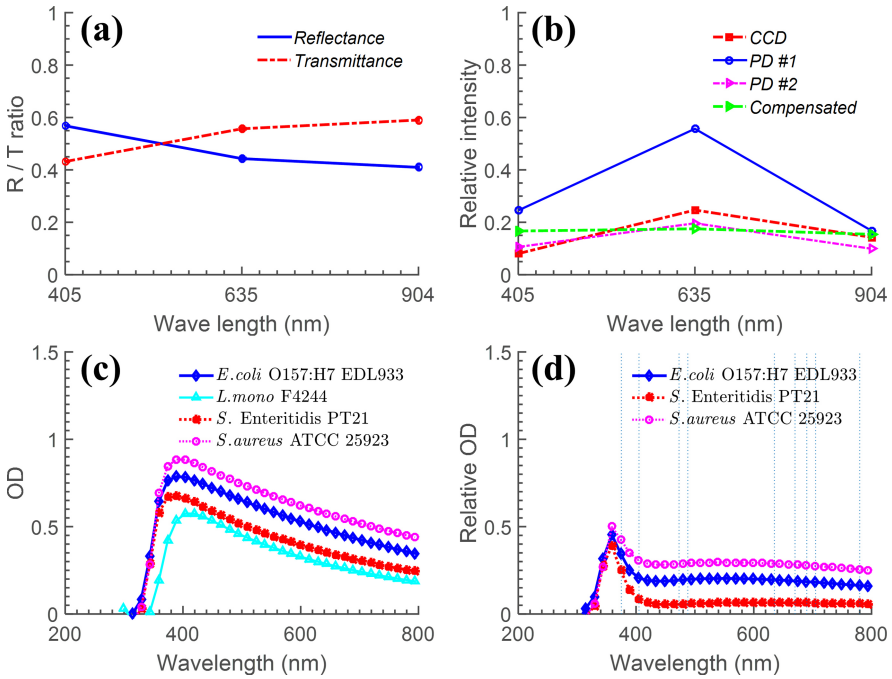


Figure 3 Wavelength-resolved OD of bacteria colony on semi-solid BHI agar. (**a**) Schematic of the measurement points for each case. (**b**) OD (BHI agar + colony) = –log 10((BHI agar + colony)/Input intensity), (**c**) OD (colony) = OD (BHI agar + colony) – OD (BHI agar), (**d**) Relative OD: #1, #2, and #3 stand for OD difference between 405 nm and 635 nm, 405 nm and 904 nm, and 635 nm and 904 nm.

which represented similar OD values for the incoming wavelength. *E. coli* and *S.* Enteritidis had similar OD difference trends at both #2 and #3 combinations.

3.3 Theoretical spectral forward-scatter model

Figure 4(a) shows the results of the multispectral forward-scattering pattern based on the spectral ESL model. *S. aureus* was selected as a target microorganism since it has a smooth convex colony shape and generates a forward-scatter pattern of concentric circles [24, 26]. For the prediction, the colony diameter was set as 1000 μm and the aspect ratio (colony center height-to-diameter ratio) was set to 1 : 6.25. Since both the amplitude and phase terms are function of the wavelength (Eq. (3)), the result is shown as a complexed form of the component by the wavelength differences. As the predicted model shows, the pattern size and the number of rings decrease while ring width and gap increase with a longer wavelength of incident light (see supplemental section and Figure S3). Figure 4(b) shows the edge region of the multispectral experimental measurement result from an *S. aureus* colony, while Figure 4(c) shows the counterpart from the model. The results show a similar trend of order of spectral peak locations; 405 nm is located at the outermost rim region, and 904 nm shows the smallest diffraction angle among the three wavelengths.

