Statistical identification of bacteria species

A. Suchwalko¹, I. Buzalewicz¹ and H. Podbielska¹

¹Bio-Optics Group, Institute of Biomedical Engineering and Instrumentation, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland

The challenge of developing fast and trustworthy method for bacteria identification has been taken by many research groups from all around the world but only two decided to develop methods based on the statistical analysis of forward light scattering on bacterial colonies grown on solid nutrient media. Optical methods contrary to other techniques offer many advantages such as non-destructive and non-contact bacteria identification, which allows application of different method on the same sample if identification confirmation is required e.g. in medical diagnostic procedures.

Both mentioned above methods use dedicated optical systems for recording diffraction patterns of the bacterial colonies, construct set of numerical features calculated from the patterns and use statistical analysis for identification of unknown samples. The methods use different optical systems, different numerical features and different statistical analysis workflows and therefore obtain different results. In this chapter we will present a brief comparison of these two methods and a more detailed description of the one developed in our group.

1. Introduction

The threat posed by pathogenic bacteria species grows worldwide at an alarming rate every year. The microbes identified as possible food and waterborne pathogens are present in the surrounding environment. Increasing bacterial resistance to well-known antibacterial agents and antibiotics is a serious threat to our health and life [1–9].

The identification of bacteria species is a difficult, time and resources consuming task. There are some attempts to identify bacteria by means of biochemical, molecular or immunological techniques. However, in spite of their high sensitivity, they are also time consuming, need high quality reagents and very pure samples, what makes them expensive. Moreover, the most sensitive methods as the PCR (Polymerase Chain Reaction) can take up to seven days to identify particular bacteria species, what for example in hospital condition may lead in extreme cases even to patient’s death. Recently, real-time PCR was demonstrated to be fast and reliable method, but it needs a priori knowledge and therefore cannot deal with unknown samples [9,10].

Various optical methods as e.g. bacterial cells detection in water suspension by forward and backward light scattering [11], infrared [12] and fluorescence spectroscopy [13–16], flow cytometry, chromatography, chemiluminescence [17], bioconjugated biomolecules [18,19] or surface plasmon resonance (SPR) [20,21] were proposed, as well. Although they offer the possibility of non-invasive, non-destructive and non-contact examination of bacterial samples, they exhibit similar disadvantages as previous techniques, including demanding time-consuming preparation of high quality samples and necessity of usage of the equipment with the high sensitivity and spectral resolution [9].

In order to be able to effectively fight with the increasingly menacing enemy, we need tools that allow fast and reliable detection and identification of the bacteria species and their strains. The challenge of developing fast and trustworthy method for bacteria identification has been taken by many research groups from all around the world, but only two decided to develop methods based on the statistical analysis of forward light scattering on bacterial colonies grown on solid nutrient media. Optical methods contrary to other techniques offer many advantages such as non-destructive and non-contact bacteria identification, which allows application of different method on the same sample, if identification confirmation is required e.g. in medical diagnostic procedures [9,10,22–38]. These two methods are as follows: a microbial high-throughput screening (HTS) system proposed by researchers from Purdue University and Bacteria Identification System (BIS) introduced by our Bio-Optics Group from the Wroclaw University of Technology. Though the HTS name appears for the first time in work from 2012 year [26] the first research on the topic was from 2006 [29]. In order to distinguish these methods in the text, they will be referred by their acronyms HTS and BIS.

Both methods use dedicated optical systems for recording diffraction patterns of the bacterial colonies, construct set of numerical features calculated form the patterns and use statistical analysis for identification of unknown samples. These methods use different optical systems, different numerical features and different statistical analysis workflows and therefore obtain different results. The methods will be compared, further in this chapter. The BIS method offers a unique optical system with converging spherical wave illumination that exhibits unique properties:

• compression of the observation plane to the finite region in the space,
• observation of Fresnel and Fraunhofer patterns in the same setup,
• diffraction patterns scaling.

HTS method is based on detection of multi-angle light-scatter patterns formed by the complex interaction of laser light with the structure of colonies to quantify differences in phenotypes of colonies [30].

© FORMATEX 2013
Additionally, in both systems there is no need for special and labor intensive preparation of bacteria samples as in case of immunological, biochemical or biomolecular techniques. Therefore, it should be pointed out, that these advantages are not offered by other optical configurations as the HTS set-up [9,32–38]. Moreover, BIS optical system for bacteria species identification can be integrated with developed in our Group method of automated bacterial colonies number evaluation by analysis of their Fourier spectra [34], which is not offered by the other system. This additionally will increase the practical application of this system in microbiological diagnosis.

