Optik 127 (2016) 1007–1010

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Auto-classification for confocal back-scattering micro-spectrum at single-cell scale using principal component analysis

Cheng Wang*, Miao Wen, Lihong Bai, Tong Zhang

Laboratory of Medical Optics & Optometry, School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai, PR China

ARTICLE INFO

Article history:
Received 11 December 2014
Accepted 9 October 2015

Keywords:
Spectrum
Single cell
Principal component analysis (PCA)
Classification

ABSTRACT

Early diagnosis and treatment for cancer is a key to reduce the mortality of cancer. Based on confocal back scatter micro-spectroscopic system, micro-spectra of normal stomach epithelial cells and cancerous cells were collected. In order to discriminate the two types of cells, the micro-spectra of cells were analyzed with principal component analysis in which the first two scores of the principal components were utilized. The results showed this method could clearly distinguish the confocal back scattering micro-spectra of the normal stomach epithelial cells and cancerous cells. This might provide a new method for early diagnosis of stomach cancer, which is fast, accurate and low-cost at single-cell scale.

1. Introduction

China is the country with the highest incidence of stomach cancer, and accounts for over 40% of all new stomach cancer cases in the world [1]. Stomach cancer is the third leading cause of cancer mortality in China [2,3], and is approaching to young people. Early diagnosis and early treatment is a key approach to reduce the mortality of stomach cancer. Surgery for early stomach cancer has excellent clinical curative effect with the 5-year survival rate being up to 95% [4]. Therefore, it will hold great clinical significance to develop a fast, accurate and low cost diagnostic method.

The elastic optical scattering is one of the basic forms of the interaction between light and tissue. The light-scattering signals from a nucleated cell contain cell size and morphological information [5,6]. Flow cytometry is the most representative scientific instrument to analyze cellular size and morphology by utilizing collected forward and side scattering light from a cell. Moreover, Understanding light-scattering in living cells is crucial to many modern diagnostic technologies, such as optical coherent tomography (OCT) [7], elastic scattering spectroscopy [8], and confocal microscopy [9]. These techniques are developed for non-invasive and real time optical diagnosis based on light scattering signal from the tissue. Light-scattering spectroscopy (LSS) [10] is a system that measures the scattering spectrum related to angle to obtain quantitative information of cell morphology. Wu et al. used the elastic scattering spectrum to study the light-scattering changes of uterine cancer cells treated with acetic acid [11]. Lin et al. reported an elastic back-scattering spectroscopy technique, which could obtain both the back-scattering spectrum and microscopic image on one pixel simultaneously [12]. Fang et al. developed a new technique, namely confocal light absorption and scattering spectroscopic (CLASS) microscopy, which could observe *in vivo* sub-micron cell structure [13,14].

Spectroscopic results are usually multiple dimensional and include large amounts of data. To classify or identify spectrum from different samples, these data would be processed with certain methods. For example, the smoothing algorithm is often used to preprocess the spectra, which can analyze characteristic spectrum line so that different spectrum in some wavebands can be compared with useful information obtained, *e.g.*, relative peak intensity. However, if the characteristic spectra have the same or similar feature, they will be too hard to be identified. With improvement of the speed of collecting spectrum, it is necessary to accomplish all-automatic spectrum identification. Principal Component Analysis (PCA) [15] is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set. PCA is sensitive to the relative scaling of the original variables. In the field of biomedicine, the potential ability of PCA method has been demonstrated. For example, spectra from particles labeled with each of four different SERS tags could be distinguished by principal component analysis [16]. Identification of hepatitis based on normalized data of reflective spectrum

* Corresponding author. Tel.: +86 2155271165.
E-mail address: shhwangcheng@163.com (C. Wang).

http://dx.doi.org/10.1016/j.ijleo.2015.10.066
0030-4026/

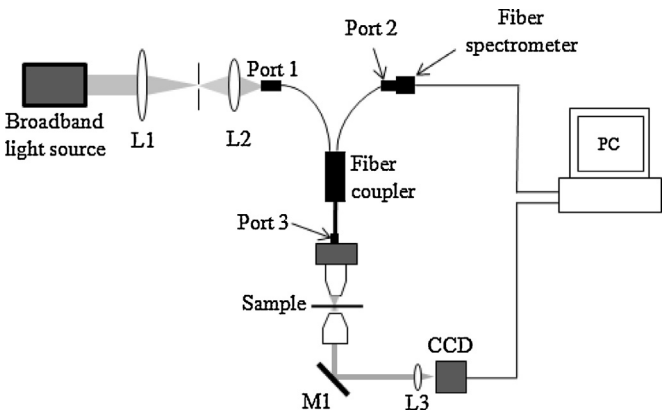

Fig. 1. Schematic of the confocal back-scattering micro-spectrometer based on fiber.

from healthy adults and viral hepatitis suffers was processed by the method of PCA with BP neural network detected non-invasively on hepatitis patients [17].