3.4 Experimental spectral forward-scatter patterns

Figure 4(d) shows the snapshot of spectral forward-scattering patterns for the four bacterial genera on semi-solid BHI agar measured with MS-BARDOT (also see supplemental Figure S4). Visual inspection of the spectral scatter patterns suggested differences in the characteristics. For *E. coli* and *S.* Enteritidis samples, the 405 nm patterns show fine structures of spokes, speckles, and rings unlike the patterns collected at the other wavelengths. For *L. mono* samples, the 405 nm pattern has the largest diameter, and the central portion of the pattern indicates higher signal intensity compared with patterns at 635 nm and 904 nm. These can be further confirmed with the patterns transformed and plotted in polar representation. Figure S5 shows the transformed patterns of Figure 4(d) (spectral forward-scattering patterns of each species; 405 nm, 635 nm, and 904 nm for top, middle, and bottom, respectively) in rectangular coordinates, where X and Y axes represent angle and radius, respectively. *S. aureus* patterns transformed into polar representation (Figure S5(d)) reveal a clear ripple structure in 635 nm and 904 nm patterns, while those at 405 nm show low intensity outside the central bright spots.

3.5 Multispectral image analysis

The spectral scatter patterns were analyzed as described before using GPZ moments as features.

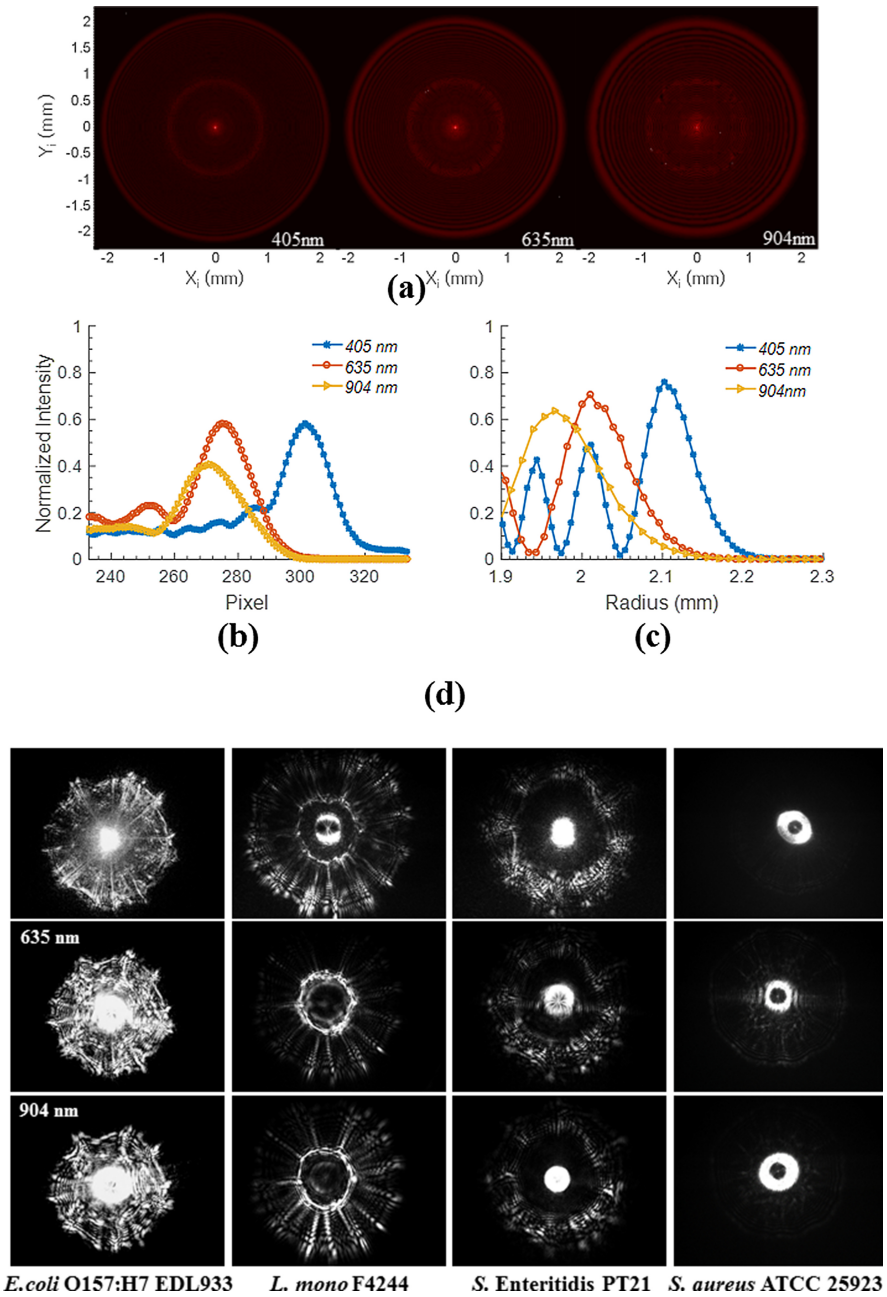


Figure 4 Theoretical simulation of the multispectral forward-scattering patterns. (**a**) The predicted scatter patterns from three wavelengths. (**b**) and (**c**) The 1-D cross-section of the intensity across the radial direction near the boundary area for experimental and simulation results, respectively. (**d**) Snapshot of the multispectral forward-scattering patterns for four different organisms: top: 405 nm, middle: 635 nm, bottom: 904 nm. Qualitative differences are observed among the wavelengths; 405 nm produces refined scatter patterns.