Results obtained by means of the BIS system have shown that colonies of different bacteria species and their strains generate unique signatures [32]. In the analysis workflow registration of the Fresnel patterns is followed by the extraction of unique classification features based on the statistical moments, which correspond to the morphological and textural characteristics of the bacterial colonies and their diffraction patterns. From many features only the best are selected for further analysis with use of statistical apparatus (ANOVA analysis of variance and Fisher divergence). Selected features are then used to build classification models (Linear Discriminant Analysis, Quadratic Discriminant Analysis and Support Vector Machine) and finally, the models are used for unknown bacteria colonies Fresnel patterns identification [9,36–38].

The workflow scheme comparing BIS and HTS systems can be found in Table 1.

**Table 1** Summary of analysis steps in the two methods described in the work

<table>
<thead>
<tr>
<th>Analysis tasks</th>
<th>Analysis subtasks</th>
<th>BIS</th>
<th>Methods used by BIS</th>
<th>HTS</th>
<th>Methods used by HTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data acquisition</td>
<td>Bacteria samples preparation</td>
<td>+</td>
<td>Incubation on Columbia agar (Oxoid)</td>
<td>+</td>
<td>Incubation on soy agar (BD, catalog #211043)</td>
</tr>
<tr>
<td></td>
<td>Diffraction patterns registration</td>
<td>+</td>
<td>Optical system with converging spherical wave illumination</td>
<td>+</td>
<td>Laser Scatterometer</td>
</tr>
<tr>
<td>2. Image Preprocessing</td>
<td>Center and edges of the colony marking</td>
<td>+</td>
<td>ImageJ macro with human interaction</td>
<td>+</td>
<td>Colony locating module</td>
</tr>
<tr>
<td></td>
<td>Normalization</td>
<td>+</td>
<td>Histogram stretching with assumption of black background</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Pattern partitioning</td>
<td>+</td>
<td>Dividing patterns into 10 rings of equal thickness</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Feature extraction</td>
<td>+</td>
<td>Textural and morphological features based on statistical moments calculated within rings</td>
<td>+</td>
<td>Features based on pseudo-Zernike moments and Haralick texture features</td>
</tr>
<tr>
<td>3. Statistical analysis</td>
<td>Feature selection</td>
<td>+</td>
<td>ANOVA, SNR</td>
<td>+</td>
<td>Sequential floating forward selection-based wrapper was used with the k-nearest neighbor, linear discriminant analysis and recursive-portioning algorithms</td>
</tr>
<tr>
<td></td>
<td>Classification</td>
<td>+</td>
<td>Best classifier was QDA</td>
<td>+</td>
<td>Best classifier was SVM</td>
</tr>
<tr>
<td></td>
<td>Classification performance assessment</td>
<td>+</td>
<td>CV, sensitivity, specificity, accuracy</td>
<td>+</td>
<td>CV, sensitivity, specificity, accuracy</td>
</tr>
</tbody>
</table>

The identification results obtained by BIS of seven bacteria species under study (Salmonella enteritidis, Staphylococcus aureus, Staphylococcus intermedius, Escherichia coli, Proteus mirabilis, Pseudomonas aeruginosa and Citrobacter freundii) depending on the selected features show the identification accuracy varying from 98.58% to 99.143% [9,36–38].
2. Sample preparation

2.1. Sample choosing

The BIS method was developed and tested on the 7 various species of pathogen bacteria: *Salmonella enteritidis*, *Staphylococcus aureus*, *Staphylococcus intermedius*, *Escherichia coli*, *Proteus mirabilis*, *Pseudomonas aeruginosa* and *Citrobacter freundii*. The samples of bacterial species were chosen according to difficulty of their identification with the use of classical methods.

The HTS was tested on the most common pathogenic species of *Vibrio*, *Vibrio cholerae*, *Vibrio parahaemolyticus* and *Vibrio vulnificus* and *Listeria monocytogenes* and other *Listeria* species [27,30].

