様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成　２９　年　　５　月　２６　日現在

機関番号：　１０１０１
研究種目：　若手研究(B)
研究期間：　2015～2016
課題番号：　１５Ｋ１８５１１

研究課題名（和文）New Data-Driven Tools to Quantify Heterogeneous Microenvironments in Live Cell Images

研究課題名（英文）New Data-Driven Tools to Quantify Heterogeneous Microenvironments in Live Cell Images

研究代表者

Ｔａｙｌｏｒ　Ｎｉｃｈｏｌａｓ（Taylor, Nicholas）

北海道大学・電子科学研究所・特任助教

研究者番号：５０７５０８２４

交付決定額（研究期間全体）：（直接経費）　　2,800,000円

研究成果の概要（和文）：ラマン顕微鏡は、生きた細胞について豊富な画像を提供してくれます。ラマン信号というのはそもそもが弱い信号であるため、その処理および分析には十分に注意を払わなければいけません。ここでは、データの汚染を詳細かつ正確に処理することに焦点を当てることで、さらなる分析を行った際に適切な特性を示すことに成功しました。

研究成果の概要（英文）：Raman microscopy provides rich images of living cells. Because Raman signal is inherently weak, care must be taken in its processing and analysis. Here, I focused on precisely and accurately processing contamination in the data so that further analyses provided appropriate characterization.

研究分野：　Biophysics

キーワード：　Raman Microscopy　Statistics　Data Science　Live-Cell Imaging

１．研究開始当初の背景

Reduction of Background contamination

Removal of background contamination from Raman spectroscopic and microscopic measurements is a notoriously difficult task due to the origin of the contamination. Experimental substrates such as quartz glass produce a Raman spectrum of their own, which is variable along the Raman wavenumber dimension. Commonly used techniques for background estimation, including recursive polynomial fitting, are plagued by subjective choices and poor extraction of the actual background spectrum.

During the course of this research, a method based in statistics and utilizing a Gaussian mixture model(GMM) was developed and tested on synthetic and experimental data. The method relies on the statistics generated by the large number of pixels in each image, and that for a particular contamination will be predictable across the spatial dimensions of the image frame. Next, the background intensity is taken to be the extracted distribution with the smallest mean intensity, and spatially-dependent probabilities of an intensity falling within the background distribution are assigned to each pixel in the image frame. This procedure is repeated for all values of the Raman shift, producing a background spectrum. Fig. 1 compares average background spectra extracted from an experimental image having a quartz glass substrate that were calculated with the GMM-based method and the recursive polynomial fitting algorithm, but are recognized by the GMM-based algorithm.

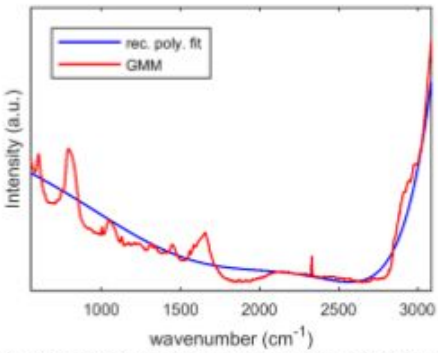

Figure 1. Comparison of background spectra for the recursive polynomial fit method (blue) and the GMM method (red).

２．研究の目的

Spatial Recognition of Signal

In the classification of spectra acquired from Raman images, it is important to exclude those spectra containing only contributions from the experimental substrate; otherwise classification analysis is adversely impacted by their presence. Commonly, clustering algorithms are involved as a preliminary step in identifying such spectra. During the development of the GMM background recognition method, it was realized that the algorithm also serves as a spatial recognition algorithm, providing the ability to simultaneously identify the background spectrum as well as identify spectra that contain mostly background contamination so that they may be removed from the classification analysis. Furthermore, because exclusion of these spectra with the GMM is quantitatively based on identifying background contamination rather than a simple partitioning as in clustering-based methods, the GMM-based algorithm. The Fig. 2 compares an experimental image of a cell culture containing Human thyroid carcinoma along with the separation of background pixels from those pixels containing signal from the cells, displaying good visual agreement.