Since three separate laser wavelengths were used, the number of extracted features per colony was three times larger than in the standard setup described previously. This brings a serious challenge, since the increase in the feature space leads to a much higher complexity of classifiers. Although it may be argued that increasing the number of features should lead to better (i.e., more accurate) classification, it is known that a naïve increase in the dimensionality may also lead to overtraining and result in lower classification accuracy in practical implementations. To deal with this issue and maximize the classification efficiency, we employed feature selection based on a random-forest (RF) algorithm. Essentially, in every run the RF selects a random feature subset and generates a classification tree. The importance of the analyzed features is determined by the accuracy of these trees.

We conducted our further analysis using 15 feature selections. After determining the subset of best features in single wavelength and mixed wavelength settings we performed the final classification using standard SVM implementation with a linear kernel. The performance of the classifiers was determined by 10×2 cross-validation, i.e. the 2-fold cross-validation procedure was repeated 10 times with different

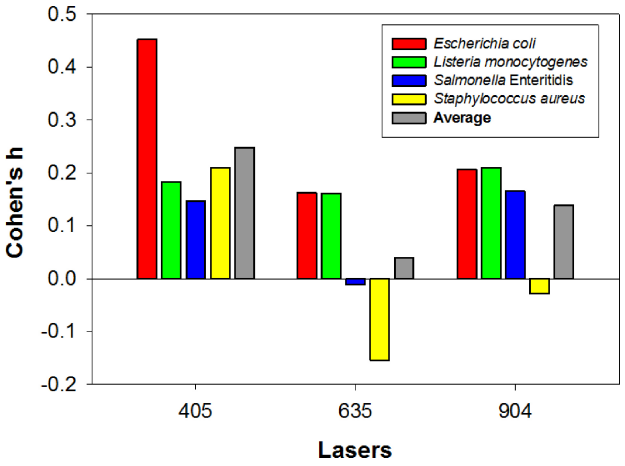


Figure 5 Improvement is classification sensitivity due to the use of multiple laser wavelengths (vs. single wavelength) expressed as Cohen's *h* effect size. The results shows that detection of *E.coli* benefits mostly from the use of multiple wavelength. On the other hand the detection of *S. aureus* is not improved. On average, the multiple wavelength set-up demonstrate an improvement over every single laser line.