2.2. Sample preparation

The seven chosen bacteria species were examined with BIS: *Salmonella Enteritidis* (ATCC 13076), *Escherichia coli* (PCM O119), *Staphylococcus aureus* (PCM 2267), *Staphylococcus intermedius* (PCM 2405), *Citrobacter freundii* (PCM 531), *Proteus mirabilis* (PCM 547) and *Pseudomonas aeruginosa* (ATCC 27853) (Table 2). The cultures were obtained from the microbiological laboratory of the Department of Epizootiology and Veterinary Administration with Clinic of Infectious Diseases of the Wroclaw University of Environmental and Life Science. Bacteria suspensions were first incubated for 18 hours at the temperature of 37°C. Respective dilutions were seeded on the surface of the solid nutrient medium in Petri dish, so as to obtain 12-20 colonies per plate, and were again incubated at 37°C for the next 11 hours. The bacteria colonies were grown on Columbia agar (Oxoid) [9].

<table>
<thead>
<tr>
<th>Bacteria species</th>
<th><em>Salmonella enteritidis</em></th>
<th><em>Staphylococcus aureus</em></th>
<th><em>Staphylococcus intermedius</em></th>
<th><em>Escherichia coli</em></th>
<th><em>Proteus mirabilis</em></th>
<th><em>Pseudomonas aeruginosa</em></th>
<th><em>Citrobacter freundii</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patterns</td>
<td>55</td>
<td>50</td>
<td>48</td>
<td>53</td>
<td>47</td>
<td>50</td>
<td>49</td>
</tr>
</tbody>
</table>

Trypscase soy agar (BD, catalog #211043) was used for sample preparation in HTS method. 40 mg agar powder was suspended in 1 litre millipore water and boiled for 1 min. After the agar cooled, 35 ml agar was dispensed onto sterile Petri dishes. Petri dishes containing test samples were incubated at 30°C for 12–18 h. Colony diameter of 1.3 - 0.2 mm was used as a fixed parameter for each test culture before analysis with the light scatterometer [26,30].

3. Pattern registration

3.1. Optical system

The configuration of the BIS (Fig. 1) dedicated optical system with converging spherical wave illumination for analysis of light diffraction on bacterial colonies includes the laser diode module (635 nm, 1 mW, collimated, Thorlabs), natural density filter (optical density: 0-4.0, Edmund Optics), linear polarizer (Thorlabs), beam expander (Edmund Optics, 4X), iris diaphragm, transforming lens (achromatic doublet, focal distance: 48.6 cm, clear aperture: 6.35 cm, Edmund Optics), XYZ sample positioning stage with Petri dish, which enables the adjustment of the uniform illumination of the single bacteria colony, CMOS camera (EO-1312, Edmund Optics) and computer. In this configuration bacterial colony is illuminated by converging spherical wave generated by transforming lens in order to control lateral scale changes of diffraction patterns. The sample was located at the back surface of transforming lens and the CMOS camera. The diameter of light beam was approximately equal to the diameter of bacterial colony [9, 32–38].

The forward scatterometer was designed and fabricated as the core of the HTS system and consists of colony locating and elastic light scatter (ELS) modules. The first component is required to obtain information on the total number of colonies and the colony centre locations, via capture of scattered light by a universal serial bus (USB) complementary metal oxide semiconductor (CMOS) camera (PL-B741U-BL, 1280 × 1024 pixels, PixeLINK, Ottawa, ON, Canada) from the back-illuminated image of the Petri dish. Compared to the previous light source with oblique illumination, back illumination by light-emitting diodes (LED) provides stable and constant illumination without flickering. The USB camera is equipped with an imaging lens (M118FM08, Tamron, Saitama, Japan) with a viewing angle of 34°×25.6° to cover the whole area of the plate within a given imaging distance. A plate diffuser is placed between the bottom of the plate and the light source. The second component is responsible for recording the ELS patterns, which are generated by passage of light from a diode laser (λ = 635 nm, Coherent 0221–698-01 REV B, CA, USA), which generates 1 mW circular beam with 1/e² diameter of 1 mm and captured by a second USB CMOS camera. Two linear stages (XN10-
0060-M02-71 with stroke 6 in. and resolution 2 mm, Velmex, Bloomfield, NY, USA) were used for 2D translation of the plate, via communication with the computer through a USB connection [26].

3.2. Pattern characteristics

Diffraction patterns of bacterial colonies exhibit unique spatial structures. They contain set of the diffraction rings, but the number and sizes of these rings depend on the bacteria species. This is correlated with the morphology of colony, as well as amplitude and phase properties that are unique for each species. The number of rings was estimated as

$$N_{\text{ring}} \cong \frac{\Delta \Phi}{2 \pi}$$

where $\Delta \Phi$ is the phase lag [23]. Morphological properties of the bacterial colonies are reflected by the morphological and texture properties of their diffraction patterns. Therefore morphological and textural properties of the diffraction patterns were examined as they are easily interpretable [9,26,38]. Exemplary Fresnel patterns of bacterial colonies of various bacteria species as captured in the BIS system are depicted in Fig. 2.