This study aimed to use principal component analysis (PCA) for accomplishing automatic identification of micro-spectra of the normal cells (GES-1) and cancerous cells (NCI-N87), which was obtained from confocal back scattering micro-spectroscopy. This technique holds potential for early diagnosis of stomach cancer at cellular level, which is fast, accurate and low-cost.

2. Material and methods

2.1. Experiment setup and sample preparing

The fiber-confocal back scattering micro-spectrometer was set up based on our previous work (Fig. 1) [18]. Briefly, broadband white light was coupled to a port of the fiber coupler with a 1×2 port and 50:50 splitting ratio *via* the SMA-905 standard interface. The fiber worked as light source and point detector and the fiber coupler worked as optical splitter to separate the light of illumination and that of the signal [19]. Moreover, the fiber coupler Port 3 replaced the pinhole in confocal system. The light was coupled by Port 3 to the optic probe, which was made up of an achromatic objective (NA = 0.25, 10×) as the collimator and an achromatic objective (NA = 0.6, 40×) as the objective. The lateral and axial resolutions of this system were 5 and 10 μm, respectively. The light from cell scattering in the objective NA angle range was all collected by the optic probe and transmitted through Port 2. The back-scattering light was transmitted through SMA-905 interface of the Port 2 to a CCD spectrograph (B&W Tek Inc., BRC112, USA) with a spectral resolution 4 nm. In order to ensure that the object in the field of view was a cell rather than impurity, the micro-spectrometer offered an observation function to help observe the location of the cell. Illuminating light was given by the same transmitted light through same optic probe.

Broadband white light was coupled to a Port 1 and transmitted to Port 3. An achromatic objective (NA = 0.25, 10×) and an achromatic objective (NA = 0.6, 40×) were used for signal and image collection, respectively. The lateral and axial resolutions of this system were 5 and 10 μm, respectively. The back-scattering light from a cell in the objective NA angle range was all collected and transmitted through Port 2 to the fiber spectrometer. CCD was used to observe the cells to ensure a single cell was detected at one time. L1, L2, L3, optical lens; M1, mirror.

A human normal stomach epithelial cell strain GES-1 and a human stomach carcinoma cell strain NCI-N87 were used in our study. They were purchased from The Cell Bank of Type Culture Collection, Shanghai Institute of Cell Biology, Chinese Academy of Sciences. Cell culture methods and environment were described in literature [18].

2.2. Spectra acquisition and preprocessing

In the experiments, the scattering light intensity from a cell was low. To improve the signal contrast for collecting cellular scattering light intensity, a silicon slice with reflectivity of 30% in both visual light and near infrared (NIR) light was used as a standard reflector in the cell back-scattering spectrum analysis. When cultured cells were steady and adhered to the petric dish, the culture medium was removed. Signals from a single cell were measured with the micro-spectrometer. Because cells of the same type have similar spectrum, twenty spectra were collected for each cell strain. All of spectra data were automatically converted and stored in ASCII format by spectrometer software.

2.3. Principal component analysis

Principal component analysis (PCA) is used to reduce the dimensionality of spectroscopic data. Principal components are guaranteed to be independent if the data set is jointly normally distributed. The principal components are sorted according to their scores which represent the correlativity with raw spectra data. Automatic identification can be performed by utilizing the first few component principals whose accumulative scores are more than 85%.

3. Results

3.1. Cellular scattering micro-spectra

Confocal back-scattering micro-spectra for normal cell (GES-1) and cancerous cell (NCI-N87) were smoothed with 50 sampling points average (Fig. 2). In Fig. 2, curve 1 represented the normal cell spectra and curve 2 represented cancerous cell spectra. The *X* axis was the spectroscopic range, and *Y* axis was the relative intensity of the cellular scattering with the reference of silicon reflective intensity. Fig. 2 showed that normal cell spectra had regular intensity changes from 800 nm to 1000 nm range. It was due to regular cell internal structure. Although heterogeneities were present in the cellular structure, they were not significant enough to affect the peaks resulting from the cytoplasm and nuclear boundaries. However, the internal structure and biochemical changes in cancerous cells were asymmetric, resulting in irregular intensity changes in the spectra [19]. Theoretic prediction had demonstrated that there

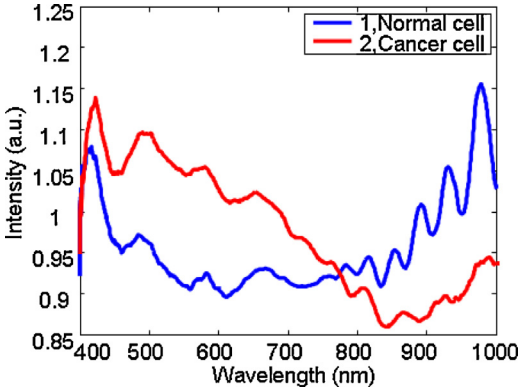

Fig. 2. Confocal back-scattering micro-spectra of normal cells (GES-1) and cancerous cells (NCI-N87), respectively. Normal cell spectra had regular intensity changes from 800 nm to 1000 nm range while cancer cells had irregular intensity changes in the spectra.