seeds for the random-number generator. The final results are reported in the confusion matrices (as percentages) with accompanying standard deviations, for 405 nm (A), 635 nm (B), 904 nm (C), and a mix of features from all colors (D) at Figure 6. For the purpose of visualization, we employed linear discriminant analysis plots, which revealed the structure of the data point clouds and 3-D and illustrated the difference between classifiers built using different feature sets (Figure S6). The improvement in classification sensitivity (true positive rates t_{all} and t_{405}, t_{635}, t_{904}) due to the use of multiple wavelength was expressed as Cohen's *h*, where

$$h = 2(\arcsin(\sqrt{t_{all}}) - \arcsin(\sqrt{t_{wav}})) \quad (9)$$

The results are shown in Figure 5, and demonstrate that *E. coli* classification benefits most from the use of multiple lasers, whereas *S. aureus* classification becomes actually less sensitive in the multi-wavelength settings. This is because the 635 nm wavelength produces the ideal feature set for the classification of this microorganisms, and every other wavelength is worse. However, on average (i.e. taking all the bacteria genus under consideration), the set-up employing multiple wavelengths is always better than a single wavelength arrangement. Even though absolute improvement levels are only small (in terms of Cohen's *h* values) for the genera level of classification, we hypothesize that this technique will provide a higher degree of improvements when operating in a more challenging setting, for instance classifying lower taxonomies samples such as *E. coli* serovars.

To test this hypothesis, MS-BARDOT was challenged with seven *E. coli* serovars (*E. coli* O103, O111, O121, O145, O157, O26, and O45). Figure 7 displays the comparison of four cases (mixed, 405 nm, 635 nm, and 904 nm) for all seven serovars. The results indicates an overall better performance

(A)

	E. coli		*L. mono*		*S.* Enteritidis		*S. aureus*	
E. coli	93.75	(0.38)	0.43	(0.49)	1.25	(0.14)	2.5	(0.17)
L. mono	3.64	(0.256)	99.16	(0.58)	0.23	(0.22)	0.13	(0.168)
***S.* Enteritidis**	2.57	(0.302)	0.43	(0.27)	96.12	(0.3)	0.66	(0)
S. aureus	0	(0)	0	(0)	2.39	(0.46)	96.73	(0)

(B)

	E. coli		*L. mono*		*S.* Enteritidis		*Staphylococcus*	
E. coli	98.85	(0.233)	0	(0)	0.66	(0)	0	(0)
L. mono	1.14	(0.23)	99.35	(0)	0.65	(0)	0	(0)
***S.* Enteritidis**	0	(0)	0.66	(0)	98.59	(0.22)	0	(0)
S. aureus	0	(0)	0	(0)	0.1	(0.22)	100	(0)

(C)

	E. coli		*L. mono*		*S.* Enteritidis		*Staphylococcus*	
E. coli	98.32	(0.423)	1.12	(0.32)	0.13	(0.23)	0.33	(0.219)
L. mono	1.27	(0.284)	98.9	(0.31)	1.82	(0.38)	0	(0)
***S.* Enteritidis**	0.39	(0.34)	0	(0)	95.76	(0.68)	0.07	(0.139)
S. aureus	0	(0)	0	(0)	2.25	(0.36)	99.61	(0.207)

(D)

	E. coli		*L. mono*		*S.* Enteritidis		*Staphylococcus*	
E. coli	99.93	(0.139)	0	(0)	0.1	(0.16)	0.59	(0.139)
L. mono	0	(0)	100	(0)	1.33	(0.42)	0	(0)
***S.* Enteritidis**	0.07	(0.139)	0	(0)	98.45	(0.47)	0	(0)
S. aureus	0	(0)	0	(0)	0.1	(0.16)	99.41	(0.138)

Figure 6 Confusion matrices with accompanying standard deviations for each wavelength, (**a**) 405 nm, (**b**) 635 nm, (**c**) 904 nm, and (**d**) a mix of features from all wavelength.