8 bits grey-scale images were analysed. This gives us 256 intensity levels per each pixel. Each image has width of 1280 pixels and height of 1024 pixels in both systems but in the HTS the bitmaps were rescaled and squares of 300 x 300 pixels representing centres of the bitmaps and were subjected to the further analysis [29,31]. The Fresnel patterns in BIS system covers at least 40% of the image space and were analysed as recorded. All diffraction patterns were recorded in the same optical system configuration under the same strictly controlled conditions. Exemplary Fresnel patterns of bacterial colonies are presented in the Fig. 2.

Fig. 1 The configuration of the optical system BIS for bacteria species classification based on Fresnel diffraction patterns
4. Image processing

4.1. Pattern normalization

In order to avoid any errors related to human and laboratory factors (e.g. recording Fresnel patterns by multiple laboratory workers or in slightly different conditions) the normalization is applied as a preliminary step of the analysis in both approaches [29,31]. Some differences in the background intensity can be observed on Fig. 2. Normalization process in the BIS system was carried out under the assumption of a black background of each pattern under controlled conditions. The mean values of the intensities of pixels belonging to the background (beyond the pattern edges) were calculated separately for each pattern and the value was set as the left edge of the stretched histograms. The histogram stretching was performed accordingly to the standard algorithm that transforms pixel intensities values according to the formula

\[ n(x,y) = \frac{o(x,y) - \min(o)}{\max(o) - \min(o)} \cdot (\max(n) - \min(n)) + \min(n) \]

where \( o \) stands for original image, \( n \) for the new (transformed) one with known range, while \( x \) and \( y \) denote pixel coordinates [9,36–40]. Exemplary Fresnel patterns after normalization as obtained in the BIS method are demonstrated on Fig. 3.

4.2. Pattern partitioning

The set of diffraction rings is correlated with the morphology of colony and with its amplitude and phase properties, which are unique for each species. Therefore, in the BIS system the characteristics of the patterns are analysed in the concentric annulus-shaped zones (Regions Of Interest - ROIs), while HTS analyses whole patterns with the use of pseudo-Zernike moments and Haralick texture features [9,26,29,36–38].

In BIS each Fresnel pattern was limited by the circle covering entire pattern, which centre and radius were manually marked. Radius of this circle was then divided into 10 equal pieces, from which rings were created. Comparison of partitioning the Fresnel patterns into 3, 5 and 10 equal thickness rings was performed and 10 ring split showed to be the most effective of fixed partitioning [9,37]. The Fresnel patterns partitioning into 10 disjoint rings was performed in the ImageJ software [41], by the dedicated macro with human interaction.

Examples of 10 rings partitioning of various Fresnel patterns are depicted on the Fig. 4. The examples of split patterns show differences in the location and size of each pattern.
5. Patterns to numbers

5.1. Features description

Features used for building classification models in both approaches are somehow based on the statistical moments: central in the BIS and pseudo-Zernike computed using pseudo-Zernike polynomials in the HTS [9,26,31]. They were chosen as it is well known that they provide good textural interpretation of the images and, in the BIS approach also, regions of interest within images. Mean value (first raw moment) and standard deviation (square root of the second central moment) of the pixel intensities within the rings denote brightness and roughness of the regions of interest, respectively. Skewness (third moment) is referring to the measure of the symmetry of the shape of the distribution of...
the pixel intensities within each ring. Greater value of the skewness means longer tail in the bright direction of the pixel intensities histogram. Kurtosis (fourth moment) is a measure of the flatness or peakedness of a distribution of the pixel intensities within each ring [38].

Apart from statistical moments other features describing textural characteristics were calculated from ring regions in the BIS. Relative smoothness, which is a measure based on the standard deviation given by the formula

\[ 1 - \frac{1}{(1 + \sigma^2)} \]

for each ring, and thus standard deviation and relative smoothness features are dependent. Uniformity was also calculated as

\[ \sum p(x_i)^2 \]

where \( p(x) \) is a probability mass function of the pixel intensities within the rings and \( i \) takes values of the available range \( i \in [0,255] \). Lastly the entropy was calculated which quantifies expected value of information carried by the pattern rings and is given by the formula

\[ -\sum p(x_i) \log_2 p(x_i) \]

The formula is called Shannon entropy. All of the features are labelled \( \text{feature} . x \) , where \( \text{feature} \) is the name of the calculated numerical feature and \( x \) is the ring number counting from the centre of the colony [38].