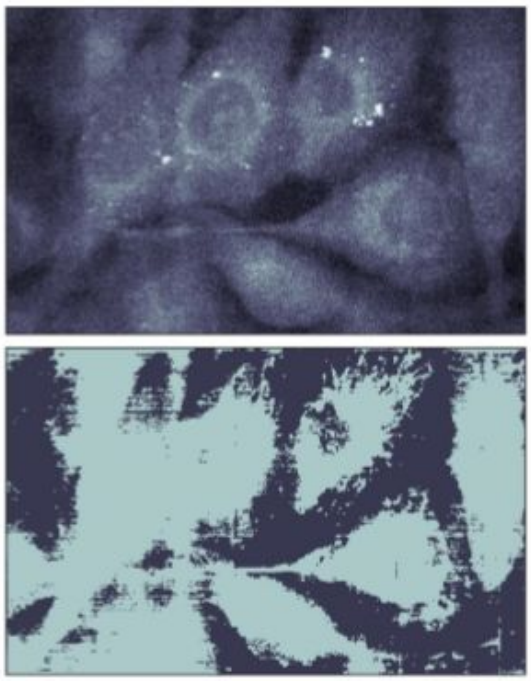

Figure 1. (top) Raman image of human thyroid carcinoma cell culture. (bottom) Dark-colored pixels were identified by the GMM as containing background spectra while light colored pixels were identified as containing Raman signal from the cells.

３．研究の方法

Reduction of Noise Contamination

Another issue in the treatment of Raman microscopic images is the prevalence and magnitude of detection noise. Due to its inherently weak signal, Raman microscopic images typically have signal-to-noise rations on the order of 0.1-1. A common method currently in use in Raman is singular value decomposition, in which all the spectra in the image are aligned in the wavenumber dimension, the decomposition is performed, singular values falling below a certain magnitude are set to zero, and then the spectra are reconstructed. This method, wavelet denoising has been used in Raman microscopy, but with little success. This in mostly due to the construction of wavelet denoising. In particular, wavelet denoising suppresses low magnitude, high frequency fluctuations. Many Raman signals are narrow peaks with small intensity. Such signals are removed my traditional wavelet denoising methods.

Towards the end of a quantitative, objective, and effective denoising algorithm was developed using the 2-dimensional wavelet transformation and the average behavior along the wavenumber dimension. Briefly, issues arise with traditional wavelet methods when the magnitude of the signal is within the fluctuation arising from the noise magnitude. However, modifying the traditional algorithm to use the average behavior over a local wavenumber range allows for estimation of the signal magnitude, and subsequent noise removal around this value. Application of this algorithm to experimental and synthetic data show that the modified wavelet algorithm performs as well as singular value decomposition while offering the advantages of being quantitative and requiring no counterpart that was produced with the modified wavelet algorithm.

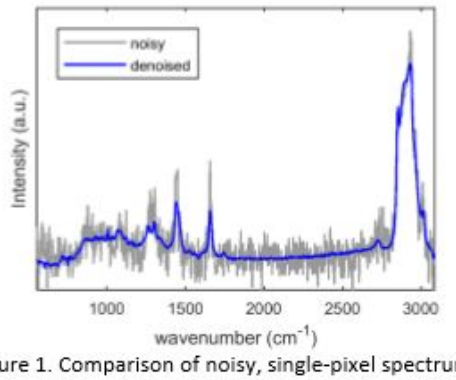

Figure 1. Comparison of noisy, single-pixel spectrum (gray) to its denoised counterpart (blue).

４．研究成果

Despite these advances in signal processing for Raman microscopy, spatial partitioning of the spectra into their cellular components remains a difficult task due to the chemical variation throughout the cell as well as within a particular organelle. Although various methods of unsupervised clustering, such as k-means clustering, agglomerative and divisive hierarchical clustering, and rate-distortion theory have been applied, positive results remain elusive. Common results include partitions of similar chemical environments from different organelles, such as nuclear membranes with mitochondrial membranes. Future efforts will need to include spatial as well as spectral relationships, perhaps venturing into the realm of supervised classification.

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔学会発表〕（計 4 件）

1. J. Nicholas Taylor, Katsumasa Fujita, Tamiki Komatsuzaki: “Data Driven Approaches to Raman Microscopic Analysis”, Biophysical Society of Japan Annual Meeting, 2016年11月25日～27日, Tsukuba International Congress Center(茨城県つくば市)
2. J. Nicholas Taylor, Katsumasa Fujita, Tamiki Komatsuzaki: “Data Driven Approaches to Raman Microscopic Analysis”, Hiroshima Workshop for Theory and Experiment, 2016年10月6日～9日, Hiroshima University(広島県東広島市)
3. J. Nicholas Taylor: “Data-driven quantification of heterogeneous microenvironments in live-cell Raman microscopic images”, 2016 Annual Meeting of Biophysical Society, 2016年2月26日～3月3日, Los Angeles, U.S.A.
4. J. Nicholas Taylor: “Data-driven quantification of heterogeneous microenvironments in live-cell Raman microscopic images”, Pacifichem, 2015年12月16日, Hawaii, U.S.A.

６．研究組織

(1)研究代表者

Ｔａｙｌｏｒ　Ｎｉｃｈｏｌａｓ

（Taylor Nicholas）

北海道大学電子科学研究所・特任助教

研究者番号：50750824