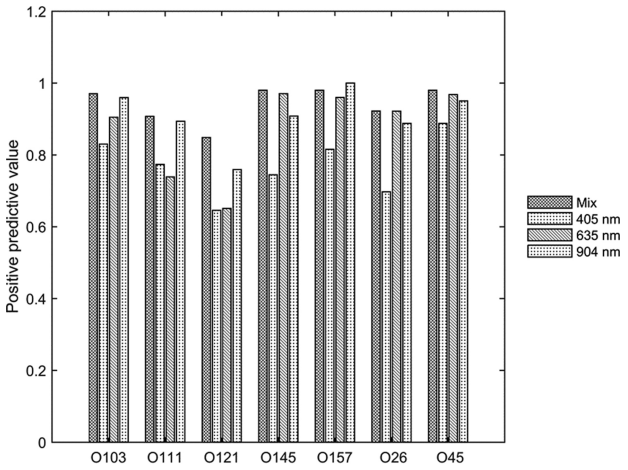


Figure 7 Comparison of the positive predictive value (PPV) of the seven *E. coli* serovars. Four bar graph represents the PPV for mixed wavelength and three single wavelengths (405 nm, 635 nm, and 904 nm). The result clearly indicates that optimal mixture of the features from different wavelengths provides overall higher performance for the single wavelength version.

for the mixed wavelength classifier for all seven serovars tested. Based on the positive predictive values, 405 nm classifier shows the lowest average PPV while the 904 nm shows the highest among the single wavelength classifier. However, the mixed wavelength classifier outperforms on all cases except for the O157 where their absolute PPV values were 0.98 and 1 for mixed and 904 nm respectively. In particular, single wavelength classifier results in 0.65–0.77 PPV values for O111 and O121 while the mixed classifier improves this value to 0.9 and 0.84 respectively.

4. Discussion

Notable improvement over the single-wavelength BARDOT was possible owing to the optical design of the multispectral structure. The overall dimension of the module that could replace the single laser source of the current system is 95 mm × 52 mm × 140 mm. The calibration of each wavelength was performed one time to accommodate different reflectance/transmission ratios from the pellicle beam splitters and spectral quantum efficiency from the CMOS camera. To deal with this issue, we measured the incoming spectral intensity and compensated each wavelength such that approximately the same intensity is perceived by the CMOS camera (Figure 2(a, b)).

Figures 2(c, d) and 3 compare the spectroscopic OD measurements for liquid and solid samples. Using the liquid sample is the standard practice for growth-curve estimation by microbiologists, but we utilized this information for the selection of the best discriminative wavelength region. For example, at 400 nm the OD of *S. aureus* is almost 1/3 higher than that of *L. mono*. Using spectroscopic absorption alone for classifying different bacterial samples has been reported [27]; for the four tested genera at least, the selected wavelengths of 405 nm, 635 nm, and 904 nm provide good separation by their spectral absorption. The ODs for solid samples provided interesting results. Compared to liquid samples, all genera except *S. aureus* showed a monotonically decreasing OD trend as the wavelength increased. By nature, liquid samples are more homogeneously spread out through the whole volume, so only the individual cell shape or other particulates can be argued as contributing factors for the differences observed. However, in solid samples, the growing microbial film (bacterial colony) itself contains more characteristic information other than the individual cell shape. For example, *E. coli*, *Listeria,* and *Bacillus* cells are all rod-shaped, but their colony characteristic show dramatic differences [28]. In addition, nutrition, agar hardness, and environmental factors have been reported to change the morphology of the solid colony. Therefore, given an automatic instrument that can capture the multimodal characteristics of a colony, a solid sample would have better opportunity to provide differentiable traits from a bacterial sample.

Figure 4(d) and Figure S5 display a snapshot of the multispectral forward-scatter patterns from three wavelengths and four genera, plotted in polar coordinates. Different bacteria have different prominent patterns. For example, some samples showed very fine structured spokes (*E. coli*) and speckles (*S.* Enteritidis) at the edge of the rims, while *L. mono* showed a difference in the center spot area. Our biophysical model establishes that a bacterial colony has two major regions; the edge regions, where cell division occurs, are generally less dense and have greater water content than the center part [11, 29]. Therefore, this pattern information can provide some understanding of how bacteria are spreading at the edge and how cells are accumulating in the center. One organism that uniquely stand out is *S. aureus*. Since the 405 nm patterns show weak rings at the edge, there is less detail information that can be extracted.