5.2. Feature extraction

The features from diffraction patterns of bacteria colonies are expressed by discrete pixel intensities. To extract interpretable features from the patterns each of them was partitioned into 10 disjoint rings of equal thickness, which was shown to be the best of fixed splits (for BIS). Pixel intensities from within each of the rings were then used to calculate features based on statistical moments, which are often used as texture features of images as they are easily interpretable. For each ring of each pattern values of the features were calculated in the BIS analysis and for whole pattern in the HTS analysis [9,26,29,38]

6. Statistical identification

6.1. Statistical analysis workflow

The data set in the BIS system consists of 70 features for each of 10 rings in each of 352 diffraction pattern images and it is just a start for the catalogue of pathogen bacteria species necessary for usage of the method in the laboratories. Since the data are large, and contain many classes (bacteria species) it is crucial to develop efficient analysis workflow that will take into consideration all of foreseeable difficulties.

The HTS method used 120 features describing each image as inputs for the analysis procedure. The order of Zernike moment-based features was high enough to include not only low-frequency shape information, but also high-frequency features details of the light-scatter patterns [29].

6.2. Feature selection

As the BIS Fresnel diffraction patterns of bacteria species data set consist of 70 features for each pattern (7 feature groups and 10 rings) choosing of those that differentiate bacteria species the best, was necessary. It is well known that too many features make classification models more complicated and less robust. For this reason we decided to choose only some groups of features. Selecting groups of features only decreases diversity of the features used and thus is a step towards better interpretation of results while maintaining good identification quality. For the task the filter class feature selection methods were applied and combinations of different feature groups and rankings of their class (species) separation capability were tested. Combination of 4 groups of the features: entropy, mean, standard deviation and uniformity gave best identification results and therefore, were chosen for the analysis [38].

To decide which features are the best for building the classification models ANOVA analysis of variance and Fisher divergence measure called also signal to noise ratio further called SNR, were used. Both tests were used to find features that differentiate bacteria species in the best way. Explicit SNR formula is \( SNR = \frac{\mu_i - \mu_j}{\sigma_i + \sigma_j} \). We used it to calculate separation measure value of the given feature by averaging separation values calculated in one-versus-all scheme. It means that for each feature under study we calculated seven separation measures using SNR formula for two series of the data

\[ SNR = \frac{1}{n} \sum \frac{\mu_i - \mu_j}{\sigma_i + \sigma_j} \]

where \( n \) is the number of bacteria species and in our case is equal seven, \( i \) is given bacteria specie while \( j \) represents complement of the specie set \( i \) in the whole data set. SNR was chosen to be better suited method as texture based features differ in mean value what was shown by the ANOVA during previous study, but show high variance, which has a great impact on differentiation possibilities of the features[38].

The HTS feature reduction was performed via linear discriminant analysis (LDA) and principal component analysis (PCA) [29], but no feature selection step was reported.
6.3. Classification and identification

Classification is a process of building statistical classifiers, basing on diffraction patterns features, and using them for prediction of unknown bacteria species, while identification is only prediction of unknown bacteria species. Therefore, identification can be assumed as final part of classification and as its part. The terms will be used alternatively when describing results.

6.4. Classification methods

In the HTS the scatter patterns represented by a number of selected numerical features are subsequently classified using Fisher’s linear discriminant, support vector machine with linear kernel (SVM-L), and support vector machine with radial-basis function kernel (SVM-RBF). The classification system follows the implementation described previously (Bayraktar et al., 2006; Banada et al., 2007) [30].

In the BIS method three classification algorithms were exploited: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Support Vector Machine (SVM). Their results were verified by a classifier performance assessment method. QDA separates classes with the quadratic surface and gave results that were affected by the least error value and therefore QDA is a main classifier in the BIS [9,38].

6.5. Performance assessment

Cross-validation (CV) was chosen as classifier performance the BIS assessment method. CV is a technique for determining how the results of the classification will work for other, independent data sets. In other words, CV estimates the unknown prediction error. CV splits the data set into two complementary subsets (learning and test sets) and performs classification analysis. These two steps are applied given number of times to the whole data set and, as a result the classification error is estimated. In our analysis, 100 repetitions for each classifier and data were performed [9,36–38].

Sensitivity and specificity are well known statistical measures of the performance of a binary (two class e.g. case and control) classification test. As we are dealing with multi-class case (7 classes) sensitivity and specificity cannot be calculated in the ordinary way. The most widely accepted method of calculating those measures in multi-class case is to treat one of classes as case class and combine the rest of the classes as the control class. In this simple way the binary definition can be applied as many times as we have classes. Averaged values of sensitivity and specificity over all of classes give measures for the whole experiment. Accuracy of the classification is equal $1 - CV_{error}$ [9,36–38].

Table 3  Best QDA classifier performance for all seven bacteria species under study and features selected with the use of SNR

<table>
<thead>
<tr>
<th></th>
<th>Error</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>QDA with 18 features selected by SNR</td>
<td>0.857%</td>
<td>0.9820</td>
<td>0.9933</td>
</tr>
</tbody>
</table>

Table 3  Best QDA classifier performance for all seven bacteria species under study and features selected with the use of SNR

In the HTS the detection capability of the best classifier (SVM-RBF) was evaluated, where a sample is represented by multiple instances (multiple colonies) presented on a plate. However, no true multiple-instance learning algorithms were used in this approach. To evaluate the classifier performance, an extended version of a cross-validation procedure was designed. 100 virtual plates were generated in silico containing random combination of colonies for each of the tested detection tasks. Out of 100, the first group of 50 plates contained one or more colonies representing the class of interest, and another 50 did not. The test was repeated six times for each species of interest. In the described setting, a sample (represented by a single plate) was considered positive, if at least one colony in that sample belonged to the class of interest, and negative, if all colonies on the plate were something other than the target. Separate accuracy scores were computed for each classification system individually optimized for detection of each of the classes [30].

The results demonstrate very high sensitivity (96–100%) and specificity (100%) of the proposed method. Out of 50 plates containing one or more colonies of *V. cholerae* and 50 plates containing *V. vulnificus*, all were tagged as positive for *V. cholerae* and *V. vulnificus* respectively in each of the six repeats. In the case of *V. parahaemolyticus*, 96% of the contaminated plates were detected as positive. In all cases the plates which did not contain contamination were correctly labeled as negative [30].
Table 4 Best SVM-RBF, support vector machine with radial-basis function kernel classifier performance assessment where mixture represents colonies of *Vibrio alginolyticus*, *Vibrio anguillarum*, *Vibrio mimicus* and *Vibrio orientalis*. AUC, area under receiver operating curve [30]

<table>
<thead>
<tr>
<th>Class</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixture</td>
<td>0.96</td>
<td>0.99</td>
<td>0.98</td>
<td>1.0</td>
</tr>
<tr>
<td><em>Vibrio cholerae</em></td>
<td>0.94</td>
<td>0.99</td>
<td>0.98</td>
<td>1.0</td>
</tr>
<tr>
<td><em>Vibrio parahaemolyticus</em></td>
<td>0.97</td>
<td>0.96</td>
<td>0.97</td>
<td>1.0</td>
</tr>
<tr>
<td><em>Vibrio vulnificus</em></td>
<td>0.99</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

7. Conclusions
Fresnel diffraction patterns of bacterial colonies exhibit unique spatial structures that can be used for the microbial identification.

Easily interpretable features based on the morphological and textural properties of bacterial colonies diffraction patterns and verification of their bacterial species differentiation capabilities with the use of statistical analysis workflow is possible with the BIS method.

Providing morphological and textural features for the bacteria identification resulted in building models of very high efficiency. With use of the features and the feature selection algorithms it was possible to obtain identification error as small as 0.857% for QDA classifier in the BIS method.

It is possible to build models of very good quality using small number of features based on the statistical moments calculated within rings. This suggests that constructing new features, e.g. related with the optical system modification that differentiate bacteria species very well can lead to building models based on as few as 5 or 6 features with an error of less than 1%. Such situation would allow the visually assisted identification with use of the BIS method to be fully interpretable by the users. It means that the BIS method in a laboratory can provide and show to its user additional information: what are similarities and differences between the sample being identified and selected samples from the database. Constraining the number of features to several only will make number of hints lower and possible to handle by a human being.

Both presented methods are open to identification of new mutations of existing species within just hours and therefore are suitable for fast and reliable diagnosis. They are also of low cost and low laboratory labour.

Optical methods for microbial identification are rapidly developing and soon will be very important apparatus in every microbiological laboratory.

Acknowledgements The support by European Union under the European Social Fund is gratefully acknowledged.

References