It has previously been reported that systems based on monochromatic elastic light scatter produce features that lead to high classification accuracies, comparable with well-established biochemical or immuno-based methods [12]. However, owing to the danger posed by many of the pathogens classified, it is imperative that the classification procedures provide the highest possible level of accuracy. Therefore, the only relevant benchmark for the systems tested is absolutely perfect sensitivity and spe-

cificity. Consequently, we consider robust increases in classification success in the range of 1–2% to be very important in the practical setting. We examined an implementation of a feature-selection strategy that allows us to maximize the potential of a multispectral system without the disadvantage of overwhelming the classification system with an increased number of features. We also demonstrated that careful selection of features may lead to increase in already impressive classification accuracies. Adopting a random-forest algorithm for feature selection, we verified that a range of 10–20 features was the optimal for the application, and used 15 features for the analysis to maximize the classification ratio. As Figure 5 shows, the optimized selection of features from all-wavelength case always provide better results comparing to the single wavelength arrangement. Figure 6 shows the cross-validation matrices for single and multi-wavelength cases where, on average, selection of the best features from multi-wavelength provides the highest classification rate.

In general, hierarchical taxonomy decides the differentiating power among the species of the bacterial sample. Except for mutation of certain genes that dramatically changes the growth morphology, moving down to lower hierarchical level of the taxonomy generally shares more common morphological traits. Since serovars are the subspecies class that only differs in surface antigen, their bacterial colony morphology and respective scatter patterns displays similar characteristics. Figure 7 result clearly delivers the merit of the multispectral approach of the bacterial colony classification where optimal selection of features from each wavelength component ensures the overall improvement in classification rates compared to single wavelength version.

5. Conclusion

We propose a bacterial phenotyping technique based on multispectral forward-scattering patterns and ODs. The system provides simultaneous measurement of forward-scattering patterns and ODs from a target bacterial colony at three different wavelengths (405 nm, 635 nm, and 904 nm). Utilizing a stackable pellicle beam-splitter structure, the system minimizes stray light and the ghost effect and can be further expanded to include additional wavelengths. Pseudo-Zernike (GPZ) polynomials/moments are used for analysis and classification. The performance of the proposed technique was verified using four different bacterial genera. Compared to the previous single-wavelength instrument, a multispectral system can provide enhanced performance with the same optimal number of features. The resulting enhancement of accuracy is critical in biosafety- and biosurvelliance-related fields.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website.

Acknowledgement This material is based upon work supported by the U.S. Department of Agriculture, Agricultural Research Service, under Agreement No. 1935-42000-072. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

Author biographies Please see Supporting Information online.

References

[1] Foodborne Outbreak Online Database (FOOD), CDC – NORS (Centers for Disease Control and Preventation, U.S.), http://wwwn.cdc.gov/foodborneoutbreaks/ (2015).
[2] G. Lu and B. Fei J. Biomed. Opt. **19**, 10901 (2014).
[3] H. Huang, L. Liu, and M. O. Ngadi, Sensors (Basel). **14**, 7248–7276 (2014).
[4] U. Khulal, J. Zhao, W. Hu, and Q. Chen, Food Chemistry. **197**, Part B, 1191–1199 (2016).
[5] Q. Huang, Q. Chen, H. Li, G. Huang, Q. Ouyang, and J. Zhao, J. Food Eng. **154**, 69–75 (2015).
[6] Q. Chen, C. Zhang, J. Zhao, and Q. Ouyang, TrAC Trends in Analytical Chemistry. **52**, 261–274 (2013).
[7] U. Siripatrawan, Y. Makino, Y. Kawagoe, and S. Oshita, Talanta. **85**, 276–281 (2011).
[8] F. Mendoza, R. Lu, and H. Cen, J. Food Eng. **125**, 59–68 (2014).
[9] G. Elmasry, D. F. Barbin, D.-W. Sun, and P. Allen, CRC CR. Rev. Food Sci. **52**, 689–711 (2012).
[10] Y. Liu, W. R. Windham, K. C. Lawrence, and B. Park, Appl. Spectrosc. **57**, 1609–1612 (2003).
[11] E. Bae, P. P. Banada, K. Huff, A. K. Bhunia, J. P. Robinson, and E. D. Hirleman, Appl. Optics. **46**, 3639–3648 (2007).
[12] P. P. Banada, K. Huff, E. Bae, B. Rajwa, A. Aroonnual, B. Bayraktar, A. Adil, J. P. Robinson, E. D. Hirleman, and A. K. Bhunia, Biosens. Bioelectron. **24**, 1685–1692 (2009).
[13] P. P. Banada, S. Guo, B. Bayraktar, E. Bae, B. Rajwa, J. P. Robinson, E. D. Hirleman, and A. K. Bhunia, Biosens. Bioelectron. **22**, 1664–1671 (2007).
[14] Y. Tang, H. Kim, A. K. Singh, A. Aroonnual, E. Bae, B. Rajwa, P. M. Fratamico, and A. K. Bhunia, PLoS ONE. **9**, e105272 (2014).
[15] A. K. Singh, A. M. Bettasso, E. Bae, B. Rajwa, M. M. Dundar, M. D. Forster, L. Liu, B. Barrett, J. Lovchik, J. P. Robinson, E. D. Hirleman, and A. K. Bhunia, mBio. **5**, e01019-13 (2014).
[16] K. Huff, A. Aroonnual, A. E. F. Littlejohn, B. Rajwa, E. Bae, P. P. Banada, V. Patsekin, E. D. Hirleman, J.

P. Robinson, G. P. Richards, and A. K. Bhunia, Microb. Biotechnol. **5**, 607–620 (2012).
[17] Y. He, S. Reed, A. K. Bhunia, A. Gehring, L.-H. Nguyen, and P. L. Irwin, Food Microbiol. **47**, 28–35 (2015).
[18] A. Suchwalko, I. Buzalewicz, A. Wieliczko, and H. Podbielska, Opt Express. **21**, 11322–11337 (2013).
[19] C.-S. Chen, M. Biasca, C. Le, E. Y.-T. Chen, E. D. Hirleman, and W.-C. Chin, Biol. Proced. Online. **15**, 2 (2013).
[20] A. K. Singh, X. Sun, X. Bai, H. Kim, M. U. Abdalhaseib, E. Bae, and A. K. Bhunia, J Microbiol Methods. **109**, 56–66 (2015).
[21] U. Minoni, A. Signoroni, and G. Nassini, Biosens. Bioelectron. **68**, 536–543 (2015).
[22] P. R. Marcoux, M. Dupoy, A. Cuer, J. L. Kodja, A. Lefebvre, F. Licari, R. Louvet, A. Narassiguin, and F. Mallard, Appl Microbiol Biotechnol. **98**, 2243–2254 (2014).
[23] W. Pan, J. Zhao, and Q. Chen, Sci Rep. **5**, 9524 (2015).
[24] H. Kim, I.-J. Doh, A. K. Bhunia, G. B. King, and E. Bae, Opt Express. **23**, 8545–8554 (2015).
[25] B. Rajwa, M. M. Dundar, F. Akova, A. Bettasso, V. Patsekin, E. D. Hirleman, A. K. Bhunia, and J. P. Robinson, Cytom. Part A. **77**, 1103–1112 (2010).
[26] J. M. Smith, D. E. Huffman, D. Acosta, Y. Serebrennikova, L. García-Rubio, and G. F. Leparc, J. Biomed. Opt. **17**, 1070021–1070029 (2012).
[27] H. Kim, N. Bai, A. K. Bhunia, G. B. King, E. D. Hirleman, and E. Bae, J. Biophotonics. **6**, 929–937 (2013).
[28] J. A. Shapiro, Sci. Prog. **76**, 399–424 (1992).
[29] E. Bae, N. Bai, A. Aroonnual, J. P. Robinson, A. K. Bhunia, and E. D. Hirleman, J. Biomed. Opt. **15**, 045001 (2010).